(19) Deutsches Patent- und Markenamt

(10) **DE 11 2020 005 696 T5** 2022.09.15

(12) # Veröffentlichung

der internationalen Anmeldung mit der
(87) Veröffentlichungs-Nr.: **WO 2021/101907**
in der deutschen Übersetzung (Art. III § 8 Abs. 2 IntPatÜbkG)
(21) Deutsches Aktenzeichen: **11 2020 005 696.1**
(86) PCT-Aktenzeichen: **PCT/US2020/060917**
(86) PCT-Anmeldetag: **17.11.2020**
(87) PCT-Veröffentlichungstag: **27.05.2021**
(43) Veröffentlichungstag der PCT Anmeldung in deutscher Übersetzung: **15.09.2022**

(51) Int Cl.: ***G06T 7/70*** (2017.01)

(30) Unionspriorität:
16/690,015 **20.11.2019** **US**

(71) Anmelder:
NVIDIA Corporation, Santa Clara, CA, US

(74) Vertreter:
Kraus & Weisert Patentanwälte PartGmbB, 80539 München, DE

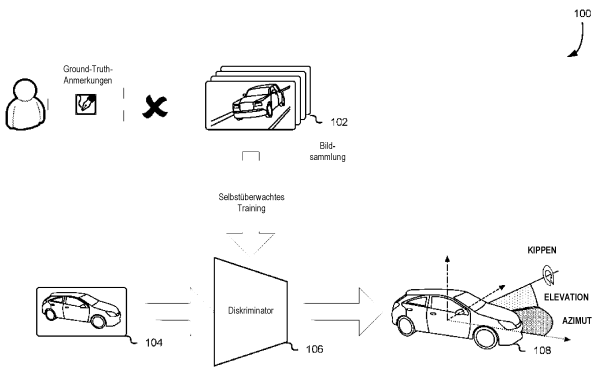
(72) Erfinder:
Mustikovela, Siva Karthik, Santa Clara, CA, US; Jampani, Varun, Santa Clara, CA, US; De Mello, Shalini, Santa Clara, CA, US; Liu, Sifei, Santa Clara, CA, US; Iqbal, Umar, Santa Clara, CA, US; Kautz, Jan, Santa Clara, CA, US

Prüfungsantrag gemäß § 44 PatG ist gestellt.

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen.

(54) Bezeichnung: **TRAINING UND ABLEITEN UNTER VERWENDUNG EINES NEURONALEN NETZWERKES ZUM VORHERSAGEN DER AUSRICHTUNG VON OBJEKTEN IN BILDERN**

(57) Zusammenfassung: Einrichtungen, Systeme und Techniken zum Identifizieren von Objektorientierungen in Bildern. Bei mindestens einer Ausführungsform werden eines oder mehrere neuronale Netzwerke dazu trainiert, Ausrichtungen eines oder mehrerer Objekte mindestens zum Teil basierend auf einer oder mehreren anderen Charakteristiken des Objekts als die Ausrichtung des Objekts zu identifizieren.



Beschreibung

QUERVERWEIS AUF VERWANDTE ANMELDUNG

[0001] Diese Anmeldung beansprucht die Priorität der US-Patentanmeldung Nr. 16/690 015, eingereicht am 20. November 2019, mit dem Titel „TRAINING AND INFERENCING USING A NEURAL NETWORK TO PREDICT ORIENTATIONS OF OBJECTS IN IMAGES“, deren gesamter Inhalt hierin durch Bezugnahme vollständig und für alle Zwecke aufgenommen wird.

TECHNISCHES GEBIET

[0002] Mindestens eine Ausführungsform betrifft das Verarbeiten von Ressourcen, die dazu verwendet werden, ein neuronales Netzwerk zu trainieren, um Sichtpunkte von Objekten in Bildern vorherzusagen. Beispielsweise betrifft mindestens eine Ausführungsform Prozessoren oder Computersysteme, die dazu verwendet werden, neuronale Netzwerke gemäß verschiedenen hierin beschriebenen neuartigen Techniken zu trainieren.

HINTERGRUND

[0003] Das Trainieren neuronaler Netzwerke kann erhebliche Speicher-, Zeit- oder Rechenressourcen beanspruchen. Das Trainieren neuronaler Netzwerke, die Ground-Truth-Anmerkungen erfordern, kann eine größere Herausforderung sein als das Trainieren neuronaler Netzwerke, die nicht erfordern, dass einige oder alle Trainingsdaten mit Ground-Truth-Anmerkungen versehen werden, mindestens weil Ground-Truth-Anmerkungen möglicherweise nicht immer verfügbar und/oder schwierig zu erhalten sind. Speicher-, Zeit- und/oder Rechenressourcen, die zum Trainieren neuronaler Netzwerke verwendet werden, können verbessert werden.

Figurenliste

Fig. 1 veranschaulicht ein Diagramm, das das Vorhersagen eines Sichtpunkts eines Objekts unter Verwendung eines auf selbstüberwachte Weise trainierten neuronalen Netzwerkes gemäß mindestens einer Ausführungsform abbildet;

Fig. 2 veranschaulicht ein Diagramm, das Verlustfunktionen abbildet, gemäß mindestens einer Ausführungsform;

Fig. 3 veranschaulicht ein Diagramm, das ein generatives kontradiktorisches Netzwerk abbildet, gemäß mindestens einer Ausführungsform;

Fig. 4 veranschaulicht ein Diagramm, das eine Diskriminatoraktualisierung abbildet, gemäß mindestens einer Ausführungsform;

Fig. 5 veranschaulicht ein Diagramm, das eine Diskriminatoraktualisierung abbildet, gemäß mindestens einer Ausführungsform;

Fig. 6 veranschaulicht ein Diagramm, das eine Generatoraktualisierung abbildet, gemäß mindestens einer Ausführungsform;

Fig. 7 veranschaulicht ein Diagramm, das eine Generatoraktualisierung abbildet, gemäß mindestens einer Ausführungsform;

Fig. 8 veranschaulicht ein Diagramm, das einen Symmetrieverlust abbildet, gemäß mindestens einer Ausführungsform;

Fig. 9 veranschaulicht ein Diagramm, das den Verlust des nächsten Nachbarn und des am weitesten entfernten Nachbarn abbildet, gemäß mindestens einer Ausführungsform;

Fig. 10 veranschaulicht ein Diagramm, das den Entflechtungsverlust abbildet, gemäß mindestens einer Ausführungsform;

Fig. 11 veranschaulicht ein Diagramm, das das Kalibrieren eines neuronalen Netzwerks abbildet, gemäß mindestens einer Ausführungsform;

Fig. 12 veranschaulicht ein Diagramm, das eine Inferenz abbildet, gemäß mindestens einer Ausführungsform;

Fig. 13 zeigt ein veranschaulichendes Beispiel eines Prozesses zum Trainieren eines neuronalen Netzwerks zum Vorhersagen eines Sichtpunkts eines Objekts innerhalb eines Bildes gemäß mindestens einer Ausführungsform;

Fig. 14 zeigt ein veranschaulichendes Beispiel eines Prozesses zum Trainieren eines neuronalen Netzwerks zum Vorhersagen eines Sichtpunkts eines Objekts innerhalb eines Bildes gemäß mindestens einer Ausführungsform;

Fig. 15A zeigt ein veranschaulichendes Beispiel eines Prozesses zum Berechnen des generativen Konsistenzverlusts gemäß mindestens einer Ausführungsform;

Fig. 15B zeigt ein veranschaulichendes Beispiel eines Prozesses zum Berechnen des Sichtpunktkonsistenzverlusts gemäß mindestens einer Ausführungsform;

Fig. 16A zeigt ein veranschaulichendes Beispiel eines Prozesses zum Berechnen eines Symmetrieverlusts gemäß mindestens einer Ausführungsform;

Fig. 16B zeigt ein veranschaulichendes Beispiel eines Prozesses zum Berechnen eines Symmetrieverlusts gemäß mindestens einer Ausführungsform;

Fig. 17 zeigt ein veranschaulichendes Beispiel eines Prozesses zum Berechnen des Verlustes des nächsten Nachbarn und des am weitesten entfernten Nachbarn gemäß mindestens einer Ausführungsform;

Fig. 18A veranschaulicht Ableitungs- und/oder Trainingslogik gemäß mindestens einer Ausführungsform;

Fig. 18B veranschaulicht Ableitungs- und/oder Trainingslogik gemäß mindestens einer Ausführungsform;

Fig. 19 veranschaulicht Training und Einsatz eines neuronalen Netzwerks gemäß mindestens einer Ausführungsform;

Fig. 20 veranschaulicht ein Beispiel für ein Rechenzentrumssystem gemäß mindestens einer Ausführungsform;

Fig. 21A veranschaulicht ein Beispiel eines autonomen Fahrzeugs gemäß mindestens einer Ausführungsform;

Fig. 21B veranschaulicht ein Beispiel von Kameralagen und -sichtfeldern für das autonome Fahrzeug der **Fig. 21A** gemäß mindestens einer Ausführungsform;

Fig. 21C ist ein Blockdiagramm, das eine beispielhafte Systemarchitektur für das autonome Fahrzeug der **Fig. 21A** gemäß mindestens einer Ausführungsform veranschaulicht;

Fig. 21D ist ein Diagramm, das ein System zur Kommunikation zwischen einem oder mehreren cloudbasierten Servern und dem autonomen Fahrzeug der **Fig. 21A** gemäß mindestens einer Ausführungsform veranschaulicht;

Fig. 22 ist ein Blockdiagramm, das ein Computersystem gemäß mindestens einer Ausführungsform darstellt;

Fig. 23 ist ein Blockdiagramm, das ein Computersystem gemäß mindestens einer Ausführungsform darstellt;

Fig. 24 veranschaulicht ein Computersystem gemäß mindestens einer Ausführungsform;

Fig. 25 veranschaulicht ein Computersystem gemäß mindestens einer Ausführungsform;

Fig. 26A veranschaulicht ein Computersystem gemäß mindestens einer Ausführungsform;

Fig. 26B veranschaulicht ein Computersystem gemäß mindestens einer Ausführungsform;

Fig. 26C veranschaulicht ein Computersystem gemäß mindestens einer Ausführungsform;

Fig. 26D veranschaulicht ein Computersystem gemäß mindestens einer Ausführungsform;

Fig. 26E und **Fig. 26F** veranschaulichen ein gemeinsames Programmiermodell gemäß mindestens einer Ausführungsform;

Fig. 27 veranschaulicht beispielhafte integrierte Schaltungen und zugehörige Grafikprozessoren gemäß mindestens einer Ausführungsform;

die **Fig. 28A** und **Fig. 28B** veranschaulichen beispielhafte integrierte Schaltungen und assoziierte Grafikprozessoren gemäß mindestens einer Ausführungsform;

die **Fig. 29A** und **Fig. 29B** veranschaulichen zusätzliche beispielhafte Grafikprozessorlogik gemäß mindestens einer Ausführungsform;

Fig. 30 veranschaulicht ein Computersystem gemäß mindestens einer Ausführungsform;

Fig. 31A veranschaulicht einen Parallelprozessor gemäß mindestens einer Ausführungsform;

Fig. 31B veranschaulicht eine Partitionseinheit gemäß mindestens einer Ausführungsform;

Fig. 31C veranschaulicht einen Verarbeitungscluster gemäß mindestens einer Ausführungsform;

Fig. 31D veranschaulicht einen Grafik-Multiprozessor gemäß mindestens einer Ausführungsform;

Fig. 32 veranschaulicht ein Multi-Grafikprozessoreinheits- (GPU-) System gemäß mindestens einer Ausführungsform;

Fig. 33 veranschaulicht einen Grafikprozessor gemäß mindestens einer Ausführungsform;

Fig. 34 ist ein Blockdiagramm, das eine Prozessor-Mikroarchitektur für einen Prozessor gemäß mindestens einer Ausführungsform illustriert;

Fig. 35 veranschaulicht einen Deep-Learning-Anwendungsprozessor gemäß mindestens einer Ausführungsform;

Fig. 36 ist ein Blockdiagramm, das einen beispielhaften neuromorphen Prozessor gemäß mindestens einer Ausführungsform veranschaulicht;

Fig. 37 veranschaulicht mindestens Abschnitte eines Grafikprozessors gemäß einer oder mehrerer Ausführungsformen;

Fig. 38 veranschaulicht mindestens Abschnitte eines Grafikprozessors gemäß einer oder mehrerer Ausführungsformen;

Fig. 39 veranschaulicht mindestens Abschnitte eines Grafikprozessors gemäß einer oder mehrerer Ausführungsformen;

Fig. 40 ist ein Blockdiagramm einer Grafikverarbeitungsmaschine 4010 eines Grafikprozessors gemäß mindestens einer Ausführungsform.

Fig. 41 ist ein Blockdiagramm von mindestens Abschnitten eines Grafikprozessorkerns gemäß mindestens einer Ausführungsform;

die **Fig. 42A** und **Fig. 42B** veranschaulichen eine Thread-Ausführungslogik 4200, die eine Anordnung von Verarbeitungselementen eines Grafikprozessorkerns beinhaltet, gemäß mindestens einer Ausführungsform

Fig. 43 veranschaulicht eine Parallelverarbeitungseinheit („PPU“) gemäß mindestens einer Ausführungsform;

Fig. 44 veranschaulicht einen allgemeinen Verarbeitungscluster („GPC“) gemäß mindestens einer Ausführungsform;

Fig. 45 veranschaulicht eine Speicherpartitionseinheit einer Parallelverarbeitungseinheit („PPU“) gemäß mindestens einer Ausführungsform; und

Fig. 46 veranschaulicht einen Streaming-Multiprozessor gemäß mindestens einer Ausführungsform.

AUSFÜHRLICHE BESCHREIBUNG

[0004] Bei mindestens einer Ausführungsform wird ein neuronales Netzwerk dazu trainiert, eine Ausrichtung eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise mit einer Sammlung von Bildern, wie sie an anderer Stelle in dieser Offenbarung beschrieben sind, zu identifizieren. Bei mindestens einer Ausführungsform wird ein neuronales Netzwerk dazu trainiert, eine Ausrichtung eines Objekts in einem Bild auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes (z. B. Sammlung von Bildern) bewerten, als Teil des Trainings berechnet werden. Bei mindestens einer Ausführungsform wird ein neuronales Netzwerk mit einer Sammlung von Bildern trainiert, der Ground-Truth-Anmerkungen fehlen oder wobei Ground-Truth-

Anmerkungen anderswie nicht verfügbar sind (z. B. werden derartige Daten einem neuronalen Netzwerk während des Trainings vorenthalten). Bei mindestens einer Ausführungsform wird ein neuronales Netzwerk dazu trainiert, aus einem Objekt innerhalb eines ersten Bildes, das eine vorhergesagte Ausrichtung aufweist, ein zweites Bild, das dieselbe Ausrichtung aufweist, zu erzeugen. Bei mindestens einer Ausführungsform wird eine vorhergesagte Ausrichtung oder ein vorhergesagter Sichtpunkt als Azimut-, Elevations- und Kippparameter codiert.

[0005] Bei mindestens einer Ausführungsform werden ein oder mehrere neuronale Netzwerke selbstüberwacht mit einer Sammlung von Bildern unterschiedlicher Objekte derselben Kategorie als ein Objekt eines Bildes, das abgeleitet werden soll, trainiert. Bei mindestens einer Ausführungsform können unterschiedliche Objekte derselben Kategorie unterschiedliche Bilder betreffen, die ein oder mehrere Bilder eines ersten Autos in einer oder mehreren Ausrichtungen, ein oder mehrere Bilder eines unterschiedlichen zweiten Autos in einer oder mehreren Ausrichtungen sein können, und so fort. Bei mindestens einer Ausführungsform ist ein Bild eines Objekts, das abgeleitet werden soll, in einer Sammlung von Bildern enthalten, die verwendet wird, um ein oder mehrere neuronale Netzwerke auf Ableitungsausrichtungen zu trainieren. Bei mindestens einer Ausführungsform werden ein oder mehrere neuronale Netzwerke auf selbstüberwachte Weise trainiert, indem mindestens ein Satz von Verlustfunktionen verwendet wird, um eine oder mehrere Charakteristiken von Objekten in Bildern zu bewerten. Bei mindestens einer Ausführungsform betreffen eine oder mehrere Charakteristiken von Objekten Eigenschaften von Objekten, die verwendet werden können, um Ausrichtungen abzuleiten. Bei mindestens einer Ausführungsform wird ein neuronales Netzwerk auf selbstüberwachte Weise trainiert, um synthetische Bilder von Objekten mit einer spezifischen Ausrichtung, die dieselbe Ausrichtung wie eine vorhergesagte Ausrichtung eines Eingabebildes sein kann, anzulegen. Bei mindestens einer Ausführungsform wird ein synthetisches Bild unter Verwendung eines tiefen generativen Modells, wie eines Variational Autoencoders (VAE), eines differenzierbaren Renderers oder eines generativen kontradiktorischen Netzwerks (GAN) oder über einen Renderer anlegt. Bei mindestens einer Ausführungsform kann ein Objekt, dessen Ausrichtung abgeleitet werden soll, ein Fahrzeug, ein Flugzeug, eine Drohne, ein Mensch, ein Gesicht (z. B. eines Menschen oder eines Tieres) und mehr sein.

[0006] Bei mindestens einer Ausführungsform betrifft selbstüberwachtes Lernen (z. B. Training) eine Form des Lernens, bei der ein neuronales Netzwerk auf einem Trainingssatz trainiert wird, bei dem Daten des Trainingssatzes keine Ground-Truth-Anmerkungen umfassen, sondern dass Daten des Trainingssatzes zum Teil gekennzeichnet sind (z. B. halbüberwachtes Lernen). Bei mindestens einer Ausführungsform verwendet ein auf selbstüberwachte Weise trainiertes neuronales Netzwerk zum Identifizieren einer Ausrichtung eines Objekts innerhalb eines Bildes einen Trainingssatz von Bildern zum Trainieren, wobei Bilder des Trainingssatzes keine Ground-Truth-Anmerkungen umfassen, die Ausrichtungen von Objekten innerhalb der Bilder angeben, sondern Labels oder anderswie andere Informationen umfassen, die verschiedene Objekte der Bilder identifizieren (z. B. umfasst ein Bild der Bilder Labels oder anderswie andere Informationen, die Objekte des Bildes identifizieren, umfasst aber keine Anmerkungen, die Ausrichtungen der Objekte angeben).

[0007] Bei mindestens einer Ausführungsform betrifft halbüberwachtes Lernen eine Form des Lernens, bei der ein neuronales Netzwerk auf einem Trainingssatz trainiert wird, in dem nur ein Teil der Daten des Trainingssatzes Ground-Truth-Anmerkungen umfasst. Bei mindestens einer Ausführungsform betrifft vollständig überwachtes Lernen eine Form des Lernens, bei der ein neuronales Netzwerk auf einem Trainingssatz trainiert wird, in dem alle Daten des Trainingssatzes Ground-Truth-Anmerkungen umfassen. Bei mindestens einer Ausführungsform betrifft unüberwachtes Lernen eine Form des Lernens, bei der ein neuronales Netzwerk auf einem Trainingssatz trainiert wird, in dem keine Daten des Trainingssatzes Ground-Truth-Anmerkungen umfassen.

[0008] Fig. 1 veranschaulicht ein Diagramm 100, das das Vorhersagen eines Sichtpunkts eines Objekts unter Verwendung eines neuronalen Netzwerks, das auf selbstüberwachte Weise trainiert wird, gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform wird das Diagramm 100 durch ein oder mehrere Systeme umgesetzt, wie z. B. ein in den **Fig. 18** - **Fig. 46** beschriebenes System. Bei mindestens einer Ausführungsform beinhaltet das Diagramm 100 ein oder mehrere neuronale Netzwerke, die mit einem Diskriminator 106 assoziiert sind, der unter Verwendung von selbstüberwachtem Lernen mit einer Sammlung von Bildern einer Kategorie trainiert wird, um Sichtpunkte von Objekten innerhalb anderer Bilder dieser Kategorie abzuleiten. Bei mindestens einer Ausführungsform wird ein Bild als eine Eingabe für ein neuronales Netzwerk bereitgestellt, um eine Ausrichtung eines Objekts einer Kategorie zu erfassen. Bei mindestens einer Ausführungsform wird ein Eingabebild mehreren neuronalen Netzwerken bereitgestellt, die unter Verwendung von hierin beschriebenen selbstüberwachten Lerntechniken trainiert sind, um Ausrichtungen oder Sichtpunkte unterschiedlicher Objekte in dem Eingabebild zu identifizieren.

[0009] Bei mindestens einer Ausführungsform betrifft ein Sichtpunkt eines Bildes eine Ausrichtung eines Objekts innerhalb eines Bildes, das eine dreidimensionale Ausrichtung des Objekts, das innerhalb eines zweidimensionalen Bild erfasst wurde, betrifft. Bei mindestens einer Ausführungsform wird eine Kamera verwendet, um ein zweidimensionales Bild eines Objekts der realen Welt, wie eines Autos, das sich in einer spezifischen Ausrichtung in Bezug auf die Kamera befindet, aufzunehmen. Bei mindestens einer Ausführungsform wird die Ausrichtung eines Objekts (z. B. Sichtpunkt) in einem Satz von Parametern codiert, der einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfasst. Bei mindestens einer Ausführungsform wird eine Ausrichtung eines Objekts innerhalb eines Bildes als ein Satz zu drei Vektoren codiert, die eine Richtung des Objekts in Bezug auf eine kanonische x-, y- und z-Achse definieren.

[0010] Bei mindestens einer Ausführungsform wird eine Bildsammlung 102 erhalten. Bei mindestens einer Ausführungsform ist die Bildsammlung 102 eine Sammlung von einem oder mehreren Bildern eines Objekttyps. Bei mindestens einer Ausführungsform wird die Bildsammlung 102 verwendet, um ein oder mehrere neuronale Netzwerke zu trainieren, um Ausrichtungen von Objekten innerhalb von Bildern zu identifizieren. Bei mindestens einer Ausführungsform ist die Bildsammlung 102 derart kategorisiert oder gekennzeichnet, dass jede denselben Typ oder dieselbe Kategorie von Objekten anzeigt. Bei mindestens einer Ausführungsform ist die Bildsammlung 102 eine Sammlung von Bildern von Autos, die unterschiedliche Arten von Autos in unterschiedlichen Ausrichtungen, bei unterschiedlichem Wetter, unter unterschiedlicher Beleuchtung und so fort beinhalten können. Bei mindestens einer Ausführungsform beinhaltet die Bildsammlung 102 Bilder des gleichen Autos oder des gleichen Autotyps in unterschiedlichen Ausrichtungen. Bei mindestens einer Ausführungsform fehlen mindestens einem Teil der Bildsammlung 102 Ground-Truth-Anmerkungen, die die Ausrichtung von Objekten innerhalb derartiger Trainingsbilder spezifizieren. Bei mindestens einer Ausführungsform fehlen allen Bildern der Bildsammlung 102 Ground-Truth-Anmerkungen, die Azimut, Elevation und Kippen von Objekten in Bildern innerhalb der Sammlung spezifizieren. Bei mindestens einer Ausführungsform betrifft eine Ground-Truth-Anmerkung für ein oder mehrere neuronale Netzwerke, die dazu konfiguriert sind, eine oder mehrere Charakteristiken eines Bildes zu bestimmen, in dem das eine oder die mehreren neuronalen Netzwerke mit einem Trainingssatz von Bildern trainiert werden, eine Anmerkung, die ein Bild des Trainingssatzes von Bildern umfassen kann, das eine oder mehrere erwartete Charakteristiken des Bildes angibt. Bei mindestens einer Ausführungsform beinhaltet die Bildsammlung 102 ein oder mehrere synthetische Bilder, wie beispielsweise ein Bild, das von einem Variational Autoencoder (VAE), Generative Adversarial Network (GAN) oder einem Renderer angelegt wird. Bei mindestens einer Ausführungsform sind alle Bilder der Bildsammlung 102 echte Bilder im Gegensatz zu jenen, die aus einem generativen Modell, wie einem Variational Autoencoder (VAE), einem Renderer oder einem generativen kontradiktorischen Netzwerk synthetisiert oder erstellt werden. Bei mindestens einer Ausführungsform wird die Bildsammlung 102 von einer Website, die Bilder nach Kategorie sortiert, gesammelt und aggregiert.

[0011] Bei mindestens einer Ausführungsform ist der Diskriminator 106 dazu trainiert, eine Ausrichtung eines Objekts innerhalb eines Bildes 104 mindestens zum Teil basierend auf einer oder mehreren anderen Charakteristiken des Objekts als der Ausrichtung des Objekts zu identifizieren. Bei mindestens einer Ausführungsform ist der Diskriminator 106 ein Klassifikator innerhalb eines oder mehrerer neuronaler Netzwerke. Bei mindestens einer Ausführungsform ist der Diskriminator 106 eine Komponente eines oder mehrerer neuronaler Netzwerke und umfasst andere neuronale Netzwerke, Klassifikatoren und verschiedene andere Maschinenlernkomponenten. Bei mindestens einer Ausführungsform ist der Diskriminator 106 ein diskriminatives Netzwerk eines generativen kontradiktorischen Netzwerks. Bei mindestens einer Ausführungsform ist der Diskriminator 106 Teil eines oder mehrerer neuronaler Netzwerke und ist dazu trainiert, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird der Diskriminator 106 mit einer Sammlung von Bildern einer Kategorie (z. B. Autos) trainiert, um Ausrichtungen anderer Objekte derselben Kategorie, die in anderen Bildern aufgenommen wurden, abzuleiten. Bei mindestens einer Ausführungsform wird der Diskriminator 106 auf eine selbstüberwachte Weise mit der Bildsammlung 102 trainiert. Bei mindestens einer Ausführungsform wird der Diskriminator 106 dazu trainiert, eine Ausrichtung eines Objekts innerhalb des Bildes 104 auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes (z. B. Bildsammlung 102) bewerten, als Teil des Trainings berechnet werden. Bei mindestens einer Ausführungsform wird ein neuronales Netzwerk, das mit dem Diskriminator 106 assoziiert wird, mindestens zum Teil basierend auf dem Berechnen eines generativen Konsistenzverlusts, eines Symmetrieverlusts, eines Verlusts eines nächsten und eines am weitesten entfernten Nachbarn und eines Entflechtungsverlusts trainiert. Bei mindestens einer Ausführungsform können neuronale Netzwerke zum Identifizieren von Ausrichtungen von Objekten gemäß Techniken, die in Verbindung mit den **Fig. 2-10** beschrieben sind, trainiert werden. Bei mindestens einer Ausführungsform wird der Diskriminator 106 mit einer Sammlung

von Bildern trainiert, denen Ground-Truth-Anmerkungen fehlen oder wobei die Ground-Truth-Anmerkungen anderswie nicht verfügbar sind (z. B. werden derartige Daten während des Trainings vorenthalten).

[0012] Bei mindestens einer Ausführungsform wird das Bild 104 für den Diskriminator 106 erhalten. Bei mindestens einer Ausführungsform ist ein Objekt innerhalb des Bildes 104 von dem gleichen Typ wie Objekte innerhalb der Bilder der Bildsammlung 102, die dazu verwendet werden, ein oder mehrere neuronale Netzwerke zu trainieren. Bei mindestens einer Ausführungsform wird das Bild 104 einem neuronalen Netzwerk zur Schlussfolgerung bereitgestellt, um eine Ausrichtung vorherzusagen. Bei mindestens einer Ausführungsform trainiert ein erstes System ein oder mehrere neuronale Netzwerke, und ein zweites anderes System verwendet dieses eine oder diese mehreren neuronalen Netzwerke, um eine Ableitung auszuführen, um Ausrichtungen von Objekten innerhalb von Bildern zu identifizieren. Bei mindestens einer Ausführungsform wird der Diskriminator 106 auf selbstüberwachte Weise mit der Bildsammlung 102 von Objekten einer spezifischen Kategorie trainiert, um Ausrichtungen anderer Objekte dieser Kategorie (z. B. Objekt innerhalb des Bildes 104) abzuleiten. Bei mindestens einer Ausführungsform werden ein oder mehrere neuronale Netzwerke, die mit dem Diskriminator 106 assoziiert sind, mit einer Sammlung von Bildern von Autos trainiert und verwendet, um Ausrichtungen von Autos abzuleiten, die in Echtzeit von einer Kamera oder einer anderen Video-/Bildaufnahmevorrichtung, die an einem Fahrzeug angebracht ist, aufgenommen werden. Bei mindestens einer Ausführungsform wird der Diskriminator 106 auf selbstüberwachte Weise mit der Bildsammlung 102 trainiert, um eine Ausrichtung 108 eines Objekts, das in dem Bild 104 abgebildet ist, zu bestimmen. Bei mindestens einer Ausführungsform bestimmt der Diskriminator 106 die Ausrichtung 108 eines in dem Bild 104 abgebildeten Autos.

[0013] **Fig. 2** veranschaulicht ein Diagramm 200, das Verlustfunktionen abbildet, gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform wird das Diagramm 200 durch ein oder mehrere Systeme umgesetzt, wie z. B. ein in den **Fig. 18** - **Fig. 46** beschriebenes System. Bei mindestens einer Ausführungsform ist ein Diskriminator 204 mit einem oder mehreren neuronalen Netzwerken assoziiert und wird unter Verwendung von mindestens einem von einem generativen Konsistenzverlust 208 eines echten Bildes, einem Verlust 210 des nächsten und am weitesten entfernten Nachbarn, einem Symmetrieverlust 212 und einem Real Echt-/Falsch-Klassifikationsverlust 214 trainiert.Bei mindestens einer Ausführungsform ist der Diskriminator 204 Teil eines oder mehrerer neuronaler Netzwerke, die dazu trainiert sind, Sichtpunkte aus einem Eingabebild abzuleiten, und das eine oder die mehreren neuronalen Netzwerke umfassen verschiedene Parameter, die mit einem oder mehreren Prozessen des einen oder der mehreren neuronalen Netzwerke assoziiert sind, und mindestens zum Teil basierend auf dem Verlust 208 der generativen Konsistenz des echten Bildes, dem Verlust 210 des nächsten und am weitesten entfernten Nachbarn und dem Symmetrieverlust 212 aktualisiert werden.

[0014] Bei mindestens einer Ausführungsform wird eine Objektbildsammlung 202 eines Objekttyps für den Diskriminator 204 erhalten. Bei mindestens einer Ausführungsform umfasst die Objektbildsammlung 202 Bilder, die alle denselben Objekttyp beinhalten. Bei mindestens einer Ausführungsform umfasst die Objektbildsammlung 202 Bilder, die Autos in unterschiedlichen Ausrichtungen, bei unterschiedlichem Wetter, unter unterschiedlicher Beleuchtung und so weiter beinhalten. Bei mindestens einer Ausführungsform erhält ein System eine Objektbildsammlung 202 gemäß Techniken, die an anderer Stelle in dieser Offenbarung, wie in **Fig. 13**, beschrieben sind.

[0015] Bei mindestens einer Ausführungsform wird ein Bild der Objektbildsammlung 202 als Eingabebild für den Diskriminator 204 ausgewählt. Bei mindestens einer Ausführungsform werden Bilder einer Sammlung in irgendeiner geeigneten Weise zum Lernen, die zufällig oder pseudozufällig aus einem Trainingssatz ausgewählt werden können, ausgewählt. Bei mindestens einer Ausführungsform sagt der Diskriminator 204 einen Sichtpunkt 206 eines Eingabebildes vorher. Bei mindestens einer Ausführungsform wird der Sichtpunkt 206 eines Eingabebildes durch den Diskriminator 204 durch einen oder mehrere Prozesse abgeleitet, die ein oder mehrere neuronale Netzwerke involvieren, die einen oder mehrere Eingabeparameter umfassen, die einen oder mehrere Prozesse, die das eine oder die mehreren neuronalen Netzwerke involvieren, vorschreiben. Bei mindestens einer Ausführungsform wird der Sichtpunkt 206 basierend auf Ground-Truth-Anmerkungen bestimmt, die als ein Teil des Trainings für mindestens einen Teil der Objektbildsammlung 202 bereitgestellt werden. Bei mindestens einer Ausführungsform entspricht der Sichtpunkt 206 einer Vorhersage einer Ausrichtung eines Objekts innerhalb eines Bildes, das in den Diskriminator 204 eingegeben wird. Bei mindestens einer Ausführungsform wird die Ausrichtung eines Objekts (z. B. Sichtpunkt) in einem Satz von Parametern codiert, der einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfasst.

[0016] Bei mindestens einer Ausführungsform wird der generative Konsistenzverlust 208 für den Diskriminator 204 berechnet. Bei mindestens einer Ausführungsform wird der generative Konsistenzverlust 208 mindestens zum Teil basierend auf einem Bildkonsistenzverlust beim Vergleichen eines ausgewählten Bildes mit einem Bild berechnet, das von einem tiefen generativen Modell erzeugt wird, und auf einem Sichtpunktkonsistenzverlust beim Vergleichen des eingegebenen Sichtpunkts mit einem generativen Modell und seinem von den Diskriminator vorhergesagten Wert. Bei mindestens einer Ausführungsform wird der generative Konsistenzverlust 208 unter Verwendung von Techniken berechnet, wie denjenigen, die in Verbindung mit den **Fig. 4-7** erörtert sind. Bei mindestens einer Ausführungsform beinhaltet der generative Konsistenzverlust 208 mindestens zwei Komponenten: einen Konsistenzverlust des Sichtpunkts eines synthetischen Bildes und einen Konsistenzverlust eines echten Bildes. Bei mindestens einer Ausführungsform kann ein Sichtpunktkonsistenzverlust als ein Ausrichtungskonsistenzverlust bezeichnet werden. Bei mindestens einer Ausführungsform wird der generative Konsistenzverlust auf echte Bilder (z. B. Bilder aus der Objektbildsammlung 202) angewandt, im Gegensatz zu synthetisierten Bildern, die von einem Generator angelegt werden. Bei mindestens einer Ausführungsform werden der Konsistenzverlust des Sichtpunkts und der Bildkonsistenzverlust verwendet, um den generativen Konsistenzverlust zu bestimmen. Bei mindestens einer Ausführungsform ist der generative Konsistenzverlust eine Kombination aus Sichtpunktkonsistenzverlust und Bildkonsistenzverlust. Bei mindestens einer Ausführungsform wird der generative Konsistenzverlust durch eine folgende symbolische mathematische Gleichung bestimmt:

$$L_{gc} = L_{vc} + L_{ic}$$

wobei L_{gc} dem generativen Konsistenzverlust entspricht, L_{vc} dem Konsistenzverlust des Sichtpunkts entspricht, und L_{ic} einem Bildkonsistenzverlust entspricht.

[0017] Bei mindestens einer Ausführungsform wird der Bildkonsistenzverlust mindestens zum Teil basierend auf einem Bild der Objektbildsammlung 202 berechnet, das in den Diskriminator 204 eingegeben wird, der mindestens zwei Eigenschaften aus dem Eingabebild bestimmt: Sichtpunkt 206 und einen Satz von Erscheinungsbildparametern. Bei mindestens einer Ausführungsform werden der Sichtpunkt 206 und ein Satz von Erscheinungsbildparametern einem Generator bereitgestellt, um ein synthetisiertes Bild zu erzeugen. Bei mindestens einer Ausführungsform empfängt ein Generative Adversarial Network (GAN) den Sichtpunkt 206 und einen Satz von Erscheinungsbildparametern und erzeugt ein synthetisches (z. B. falsches) Bild, das dem Sichtpunkt 206 und dem Satz von Erscheinungsbildparametern entspricht. Bei mindestens einer Ausführungsform werden ein synthetisiertes Bild und ein Eingabebild verglichen, um den Bildkonsistenzverlust zu bestimmen. Bei mindestens einer Ausführungsform wird ein Kosinus-Abstand zwischen einem Eingabebild und einem synthetisierten Bild verglichen, um Merkmalsähnlichkeiten zu bestimmen, wobei eine größere Ähnlichkeit einem geringeren Verlust entspricht. Bei mindestens einer Ausführungsform werden L1-, L2- oder Kosinus-Abstände verwendet, um den Bildkonsistenzverlust zwischen zwei Bildern zu bestimmen.

[0018] Bei mindestens einer Ausführungsform wird ein Sichtpunktkonsistenzverlust mindestens zum Teil basierend auf einem Sichtpunkt (z. B. Sichtpunkt 206) eines Eingabebildes berechnet. Bei mindestens einer Ausführungsform wird ein Generator verwendet, um ein synthetisches Bild aus dem Sichtpunkt 206, der von dem Diskriminator 204 vorhergesagt wird, aus einem Eingabebild anzulegen. Bei mindestens einer Ausführungsform wird ein synthetisches Bild, das vom Sichtpunkt 206 erzeugt wird, dem Diskriminator 204 bereitgestellt, der einen zweiten Sichtpunkt des synthetischen Bildes bestimmt. Bei mindestens einer Ausführungsform wird der Sichtpunkt 206 mit einem zweiten Sichtpunkt eines synthetischen Bildes, das mindestens zum Teil basierend auf dem Sichtpunkt 206 angelegt wird, verglichen. Bei mindestens einer Ausführungsform wird ein Abstand zwischen dem Sichtpunkt 206 und einem zweiten Sichtpunkt eines synthetischen Bildes verwendet, um einen Sichtpunktkonsistenzverlust zu berechnen, wobei nähere Sichtpunkte einem geringeren Verlust entsprechen. Bei mindestens einer Ausführungsform wird der generative Konsistenzverlust gemäß Techniken berechnet, die in Verbindung mit **Fig. 15** beschrieben sind.

[0019] Bei mindestens einer Ausführungsform wird der Echt/Falsch-Klassifizierungsverlust 214 basierend darauf berechnet, ob der Diskriminator 204 in der Lage ist, korrekt vorherzusagen, ob ein Eingabebild in den Diskriminator 204 ein echtes Bild oder ein synthetisiertes Bild ist. Bei mindestens einer Ausführungsform wird der Echt/Falsch-Klassifizierungsverlust 214 basierend darauf berechnet, ob der Diskriminator 204 in der Lage ist, korrekt vorherzusagen, ob Sätze von Eingabebildern echt oder falsch sind, wobei dem Diskriminator 204 entweder echte oder falsche (z. B. synthetische) Bilder bereitgestellt werden können und vorhersagen soll, ob diese Bilder echt oder falsch sind. Als Teil des Trainings des Diskriminators 204 ist Ground-Truth darüber, ob ein Bild, das dem Diskriminator 204 bereitgestellt wird, Bild echt oder falsch ist, als Teil des Trainings verfügbar (z. B. um Verlust zu berechnen).

[0020] Bei mindestens einer Ausführungsform wird der Symmetrieverlust 212 berechnet, indem mindestens ein Eingabebild mit einer transformierten Version dieses Eingabebildes verglichen wird. Bei mindestens einer Ausführungsform wird ein Eingabebild aus der Objektbildsammlung 202 ausgewählt. Bei mindestens einer Ausführungsform wird eine Transformation auf ein Eingabebild angewandt, um ein transformiertes Bild zu erzeugen. Bei mindestens einer Ausführungsform wird ein Eingabebild horizontal gespiegelt, um ein transformiertes Bild zu erzeugen. Bei mindestens einer Ausführungsform wird der Diskriminator 204 verwendet, um den Sichtpunkt 206 eines Eingabebildes und einen zweiten Sichtpunkt eines transformierten Bildes vorherzusagen. Bei mindestens einer Ausführungsform wird der Sichtpunkt 206 für ein Eingabebild vorhergesagt, und ein zweiter Sichtpunkt wird für eine horizontal gespiegelte Version dieses Eingabebildes vorhergesagt. Bei mindestens einer Ausführungsform wird Verlust basierend darauf berechnet, ob bestimmte Eigenschaften zutreffen. Bei mindestens einer Ausführungsform wird eine Transformation oder Umkehrung davon auf einen vorhergesagten Sichtpunkt einer transformierten Version eines Eingabebildes angewandt. Falls bei mindestens einer Ausführungsform ein Eingabebild um Winkel (φ,θ,ψ) gedreht wird, um ein transformiertes Bild zu erzeugen, dann kann ein abgeleiteter Sichtpunkt dieses transformierten Bildes umgekehrt um Winkel $(-\varphi,-\theta,-\psi)$ gedreht werden. Bei mindestens einer Ausführungsform wird Verlust berechnet, indem die Größen von Azimut, Elevation und Kippen des Sichtpunkts 206 eines Eingabebildes mit einem zweiten Sichtpunkt eines transformierten Bildes verglichen werden, wobei sich Nullverlust ergibt, wenn die Größen aller Ausrichtungsparameter gleich sind. Bei mindestens einer Ausführungsform wird der Symmetrieverlust gemäß Techniken berechnet, die an anderer Stelle in dieser Offenbarung beschrieben sind, wie jenen, die in Verbindung mit den **Fig. 8** und **Fig. 16** erläutert sind. Bei mindestens einer Ausführungsform wird Verlust berechnet, indem bestimmt wird, wie genau ein erster Satz von Erscheinungsbildparametern, der von dem Diskriminator 204 für ein Bild vorhergesagt wird, mit einem zweiten Satz von Erscheinungsbildparametern, der von dem Diskriminator 204 für eine transformierte Version dieses Bildes vorhergesagt wird, übereinstimmt.

[0021] Bei mindestens einer Ausführungsform wird Verlust des nächsten Nachbarn und des am weitesten entfernten Nachbarn 210 berechnet, indem mindestens ein Eingabebild der Objektbildsammlung 202 mit seinen nächsten und am weitesten entfernten Nachbarn mindestens zum Teil basierend auf einem Sichtpunktgraph der Objektbildsammlung 202 verglichen wird. Bei mindestens einer Ausführungsform wird Verlust 210 des nächsten Nachbarn und des am weitesten entfernten Nachbarn gemäß Techniken berechnet, die in Verbindung mit den **Fig. 4**, **Fig. 9** und **Fig. 17** beschrieben sind. Bei mindestens einer Ausführungsform wird die Objektbildsammlung 202 verwendet, um einen Sichtpunktgraphen zu erzeugen, wobei Knoten eines derartigen Graphen Bildern entsprechen und Kanten ihren Sichtpunkt-Äquivariant-Abständen (z. B. Kosinus-Abständen) entsprechen. Bei mindestens einer Ausführungsform werden Kosinus-Distanzen basierend auf Merkmalsähnlichkeiten von Bildpaaren unter Verwendung eines neuronalen Faltungsnetzwerks (Convolutional Neural Network, CNN) berechnet. Bei mindestens einer Ausführungsform wird ein Ankerbild aus der Objektbildsammlung 202 ausgewählt. Bei mindestens einer Ausführungsform wird ein Ankerbild aus einem Sichtpunktgraphen lokalisiert und ein nächster Nachbar und ein am weitesten entfernter Nachbar werden basierend auf Kantengewichten ausgewählt. Bei mindestens einer Ausführungsform weist ein nächster Nachbar eine kürzeste Kante, die mit einem Ankerbild verbunden ist, auf. Bei mindestens einer Ausführungsform weist ein am weitesten entfernter Nachbar eine am weitesten entfernte Kante auf, die mit einem Ankerbild verbunden ist. Bei mindestens einer Ausführungsform sagt der Diskriminator 204 einen ersten Sichtpunkt für ein Ankerbild vorher (z. B. den Sichtpunkt 206, der für das Ankerbild vorhergesagt wird), und sagt einen zweiten Sichtpunkt für ein Bild des nächsten Nachbarn vorher (z. B. den Sichtpunkt 206, der für das Bild des nächsten Nachbarn vorhergesagt wird), und Verlust wird derart berechnet, dass ein geringerer Abstand zwischen diesen Sichtpunkten einem geringeren Verlust entspricht. Bei mindestens einer Ausführungsform sagt ein neuronales Netzwerk des Diskriminators 204 einen ersten Sichtpunkt für ein Ankerbild vorher, und sagt einen dritten Sichtpunkt für ein Bild des am weitesten entfernten Nachbarn vorher (z. B. den Sichtpunkt 206 für das von einem Bild des am weitesten entfernten Nachbarn), und der Verlust wird derart berechnet, dass der größere Abstand zwischen diesen Sichtpunkten geringerem Verlust entspricht.

[0022] Bei mindestens einer Ausführungsform werden berechnete Verluste (z. B. Verlust der generativen Konsistenz 208, Verlust des nächsten und am weitesten entfernten Nachbarn 210, Symmetrieverlust 212 und Echt/Falsch-Klassifizierungsverlust 214) eingesetzt, um Parameter eines oder mehrerer neuronaler Netzwerke zu aktualisieren, die mit dem Diskriminator 204, der mit der Objektbildsammlung 202 trainiert wird, assoziiert sind. Bei mindestens einer Ausführungsform beinhaltet ein Systemumsetzungsdiagramm 200 ausführbaren Code zum kontinuierlichen Aktualisieren von Parametern eines oder mehrerer neuronaler Netzwerke, die mit dem Diskriminator 204 assoziiert sind, so dass das eine oder die mehreren neuronalen Netzwerke und der Diskriminator 204 dazu trainiert werden, einen Sichtpunkt und andere Charakteristiken eines Eingabebildes abzuleiten. Bei mindestens einer Ausführungsform wird das Training gemäß einer beliebigen geeigneten Technik ausgeführt und kann das Auswählen und Verwenden verschiedener zusätzlicher Bilder

der Objektbildsammlung 202 beinhalten, um Verluste zu berechnen und Parameter für ein oder mehrere neuronale Netzwerke, die trainiert werden, um Sichtpunkte abzuleiten, zu verfeinern. Bei mindestens einer Ausführungsform wird, sobald das Training abgeschlossen ist, ein trainiertes neuronales Netzwerk für das Ableiten verfügbar gemacht (z. B. ein neuronales Netzwerk oder Parameter davon, das/der an ein anderes System übertragen wird).

[0023] **Fig. 3** veranschaulicht ein Diagramm 300, das ein generatives kontradiktorisches Netzwerk gemäß mindestens einer Ausführungsform abbildet. Bei mindestens einer Ausführungsform wird das Diagramm 300 durch ein oder mehrere Systeme umgesetzt, wie z. B. ein in den **Fig. 18** - **Fig. 46** beschriebenes System. Bei mindestens einer Ausführungsform beinhaltet das Diagramm 300 einen Generator 306, der einen Eingabesichtpunkt 302 und einen Eingabesatz von Erscheinungsbildparametern 304, ein synthetisches Bild 308, verwendet. Bei mindestens einer Ausführungsform veranschaulicht das Diagramm 300 einen Diskriminator 310, der ein Eingabebild 318 verwendet und einen Ausgabesichtpunkt 312, eine Ausgabebestimmung 314 und einen Ausgabesatz von Erscheinungsbildparametern 316 ausgibt. Bei mindestens einer Ausführungsform werden Parameter des Generators 306 und/oder des Diskriminators 310 unter Verwendung von Techniken ausgewählt, die in Verbindung mit den **Fig. 4-7** beschrieben sind.

[0024] Bei mindestens einer Ausführungsform entspricht der Eingabesichtpunkt 302 einer Ausrichtung eines Objekts innerhalb eines Bildes, das eine dreidimensionale Ausrichtung eines in einem zweidimensionalen Bild erfassten Objekts betrifft. Bei mindestens einer Ausführungsform wird die Ausrichtung eines Objekts (z. B. Sichtpunkt) in einem Satz von Parametern codiert, der einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfasst. Bei mindestens einer Ausführungsform entspricht der Eingabesichtpunkt 302 einer spezifischen Ausrichtung eines Objekts und umfasst spezifische Werte für einen Satz von Parametern, der einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfasst. Bei mindestens einer Ausführungsform gibt der Eingabesichtpunkt 302 eine 3D-Drehung eines Objekts an (z. B. kann der Eingabesichtpunkt 302 eine Drehung eines Objekts um eine spezifizierte Anzahl von Graden auf einer spezifizierten Achse und Variationen davon angeben). Bei mindestens einer Ausführungsform sind der Eingabesatz von Erscheinungsbildparametern 304 Parameter, die ein Erscheinungsbild eines Objekts definieren. Bei mindestens einer Ausführungsform beinhaltet ein Objekt ein Fahrzeug, ein Flugzeug, eine Drohne, einen Menschen, ein Gesicht (z. B. eines Menschen oder eines Tieres) und mehr. Bei mindestens einer Ausführungsform entspricht der Eingabesatz von Erscheinungsbildparametern 304 Erscheinungsbildparametern eines Autos, wie Farbe, Größe, Radtyp und verschiedenen anderen Parametern, die das Erscheinungsbild eines Autos definieren.

[0025] Bei mindestens einer Ausführungsform werden der Eingabeansichtspunkt 302 und der Eingabesatz von Erscheinungsbildparametern 304 dem Generator 306 bereitgestellt, um das Bild 308 anzulegen. Bei mindestens einer Ausführungsform sind der Generator 306 und der Diskriminator 310 Teil eines Generative Adversarial Network (GAN). Bei mindestens einer Ausführungsform ist der Generator 306 ein generatives Netzwerk in einem generativen kontradiktorischen Netzwerk. Bei mindestens einer Ausführungsform ist der Generator 306 Teil eines oder mehrerer neuronaler Netzwerke und ist dazu trainiert, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform empfängt der Generator 306 den eingegebenen Sichtpunkt 302 und den eingegebenen Satz von Erscheinungsbildparametern 304 und erzeugt das Bild 308, das ein synthetisches (z. B. falsches) Bild ist, das dem eingegebenen Sichtpunkt 302 und dem eingegebenen Satz von Erscheinungsbildparametern 304 entspricht. Bei mindestens einer Ausführungsform akzeptiert der Generator zwei getrennte (z. B. unabhängige) Parameter, die dazu verwendet werden, das Bild 308 anzulegen - den Eingabesichtpunkt 302, der einen bestimmten Sichtpunkt (z. B. codierte Azimut-, Elevations- und Kippparameter) angibt, wobei das Bild 308 damit erzeugt werden soll, und Erscheinungsbildparameter 304, die Erscheinungsbildeigenschaften des Bildes 308 codieren (z. B. können derartige Eigenschaften für ein Auto Farbe, Marke, Modell, Herstellungsjahr und mehr beinhalten). Bei mindestens einer Ausführungsform erzeugt der Generator 306 ein Bild 308, das ein Objekt umfasst, das gemäß dem Eingabesatz von Erscheinungsbildparametern 304 erzeugt wird, das gemäß dem Eingabesichtpunkt 302 ausgerichtet ist. Bei mindestens einer Ausführungsform ist das Bild 308 ein synthetisches Bild, das ein Auto umfasst, in dem das Erscheinungsbildbild des Autos dem Eingabesatz von Erscheinungsbildbildparametern 304 entspricht und die Ausrichtung des Autos dem Eingabesichtpunkt 302 entspricht.

[0026] Bei mindestens einer Ausführungsform beinhaltet ein Generative Adversarial Network (GAN) einen Diskriminator 310. Bei mindestens einer Ausführungsform akzeptiert der Diskriminator 310 das Eingabebild 318 und erzeugt den Ausgabesichtpunkt 312, die Ausgabebestimmung 314 und den Ausgabesatz von Erscheinungsbildparametern 316. Bei mindestens einer Ausführungsform kann das Eingabebild 318 ein ech-

tes Bild oder ein synthetisches Bild sein. Bei mindestens einer Ausführungsform wird das Eingabebild 318 von einer oder mehreren anderen Quellen abgerufen, wie von einer Bilddatenbank, einer oder mehreren Kameras und/oder Variationen davon. Bei mindestens einer Ausführungsform verarbeitet der Diskriminator 310 das Bild 308. Bei mindestens einer Ausführungsform umfasst der Diskriminator 310 verschiedene neuronale Netzwerke und maschinelle Lernprozesse. Bei mindestens einer Ausführungsform wird der Diskriminator 310 gemäß denjenigen umgesetzt, die an anderer Stelle in dieser Offenbarung beschrieben sind, wie diejenigen, die in Verbindung mit **Fig. 2** erörtert sind. Bei mindestens einer Ausführungsform ist der Diskriminator 310 mit einem oder mehreren neuronalen Netzwerken assoziiert, die dazu trainiert sind, einen Sichtpunkt sowie andere Charakteristiken eines Eingabebildes abzuleiten. Bei mindestens einer Ausführungsform wird der Diskriminator 310 durch verschiedene Prozesse verfeinert, die Berechnungen verschiedener Verlustfunktionen involvieren, die dazu verwendet werden, verschiedene Parameter, die mit dem Diskriminator 310 assoziiert sind, zu aktualisieren.

[0027] Bei mindestens einer Ausführungsform empfängt der Diskriminator 310 das Eingabebild 318 und erzeugt den Ausgabesichtpunkt 312, die Ausgabebestimmung 314 und den Ausgabesatz von Erscheinungsbildparametern 316. Bei mindestens einer Ausführungsform ist der Ausgabesichtpunkt 312 ein vorhergesagter Sichtpunkt des Eingabebildes 318, der von einem oder mehreren Prozessen des Diskriminators 310 erzeugt wird. Bei mindestens einer Ausführungsform ist die Ausgabebestimmung 314 eine Bestimmung, die von einem oder mehreren Prozessen des Diskriminators 310 erzeugt wird, die angibt, ob das Eingabebild 318 ein echtes Bild oder ein synthetisches (z. B. falsches) Bild ist. Bei mindestens einer Ausführungsform ist die Bestimmung 314 eine binäre Ausgabe (z. B. ein ECHT/FALSCH-Indikator dafür, ob der Diskriminator 310 glaubt, dass das Eingabebild 318 ein echtes Bild oder ein synthetisches Bild ist). Bei mindestens einer Ausführungsform ist die Bestimmung 314 ein Zahlenwert zwischen 0 und 1 (einschließlich oder ausschließlich eines oder beider Endpunkte), der einen Konfidenzwert dafür codiert, ob der Diskriminator 310 denkt, dass das eingegebene Bild 318 echt oder falsch ist (z. B. gibt 0,5 an, dass es gleich wahrscheinlich ist, ob ein Bild echt oder falsch ist; 0 eine hohe Wahrscheinlichkeit angibt, dass ein Bild falsch ist). Bei mindestens einer Ausführungsform ist der Ausgabesatz von Erscheinungsbildparametern 316 ein vorhergesagter Satz von Erscheinungsbildparametern des Eingabebildes 318, das von einem oder mehreren Prozessen des Diskriminators 310 erzeugt wird.

[0028] Falls bei mindestens einer Ausführungsform der Diskriminator 310 genau kalibriert ist (z. B. der Diskriminator 310 für einen gewünschten Grad an Genauigkeit oder einen gewünschten Grad an akzeptablem Verlust trainiert ist) und das Bild 308 von dem Generator 306 erzeugt wird, gibt die Ausgangsbestimmung 314 an, dass das Bild 308 falsch ist, und der Ausgangssichtpunkt 312 und der Ausgangssatz von Erscheinungsbildparametern 316 sind jeweils mit dem Eingangssichtpunkt 302 und dem Eingangssatz von Erscheinungsbildparametern 304 identisch. Falls der Diskriminator 310 bei mindestens einer Ausführungsform nicht genau kalibriert ist (z. B. der Diskriminator 310 nicht vollständig für einen gewünschten Grad an Genauigkeit oder einen gewünschten Grad an akzeptablem Verlust trainiert ist) und das Bild 308 von dem Generator 306 erzeugt wird, gibt die Ausgabebestimmung 314 eine falsche Bestimmung an (falls z. B. das Bild 308 synthetisch ist, würde die Ausgabebestimmung 314 angeben, dass das Bild 308 echt ist), und der Ausgabesichtpunkt 312 und der Ausgabesatz von Erscheinungsbildparametern 316 unterscheiden sich jeweils von dem Eingabeansichtspunkt 302 und dem Eingabesatz von Erscheinungsbildparametern 304. Bei mindestens einer Ausführungsform wird ein Vergleich zwischen dem Ausgangssichtpunkt 312 und dem Ausgangssatz von Erscheinungsbildparametern 316 und dem Eingangssichtpunkt 302 und dem Eingangssatz von Erscheinungsbildparametern 304 verwendet, um den Diskriminator 310 und den Generator 306 zu bewerten und weiter zu verarbeiten, zu trainieren und/oder zu kalibrieren.

[0029] **Fig. 4** veranschaulicht ein Diagramm 400, das eine Diskriminatoraktualisierung gemäß mindestens einer Ausführungsform abbildet. Bei mindestens einer Ausführungsform werden Verlustfunktionen berechnet und verwendet, um Parameter des Diskriminators 404 zu aktualisieren, die dazu verwendet werden, verschiedene Ausgaben von einem Eingabebild vorherzusagen. Bei mindestens einer Ausführungsform veranschaulicht **Fig. 4** ein Eingabebild 402; einen Diskriminator 404; einen vorhergesagten Sichtpunkt 406; eine vorhergesagte Bestimmung 408, ob das eingegebene Bild 402 echt oder falsch ist; einen Satz von Erscheinungsbildparametern 410; einen Generator 412; ein erzeugtes Bild 414; Echt/Falsch-Klassifizierungsverlust 416; den Bildkonsistenzverlust 418; Verlust des nächsten und des am weitesten entfernten Nachbarn 420; und Symmetrieverlust 422. Bei mindestens einer Ausführungsform veranschaulicht **Fig. 4** eine Diskriminatoraktualisierung unter Verwendung eines echten Bildes (z. B. eines Bildes, das nicht von einem Generator synthetisiert wurde). Bei mindestens einer Ausführungsform sind Techniken, die in Verbindung mit **Fig. 4** beschrieben sind, die gleiche Ausdehnung aufweisen wie die in Verbindung mit den **Fig. 5-7** beschriebenen, um Generatoren und/oder Diskriminatoren zu trainieren.

[0030] Bei mindestens einer Ausführungsform verarbeitet der Diskriminator 404 das Eingabebild 402. Bei mindestens einer Ausführungsform ist der Diskriminator 404 mit einem oder mehreren neuronalen Netzwerken assoziiert, die dazu trainiert sind, einen Sichtpunkt sowie andere Charakteristiken eines Eingabebildes abzuleiten. Bei mindestens einer Ausführungsform empfängt der Diskriminator 404 das Eingabebild 402 und erzeugt einen vorhergesagten Sichtpunkt 406, eine Bestimmung 408, ob das Eingabebild 402 echt oder falsch ist, und einen Satz von Erscheinungsbildparametern 410. Bei mindestens einer Ausführungsform ist der Sichtpunkt 406 ein vorhergesagter Sichtpunkt des Eingabebildes 402, der von einem oder mehreren Prozessen des Diskriminators 404 bestimmt wird. Bei mindestens einer Ausführungsform entspricht der Sichtpunkt 406 einer vorhergesagten spezifischen Ausrichtung eines Objekts innerhalb eines Bildes und umfasst spezifische Werte für einen Satz von Parametern, der einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfasst. Bei mindestens einer Ausführungsform gibt der Sichtpunkt 406 eine 3D-Drehung eines Objekts an (z. B. kann der Sichtpunkt 406 eine Drehung eines Objekts um eine spezifizierte Anzahl von Graden auf einer spezifizierten Achse und Variationen davon angeben). Bei mindestens einer Ausführungsform umfasst der Sichtpunkt 406 eine vorhergesagte Ausrichtung eines Autos, das in dem Eingabebild 402 abgebildet ist. Bei mindestens einer Ausführungsform ist die Bestimmung 408 eine Bestimmung davon, ob das Eingabebild 402 ein echtes Bild oder ein falsches Bild ist. Bei mindestens einer Ausführungsform betrifft ein falsches Bild ein synthetisiertes Bild, das von einem generativen kontradiktorischen Netzwerk angelegt wurde. Bei mindestens einer Ausführungsform ist die Bestimmung 408 ein Binärwert (z. B. ein ECHT/FALSCH-Wert, der eine Vorhersage darüber angibt, ob das eingegebene Bild 402 echt oder falsch ist). Bei mindestens einer Ausführungsform ist die Bestimmung 408 ein nicht-binärer Wert, der einen Konfidenzgrad angibt, ob das Eingabebild 402 echt oder falsch ist. Bei mindestens einer Ausführungsform ist der Satz von Erscheinungsbildparametern 410 ein vorhergesagter Satz von Erscheinungsbildparametern des Eingabebildes 402, das von einem oder mehreren Prozessen des Diskriminators 404 erzeugt wird. Bei mindestens einer Ausführungsform besteht der Satz von Erscheinungsbildparametern 410 aus vorhergesagten Parametern, die ein Erscheinungsbild eines Bildes definieren, das in dem Eingabebild 402 abgebildet ist. Bei mindestens einer Ausführungsform entspricht der Satz von Erscheinungsbildparametern 410 einem Satz von vorhergesagten Erscheinungsbildparametern eines Autos, das in dem Eingabebild 402 abgebildet ist, wie vorhergesagte Farbe, Größe, vorhergesagter Radtyp und verschiedene andere Parameter, die ein Erscheinungsbild eines abgebildeten Autos in dem Eingabebild 402 definieren.

[0031] Bei mindestens einer Ausführungsform werden der Sichtpunkt 406 und der Satz von Erscheinungsbildparametern 410 einem Generator 412 bereitgestellt, um ein erzeugtes Bild 414 zu erzeugen. Bei mindestens einer Ausführungsform legt der Generator 412 ein synthetisiertes Bild an. Bei mindestens einer Ausführungsform ist der Generator 412 Teil eines Generative Adversarial Network (GAN). Bei mindestens einer Ausführungsform empfängt der Generator 412 den Sichtpunkt 406 und den Satz von Erscheinungsbildparametern 410 und erzeugt das erzeugte Bild 414, das ein synthetisches (z. B. falsches) Bild ist, das dem Sichtpunkt 406 und dem Satz von Erscheinungsbildparametern 410 entspricht. Bei mindestens einer Ausführungsform erzeugt der Generator 412 ein erzeugtes Bild 414, das ein Objekt umfasst, das gemäß dem Satz von Erscheinungsbildparametern 410 erzeugt und gemäß dem Sichtpunkt 406 ausgerichtet ist. Bei mindestens einer Ausführungsform ist das erzeugte Bild 414 ein synthetisches Bild, das ein Auto umfasst, in dem das Erscheinungsbildbild des Autos dem Satz von Erscheinungsbildbildparametern 410 entspricht und die Ausrichtung des Autos dem Sichtpunkt 406 entspricht.

[0032] Bei mindestens einer Ausführungsform wird die Bestimmung 408 verwendet, um einen Klassifizierungsverlust, wie einen Echt/Falsch-Klassifizierungsverlust 416 zu bestimmen. Bei mindestens einer Ausführungsform wird der Echt/Falsch-Klassifizierungsverlust 416 basierend darauf berechnet, ob der Diskriminator 404 in der Lage ist, korrekt vorherzusagen, ob ein in den Diskriminator 404 ein gegebenes Bild ein echtes Bild oder ein synthetisiertes Bild ist. Bei mindestens einer Ausführungsform wird der Echt/Falsch-Klassifizierungsverlust 416 basierend darauf berechnet, ob der Diskriminator 404 in der Lage ist, korrekt vorherzusagen, ob Sätze von Eingabebildern echt oder falsch sind, wobei dem Diskriminator 404 entweder echte oder falsche (z. B. synthetische) Bilder bereitgestellt werden können, und vorhersagen soll, ob diese Bilder echt oder falsch sind. Als Teil des Trainings des Diskriminators 404 ist Ground-Truth darüber, ob ein Bild, das dem Diskriminator 404 bereitgestellt wird, echt oder falsch ist, als Teil des Trainings (um z. B. Verlust zu berechnen).

[0033] Bei mindestens einer Ausführungsform werden das erzeugte Bild 414 und das Eingabebild 402 verglichen, um einen Bildkonsistenzverlust 418 zu bestimmen. Bei mindestens einer Ausführungsform wird ein Kosinus-Abstand zwischen dem Eingabebild 402 und dem erzeugten Bild 414 verglichen, um Merkmalsähnlichkeiten zu bestimmen, wobei eine größere Ähnlichkeit einem geringeren Verlust entspricht. Bei mindestens einer Ausführungsform wird mindestens einer von L1-, L2- oder Kosinus-Abständen verwendet, um den Bild-

konsistenzverlust 418 zwischen dem Eingabebild 402 und dem erzeugten Bild 414 zu bestimmen. Bei mindestens einer Ausführungsform wird der LI-Abstand durch eine folgende symbolische mathematische Gleichung bestimmt:

$$L_1\ Abstand = I_{in} - I_{gen}$$

wobei I_{in} einer Darstellung eines Eingabebildes entspricht, und I_{gen} einer Darstellung eines erzeugten Bildes entspricht.

[0034] Bei mindestens einer Ausführungsform wird der L2-Abstand durch eine folgende symbolische mathematische Gleichung bestimmt:

$$L_2\ Abstand = \sum_{j=1}^{3}\left(I_{in}^{j} - I_{gen}^{j}\right)^2$$

wobei I_{in} einer Darstellung eines Eingabebildes entspricht, und I_{gen} einer Darstellung eines erzeugten Bildes entspricht.

[0035] Bei mindestens einer Ausführungsform wird der Kosinus-Abstand durch eine folgende symbolische mathematische Gleichung bestimmt:

$$Kosinus - Abstand = f_{in} - f_{gen}$$

wobei f_{in} einer Darstellung von Merkmalen eines Eingabebildes entspricht, und f_{gen} einer Darstellung von Merkmalen eines erzeugten Bildes entspricht.

[0036] Bei mindestens einer Ausführungsform werden zusätzliche Verlustfunktionen als Teil von Diskriminatoraktualisierungen, die in Verbindung mit **Fig. 4** beschrieben wurden, berechnet. Bei mindestens einer Ausführungsform wird ein Verlust 420 eines nächsten und eines am weitesten entfernten Nachbarn berechnet. Bei mindestens einer Ausführungsform wird ein Symmetrieverlust 422 berechnet. Bei mindestens einer Ausführungsform werden der Verlust 420 des nächsten und des am weitesten entfernten Nachbarn und/oder der Symmetrieverlust 422 gemäß an anderer Stelle beschriebenen Techniken berechnet, wie in den **Fig. 8-10** erörtert. Bei mindestens einer Ausführungsform werden berechnete Verluste (z. B. die in **Fig. 4** veranschaulichten) verwendet, um Gradienten zu berechnen und Parameter für den Diskriminator 410 zu aktualisieren, während die Parameter des Generators 412 unter Verwendung einer beliebigen geeigneten Technik, wie z. B. Gradientenabstieg, konstant gehalten werden.

[0037] Fig. 5 veranschaulicht ein Diagramm 500, das eine Diskriminatoraktualisierung gemäß mindestens einer Ausführungsform abbildet. Bei mindestens einer Ausführungsform werden Verlustfunktionen berechnet und verwendet, um Parameter des Diskriminators 504 zu aktualisieren, die verwendet werden, um verschiedene Ausgaben von einem Eingabebild vorherzusagen. Bei mindestens einer Ausführungsform veranschaulicht **Fig. 5** einen Sichtpunkt 502; einen Satz von Erscheinungsbildparametern 504; einen Generator 506; ein erzeugtes Bild 508; einen Diskriminator 510; einen vorhergesagten Sichtpunkt 512; eine vorhergesagte Bestimmung 514, ob das erzeugte Bild 508 echt oder falsch ist; einen vorhergesagten Satz von Erscheinungsbildparametern 516; Sichtpunktkonsistenzverlust 518; Z-Rekonstruktionsverlust 520; und Echt/Falsch-Klassifizierungsverlust 522. Bei mindestens einer Ausführungsform veranschaulicht **Fig. 5** eine Diskriminatoraktualisierung unter Verwendung eines synthetisierten Bildes (z. B. eines Bildes, das von einem generativen gegnerischen Netzwerk angelegt wird). Bei mindestens einer Ausführungsform haben Techniken, die in Verbindung mit **Fig. 5** sind, die gleiche Ausdehnung wie die in Verbindung mit den **Fig. 4**, **Fig. 6** und **Fig. 7** beschriebenen, um Generatoren und/oder Diskriminatoren zu trainieren.

[0038] Bei mindestens einer Ausführungsform werden ein Sichtpunkt 502 und ein Satz von Erscheinungsbildparametern 504 auf jede geeignete Weise ausgewählt, was eine zufällige Auswahl von Parameterwerten, eine gewichtete zufällige Auswahl und mehr beinhalten kann. Bei mindestens einer Ausführungsform sind der Sichtpunkt 502 und der Satz von Erscheinungsbildparametern 504 entflochtene Parameter, die unabhängig ausgewählt werden können. Bei mindestens einer Ausführungsform akzeptiert der Generator 506 den Sichtpunkt 502 und den Satz von Erscheinungsbildparametern 504 als Eingaben und legt ein erzeugtes Bild 508 an. Bei mindestens einer Ausführungsform ist das erzeugte Bild 508 ein synthetisches Bild, das basierend auf dem Satz von Erscheinungsbildparametern 504 erscheint und gemäß dem Sichtpunkt 502 ausgerichtet ist.

[0039] Bei mindestens einer Ausführungsform wird ein Satz von Bildern (z. B. das erzeugte Bild 508) einem Diskriminator 510 als eine Eingabe bereitgestellt, und der Diskriminator sagt verschiedene Eigenschaften dieses Satzes von Bildern vorher. Bei mindestens einer Ausführungsform empfängt der Diskriminator 510 das erzeugte Bild 508 und erzeugt einen vorhergesagten Sichtpunkt 512; eine vorhergesagte Bestimmung 514, ob das erzeugte Bild 508 echt oder falsch ist, und einen vorhergesagten Satz von Erscheinungsbildparametern 516. Bei mindestens einer Ausführungsform hat der Diskriminator keinen Zugriff auf den Sichtpunkt 502, und ein Satz von Erscheinungsbildparametern 504 wird verwendet, um das erzeugte Bild 508 anzulegen (z. B. werden derartige Informationen dem Diskriminator 510 während der Vorhersage vorenthalten). Bei mindestens einer Ausführungsform werden Ausgaben des Diskriminators 510 verwendet, um Verlust zu berechnen. Bei mindestens einer Ausführungsform werden Verlustfunktionen verwendet, um Gradienten zu berechnen (z. B. unter Verwendung von Gradientenabstieg) und Parameter für den Diskriminator 510 zu aktualisieren, während Parameter des Generators 506 konstant gehalten werden.

[0040] Bei mindestens einer Ausführungsform wird der Sichtpunktkonsistenzverlust 518 berechnet. Bei mindestens einer Ausführungsform betrifft Sichtpunktkonsistenzverlust eine Verlustfunktion, die basierend darauf berechnet wird, wie genau der Diskriminator 510 Sichtpunkte vorhersagt. Bei mindestens einer Ausführungsform wird der Sichtpunktkonsistenzverlust 518 als eine Differenz oder ein Abstand zwischen dem eingegebenen Sichtpunkt 502 und dem vorhergesagten Sichtpunkt 512 berechnet. Bei mindestens einer Ausführungsform ist der Sichtpunktkonsistenzverlust eine Komponente des generativen Konsistenzverlusts. Bei mindestens einer Ausführungsform wird ein Abstand (z. B. L1-Abstand, L2-Abstand, Kosinus-Abstand und/oder Variationen davon) zwischen dem eingegebenen Sichtpunkt 502 und dem vorhergesagten Sichtpunkt 512 verwendet, um einen Sichtpunktkonsistenzverlust 518 zu berechnen, wobei nähere Sichtpunkte (z. B. Sichtpunkte mit kürzeren Abständen voneinander) einem geringeren Verlust entsprechen.

[0041] Bei mindestens einer Ausführungsform wird der Z-Rekonstruktionsverlust 520 berechnet. Bei mindestens einer Ausführungsform betrifft der Z-Rekonstruktionsverlust eine Differenz oder einen Abstand zwischen einem Eingabesatz von Erscheinungsbildparametern 504 und einem vorhergesagten Satz von Erscheinungsbildparametern 516. Bei mindestens einer Ausführungsform betrifft der Z-Rekonstruktionsverlust eine Verlustfunktion, die basierend darauf berechnet ist, wie genau der Diskriminator 510 bei dem Vorhersagen von Erscheinungsbildparametern oder Erscheinungsbildeigenschaften eines Bildes ist.

[0042] Bei mindestens einer Ausführungsform wird die Bestimmung 514 verwendet, um einen Klassifizierungsverlust, wie einen Echt/Falsch-Klassifizierungsverlust 522 zu bestimmen. Bei mindestens einer Ausführungsform wird der Echt/Falsch-Klassifizierungsverlust 522 basierend darauf berechnet, ob der Diskriminator 514 in der Lage ist, korrekt vorherzusagen, ob das erzeugte Bild 508, das dem Diskriminator 510 unterbreitet wird, ein echtes Bild oder ein synthetisiertes Bild ist. Bei mindestens einer Ausführungsform wird der Echt/-Falsch-Klassifizierungsverlust 522 basierend darauf berechnet, ob der Diskriminator 510 in der Lage ist, korrekt vorherzusagen, ob Sätze von Eingabebildern echt oder falsch sind, wobei dem Diskriminator 510 entweder echte oder falsche (z. B. synthetische) Bilder bereitgestellt werden können, und er vorhersagen soll, ob diese Bilder echt oder falsch sind.

[0043] **Fig. 6** veranschaulicht ein Diagramm 600, das eine Generatoraktualisierung abbildet, gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform werden Verlustfunktionen berechnet und verwendet, um Parameter eines Generators, die dazu verwendet werden, synthetische Bilder anzulegen, zu aktualisieren. Bei mindestens einer Ausführungsform zeigt **Fig. 6** veranschaulicht das Eingabebild 602; den Diskriminator 604; einen vorhergesagten Sichtpunkt 606; eine vorhergesagte Bestimmung 608, ob das eingegebene Bild 602 echt oder falsch ist; einen vorhergesagten Satz von Erscheinungsbildparametern 610; den Generator 612; das erzeugte Bild 614; den Bildkonsistenzverlust 616; und den Echt/Falsch-Klassifizierungsverlust 618. Bei mindestens einer Ausführungsform veranschaulicht **Fig. 6** eine Generatoraktualisierung unter Verwendung eines echten Bildes (z. B. eines Bildes, das nicht durch ein generatives gegnerisches Netzwerk angelegt wird). Bei mindestens einer Ausführungsform haben Techniken, die in Verbindung mit **Fig. 6** sind, die gleiche Ausdehnung wie die in Verbindung mit den **Fig. 4**, **Fig. 5** und **Fig. 7** beschrieben, um Generatoren und/oder Diskriminatoren zu trainieren.

[0044] Bei mindestens einer Ausführungsform wird das Eingabebild 602 in den Diskriminator 604 eingegeben. Bei mindestens einer Ausführungsform ist das Eingabebild 602 ein echtes Bild, das aus einem Satz von Trainingsdaten ausgewählt wird. Bei mindestens einer Ausführungsform wird das Eingabebild 602 aus einer Objektbildsammlung ausgewählt, wie z. B. denjenigen, die in Verbindung mit **Fig. 2** beschrieben wurden. Bei mindestens einer Ausführungsform verarbeitet der Diskriminator 604 das Eingabebild 602. Bei mindestens einer Ausführungsform empfängt der Diskriminator 604 das Eingabebild 602 und erzeugt einen vorhergesag-

ten Sichtpunkt 606, eine vorhergesagte Bestimmung, ob das Eingabebild 602 ein echtes oder ein falsches Bild ist, und einen vorhergesagten Satz von Erscheinungsbildparametern 608. Bei mindestens einer Ausführungsform ist der Sichtpunkt 606 ein vorhergesagter Sichtpunkt des Eingabebildes 602, der von einem oder mehreren Prozessen des Diskriminators 604 bestimmt wird. Bei mindestens einer Ausführungsform ist die Bestimmung 608 eine Vorhersage darüber, ob das Eingabebild 602 ein echtes Bild oder ein falsches Bild ist (z. B. ein synthetisches Bild, das von einem generativen kontradiktorischen Netzwerk angelegt wird). Bei mindestens einer Ausführungsform ist die Bestimmung 608 ein Binärwert oder ein Konfidenzwert über einen nicht-binären Bereich (z. B. 0-100). Bei mindestens einer Ausführungsform ist der Satz von Erscheinungsbildparametern 610 ein vorhergesagter Satz von Erscheinungsbildparametern des Eingabebildes 602, das von einem oder mehreren Prozessen des Diskriminators 604 erzeugt wird. Bei mindestens einer Ausführungsform besteht der Satz von Erscheinungsbildparametern 610 aus vorhergesagten Parametern, die ein Erscheinungsbild eines Bildes definieren, das in dem Eingabebild 602 abgebildet ist.

[0045] Bei mindestens einer Ausführungsform werden der Sichtpunkt 606 und der Satz von Erscheinungsbildparametern 610 als Eingaben für den Generator 612 bereitgestellt, um ein erzeugtes Bild 614 zu erzeugen. Bei mindestens einer Ausführungsform ist das erzeugte Bild 614 ein synthetisches (z. B. falsches) Bild, das von dem Generator 612 mit einer Ausrichtung und einem Erscheinungsbildbild gemäß dem Sichtpunkt 606 und einem Satz von Erscheinungsbildbildparametern 610 angelegt wird, die wie in **Fig. 6** veranschaulicht auch ein vorhergesagter Sichtpunkt und ein Erscheinungsbild des Eingabebildes 602 sind.

[0046] Bei mindestens einer Ausführungsform wird der Bildkonsistenzverlust 616 mindestens zum Teil basierend auf dem Eingabebild 602 berechnet, das dem Diskriminator 604 bereitgestellt wird, der verwendet wird, um mindestens zwei Eigenschaften aus dem Eingabebild 602 zu disaggregieren: einen vorhergesagten Sichtpunkt 606 und einen vorhergesagten Satz von Erscheinungsbildparameter 610. Bei mindestens einer Ausführungsform werden der vorhergesagte Sichtpunkt 606 und der vorhergesagte Satz von Erscheinungsbildparametern 610 dem Generator 612 bereitgestellt, um ein erzeugtes Bild 614 anzulegen. Bei mindestens einer Ausführungsform empfängt ein Generative Adversarial Network (GAN) einen Sichtpunkt und einen Satz von Erscheinungsbildparametern und erzeugt ein synthetisches (z. B. falsches) Bild, das dem jeweiligen Sichtpunkt und Satz von Erscheinungsbildparametern entspricht, der bereitgestellt wurde. Bei mindestens einer Ausführungsform werden das Eingabebild 602 und das erzeugte Bild 614 verglichen, um den Bildkonsistenzverlust 616 zu bestimmen. Bei mindestens einer Ausführungsform wird ein Kosinus-Abstand zwischen einem Eingabebild und einem synthetisierten Bild verglichen, um Merkmalsähnlichkeiten zu bestimmen, wobei eine größere Ähnlichkeit einem geringeren Verlust entspricht. Bei mindestens einer Ausführungsform werden L1-, L2- oder Kosinus-Abstände verwendet, um den Bildkonsistenzverlust 616 zwischen zwei Bildern zu bestimmen.

[0047] Bei mindestens einer Ausführungsform wird die Bestimmung 608 verwendet, um einen Klassifizierungsverlust, wie etwa einen Echt/Falsch-Klassifizierungsverlust 618, zu bestimmen. Bei mindestens einer Ausführungsform wird der Echt/Falsch-Klassifizierungsverlust 618 basierend darauf berechnet, ob der Diskriminator 604 in der Lage ist, korrekt vorherzusagen, ob das dem Diskriminator 604 unterbreitete Eingabebild 602 ein echtes Bild oder ein synthetisiertes Bild ist. Bei mindestens einer Ausführungsform wird der Echt/-Falsch-Klassifizierungsverlust 618 basierend darauf berechnet, ob der Diskriminator 604 in der Lage ist, korrekt vorherzusagen, ob Sätze von Eingabebildern echt oder falsch sind, wobei dem Diskriminator 604 entweder echte oder falsche (z. B. synthetische) Bilder bereitgestellt werden können und er vorhersagen soll, ob diese Bilder echt oder falsch sind. Bei mindestens einer Ausführungsform werden eine oder mehrere Verlustfunktionen berechnet. Bei mindestens einer Ausführungsform werden Parameter des Generators 612 aktualisiert, indem mindestens Gradienten berechnet werden (z. B. durch Ausführen eines stochastischen Gradientenabstiegs), während Diskriminatorparameter festgelegt werden.

[0048] **Fig. 7** veranschaulicht ein Diagramm 700, das eine Generatoraktualisierung abbildet, gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform werden Verlustfunktionen berechnet und verwendet, um Parameter eines Generators, die dazu verwendet werden, synthetische Bilder anzulegen, zu aktualisieren. Bei mindestens einer Ausführungsform veranschaulicht **Fig. 7** einen ersten Sichtpunkt 702; einen Satz von Erscheinungsbildparametern 704; einen Generator 706; ein erstes erzeugtes Bild 708; einen Diskriminator 710; einen vorhergesagten ersten Sichtpunkt 712; eine vorhergesagte Bestimmung 714, ob das erzeugte Bild 708 echt oder falsch ist, einen vorhergesagten ersten Satz von Erscheinungsbildparametern 716; den Sichtpunktkonsistenzverlust 718; den Z-Rekonstruktionsverlust 720; den Echt/Falsch-Klassifizierungsverlust 722; den zweiten Sichtpunkt 724; ein zweites erzeugtes Bild 726; einen vorhergesagten zweiten Sichtpunkt 728; einen vorhergesagten zweiten Satz von Erscheinungsbildparametern 730; den Sichtpunktkonsistenzverlust 732; den Z-Rekonstruktionsverlust 734; und den Symmetrieverlust 736. Bei min-

destens einer Ausführungsform veranschaulicht **Fig. 6** eine Generatoraktualisierung unter Verwendung eines falschen Bildes (z. B. eines synthetischen oder erzeugten Bildes, das von einem generativen gegnerischen Netzwerk angelegt wurde). Bei mindestens einer Ausführungsform sind Techniken, die in Verbindung mit **Fig. 7** beschrieben sind, die gleiche Ausdehnung aufweisen wie die in Verbindung mit den **Fig. 4-6** beschriebenen, um Generatoren und/oder Diskriminatoren zu trainieren.

[0049] Bei mindestens einer Ausführungsform sind der erste Sichtpunkt 702 und der Satz von Erscheinungsbildparametern 704 entflochtene Parameter, die unabhängig ausgewählt werden können. Bei mindestens einer Ausführungsform akzeptiert der Generator 706 den ersten Sichtpunkt 702 und den Satz von Erscheinungsbildparametern 704 als Eingaben und legt ein erstes erzeugtes Bild 708 an. Bei mindestens einer Ausführungsform ist das erste erzeugte Bild 708 ein synthetisches Bild, das basierend auf dem Satz von Erscheinungsbildparametern 704 erscheint und gemäß dem ersten Sichtpunkt 702 ausgerichtet ist. Der Generator 706 kann jenen entsprechen, die an anderer Stelle in dieser Offenbarung beschrieben sind, wie jenen, die in Verbindung mit **Fig. 2** erörtert wurden.

[0050] Bei mindestens einer Ausführungsform wird ein Satz von Bildern (z. B. das erste erzeugte Bild 708) einem Diskriminator 710 als eine Eingabe bereitgestellt, und der Diskriminator sagt verschiedene Eigenschaften dieses Satzes von Bildern vorher. Bei mindestens einer Ausführungsform empfängt der Diskriminator 710 das erste erzeugte Bild 708 und erzeugt einen vorhergesagten ersten Sichtpunkt 712; eine vorhergesagte Bestimmung 714, ob das erzeugte Bild 708 echt oder falsch ist, und einen vorhergesagten ersten Satz von Erscheinungsbildparametern 716. Bei mindestens einer Ausführungsform hat der Diskriminator 710 keinen Zugriff auf den ersten Sichtpunkt 702 und den Satz von Erscheinungsbildparametern 704, die dazu verwendet werden, das erste erzeugte Bild 708 anzulegen (z. B. werden derartige Informationen dem Diskriminator 710 während der Vorhersage vorenthalten). Bei mindestens einer Ausführungsform werden Ausgaben des Diskriminators 710 verwendet, um Verlust zu berechnen. Bei mindestens einer Ausführungsform werden Verlustfunktionen verwendet, um Gradienten zu berechnen (z. B. unter Verwendung von Gradientenabstieg) und Parameter für den Diskriminator 710 zu aktualisieren, während Parameter des Generators 706 konstant gehalten werden.

[0051] Bei mindestens einer Ausführungsform wird der Sichtpunktkonsistenzverlust 718 berechnet. Bei mindestens einer Ausführungsform betrifft Sichtpunktkonsistenzverlust eine Verlustfunktion, die basierend darauf berechnet wird, wie genau der Diskriminator 710 Sichtpunkte vorhersagt. Bei mindestens einer Ausführungsform wird der Sichtpunktkonsistenzverlust 718 als eine Differenz oder ein Abstand zwischen dem ersten Sichtpunkt 702 und dem vorhergesagten ersten Sichtpunkt 712 berechnet. Bei mindestens einer Ausführungsform ist der Sichtpunktkonsistenzverlust eine Komponente des generativen Konsistenzverlusts. Bei mindestens einer Ausführungsform wird ein Abstand (z. B. L1-Abstand, L2-Abstand, Kosinus-Abstand und/oder Variationen davon) zwischen dem ersten Sichtpunkt 702 und dem vorhergesagten ersten Sichtpunkt 712 verwendet, um einen Sichtpunktkonsistenzverlust 718 zu berechnen, wobei nähere Sichtpunkte (z. B., Sichtpunkte mit kürzeren Abständen voneinander) einem geringeren Verlust entsprechen.

[0052] Bei mindestens einer Ausführungsform wird der Z-Rekonstruktionsverlust 720 berechnet. Bei mindestens einer Ausführungsform betrifft der Z-Rekonstruktionsverlust eine Differenz oder einen Abstand zwischen dem Satz von Erscheinungsbildparametern 704 und einem vorhergesagten ersten Satz von Erscheinungsbildparametern 716. Bei mindestens einer Ausführungsform betrifft der Z-Rekonstruktionsverlust eine Verlustfunktion, die basierend darauf berechnet ist, wie genau der Diskriminator 710 bei dem Vorhersagen von Erscheinungsbildparametern oder Erscheinungsbildeigenschaften eines Bildes ist.

[0053] Bei mindestens einer Ausführungsform wird die Bestimmung 714 verwendet, um einen Klassifizierungsverlust, wie etwa einen Echt/Falsch-Klassifizierungsverlust 722 zu bestimmen. Bei mindestens einer Ausführungsform wird der Echt/Falsch-Klassifizierungsverlust 722 basierend darauf berechnet, ob der Diskriminator 714 in der Lage ist, korrekt vorherzusagen, ob das erste erzeugte Bild 708, das dem Diskriminator 710 unterbreitet wird, ein echtes Bild oder ein synthetisiertes Bild ist. Bei mindestens einer Ausführungsform wird der Echt/Falsch-Klassifizierungsverlust 722 basierend darauf berechnet, ob der Diskriminator 710 in der Lage ist, korrekt vorherzusagen, ob Sätze von Eingabebildern echt oder falsch sind, wobei dem Diskriminator 710 entweder echte oder falsche (z. B. synthetische) Bilder bereitgestellt werden können, und er vorhersagen soll, ob diese Bilder echt oder falsch sind.

[0054] Bei mindestens einer Ausführungsform ist der zweite Sichtpunkt 724 eine Transformation des ersten Sichtpunkts 702. Bei mindestens einer Ausführungsform wird der erste Sichtpunkt 702 horizontal gespiegelt, um den zweiten Sichtpunkt 724 zu erzeugen. Bei mindestens einer Ausführungsform erzeugt jede geeignete

Transformation von Azimut-, Kipp- und Elevationsparametern an dem ersten Sichtpunkt 702 den zweiten Sichtpunkt 724. Bei mindestens einer Ausführungsform wird der zweite Sichtpunkt 724 gemäß Techniken bestimmt, die an anderer Stelle in dieser Offenbarung beschrieben sind, wie die in Verbindung mit **Fig. 8** erläuterten. Falls bei mindestens einer Ausführungsform der erste Sichtpunkt 702 jeweils Azimut-, Elevations- und Kippparameter wie φ,θ,ψ aufweist, weist der zweite Sichtpunkt 724 Azimut-, Elevations- und Kippparameter wie -φ,θ,-ψ auf.

[0055] Bei mindestens einer Ausführungsform wird der Satz von Erscheinungsbildparametern 704 verwendet, um ein zweites Bild zu erzeugen. Bei mindestens einer Ausführungsform werden der zweite Sichtpunkt 724 und der Satz von Erscheinungsbildparametern 704 als Eingaben für den Generator 706 bereitgestellt, um ein zweites erzeugtes Bild 726 anzulegen. Falls bei mindestens einer Ausführungsform das zweite erzeugte Bild 726 horizontal gespiegelt wird, erzeugt er das erste erzeugte Bild 708.

[0056] Bei mindestens einer Ausführungsform wird der Symmetrieverlust 736 basierend auf dem ersten erzeugten Bild 708 und dem zweiten erzeugten Bild 726 berechnet. Bei mindestens einer Ausführungsform wird Symmetrieverlust 736 berechnet, indem die Größen von Azimut, Elevation und Kippen des vorhergesagten ersten Sichtpunkts 712 des erzeugten Bildes 708 mit dem vorhergesagten zweiten Sichtpunkt 728 des zweiten erzeugten Bildes 726 verglichen werden, wobei sich Nullverlust ergibt, wenn die Größen jedes Sichtpunktparameters gleich sind, und der Verlust mit dem zunehmenden Unterschied zwischen den Größen jedes Sichtpunktparameters zunimmt. Bei mindestens einer Ausführungsform wird Symmetrieverlust 736 gemäß Techniken berechnet, die an anderer Stelle in dieser Offenbarung beschrieben sind, wie etwa jenen, die in Verbindung mit **Fig. 8** erörtert sind. Bei mindestens einer Ausführungsform wird der Entflechtungsverlust innerhalb des Kontextes der **Fig. 7** angewandt. Bei mindestens einer Ausführungsform werden ein Sichtpunkt V_1 und Satz von Erscheinungsbildparametern Z_1 ausgewählt und einem Generator bereitgestellt, um ein erstes synthetisches Bild I_1 zu erzeugen. Bei mindestens einer Ausführungsform wird ein zweites Bild I_2 erzeugt, indem der Sichtpunkt konstant gehalten wird (z. B. mit V_1) und störende Erscheinungsbildparameter durch Verwendung eines zweiten Satzes von Erscheinungsbildparametern Z_2 anders als Z_1 dazu verwendet werden, ein erstes Bild I_1 zu erzeugen. Bei mindestens einer Ausführungsform wird ein drittes Bild I_3 erzeugt, indem Erscheinungsbildparameter relativ zu Bild I_1 (z. B. mit Z_1), und ein störender Sichtpunkt durch Verwendung eines zweiten Sichtpunkts V_2 anders als Sichtpunkt V_1, konstant gehalten werden, um ein drittes Bild zu erzeugen I_3. Bei mindestens einer Ausführungsform wird ein synthetisches Bild I_1, das unter Verwenden eines spezifischen Sichtpunkts V_1 und dem Satz von Erscheinungsbildparametern Z_1 erzeugt wird, mit dem synthetischen Bild I_2 verglichen, das mit dem Sichtpunkt V_1 und einem zweiten Satz von Erscheinungsbildparametern Z_2 erzeugt wird, und/oder einem synthetischen Bild I_3, das unter Verwendung eines zweiten Sichtpunkts V_2 und Satz von Erscheinungsbildparametern Z_1 erzeugt wird. Techniken, die in Verbindung mit **Fig. 10** beschrieben sind, können auf Entflechtungsverlust, der in Verbindung mit **Fig. 7** beschrieben ist, gemäß mindestens einer Ausführungsform, anwendbar sein.

[0057] Fig. 8 veranschaulicht ein Diagramm 800, das das Berechnen von Symmetrieverlusten abbildet, gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform wird Symmetrieverlust verwendet, um ein oder mehrere neuronale Netzwerke zu trainieren, die mit einem Diskriminator wie etwa einem Diskriminator 806 assoziiert sind. Bei mindestens einer Ausführungsform wird Symmetrieverlust mit einer oder mehreren anderen Verlustfunktionen verwendet, um Parameter zu verfeinern, die mit einem Diskriminator assoziiert sind.

[0058] Bei mindestens einer Ausführungsform beinhaltet das Diagramm 800 ein Eingabebild 802. Bei mindestens einer Ausführungsform ist das Eingabebild 802 Teil einer Sammlung von einem oder mehreren Bildern eines Objekttyps. Bei mindestens einer Ausführungsform ist das Eingabebild 802 ein Bild, das ein Auto in einer spezifischen Ausrichtung, die spezifische Erscheinungsbildmerkmale umfasst, abbildet. Bei mindestens einer Ausführungsform wird das Eingabebild 802 aus einer Sammlung von einem oder mehreren Bildern ausgewählt. Bei mindestens einer Ausführungsform werden Bilder einer Sammlung in irgendeiner geeigneten Weise zum Lernen, die zufällig oder pseudozufällig aus einem Trainingssatz ausgewählt werden können, ausgewählt.

[0059] Bei mindestens einer Ausführungsform wird eine Transformation auf das Eingabebild 802 angewandt, um ein transformiertes Bild 804 zu erzeugen. Bei mindestens einer Ausführungsform wird das Eingabebild 802 horizontal gespiegelt, um das transformierte Bild 804 zu erzeugen. Bei mindestens einer Ausführungsform wird eine Transformation von einem oder mehreren Systemen angewandt, die mit dem Diskriminator 806 assoziiert sind. Bei mindestens einer Ausführungsform wendet der Diskriminator 806 eine oder mehrere Bildverarbeitungstechniken auf das Eingabebild 802 an, um das transformierte Bild 804 zu erzeugen. Bei

mindestens einer Ausführungsform werden Azimut- und Kippwinkel (in dem Diagramm 800 als „az“ und „ti“ abgebildet) eines Sichtpunkts eines Objekts innerhalb des Eingabebildes 802 umgekehrt, wenn das Eingabebild 802 gespiegelt und/oder transformiert wird, um ein transformiertes Bild 804 zu erzeugen, während Elevationswinkel (in dem Diagramm 800 als „el“ abgebildet) über das Eingabebild 802 und das transformierte Bild 804 gleich bleiben. Bei mindestens einer Ausführungsform wird das transformierte Bild 804 durch Anwenden einer oder mehrerer Bildtransformationen auf das Eingabebild 802 erzeugt. Bei mindestens einer Ausführungsform wird das transformierte Bild 804 erzeugt, indem das Eingabebild 802 mindestens horizontal gespiegelt wird, das Eingabebild 802 horizontal gespiegelt wird, das Eingabebild 802 gemäß einer spezifizierten Achse gespiegelt wird, das Eingabebild 802 um eine spezifizierte Anzahl von Graden gedreht wird und/oder verschiedene andere 2D-Transformationen, die auf das Eingabebild 802 angewandt werden.

[0060] Bei mindestens einer Ausführungsform verarbeitet der Diskriminator 806 das Eingabebild 802. Bei mindestens einer Ausführungsform ist der Diskriminator 806 mit einem oder mehreren neuronalen Netzwerken assoziiert, die dazu trainiert sind, einen Sichtpunkt sowie andere Charakteristiken eines Eingabebildes abzuleiten. Bei mindestens einer Ausführungsform empfängt der Diskriminator 806 das Eingabebild 802 und erzeugt eine erste Vorhersage 808. Bei mindestens einer Ausführungsform entspricht die erste Vorhersage 808 einer vorhergesagten spezifischen Ausrichtung eines Objekts innerhalb eines Bildes und umfasst spezifische Werte für einen Satz von Parametern, der einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfasst. Bei mindestens einer Ausführungsform umfasst eine erste Vorhersage 808 einen ersten vorhergesagten Sichtpunkt V_1 und einen ersten vorhergesagten Satz von Erscheinungsbildparametern Z_1 des Eingabebildes 802. Bei mindestens einer Ausführungsform umfasst die erste Vorhersage 808 eine vorhergesagte Ausrichtung eines Autos, das in dem Eingabebild 802 abgebildet ist. Bei mindestens einer Ausführungsform verarbeitet der Diskriminator 806 das transformierte Bild 804. Bei mindestens einer Ausführungsform empfängt der Diskriminator 806 das transformierte Bild 804 und erzeugt eine zweite Vorhersage 810. Bei mindestens einer Ausführungsform entspricht die zweite Vorhersage 810 einer vorhergesagten spezifischen Ausrichtung eines Objekts innerhalb eines Bildes und umfasst spezifische Werte für einen Satz von Parametern, der einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfasst. Bei mindestens einer Ausführungsform umfasst eine zweite Vorhersage 810 einen zweiten vorhergesagten Sichtpunkt V_2 und einen zweiten vorhergesagten Satz von Erscheinungsbildparametern Z_2 des transformierten Bildes 804. Bei mindestens einer Ausführungsform umfasst die zweite Vorhersage 810 eine vorhergesagte Ausrichtung eines Autos, das in dem transformierten Bild 804 abgebildet ist.

[0061] Bei mindestens einer Ausführungsform wird die erste Vorhersage 808 für das Eingabebild 802 vorhergesagt, und die zweite Vorhersage 810 wird für das transformierte Bild 804, das eine horizontal gespiegelte Version des Eingabebildes 802 ist, vorhergesagt. Bei mindestens einer Ausführungsform wird Verlust basierend darauf berechnet, ob bestimmte Eigenschaften zutreffen. Bei mindestens einer Ausführungsform wird eine Transformation oder Inverse davon auf die zweite Vorhersage 810 angewandt. Falls bei mindestens einer Ausführungsform das Eingabebild 802 um Winkel (φ,θ,ψ) gedreht wird, um das transformierte Bild 804 zu erzeugen, kann die zweite Vorhersage 810 um Winkel $(-\varphi,-\theta,-\psi)$ umgekehrt gedreht werden. Bei mindestens einer Ausführungsform wird Symmetrieverlust durch Vergleichen von Größen von Azimut, Elevation und Kippen der ersten Vorhersage 808 des Eingabebildes 802 mit der zweiten Vorhersage 810 des transformierten Bildes 804 berechnet, wobei sich ein Nullverlust ergibt, wenn die Größen jedes Sichtpunktparameters gleich sind/oder die für das Eingabebild 802 vorhergesagten Erscheinungsbildparameter mit denen, die für das transformierte Bild 804 vorhergesagt wurden, übereinstimmen. Bei mindestens einer Ausführungsform nimmt der Verlust zu, wenn eine Differenz zwischen Größen jedes Sichtpunktparameters zunimmt. Bei mindestens einer Ausführungsform wird Symmetrieverlust gemäß Techniken berechnet, die an anderer Stelle in dieser Offenbarung beschrieben sind, wie etwa jenen, die in Verbindung mit **Fig. 16** erörtert sind. Bei mindestens einer Ausführungsform wird symmetrischer Verlust mindestens zum Teil basierend darauf berechnet, wie genau die vorhergesagten Erscheinungsbildparameter für das Eingabebild 802 mit denen übereinstimmen, die für das transformierte Bild 804 vorhergesagt werden. Bei mindestens einer Ausführungsform werden Gewichte und Parameter des Diskriminators 806 dazu trainiert vorherzusagen, dass Erscheinungsbildparameter für das Eingabebild 802 mit Erscheinungsbildparametern übereinstimmen, die für eine transformierte Version des Eingabebildes 802 (z. B. das transformierte Bild 804) vorhergesagt werden.

[0062] Bei mindestens einer Ausführungsform sagt der Diskriminator 806 einen ersten Satz von Erscheinungsbildparametern für das Eingabebild 802 vorher und sagt einen zweiten Satz von Erscheinungsbildparametern für das transformierte Bild 804 vorher. Bei mindestens einer Ausführungsform wird eine Verlustfunktion basierend darauf berechnet, wie ähnlich ein erster Satz von Erscheinungsbildparametern, der für das Eingabebild 802 vorhergesagt wird, einem zweiten Satz von Erscheinungsbildparametern, der für das transformierte Bild 804 vorhergesagt sind, ist. Bei mindestens einer Ausführungsform wird der Diskriminator 806

damit trainiert, dass vorhergesagte Parameter für das Eingabebild 802 und das transformierte Bild 804 äquivalent sind.

[0063] **Fig. 9** veranschaulicht ein Diagramm 900, das einen Sichtpunktgraphen abbildet, gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform wird ein Sichtpunktgraph 908 verwendet, um einen Verlust des nächsten Nachbarn und des am weitesten entfernten Nachbarn zu bestimmen, der mit einer oder mehreren anderen Verlustfunktionen eingesetzt wird, um Parameter, die einem Diskriminator assoziiert sind, zu verfeinern. Bei mindestens einer Ausführungsform wird der Sichtpunktgraph 908 von einem oder mehreren Prozessen und Systemen, die mit einem Diskriminator assoziiert sind, konstruiert. Bei mindestens einer Ausführungsform wird der Sichtpunktgraph 908 basierend auf einer Sammlung von einem oder mehreren Bildern eines Objekttyps erzeugt. Bei mindestens einer Ausführungsform sind ein erstes Bild 902 und ein zweites Bild 904 Teil einer Sammlung von einem oder mehreren Bildern eines Objekttyps. Bei mindestens einer Ausführungsform sind das erste Bild 902 und das zweite Bild 904 Teil einer Sammlung von einem oder mehreren Bildern, die Autos beinhalten. Bei mindestens einer Ausführungsform sind das erste Bild 902 und das zweite Bild 904 Bilder, die Autos in spezifischen Ausrichtungen darstellen.

[0064] Bei mindestens einer Ausführungsform wird der Sichtpunktgraph 908 basierend auf Sichtpunkt-Äquivariant-Abständen (z. B. Kosinus-Abständen) zwischen Bildern einer Sammlung von Bildern erzeugt. Bei mindestens einer Ausführungsform ist der Kosinus-Abstand ein mathematisches Komplement der Kosinusähnlichkeit (z. B. Kosinus-Abstand = 1 - Kosinusähnlichkeit). Bei mindestens einer Ausführungsform ist Kosinusähnlichkeit ein Maß für die Ähnlichkeit zwischen zwei Vektoren, die Bilder, Text, Daten und/oder Variationen, basierend auf einem Kosinus eines Winkels zwischen ihnen, davon darstellen können. Bei mindestens einer Ausführungsform ist, wenn zwei Bilder Merkmale umfassen, die Objekten entsprechen, die ähnliche Sichtpunkte aufweisen, ein Kosinus-Abstand, der für die Bilder berechnet wird, gering. Bei mindestens einer Ausführungsform ist, wenn zwei Bilder Merkmale umfassen, die Objekten mit unterschiedlichen Sichtpunkten entsprechen, ein Kosinus-Abstand, der für die Bilder berechnet wird, hoch.

[0065] Bei mindestens einer Ausführungsform wird eine Sammlung von Bildern verwendet, um den Sichtpunktgraph 908 zu erzeugen, wobei Knoten des Sichtpunktgraphs 908 Bildern entsprechen und Kanten ihren Kosinus-Abständen entsprechen. Bei mindestens einer Ausführungsform werden Kosinus-Abstände basierend auf Merkmalsähnlichkeiten von Bildpaaren unter Verwendung eines faltenden neuronalen Netzwerks berechnet. Bei mindestens einer Ausführungsform werden Kanten des Sichtpunktgraphen 908 derart gewichtet, dass dickere Kanten zwischen zwei Bildern einem höheren Grad an Ähnlichkeit zwischen den zwei Bildern entsprechen und dünnere Kanten zwischen zwei Bildern einem geringeren Grad an Ähnlichkeit zwischen den zwei Bildern entsprechen. Bei mindestens einer Ausführungsform wird ein Kosinus-Abstand 910 zwischen dem ersten Bild 902 und dem zweiten Bild 904 berechnet. Bei mindestens einer Ausführungsform werden das erste Bild 902 und das zweite Bild 904 als Teil des Sichtpunktgraphs 908 eingesetzt und sind mit einer Kante, die dem berechneten Kosinus-Abstand 910 entspricht, verbunden. Bei mindestens einer Ausführungsform umfassen das erste Bild 902 und das zweite Bild 904 Autos, die in ähnlichen Ausrichtungen und/oder Sichtpunkten ausgerichtet sind, und eine dicke Kante, die zwischen dem ersten Bild 902 und dem zweiten Bild 904 innerhalb des Sichtpunktgraphs 908 verwendet wird, spiegelt derartige Ähnlichkeiten wider.

[0066] **Fig. 9** veranschaulicht ein Diagramm 900, das die Berechnung von Verlusten an dem nächsten und an dem weitesten entfernten Nachbarn gemäß mindestens einer Ausführungsform abbildet. Bei mindestens einer Ausführungsform wird ein Verlust des nächsten Nachbarn und des am weitesten entfernten Nachbarn verwendet, um ein oder mehrere neuronale Netzwerke zu trainieren, die mit einem Diskriminator, wie etwa einem Diskriminator 912, assoziiert sind. Bei mindestens einer Ausführungsform wird Verlust des nächsten Nachbarn und des am weitesten entfernten Nachbarn mit einer oder mehreren anderen Verlustfunktionen verwendet, um Parameter zu verfeinern, die mit einem Diskriminator assoziiert sind. Bei mindestens einer Ausführungsform kann der Verlust des nächsten und des am weitesten entfernten Nachbarn als eine Art von Sichtpunkt-Äquivarianz-Verlust betrachtet werden. Hierin beschriebene Techniken, die für den Verlust des nächsten und des am weitesten entfernten Nachbarn gelten (z. B. in Verbindung mit **Fig. 9** und an anderer Stelle beschrieben), sind bei mindestens einer Ausführungsform auch auf den Äquivarianzverlust des Sichtpunkts anwendbar.

[0067] Bei mindestens einer Ausführungsform wird ein Sichtpunktgraph 908 durch einen oder mehrere Prozesse und Systeme, die mit dem Diskriminator 912 assoziiert sind, konstruiert. Bei mindestens einer Ausführungsform wird der Sichtpunktgraph 908 basierend auf einer Sammlung von einem oder mehreren Bildern eines Objekttyps erzeugt. Bei mindestens einer Ausführungsform sind die Bilder 902-906 Teil einer Samm-

lung von einem oder mehreren Bildern eines Objekttyps. Bei mindestens einer Ausführungsform sind die Bilder 902-906 Teil einer Sammlung von einem oder mehreren Bildern, die Autos beinhalten. Bei mindestens einer Ausführungsform sind die Bilder 902-906 Bilder, die Autos in spezifischen Ausrichtungen abbilden.

[0068] Bei mindestens einer Ausführungsform verarbeitet der Diskriminator 912 die Bilder 902-906. Bei mindestens einer Ausführungsform ist der Diskriminator 912 mit einem oder mehreren neuronalen Netzwerken assoziiert, die dazu trainiert sind, einen Sichtpunkt sowie andere Charakteristiken eines Eingabebildes abzuleiten. Bei mindestens einer Ausführungsform empfängt der Diskriminator 912 das Bild 902 und erzeugt einen Sichtpunkt 914. Bei mindestens einer Ausführungsform empfängt der Diskriminator 912 das Bild 904 und erzeugt einen Sichtpunkt 916. Bei mindestens einer Ausführungsform empfängt der Diskriminator 912 das Bild 906 und erzeugt einen Sichtpunkt 918. Bei mindestens einer Ausführungsform entsprechen die Sichtpunkte 914-918 vorhergesagten spezifischen Ausrichtungen von Objekten, die in den Bildern 902-906 abgebildet sind, und umfassen spezifische Werte für Sätze von Parametern, die Azimutparameter, Elevationsparameter und Kippparameter umfassen. Bei mindestens einer Ausführungsform umfassen die Sichtpunkte 914-918 vorhergesagte Ausrichtungen von Autos, die jeweils in den Bildern 902-906 abgebildet sind.

[0069] Bei mindestens einer Ausführungsform wird Verlust des nächsten und des am weitesten entfernten Nachbarn berechnet, indem mindestens ein ausgewähltes Bild mit seinen nächsten und am weitesten entfernten Nachbarn, mindestens zum Teil basierend auf dem Sichtpunktgraph 908 einer Sammlung von Bildern verglichen wird. Bei mindestens einer Ausführungsform umfasst der Verlust des nächsten Nachbarn und des am weitesten entfernten Nachbarn einen Verlust des nächsten Nachbarn und einen Verlust des am weitesten entfernten Nachbarn. Bei mindestens einer Ausführungsform wird ein Ankerbild aus einer Sammlung von Trainingsbildern ausgewählt. Bei mindestens einer Ausführungsform wird das Bild 902 als ein Ankerbild ausgewählt. Bei mindestens einer Ausführungsform wird das Bild 902 von dem Sichtpunktgraphen 908 aus lokalisiert, und ein nächster Nachbar und ein am weitesten entfernter Nachbar werden basierend auf Kantengewichten ausgewählt. Bei mindestens einer Ausführungsform weist ein nächster Nachbar, wie etwa das Bild 904, eine kürzeste Kante, die mit dem Bild 902 verbunden ist, auf. Bei mindestens einer Ausführungsform weist ein am weitesten entfernter Nachbar, wie etwa das Bild 906, eine am weitesten entferne Kante auf, die mit dem Bild 902 verbunden ist.

[0070] Bei mindestens einer Ausführungsform wird das Bild 904 als ein nächster Nachbar des Bildes 902 bestimmt. Bei mindestens einer Ausführungsform wird Verlust des nächsten Nachbarn zwischen dem Sichtpunkt 914 und dem Sichtpunkt 916 berechnet. Bei mindestens einer Ausführungsform wird Verlust des nächsten Nachbarn derart berechnet, dass eine höhere Ähnlichkeit zwischen dem Sichtpunkt 914 und dem Sichtpunkt 916 einem geringeren Verlust entspricht. Bei mindestens einer Ausführungsform wird Verlust des nächsten Nachbarn basierend auf einer folgenden symbolischen mathematischen Gleichung berechnet:

$$L_{nn} = d\left(v_{l_1}, v_{l_2}\right)$$

[0071] wobei L_{nn} der Verlust des nächsten Nachbarn ist, v_{l_1} ein Sichtpunkt eines Ankerbildes ist, v_{l_2} ein Sichtpunkt eines Bildes des nächsten Nachbarn ist, und $d(v_{l_1}, v_{l_2})$ eine Funktion ist, die einen Abstand (z. B. L1-Abstand, L2-Abstand, Kosinus-Abstand und/oder Variationen davon) oder eine Differenz zwischen Sichtpunkten bestimmt.

[0072] Bei mindestens einer Ausführungsform wird das Bild 906 als der am weitesten entfernte Nachbar des Bildes 902 bestimmt. Bei mindestens einer Ausführungsform wird Verlust des am weitesten entfernte Nachbar Nachbarn zwischen dem Sichtpunkt 914 und dem Sichtpunkt 916 berechnet. Bei mindestens einer Ausführungsform wird Verlust des am weitesten entfernten Nachbarn derart berechnet, dass eine geringere Ähnlichkeit zwischen dem Sichtpunkt 914 und dem Sichtpunkt 918 einem geringeren Verlust entspricht. Bei mindestens einer Ausführungsform wird Verlust des am weitesten entfernten Nachbarn basierend auf einer folgenden symbolischen mathematischen Gleichung berechnet:

$$L_{fn} = \min\left(0, t - d\left(v_{l_1}, v_{l_3}\right)\right)$$

[0073] wobei L_{fn} der Verlust des am weitesten entfernten Nachbarn ist, v_{l_1} ein Sichtpunkt eines Ankerbildes ist, v_{l_3} ein Sichtpunkt ein Bild des am weitesten entfernten Nachbarn ist, min() eine Funktion ist, die ein Minimum zwischen zwei Werten bestimmt, $d(v_{l_1}, v_{l_3})$ eine Funktion ist, die einen Abstand (z. B. L1-Abstand, L2-Abstand, Kosinus-Abstand und/oder Variationen davon) oder eine Differenz zwischen Sichtpunkten bestimmt, und t eine Mindestschwellenwert ist, der von einem oder mehreren Prozessen bestimmt wird. Bei mindestens

einer Ausführungsform t werden von einem Diskriminator ein oder mehrere Systeme, die mit einem Diskriminator assoziiert sind, und/oder Variationen davon bestimmt. Bei mindestens einer Ausführungsform ist t ein Parameter, der von einem oder mehreren Prozessen vor Beginn des Trainings eingestellt wird, und ist ein Schwellenwert, für den der Verlust des am weitesten entfernten Nachbarn als verlustfrei angesehen wird.

[0074] Bei mindestens einer Ausführungsform werden nächste und/oder am weitesten entfernte Nachbarn nicht deterministisch ausgewählt. Bei mindestens einer Ausführungsform wird jeder Kante eine Wahrscheinlichkeit zugewiesen, um als nächster und/oder am weitesten entfernter Nachbar eines Ankerbildes ausgewählt zu werden. Bei mindestens einer Ausführungsform sind Wahrscheinlichkeiten eines nächsten Nachbarn umgekehrt proportional zu Kantengewichten (z. B. hat ein Knoten, der das niedrigste Kantengewicht aufweist, der mit einem Ankerbild verbunden ist, die höchste Wahrscheinlichkeit, ausgewählt zu werden). Bei mindestens einer Ausführungsform sind Wahrscheinlichkeiten eines am weitesten entfernten Nachbarn direkt proportional zu Kantengewichten (z. B. weist ein Knoten, der das höchste Kantengewicht aufweist, das mit einem Ankerbild verbunden ist, die höchste Wahrscheinlichkeit, ausgewählt zu werden, auf).

[0075] Fig. 10 veranschaulicht ein Diagramm 1000, das den Entflechtungsverlust abbildet, gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform wird ein Entflechtungsverlust verwendet, um ein oder mehrere neuronale Netzwerke zu trainieren, die mit einem Generator, wie etwa einem Generator 1002, assoziiert sind. Bei mindestens einer Ausführungsform wird der Entflechtungsverlust mit einer oder mehreren anderen Verlustfunktionen verwendet, um Parameter, die mit dem Generator 1002 assoziiert sind, zu verfeinern. Bei mindestens einer Ausführungsform arbeitet der Entflechtungsverlust, wenn Generatorparameter aktualisiert werden. Bei mindestens einer Ausführungsform arbeitet der Entflechtungsverlust in einer Generatoraktualisierungsstufe.

[0076] Bei mindestens einer Ausführungsform ist der Generator 1002 mit einem generativen Modell, wie etwa einem Variations-Autoencoder (VAE), einem differenzierbaren Renderer, einem generativen kontradiktorischen Netzwerk oder einem Renderer assoziiert. Bei mindestens einer Ausführungsform ist der Generator 1002 Teil eines oder mehrerer neuronaler Netzwerke, die dazu trainiert sind, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen, und das eine oder die mehreren neuronalen Netzwerke umfassen verschiedene Parameter, die mit einem oder mehreren Prozessen des einen oder der mehreren neuronalen Netzwerke assoziiert sind. Bei mindestens einer Ausführungsform ist der Diskriminator 1004 Teil eines oder mehrerer neuronaler Netzwerke, die dazu trainiert sind, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten, und das eine oder die mehreren neuronalen Netzwerke umfassen verschiedene Parameter, die mit einem oder mehreren Prozessen des einen oder der mehreren neuronalen Netzwerke assoziiert sind.

[0077] Bei mindestens einer Ausführungsform werden ein erster Sichtpunkt 1006A und ein erster Satz von Erscheinungsbildattributen 1008A erhalten. Bei mindestens einer Ausführungsform werden der erste Sichtpunkt 1006A und der erste Satz von Erscheinungsbildattributen 1008A zufällig erzeugt. Bei mindestens einer Ausführungsform werden der erste Sichtpunkt 1006A und der erste Satz von Erscheinungsbildattributen 1008A von einem oder mehreren Prozessen erhalten, die dem Generator 1002 und dem Diskriminator 1004 assoziiert sind. Bei mindestens einer Ausführungsform erzeugt der Generator 1002 ein Bild 1010A basierend auf dem ersten Sichtpunkt 1006A und dem ersten Satz von Erscheinungsbildattributen 1008A. Bei mindestens einer Ausführungsform wird das Bild 1010A in den Diskriminator 1004 eingegeben. Bei mindestens einer Ausführungsform führt der Diskriminator 1004 einen oder mehrere Prozesse aus und bestimmt einen ersten vorhergesagten Sichtpunkt 1012A und einen ersten vorhergesagten Satz von Erscheinungsbildattributen 1012B basierend auf dem Eingabebild 1010A.

[0078] Bei mindestens einer Ausführungsform wird ein zweiter Satz von Erscheinungsbildattributen 1008B erhalten. Bei mindestens einer Ausführungsform wird der zweite Satz von Erscheinungsbildattributen 1008B zufällig erzeugt. Bei mindestens einer Ausführungsform wird der zweite Satz von Erscheinungsbildattributen 1008B von einem oder mehreren Prozessen erhalten, die mit dem Generator 1002 und dem Diskriminator 1004 assoziiert sind. Bei mindestens einer Ausführungsform erzeugt der Generator 1002 ein Bild 1010B basierend auf dem ersten Sichtpunkt 1006A und dem zweiten Satz von Erscheinungsbildattributen 1008B. Bei mindestens einer Ausführungsform wird das Bild 1010B in den Diskriminator 1004 eingegeben. Bei mindestens einer Ausführungsform führt der Diskriminator 1004 einen oder mehrere Prozesse aus und bestimmt einen zweiten vorhergesagten Sichtpunkt 1014A und einen zweiten vorhergesagten Satz von Erscheinungsbildattributen 1014B basierend auf dem Eingabebild 1010B.

[0079] Bei mindestens einer Ausführungsform wird ein zweiter Sichtpunkt 1006B erhalten. Bei mindestens einer Ausführungsform wird der zweite Sichtpunkt 1006B zufällig erzeugt. Bei mindestens einer Ausführungsform wird der zweite Sichtpunkt 1006B von einem oder mehreren Prozessen erhalten, die mit dem Generator 1002 und dem Diskriminator 1004 assoziiert sind. Bei mindestens einer Ausführungsform erzeugt der Generator 1002 ein Bild 1010C basierend auf dem zweiten Sichtpunkt 1006B und dem ersten Satz von Erscheinungsbildattributen 1008A. Bei mindestens einer Ausführungsform wird das Bild 1010C in den Diskriminator 1004 eingegeben. Bei mindestens einer Ausführungsform führt der Diskriminator 1004 einen oder mehrere Prozesse aus und bestimmt einen dritten vorhergesagten Sichtpunkt 1016A und einen dritten vorhergesagten Satz von Erscheinungsbildattributen 1016B basierend auf dem Eingabebild 1010C.

[0080] Bei mindestens einer Ausführungsform umfasst der Entflechtungsverlust einen z-Rekonstruktionsverlust und einen Sichtpunkt-Rekonstruktionsverlust. Bei mindestens einer Ausführungsform wird z-Rekonstruktionsverlust berechnet, indem eine von einem Diskriminator, wie dem Diskriminator 1004, erzeugte Vorhersage eines Satzes von Erscheinungsbildparametern für ein Bild mit einem Eingabesatz von Erscheinungsbildparametern verglichen wird, die zum Erzeugen des Bildes verwendet werden, der von einem Generator, wie dem Generator 1002, erzeugt wird, wobei sich ein geringerer Verlust ergibt, wenn die Vorhersage eines Satzes von Erscheinungsbildparametern und der eingegebene Satz von Erscheinungsbildparametern ähnlicher sind, und sich ein höherer Verlust ergibt, wenn die Vorhersage eines Satzes von Erscheinungsbildparametern und der eingegebene Satz von Erscheinungsbildparametern weniger ähnlich sind. Bei mindestens einer Ausführungsform wird Sichtpunktrekonstruktionsverlust berechnet, indem eine Vorhersage, die von einem Diskriminator, wie dem Diskriminator 1004, erzeugt wird, eines Sichtpunkts für ein Bild mit einem eingegebenen Sichtpunkt verglichen wird, der eingesetzt wird, um das Bild zu erzeugen, das von einem Generator, wie dem Generator 1002, erzeugt wird, wobei sich ein geringerer Verlust ergibt, wenn die Vorhersage eines Sichtpunkts und der eingegebene Sichtpunkt ähnlicher sind, und sich ein höherer Verlust ergibt, wenn die Vorhersage eines Sichtpunkts und der eingegebene Sichtpunkt weniger ähnlich sind.

[0081] Bei mindestens einer Ausführungsform werden der z-Rekonstruktionsverlust und der Sichtpunkt-Rekonstruktionsverlust für jeden Satz vorhergesagter Sichtpunkte und Erscheinungsbildattribute (z. B. erster vorhergesagter Sichtpunkt 1012A und erster vorhergesagter Satz von Erscheinungsbildattributen 1012B basierend auf dem Eingabebild 1010A, zweiter vorhergesagter Sichtpunkt 1014A und zweiter vorhergesagter Satz von Erscheinungsbildattributen 1014B basierend auf dem Eingabebild 1010B, und dritter vorhergesagter Sichtpunkt 1016A und dritter vorhergesagter Satz von Erscheinungsbildattributen 1016B basierend auf dem Eingabebild 1010C) berechnet. Bei mindestens einer Ausführungsform werden der z-Rekonstruktionsverlust und der Sichtpunkt-Rekonstruktionsverlust verwendet, um den Entflechtungsverlust zu bestimmen. Bei mindestens einer Ausführungsform werden auch zusätzliche Verlustfunktionen als Teil des Entflechtungsverlusts berechnet. Bei mindestens einer Ausführungsform wird der Entflechtungsverlust basierend auf einer folgenden symbolischen mathematischen Gleichung berechnet:

$$\begin{aligned} &\textit{Entflechtungsverlust} \\ &\quad = \sum \textit{Sichtpunkt} - \textit{Rekonstruktionsverlust} + \textit{z reconstruction loss} \\ &\quad + \textit{andere Verlustfunktionen} \end{aligned}$$

[0082] Bei mindestens einer Ausführungsform wird der Entflechtungsverlust eingesetzt, um einen oder mehrere Parameter, die mit dem Generator 1002 assoziiert sind, zu verfeinern. Bei mindestens einer Ausführungsform werden Parameter des Generators 1002 derart aktualisiert, dass der Verlust mindestens von mindestens dem Entflechtungsverlust minimiert wird.

[0083] **Fig. 11** veranschaulicht ein Diagramm 1100, das das Kalibrieren eines neuronalen Netzwerks abbildet, gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform wird ein Diskriminator mit einem Satz von Bildern trainiert und wird mit einem Teil des Satzes von Bildern, der Ground-Truth-Anmerkungen umfasst, kalibriert. Bei mindestens einer Ausführungsform ist ein Diskriminator 1104 Teil eines oder mehrerer neuronaler Netzwerke, die dazu trainiert sind, eine Ausrichtung eines Objekts innerhalb eines Bildes mindestens zum Teil basierend auf einer oder mehreren anderen Charakteristiken des Objekts als der Ausrichtung des Objekts zu identifizieren. Bei mindestens einer Ausführungsform wird der Diskriminator 1106 mit einer Sammlung von Bildern trainiert, um Ausrichtungen anderer Objekte derselben Kategorie abzuleiten, die innerhalb anderer Bilder aufgenommen sind.

[0084] Bei mindestens einer Ausführungsform ist der Diskriminator 1104 dazu trainiert, einen Sichtpunkt eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise mit einer Sammlung von Bildern zu identifizieren, wie etwa jenen, die an anderer Stelle in dieser Offenbarung beschrieben sind. Bei mindestens einer

Ausführungsform wird der Diskriminator 1104 dazu trainiert, eine Ausrichtung eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden. Bei mindestens einer Ausführungsform wird der Diskriminator 1106 mindestens zum Teil basierend auf Berechnungsverlusten trainiert, wie etwa einem oder mehreren von: einem generativen Konsistenzverlust, einem Symmetrieverlust, einem Verlust eines nächsten oder am weitesten entfernten Nachbarn, einem Entflechtungsverlust, der mit denjenigen in Verbindung mit den **Fig. 4-10** beschriebenen übereinstimmen kann. Bei mindestens einer Ausführungsform wird der Diskriminator 1104 mit einer Sammlung von Bildern trainiert, denen Ground-Truth-Anmerkungen fehlen oder und wobei die Ground-Truth-Anmerkungen anderweitig nicht verfügbar sind.

[0085] Bei mindestens einer Ausführungsform wird eine Objektbildsammlung 1102 erhalten, um den Diskriminator 1104 zu kalibrieren. Bei mindestens einer Ausführungsform umfasst die Objektbildsammlung 1102 Bilder mit Ground-Truth-Anmerkungen. Bei mindestens einer Ausführungsform umfasst die Objektbildsammlung 1102 Bilder, die Objekte abbilden, für deren Analyse und Bestimmung von Sichtpunkten der Diskriminator 1104 trainiert ist. Bei mindestens einer Ausführungsform ist die Objektbildsammlung 1102 ein Teil einer Sammlung von Bildern, die zum Trainieren des Diskriminators 1104 verwendet werden. Bei mindestens einer Ausführungsform wird der Diskriminator 1104 mit einer Sammlung von Bildern trainiert, die sich von der Objektbildsammlung 1102 unterscheidet. Bei mindestens einer Ausführungsform erhält der Diskriminator 1104 ein Bild der Objektbildsammlung 1102 und bestimmt einen Sichtpunkt 1106 für das Bild. Obwohl der Diskriminator 1104 trainiert ist, stimmt bei mindestens einer Ausführungsform der Sichtpunkt 1106 nicht mit einem Ground-Truth-Sichtpunkt eines Bildes der Objektbildsammlung 1102 überein. Bei mindestens einer Ausführungsform ist der Sichtpunkt 1106 ein korrekter Sichtpunkt eines Bildes der Objektbildsammlung 1102, das in eine oder mehrere Dimensionen übersetzt ist.

[0086] Bei mindestens einer Ausführungsform wird ein lineares Modell 1108 bestimmt, das Sichtpunkte, die von dem Diskriminator 1104 für Bilder der Objektbildsammlung 1102 bestimmt werden, in ihre jeweiligen korrekten Positionen, die durch Ground-Truth-Anmerkungen angegeben sind, übersetzt. Bei mindestens einer Ausführungsform ist das lineare Modell 1108 eine lineare Funktion, die durch einen Vergleich von Sichtpunkten, die von dem Diskriminator 1104 für Bilder der Objektbildsammlung 1102 erzeugt wurden, mit Ground-Truth-Anmerkungen von Sichtpunkten für die Bilder der Objektbildsammlung 1102 bestimmt wird. Bei mindestens einer Ausführungsform wird das lineare Modell 1108 durch einen oder mehrere Prozesse bestimmt, wie etwa verschiedene Regressionsalgorithmen, mathematische Prozesse und/oder Variationen davon. Bei mindestens einer Ausführungsform wird das lineare Modell 1108 derart bestimmt, dass ein Sichtpunkt, der von dem Diskriminator 1104 für ein Bild bestimmt wird, übersetzt und/oder korrigiert werden kann, um mit einem Ground-Truth-Sichtpunkt für das Bild übereinzustimmen. Bei mindestens einer Ausführungsform kalibriert das lineare Modell 1108 einen Sichtpunkt, der von dem Diskriminator 1104 bestimmt wird, für ein Bild in ein Koordinatensystem eines Ground-Truth-Sichtpunkts für das Bild. Bei mindestens einer Ausführungsform übersetzt das lineare Modell 1108 eine Nullposition eines Sichtpunkts, der von dem Diskriminator 1104 für ein Bild bestimmt wird, in eine Nullposition eines Ground-Truth-Sichtpunkts für das Bild.

[0087] Bei mindestens einer Ausführungsform wird das lineare Modell 1108 eingesetzt, um den Sichtpunkt 1106 zu kalibrieren oder anderswie zu korrigieren, um einen korrigierten Sichtpunkt 1110 zu erzeugen. Bei mindestens einer Ausführungsform ist der korrigierte Sichtpunkt 1110 der Sichtpunkt 1106, der übersetzt wird, um mit einem Ground-Truth-Sichtpunkt übereinzustimmen. Bei mindestens einer Ausführungsform wird das lineare Modell 1108 eingesetzt, um Sichtpunkte, die von dem Diskriminator 1104 bestimmt/erzeugt werden, für Bilder der Objektbildsammlung 1102 zu kalibrieren oder anderswie zu korrigieren, um mit Ground-Truth-Anmerkungen von Sichtpunkten der Bilder der Objektbildsammlung 1102 übereinzustimmen.

[0088] Fig. 12 veranschaulicht ein Diagramm 1200, das eine Ableitung abbildet, gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform ist ein Diskriminator als Teil eines oder mehrerer Systeme, die mit einem Sicherheitssystem eines Kraftfahrzeugs assoziiert sind, dazu trainiert, Sichtpunkte von Bildern anderer Kraftfahrzeuge, die von einer Kamera aufgenommen werden, die mit dem Kraftfahrzeug assoziiert ist, derart abzuleiten, dass diese abgeleiteten Sichtpunkte eingesetzt werden, um eine oder mehrere Operationen in Verbindung mit dem Kraftfahrzeug auszuführen, wie beispielsweise das Bremsen des Kraftfahrzeugs, um einem anderen Kraftfahrzeug auszuweichen, oder das Lenken des Kraftfahrzeugs, um einem anderen Kraftfahrzeug auszuweichen. Bei mindestens einer Ausführungsform ist ein Diskriminator 1204 dazu trainiert, einen Sichtpunkt eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise mit einer Sammlung von Bildern, wie etwa jenen zu identifizieren, die an anderer Stelle in dieser Offenbarung beschrieben sind. Bei mindestens einer Ausführungsform wird der Diskriminator 1204 dazu trainiert, eine

Ausrichtung eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen als Teil eines Trainings, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, berechnet werden. Bei mindestens einer Ausführungsform wird der Diskriminator 1204 mindestens zum Teil basierend auf des Berechnens eines generativen Konsistenzverlusts, eines Symmetrieverlusts, eines Verlusts des nächsten und des am weitesten entfernten Nachbarn und eines Entflechtungsverlusts trainiert, die mit denjenigen übereinstimmen können, die in Verbindung mit den **Fig. 4-10** beschrieben sind. Bei mindestens einer Ausführungsform wird der Diskriminator 1204 mit einer Sammlung von Bildern von Autos trainiert. Bei mindestens einer Ausführungsform ist der Diskriminator 1204 dazu trainiert, einen Sichtpunkt eines Autos innerhalb eines Bildes zu identifizieren.

[0089] Bei mindestens einer Ausführungsform ist der Diskriminator 1204 Teil eines oder mehrerer Systeme eines Kraftfahrzeugs 1210. Bei mindestens einer Ausführungsform ist das Kraftfahrzeug 1210 ein autonomes Fahrzeug. Bei mindestens einer Ausführungsform ist das Kraftfahrzeug 1210 ein Fahrzeug, das von einem Bediener betrieben wird. Bei mindestens einer Ausführungsform umfasst das Kraftfahrzeug 1210 ein oder mehrere Systeme, die den Diskriminator 1204 umsetzen. Bei mindestens einer Ausführungsform umfasst das Kraftfahrzeug 1210 ein oder mehrere Systeme, die mit dem Diskriminator 1204 assoziiert sind. Bei mindestens einer Ausführungsform umfasst das Kraftfahrzeug 1210 ein oder mehrere Systeme, die es dem Kraftfahrzeug 1210 ermöglichen, aus der Ferne auf den Diskriminator 1204, der auf verschiedenen entfernten und/oder lokalen Systemen umgesetzt sein kann, zuzugreifen und ihn einzusetzen.

[0090] Bei mindestens einer Ausführungsform umfasst das Kraftfahrzeug 1210 eine Vielzahl von Kameras. Bei mindestens einer Ausführungsform befindet sich eine Kamera 1202 an Bord des Kraftfahrzeugs 1210. Bei mindestens einer Ausführungsform ist die Kamera 1202 eine Kamera, auf die das Kraftfahrzeug 1210 aus der Ferne zugreifen kann. Bei mindestens einer Ausführungsform nimmt die Kamera 1202 ein Bild 1202A auf oder erhält es anderswie. Bei mindestens einer Ausführungsform ist das Kraftfahrzeug 1210 ein Fahrzeug, das in einer Umgebung betrieben wird, die andere Kraftfahrzeuge umfasst, mit denen das Kraftfahrzeug 1210 eine oder mehrere Aktionen ausführen muss (z. B. einem Kraftfahrzeug das Vorbeifahren gestatten, ein Kraftfahrzeug überholen, Bremsen für ein entgegenkommendes Kraftfahrzeug, Lenken, um einem entgegenkommenden Fahrzeug auszuweichen, und/oder Variationen davon). Bei mindestens einer Ausführungsform ist das Bild 1202A ein Bild eines anderen Kraftfahrzeugs, mit dem das Kraftfahrzeug 1210 interagieren muss. Bei mindestens einer Ausführungsform ist das Bild 1202A ein Bild, das von einer Bordkamera des Kraftfahrzeugs 1210 erfasst wird, während das Kraftfahrzeug 1210 in einer Umgebung betrieben wird.

[0091] Bei mindestens einer Ausführungsform wird das Bild 1202A in den Diskriminator 1204, der einen Sichtpunkt 1206 für das Bild 1202A bestimmt, eingegeben. Bei mindestens einer Ausführungsform ist der Sichtpunkt 1206 ein Sichtpunkt eines in Bild 1202A abgebildeten Autos. Bei mindestens einer Ausführungsform ist ein Sicherheitssystem 1208 Teil des Kraftfahrzeugs 1210. Bei mindestens einer Ausführungsform umfasst das Sicherheitssystem 1208 ein oder mehrere Systeme, die dazu konfiguriert sind, dem Kraftfahrzeug 1210 Unterstützung bereitzustellen. Bei mindestens einer Ausführungsform umfasst das Sicherheitssystem 1208 ein oder mehrere Systeme, die dazu konfiguriert sind, ein oder mehrere Systeme des Kraftfahrzeugs 1210 zu betreiben, wie etwa einen Bremsaktuator 1208A und einen Lenkaktuator 1208B, die Komponenten sind, die dazu konfiguriert sind, jeweils Bremsen des Kraftfahrzeugs 1210 und Lenkung des Kraftfahrzeugs 1210 zu betreiben. Bei mindestens einer Ausführungsform sind der Bremsaktuator 1208A und der Lenkaktuator 1208B gleich wie der Bremsaktuator 2148 bzw. der Lenkaktuator 2156. Bei mindestens einer Ausführungsform kann das Kraftfahrzeug 1210 in Übereinstimmung mit an anderer Stelle beschriebenen Techniken, wie in **Fig. 21**, umgesetzt sein.

[0092] Bei mindestens einer Ausführungsform erhält das Sicherheitssystem 1208 den Sichtpunkt 1206. Bei mindestens einer Ausführungsform bestimmt das Sicherheitssystem 1208 basierend auf dem Sichtpunkt 1206 eine Fahrtrichtung eines Autos, das in Bild 1202A abgebildet ist. Bei mindestens einer Ausführungsform bestimmt das Sicherheitssystem 1208 basierend auf dem Sichtpunkt 1206, ob der Bremsaktuator 1208A oder der Lenkaktuator 1208B eingesetzt werden soll. Bei mindestens einer Ausführungsform, falls das Sicherheitssystem 1208 bestimmt, dass ein Auto, das in dem Bild 1202A abgebildet ist, in einer Richtung in Bezug auf das Kraftfahrzeug 1210 fährt, die erfordert, dass das Kraftfahrzeug 1210 bremst, aktiviert das Sicherheitssystem 1208 den Bremsaktuator 1208A, um das Kraftfahrzeug 1210 derart so zu bremsen, dass das Kraftfahrzeug 1210 jegliche potenziellen Sicherheitsprobleme vermeidet, die sich aus dem Fahren des Autos, das in dem Bild 1202A abgebildet ist, ergeben. Bei mindestens einer Ausführungsform, falls das Sicherheitssystem 1208 bestimmt, dass ein Auto, das in dem Bild 1202A abgebildet ist, in eine Richtung in Bezug auf das Kraftfahrzeug 1210 fährt, die erfordert, dass das Kraftfahrzeug 1210 lenkt, aktiviert das Sicher-

heitssystem 1208 den Lenkaktuator 1208B, um das Kraftfahrzeug 1210 derart zu lenken, dass das Kraftfahrzeug 1210 jegliche potenziellen Sicherheitsprobleme vermeidet, die sich aus dem Fahren Autos, das in dem Bild 1202A abgebildet ist, ergeben.

[0093] **Fig. 13** zeigt ein veranschaulichendes Beispiel eines Prozesses 1300 zum Trainieren eines neuronalen Netzwerks zum Vorhersagen eines Sichtpunkts eines Objekts innerhalb eines Bildes gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform werden ein Teil oder der gesamte Prozess 1300 (oder ein anderer Prozess, der hierin beschrieben ist, oder Variationen und/oder Kombinationen davon) unter der Steuerung eines oder mehrerer Computersysteme ausgeführt, die mit computerausführbaren Anweisungen konfiguriert sind, die als Code umgesetzt werden können (zum Beispiel computerausführbare Anweisungen, ein oder mehrere Computerprogramme oder eine oder mehrere Anwendungen), die gemeinsam auf einem oder mehreren Prozessoren, durch Hardware, Software oder Kombinationen davon ausführen. Bei mindestens einer Ausführungsform ist der Code auf einem computerlesbaren Speichermedium gespeichert, zum Beispiel in Form eines Computerprogramms, das eine Vielzahl von computerlesbaren Anweisungen umfasst, die von einem oder mehreren Prozessoren ausgeführt werden können. Ein computerlesbares Speichermedium ist bei mindestens einer Ausführungsform ein nichtflüchtiges computerlesbares Medium. Bei mindestens einer Ausführungsform werden mindestens einige computerlesbare Anweisungen, die zum Ausführen des Prozesses 1300 verwendbar sind, nicht ausschließlich unter Verwendung von transienten Signalen (z. B. einer sich ausbreitenden transienten elektrischen oder elektromagnetischen Übertragung) gespeichert. Ein nichtflüchtiges computerlesbares Medium beinhaltet nicht notwendigerweise nichtflüchtige Datenspeicherschaltungen (z. B. Puffer, Caches und Warteschlangen) innerhalb von Transceivern für flüchtige Signale. Bei mindestens einer Ausführungsform wird der Prozess 1300 mindestens zum Teil auf einem Computersystem, wie zum Beispiel denjenigen, die an anderer Stelle in dieser Offenbarung beschrieben sind, ausgeführt. Bei mindestens einer Ausführungsform trainiert ein erstes Computersystem ein oder mehrere neuronale Netzwerke und ein zweites Computersystem folgert (sagt z.B. einen Sichtpunkt eines Objekts innerhalb eines Bildes vorher) unter Verwendung des einen oder der mehreren neuronalen Netzwerke. Bei mindestens einer Ausführungsform sind Techniken, die in Verbindung mit den **Fig. 1**, **Fig. 12** und **Fig. 14** beschrieben sind, auf den Prozess 1300 anwendbar.

[0094] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1300 ausführt, ausführbaren Code, um eine Sammlung von einem oder mehreren Bildern eines Objekttyps zu erhalten 1302. Bei mindestens einer Ausführungsform wird eine Sammlung von einem oder mehreren Bildern verwendet, um ein oder mehrere neuronale Netzwerke zu trainieren, um Ausrichtungen von Objekten innerhalb von Bildern zu identifizieren. Bei mindestens einer Ausführungsform wird eine Sammlung von Bildern derart kategorisiert oder gekennzeichnet, dass jedes einen gleichen Typ oder dieselbe Kategorie von Objekten zeigt. Bei mindestens einer Ausführungsform ist eine Sammlung von Bildern eine Sammlung von Bildern von Autos, die unterschiedliche Arten von Autos in unterschiedlichen Ausrichtungen, bei unterschiedlichem Wetter und unter unterschiedlicher Beleuchtung beinhalten können. Bei mindestens einer Ausführungsform beinhaltet eine Sammlung von Autos Bilder desselben Autos oder desselben Autotyps in unterschiedlichen Ausrichtungen. Bei mindestens einer Ausführungsform fehlen mindestens einem Teil einer Sammlung von Trainingsbildern Ground-Truth-Anmerkungen, die die Ausrichtung von Objekten innerhalb derartiger Trainingsbilder spezifizieren. Bei mindestens einer Ausführungsform fehlen allen Bildern einer Sammlung von Bildern Ground-Truth-Anmerkungen, die Azimut, Elevation und Kippen von Objekten innerhalb von Bildern der Sammlung spezifizieren. Bei mindestens einer Ausführungsform beinhaltet eine Sammlung von Bildern ein oder mehrere synthetische Bilder, wie beispielsweise ein Bild, das von einem generativen kontradiktorischen Netzwerk (Generative Adversarial Network, GAN) angelegt wird. Bei mindestens einer Ausführungsform sind alle Bilder einer Sammlung von Bildern echte Bilder im Gegensatz zu jenen, die aus einem generativen Modell, wie einem VariationalAutoencoder (VAE), einem differenzierbaren Renderer, einem generativen kontradiktorischen Netzwerk (GAN) oder einem Renderer synthetisiert oder angelegt werden. Bei mindestens einer Ausführungsform wird eine Sammlung von Bildern von einer Website, die Bilder nach Kategorie sortiert, gesammelt und aggregiert.

[0095] Bei mindestens einer Ausführungsform betrifft die Ausrichtung eines Objekts innerhalb eines Bildes eine dreidimensionale Ausrichtung eines Objekts, das in innerhalb eines zweidimensionalen Bildes aufgenommen wird. Bei mindestens einer Ausführungsform wird eine Kamera verwendet, um ein zweidimensionales Bild eines echten Autos, das sich in einer spezifischen Ausrichtung in Bezug auf die Kamera befindet aufzunehmen. Bei mindestens einer Ausführungsform wird die Ausrichtung eines Objekts (z. B. Sichtpunkt) in einem Satz von Parametern codiert, der einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfasst. Bei mindestens einer Ausführungsform wird die Ausrichtung eines Objekts als ein Satz von drei Vektoren, die die Richtung des Objekts in Bezug auf eine x-, y- und z-Achse definieren, codiert.

[0096] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1300 ausführt, ausführbaren Code zum Trainieren 1304 eines oder mehrerer neuronaler Netzwerke zum Identifizieren einer Ausrichtung eines Objekts innerhalb eines Bildes, mindestens zum Teil basierend auf einer oder mehreren anderen Charakteristiken des Objekts als der Ausrichtung des Objekts. Bei mindestens einer Ausführungsform wird der Prozess 1300 auf einem Prozessor umgesetzt, der Folgendes umfasst: eine oder mehrere Schaltungen, um dabei zu helfen, ein oder mehrere neuronale Netzwerke dazu zu trainieren, eine Ausrichtung eines Objekts innerhalb eines Bildes mindestens zum Teil basierend auf einer oder mehreren anderen Charakteristiken des Objekt als der Ausrichtung des Objekts zu identifizieren. Bei mindestens einer Ausführungsform werden ein oder mehrere neuronale Netzwerke mit einer Sammlung von Bildern trainiert, um Ausrichtungen anderer Objekte derselben Kategorie, die innerhalb anderer Bilder aufgenommen sind, abzuleiten. Bei mindestens einer Ausführungsform wird ein neuronales Netzwerk mit einer Sammlung von Bildern von Flugzeugen trainiert und wird, sobald es trainiert ist, dazu verwendet, Ausrichtungen anderer Flugzeuge innerhalb anderer Bilder abzuleiten.

[0097] Bei mindestens einer Ausführungsform wird ein neuronales Netzwerk dazu trainiert, eine Ausrichtung eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise mit einer Sammlung von Bildern, wie sie an anderer Stelle in dieser Offenbarung beschrieben sind, zu identifizieren. Bei mindestens einer Ausführungsform wird ein neuronales Netzwerk dazu trainiert, eine Ausrichtung eines Objekts in einem Bild auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes (z. B. Sammlung von Bildern) bewerten, als Teil des Trainings berechnet werden. Bei mindestens einer Ausführungsform wird ein neuronales Netzwerk mindestens zum Teil basierend auf der Berechnung eines generativen Konsistenzverlusts, eines Symmetrieverlusts, eines Verlusts eines nächsten und eines am weitesten entfernten Nachbarn und eines Entflechtungsverlusts trainiert, die mit denjenigen, die in Verbindung mit den **Fig. 4-10** beschrieben sind, übereinstimmen können. Bei mindestens einer Ausführungsform wird ein neuronales Netzwerk mit einer Sammlung von Bildern trainiert, der Ground-Truth-Anmerkungen fehlen oder wobei Ground-Truth-Anmerkungen anderswie nicht verfügbar sind (z. B. werden derartige Daten einem neuronalen Netzwerk während des Trainings vorenthalten). Bei mindestens einer Ausführungsform wird ein neuronales Netzwerk dazu trainiert, aus einem Objekt innerhalb eines ersten Bildes, das eine vorhergesagte Ausrichtung aufweist, ein zweites Bild, das dieselbe Ausrichtung aufweist, zu erzeugen.

[0098] Bei mindestens einer Ausführungsform umfasst ein Systemumsetzungsprozess 1300 einen oder mehrere Prozessoren zum Berechnen von Parametern zum Trainieren eines oder mehrerer neuronaler Netzwerke zum Identifizieren einer Ausrichtung eines Objekts in einem Bild mindestens zum Teil basierend auf einer oder mehreren anderen Charakteristiken des Objekts als die Ausrichtung des Objekts; und einen oder mehrere Speicher zum Speichern der Parameter. Bei mindestens einer Ausführungsform werden ein oder mehrere neuronale Netzwerke trainiert, um eine Ausrichtung eines Objekts innerhalb eines Bildes unter Verwendung einer Sammlung von Bildern unterschiedlicher Objekte derselben Kategorie wie das Objekt zu identifizieren (z. B. wird ein neuronales Netzwerk zum Ableiten von Sichtpunkten von Fahrzeugen) mit einer Sammlung von Bildern, die als Fahrzeuge gekennzeichnet sind, trainiert).

[0099] Bei mindestens einer Ausführungsform werden ein oder mehrere neuronale Netzwerke selbstüberwacht mit einer Sammlung von Bildern unterschiedlicher Objekte derselben Kategorie als ein Objekt eines Bildes, das abgeleitet werden soll, trainiert. Bei mindestens einer Ausführungsform können unterschiedliche Objekte derselben Kategorie unterschiedliche Bilder betreffen, die ein oder mehrere Bilder eines ersten Autos in einer oder mehreren Ausrichtungen, ein oder mehrere Bilder eines unterschiedlichen zweiten Autos in einer oder mehreren Ausrichtungen sein können, und so fort. Bei mindestens einer Ausführungsform ist ein Bild eines Objekts, das abgeleitet werden soll, in einer Sammlung von Bildern enthalten, die verwendet wird, um ein oder mehrere neuronale Netzwerke auf Ableitungsausrichtungen zu trainieren. Bei mindestens einer Ausführungsform werden ein oder mehrere neuronale Netzwerke auf selbstüberwachte Weise trainiert, indem mindestens ein Satz von Verlustfunktionen verwendet wird, um eine oder mehrere Charakteristiken von Objekten in Bildern zu bewerten. Bei mindestens einer Ausführungsform betreffen eine oder mehrere Charakteristiken von Objekten Eigenschaften von Objekten, die verwendet werden können, um Ausrichtungen abzuleiten. Bei mindestens einer Ausführungsform wird ein neuronales Netzwerk trainiert, indem eines oder mehrere der Folgenden berechnet werden: generativer Konsistenzverlust; Symmetrieverlust; Verlust des nächsten Nachbarn und des am weitesten entfernten Nachbarn; und Entflechtungsverlust. Bei mindestens einer Ausführungsform wird ein auf selbstüberwachte Weise trainiertes neuronales Netzwerk dazu trainiert, synthetische Bilder von Objekten mit einer spezifischen Ausrichtung zu erzeugen, die dieselbe Ausrichtung wie eine vorhergesagte Ausrichtung eines Eingabebildes sein kann. Bei mindestens einer Ausführungsform wird ein synthetisches Bild unter Verwendung eines tiefen generativen Modells, wie eines

Variational Autoencoders (VAE), eines differenzierbaren Renderers, eines generativen kontradiktorischen Netzwerks (GAN) oder eines Renderers, angelegt. Bei mindestens einer Ausführungsform kann ein Objekt, dessen Ausrichtung abgeleitet werden soll, ein Fahrzeug, ein Flugzeug, eine Drohne, ein Mensch, ein Gesicht (z. B. eines Menschen oder eines Tieres) und mehr sein.

[0100] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1300 ausführt, einen ausführbaren Code, um ein zweites Bild zu erhalten 1306. Bei mindestens einer Ausführungsform ist ein zweites Objekt in einem zweiten Bild vom gleichen Typ wie eine Sammlung von Bildern, die dazu verwendet werden, ein oder mehrere neuronale Netzwerke zu trainieren. Bei mindestens einer Ausführungsform wird ein zweites Bild einem neuronalen Netzwerk zur Schlussfolgerung bereitgestellt, um eine zweite Ausrichtung vorherzusagen. Bei mindestens einer Ausführungsform werden Bilder von einer Kamera erhalten, die Standbilder und/oder ein Video aufnimmt, das aus mehreren Einzelbildern besteht, die mit einer variablen oder festen Rate aufgenommen werden. Bei mindestens einer Ausführungsform umfasst ein Video eine Anzahl von Einzelbildern (z. B. Bilder). Bei mindestens einer Ausführungsform trainiert ein erstes System ein oder mehrere neuronale Netzwerke, und ein zweites anderes System verwendet dieses eine oder diese mehreren neuronalen Netzwerke, um eine Ableitung auszuführen, um Ausrichtungen von Objekten innerhalb von Bildern zu identifizieren.

[0101] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1300 ausführt, ausführbaren Code zur Verwendung 1308 für ein oder mehrere neuronale Netzwerke (z. B. wie in Ziffer 1304 beschrieben trainiert), um eine zweite Ausrichtung des zweiten Objekts innerhalb des zweiten Bildes zu identifizieren. Bei mindestens einer Ausführungsform verwendet ein System einen Diskriminator, der auf selbstüberwachte Weise mit einer Sammlung von Bildern von Objekten einer spezifischen Kategorie trainiert ist, um Ausrichtungen anderer Objekte dieser Kategorie abzuleiten. Bei mindestens einer Ausführungsform werden ein oder mehrere neuronale Netzwerke mit einer Sammlung von Bildern von Autos trainiert und verwendet, um Ausrichtungen von Autos abzuleiten, die in Echtzeit von einer Kamera oder einem anderen geeigneten Video-/Bilderfassungsgerät, das an einem Fahrzeug angebracht ist, aufgenommen werden.

[0102] Bei mindestens einer Ausführungsform wird ein erstes neuronales Netzwerk unter Verwendung von selbstüberwachtem Lernen mit einer ersten Sammlung von Bildern einer ersten Kategorie trainiert, um Sichtpunkte von Objekten dieser ersten Kategorie abzuleiten, und ein zweites neuronales Netzwerk wird unter Verwendung ähnlicher/gleicher selbstüberwachter Lerntechniken mit einer zweiten Sammlung von Bildern einer zweiten Kategorie trainiert. Bei mindestens einer Ausführungsform wird ein Bild als Eingabe für ein erstes neuronales Netzwerk bereitgestellt, um eine erste Ausrichtung eines ersten Objekts einer ersten Kategorie zu erfassen, und auch als Eingabe für ein zweites neuronales Netzwerk bereitgestellt, um eine zweite Ausrichtung eines zweiten Objekts einer zweiten Kategorie zu erfassen. Bei mindestens einer Ausführungsform wird ein Eingabebild mehreren neuronalen Netzwerken bereitgestellt, die unter Verwendung von hierin beschriebenen selbstüberwachten Lerntechniken trainiert werden, um Ausrichtungen unterschiedlicher Objekte in dem Eingabebild zu identifizieren.

[0103] **Fig. 14** zeigt ein veranschaulichendes Beispiel eines Prozesses 1400 zum Trainieren eines neuronalen Netzwerks zum Vorhersagen eines Sichtpunkts eines Objekts innerhalb eines Bildes gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform werden ein Teil oder der gesamte Prozess 1400 (oder ein anderer Prozess, der hierin beschrieben ist, oder Variationen und/oder Kombinationen davon) unter der Steuerung eines oder mehrerer Computersysteme ausgeführt, die mit computerausführbaren Anweisungen konfiguriert sind, die als Code umgesetzt werden können (zum Beispiel computerausführbare Anweisungen, ein oder mehrere Computerprogramme oder eine oder mehrere Anwendungen), die gemeinsam auf einem oder mehreren Prozessoren, durch Hardware, Software oder Kombinationen davon ausführen. Bei mindestens einer Ausführungsform ist der Code auf einem computerlesbaren Speichermedium gespeichert, zum Beispiel in Form eines Computerprogramms, das eine Vielzahl von computerlesbaren Anweisungen umfasst, die von einem oder mehreren Prozessoren ausgeführt werden können. Ein computerlesbares Speichermedium ist bei mindestens einer Ausführungsform ein nichtflüchtiges computerlesbares Medium. Bei mindestens einer Ausführungsform werden mindestens einige computerlesbare Anweisungen, die zum Ausführen des Prozesses 1300 verwendbar sind, nicht ausschließlich unter Verwendung von transienten Signalen (z. B. einer sich ausbreitenden transienten elektrischen oder elektromagnetischen Übertragung) gespeichert. Ein nichtflüchtiges computerlesbares Medium beinhaltet nicht notwendigerweise nichtflüchtige Datenspeicherschaltungen (z. B. Puffer, Caches und Warteschlangen) innerhalb von Transceivern für flüchtige Signale. Bei mindestens einer Ausführungsform wird der Prozess 1400 mindestens zum Teil auf einem Computersystem wie zum Beispiel denjenigen, die an anderer Stelle in dieser Offenbarung beschrie-

ben sind, ausgeführt. Bei mindestens einer Ausführungsform trainiert ein erstes Computersystem ein oder mehrere neuronale Netzwerke und ein zweites Computersystem folgert (sagt z.B. einen Sichtpunkt eines Objekts innerhalb eines Bildes vorher) unter Verwendung des einen oder der mehreren neuronalen Netzwerke. Bei mindestens einer Ausführungsform sind Techniken, die in Verbindung mit den **Fig. 1**, **Fig. 12** und **Fig. 13** beschrieben sind, auf den Prozess 1400 anwendbar.

[0104] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1400 ausführt, ausführbaren Code, um eine Sammlung von einem oder mehreren Bildern eines Objekttyps zu erhalten 1402. Bei mindestens einer Ausführungsform kann ein Bildtyp bedeuten, dass alle Bilder einer Sammlung von Bildern Bilder sind, die Autos beinhalten. Bei mindestens einer Ausführungsform erhält ein System eine Sammlung von einem oder mehreren Bildern gemäß Techniken, die an anderer Stelle in dieser Offenbarung, wie **Fig. 13** beschrieben sind.

[0105] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1400 ausführt, einen ausführbaren Code, um ein erstes Bild aus einer Sammlung von Bildern auszuwählen 1404. Bei mindestens einer Ausführungsform werden Bilder einer Sammlung, die zufällig oder pseudozufällig aus einem Trainingssatz ausgewählt werden können, in beliebigen geeigneten Weisen zum Lernen, ausgewählt.

[0106] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1400 ausführt, ausführbaren Code zum Berechnen 1406 eines generativen Konsistenzverlusts mindestens zum Teil basierend auf dem Vergleichen eines ausgewählten Bildes mit einem Bild, das durch ein tiefes generatives Modell erzeugt wurde. Bei mindestens einer Ausführungsform wird ein generativer Konsistenzverlust unter Verwendung von Techniken berechnet, die an anderer Stelle in dieser Offenbarung, wie den **Fig. 4-7** beschrieben sind. Bei mindestens einer Ausführungsform beinhaltet ein generativer Konsistenzverlust mindestens zwei Komponenten: einen Sichtpunkt-Konsistenzverlust und einen Bildkonsistenzverlust. Bei mindestens einer Ausführungsform wird der generative Konsistenzverlust gemäß Techniken berechnet, die in Verbindung mit **Fig. 15** beschrieben sind.

[0107] Bei mindestens einer Ausführungsform wird ein Bildkonsistenzverlust mindestens zum Teil basierend auf einem ausgewählten Bild (z. B. Eingabebild) berechnet, das einem Diskriminator bereitgestellt wird, der verwendet wird, um mindestens zwei Eigenschaften aus dem Bild zu disaggregieren: einen vorhergesagten Sichtpunkt und Satz von Erscheinungsbildparametern. Bei mindestens einer Ausführungsform werden ein vorhergesagter Sichtpunkt und ein Satz von Erscheinungsbildparametern einem Generator bereitgestellt, um ein synthetisiertes Bild anzulegen. Bei mindestens einer Ausführungsform empfängt ein generativen kontradiktorischen Netzwerk (GAN) einen Sichtpunkt und einen Satz von Erscheinungsbildparametern und erzeugt ein synthetisches (z. B. falsches) Bild, das dem jeweiligen Sichtpunkt und Satz von Erscheinungsbildparametern entspricht, der bereitgestellt wurde. Bei mindestens einer Ausführungsform werden ein synthetisiertes Bild und ein Eingabebild verglichen, um den Bildkonsistenzverlust zu bestimmen. Bei mindestens einer Ausführungsform wird ein Kosinus-Abstand zwischen einem Eingabebild und einem synthetisierten Bild verglichen, um Merkmalsähnlichkeiten zu bestimmen, wobei eine größere Ähnlichkeit einem geringeren Verlust entspricht. Bei mindestens einer Ausführungsform werden L1-, L2- oder Kosinus-Abstände verwendet, um den Bildkonsistenzverlust zwischen zwei Bildern zu bestimmen.

[0108] Bei mindestens einer Ausführungsform wird ein Sichtpunkt-Konsistenzverlust (z. B. Ausrichtungs-Konsistenzverlust) mindestens zum Teil basierend auf einem Sichtpunkt eines Eingabebildes berechnet. Bei mindestens einer Ausführungsform wird ein Sichtpunkt eines Eingabebildes von einem Diskriminator abgeleitet. Bei mindestens einer Ausführungsform wird ein Sichtpunkt eines Eingabebildes basierend auf Ground-Truth-Anmerkungen, die als Teil des Trainings für mindestens einen Teil einer Sammlung von Trainingsbildern bereitgestellt werden, abgeleitet. Bei mindestens einer Ausführungsform wird ein Generator verwendet, um ein synthetisches Bild mit dem gleichen Sichtpunkt wie ein Eingabebild anzulegen. Bei mindestens einer Ausführungsform wird ein synthetisches Bild, das von einem Sichtpunkt eines Eingabebildes erzeugt wird, einem Diskriminator bereitgestellt, der einen zweiten Sichtpunkt des synthetischen Bildes bestimmt. Bei mindestens einer Ausführungsform wird ein erster Sichtpunkt eines Eingabebildes mit einem zweiten Sichtpunkt eines synthetischen Bildes, das mindestens zum Teil basierend auf des Eingabebildes erzeugt wird, verglichen. Bei mindestens einer Ausführungsform wird ein Abstand zwischen einem ersten Sichtpunkt eines Eingabebildes und einem zweiten Sichtpunkt eines synthetischen Bildes verwendet, um einen Sichtpunktkonsistenzverlust zu berechnen, wobei nähere Sichtpunkte einem geringeren Verlust entsprechen.

[0109] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1400 ausführt, ausführbaren Code zum Berechnen 1408 eines Symmetrieverlusts durch mindestens Vergleichen eines ausgewählten Bildes mit einer transformierten Version dieses ausgewählten Bildes. Bei mindestens einer Ausführungsform wird der Symmetrieverlust gemäß Techniken, die in Verbindung mit **Fig. 15** beschrieben sind, berechnet. Bei mindestens einer Ausführungsform wird ein Eingabebild aus einer Sammlung von Trainingsbildern ausgewählt. Bei mindestens einer Ausführungsform wird eine Transformation auf ein Eingabebild angewandt, um ein transformiertes Bild zu erzeugen. Bei mindestens einer Ausführungsform wird ein Eingabebild horizontal gespiegelt, um ein gespiegeltes Bild zu erzeugen. Bei mindestens einer Ausführungsform werden ein oder mehrere neuronale Netzwerke verwendet, um eine erste Ausrichtung eines Eingabebildes und eine zweite Ausrichtung eines transformierten Bildes vorherzusagen. Bei mindestens einer Ausführungsform wird eine erste Ausrichtung für ein Eingabebild vorhergesagt, und eine zweite Ausrichtung wird für eine horizontal gespiegelte Version dieses Eingabebildes vorhergesagt. Bei mindestens einer Ausführungsform wird Verlust basierend darauf berechnet, ob bestimmte Eigenschaften zutreffen. Bei mindestens einer Ausführungsform wird eine Transformation oder Inverse davon auf eine vorhergesagte Ausrichtung einer transformierten Version eines Eingabebildes angewandt. Falls bei mindestens einer Ausführungsform ein Eingabebild um Winkel (φ,θ,ψ) gedreht wird, um ein transformiertes Bild zu erzeugen, dann kann eine abgeleitete Ausrichtung dieses transformierten Bildes um Winkel $(-\varphi,-\theta,-\psi)$ umgekehrt gedreht werden. Bei mindestens einer Ausführungsform wird Verlust durch Vergleichen von Größen von Azimut, Elevation und Kippen einer ersten Ausrichtung eines Eingabebildes mit einer zweiten Ausrichtung eines transformierten Bildes berechnet, wobei sich ein Nullverlust ergibt, wenn die Größen aller Ausrichtungsparameter gleich sind. Bei mindestens einer Ausführungsform ergibt sich ein Nullverlust, wenn für ein Eingabebild vorhergesagte Erscheinungsbildparameter und eine transformierte Version dieses Eingabebildes übereinstimmen. Bei mindestens einer Ausführungsform wird Symmetrieverlust gemäß Techniken berechnet, die an anderer Stelle in dieser Offenbarung beschrieben sind, wie etwa jenen, die in Verbindung mit **Fig. 16** erörtert sind.

[0110] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1400 ausführt, ausführbaren Code zum Berechnen 1410 des Verlusts des nächsten Nachbarn und des am weitesten entfernten Nachbarn durch mindestens Vergleichen eines ausgewählten Bildes mit seinem nächsten und dem am weitesten entfernten Nachbarn mindestens zum Teil basierend auf einem Sichtpunktgraph einer Bildersammlung. Bei mindestens einer Ausführungsform werden der Verlust des nächsten Nachbarn und des am weitesten entfernten Nachbarn gemäß der Technik, die in Verbindung mit **Fig. 17** beschriebenen ist, berechnet. Bei mindestens einer Ausführungsform wird eine Sammlung von Bildern verwendet, um einen Sichtpunktgraphen zu erzeugen, wobei Knoten eines derartigen Graphen Bildern entsprechen und Kanten ihren Sichtpunkt-Äquivariant-Abständen entsprechen. Bei mindestens einer Ausführungsform werden Sichtpunkt-Äquivariant-Abstände basierend auf Merkmalsähnlichkeiten von Bildpaaren unter Verwendung eines neuronalen Faltungsnetzwerks (CNN) berechnet. Bei mindestens einer Ausführungsform wird ein Ankerbild aus einer Sammlung von Trainingsbildern ausgewählt. Bei mindestens einer Ausführungsform wird ein Ankerbild aus einem Sichtpunktgraphen lokalisiert und ein nächster Nachbar und ein am weitesten entfernter Nachbar werden basierend auf Kantengewichten ausgewählt. Bei mindestens einer Ausführungsform weist ein nächster Nachbar eine kürzeste Kante, die mit einem Ankerbild verbunden ist, auf Bei mindestens einer Ausführungsform weist ein am weitesten entfernter Nachbar eine am weitesten entfernte Kante auf, die mit einem Ankerbild verbunden ist. Bei mindestens einer Ausführungsform sagt ein neuronales Netzwerk einen ersten Sichtpunkt für ein Ankerbild vorher und sagt einen zweiten Sichtpunkt für ein Bild des nächsten Nachbarn vorher, und der Verlust wird derart berechnet, dass ein geringerer Abstand zwischen diesen Sichtpunkten einem geringeren Verlust entspricht. Bei mindestens einer Ausführungsform sagt ein neuronales Netzwerk einen ersten Sichtpunkt für ein Ankerbild vorher und sagt einen dritten Sichtpunkt für ein Bild des am weitesten entfernten Nachbarn vorher, und der Verlust wird derart berechnet, dass ein längerer Abstand zwischen diesen Sichtpunkten einem geringeren Verlust entspricht.

[0111] Bei mindestens einer Ausführungsform werden nächste und/oder am weitesten entfernte Nachbarn nicht deterministisch ausgewählt. Bei mindestens einer Ausführungsform wird jeder Kante eine Wahrscheinlichkeit zugewiesen, um als nächster und/oder am weitesten entfernter Nachbar eines Ankerbildes ausgewählt zu werden. Bei mindestens einer Ausführungsform sind Wahrscheinlichkeiten eines nächsten Nachbarn umgekehrt proportional zu Kantengewichten (z. B. hat ein Knoten, der das niedrigste Kantengewicht aufweist, der mit einem Ankerbild verbunden ist, die höchste Wahrscheinlichkeit, ausgewählt zu werden). Bei mindestens einer Ausführungsform sind Wahrscheinlichkeiten eines am weitesten entfernten Nachbarn direkt proportional zu Kantengewichten (z. B. weist ein Knoten, der das höchste Kantengewicht aufweist, das mit einem Ankerbild verbunden ist, die höchste Wahrscheinlichkeit, ausgewählt zu werden, auf).

[0112] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1400 ausführt, ausführbaren Code zur Verwendung 1412 berechneter Verluste (z. B. aus den Zahlen 1406-1410), um Parameter eines oder mehrerer neuronaler Netzwerke zu aktualisieren, die in einer Sammlung von Bildern trainiert werden. Bei mindestens einer Ausführungsform wird ein Generator auf Symmetrieverlust, Sichtpunktkonsistenzverlust, Echt/Falsch-Klassifizierungsverlust, Entflechtungsverlust oder eine beliebige Kombination davon trainiert. Bei mindestens einer Ausführungsform können Techniken, die in Verbindung mit den **Fig. 4-7** beschrieben sind, dazu verwendet werden, Netzwerke gemäß Prozess 1400 zu trainieren.

[0113] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1400 ausführt, ausführbaren Code, um zu bestimmen, ob 1410 mehr trainiert werden soll. Bei mindestens einer Ausführungsform wird das Training gemäß einer beliebigen geeigneten Technik ausgeführt und kann das Auswählen eines zweiten Bildes und das Ausführen der Schritte 1406-1412 unter Verwendung eines zweiten ausgewählten Bildes beinhalten, um Verluste zu berechnen und Parameter für ein oder mehrere neuronale Netzwerke, die zum Ableiten von Sichtpunkten trainiert werden, zu verfeinern. Sobald das Training abgeschlossen ist, kann ein trainiertes neuronales Netzwerk zur Schlussfolgerung verfügbar gemacht werden (z. B. wird ein neuronales Netzwerk oder Parameter davon an ein anderes System übertragen).

[0114] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1400 ausführt, ausführbaren Code zum Empfangen 1416 eines Bildes desselben Typs wie eine Sammlung von Bildern, die dazu verwendet werden, ein oder mehrere neuronale Netzwerke zu trainieren. Bei mindestens einer Ausführungsform wird ein Bild von einer Kamera oder einer anderen Art von Aufzeichnungsvorrichtung empfangen, die Bilder der Umgebung oder Umgebung eines Systems aufgezeichnet. Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1400 ausführt, ausführbaren Code, um ein trainiertes neuronales Netzwerk zu verwenden 1418, um einen Sichtpunkt eines Objekts in einem Bild zu abzuleiten. Bei mindestens einer Ausführungsform beinhaltet ein Fahrzeug eine Kamera, die ein Bild aufnimmt und dieses Bild einem neuronalen Netzwerk bereitstellt, das mit einer Sammlung von Bildern von Autos trainiert ist, um zu bestimmen, ob das erfasste Bild ein Auto und/oder Ausrichtungen von Autos, die in aufgenommenen Bildern beinhaltet sind, beinhaltet.

[0115] Fig. 15A zeigt ein veranschaulichendes Beispiel eines Prozesses 1500A zum Berechnen von Bildkonsistenzverlust gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform wird der Prozess 1500A (oder ein anderer hierin beschriebene Prozess oder Variationen und/oder Kombinationen davon) unter der Steuerung eines oder mehrerer Computersysteme ausgeführt, die mit computerausführbaren Anweisungen konfiguriert sind und als Code (zum Beispiel Computer ausführbare Anweisungen, ein oder mehrere Computerprogramme oder eine oder mehrere Anwendungen) umgesetzt werden, die gemeinsam auf einem oder mehreren Prozessoren durch Hardware, Software oder Kombinationen davon ausführen. Bei mindestens einer Ausführungsform ist der Code auf einem computerlesbaren Speichermedium gespeichert, zum Beispiel in Form eines Computerprogramms, das eine Vielzahl von computerlesbaren Anweisungen umfasst, die von einem oder mehreren Prozessoren ausgeführt werden können. Ein computerlesbares Speichermedium ist bei mindestens einer Ausführungsform ein nichtflüchtiges computerlesbares Medium. Bei mindestens einer Ausführungsform werden mindestens einige computerlesbare Anweisungen, die zum Ausführen des Prozesses 1500A verwendbar sind, nicht ausschließlich unter Verwendung von transienten Signale (z. B. einer sich ausbreitenden transienten elektrischen oder elektromagnetischen Übertragung) gespeichert. Ein nichtflüchtiges computerlesbares Medium beinhaltet nicht notwendigerweise nichtflüchtige Datenspeicherschaltungen (z. B. Puffer, Caches und Warteschlangen) innerhalb von Transceivern für flüchtige Signale. Bei mindestens einer Ausführungsform wird der Prozess 1500A mindestens zum Teil auf einem Computersystem ausgeführt, wie etwa den an anderer Stelle in dieser Offenbarung beschriebenen. Bei mindestens einer Ausführungsform berechnet ein erstes Computersystem den generativen Konsistenzverlust. Bei mindestens einer Ausführungsform können Techniken, die in Verbindung mit **Fig. 4-7** gelten, für den Prozess 1500A anwendbar sein. Bei mindestens einer Ausführungsform beschreibt der Prozess 1500A einen Prozess zum Berechnen eines oder mehrerer Verluste (z. B. Bildkonsistenzverlust), die verwendet werden können, um Parameter eines Diskriminators als Teil eines Trainingsprozesses zu aktualisieren.

[0116] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1500A ausführt, ausführbaren Code, um ein Eingabebild einer Sammlung von einem oder mehreren Bildern eines Objekttyps zu erhalten 1502. Bei mindestens einer Ausführungsform kann ein Objekttyp bedeuten, dass alle Bilder einer Sammlung von Bildern Bilder sind, die Autos beinhalten. Bei mindestens einer Ausführungsform stellt ein Eingabebild ein Objekt in einer spezifischen Ausrichtung dar, die spezifische

Erscheinungsbildmerkmale (z. B. Attribute) umfasst. Bei mindestens einer Ausführungsform ist ein Eingabebild ein echtes Bild (z. B. im Gegensatz zu synthetischen Bildern), das gemäß den in Verbindung mit **Fig. 4** beschriebenen erhalten wird.

[0117] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1500A ausführt, ausführbaren Code zur Verwendung 1504 eines Diskriminators, um für ein Eingabebild einen Satz von Erscheinungsbildattributen, eine Bestimmung, ob das Eingabebild echt oder falsch ist, und einen Sichtpunkt vorherzusagen. Bei mindestens einer Ausführungsform ist ein Diskriminator mit einem oder mehreren neuronalen Netzwerken assoziiert, die dazu trainiert sind, einen Sichtpunkt sowie andere Charakteristiken aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform ist ein Sichtpunkt ein vorhergesagter Sichtpunkt eines Eingabebildes und entspricht einer vorhergesagten spezifischen Ausrichtung eines in dem Eingabebild abgebildeten Objekts, das spezifische Werte für einen Satz von Parametern umfasst, die einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfassen. Bei mindestens einer Ausführungsform ist ein Satz von Erscheinungsbildattributen oder Parametern ein vorhergesagter Satz von Erscheinungsbildattributen eines Eingabebildes und definiert ein Erscheinungsbild eines Bildes, das in dem Eingabebild abgebildet ist. Bei mindestens einer Ausführungsform ist eine Bestimmung, ob das eingegebene Bild echt oder falsch ist, ein binärer Wert (z. B. wahr oder falsch), der angibt, ob der Diskriminator vorhergesagt hat, dass das eingegebene Bild echt oder falsch ist. Bei mindestens einer Ausführungsform wird die vorhergesagte Bestimmung, ob das eingegebene Bild echt oder falsch ist, verwendet, um einen Echt/Falsch-Klassifizierungsverlust zu berechnen, der den an anderer Stelle in dieser Offenbarung beschriebenen entsprechen kann, einschließlich, aber nicht beschränkt auf die in Verbindung mit den **Fig. 4** und/oder **Fig. 6** erörterten.

[0118] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1500A ausführt, ausführbaren Code, um einen Generator zu verwenden 1506, um ein synthetisches Bild anzulegen, das mindestens zum Teil auf einem vorhergesagten Satz von Erscheinungsbildattributen und einem vorhergesagten Sichtpunkt basiert. Bei mindestens einer Ausführungsform werden ein vorhergesagter Satz von Erscheinungsbildattributen und ein vorhergesagter Sichtpunkt einem Generator bereitgestellt, um ein synthetisches Bild zu erzeugen. Bei mindestens einer Ausführungsform ist ein Generator Teil eines generativen kontradiktorischen Netzwerks. Bei mindestens einer Ausführungsform werden ein vorhergesagter Satz von Erscheinungsbildattributen und ein vorhergesagter Sichtpunkt eingesetzt, um ein synthetisches Bild zu erzeugen, das dem vorhergesagten Satz von Erscheinungsbildattributen und dem vorhergesagten Sichtpunkt entspricht. Bei mindestens einer Ausführungsform erzeugt ein Generator ein synthetisches Bild, das ein Objekt umfasst, das gemäß einem vorhergesagten Satz von Erscheinungsbildattributen erzeugt wird und gemäß einem vorhergesagten ersten Sichtpunkt ausgerichtet ist.

[0119] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1500A ausführt, ausführbaren Code zum Berechnen 1508 von Bildkonsistenzverlust basierend auf einem Eingabebild und einem synthetischen Bild. Bei mindestens einer Ausführungsform werden ein Eingabebild und ein synthetisches Bild verglichen, um Bildkonsistenzverlust zu bestimmen. Bei mindestens einer Ausführungsform wird ein Kosinus-Abstand zwischen einem Eingabebild und einem synthetischen Bild verglichen, um Merkmalsähnlichkeiten zu bestimmen, wobei eine größere Ähnlichkeit einem geringeren Verlust entspricht. Bei mindestens einer Ausführungsform werden L1-, L2- oder Kosinus-Abstände verwendet, um Bildkonsistenzverlust zwischen einem Eingabebild und einem synthetischen Bild zu bestimmen.

[0120] Bei mindestens einer Ausführungsform wird ein Verlust des nächsten und des am weitesten entfernten Nachbarn mindestens zum Teil basierend auf dem Eingabebild und dem synthetischen Bild, die in Verbindung mit dem Prozess 1500A beschrieben sind, berechnet. Bei mindestens einer Ausführungsform wird Verlust des nächsten und des am weitesten entfernten Nachbarn unter Verwendung von Techniken berechnet, die in Verbindung mit den **Fig. 4** und **Fig. 9** beschrieben sind. Bei mindestens einer Ausführungsform wird ein Symmetrieverlust mindestens zum Teil basierend auf dem Eingabebild berechnet.

[0121] Bei mindestens einer Ausführungsform veranschaulicht **Fig. 15B** einen Prozess 1500B zum Berechnen eines Sichtpunktkonsistenzverlusts. Der Prozess 1500B kann durch jedes geeignete System umgesetzt werden, wie z. B. diejenigen, die in Verbindung mit **Fig. 5** beschrieben sind. Bei mindestens einer Ausführungsform werden der Prozess 1500A und der Prozess 1500B von einem Computersystem als Teil eines Trainingsprozesses ausgeführt, der Parameter von Diskriminatoren und Generatoren anpasst, die dazu verwendet werden, Sichtpunkte von Objekten vorherzusagen. Bei mindestens einer Ausführungsform beinhaltet ein Computersystem, das den Prozess 1500B ausführt, ausführbaren Code, der bewirkt, dass das Computer-

system einen Sichtpunkt und einen Satz von Erscheinungsbildparametern erhält 1512. Bei mindestens einer Ausführungsform werden Sichtpunkte und/oder Sätze von Erscheinungsbildparametern zufällig ausgewählt.

[0122] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1500B ausführt, ausführbaren Code, um einen Generator zu verwenden 1514, um ein synthetisches Bild aus einem Sichtpunkt und einen Satz von Erscheinungsbildparametern anzulegen. Bei mindestens einer Ausführungsform werden synthetische Bilder gemäß den oben in Verbindung mit **Fig. 2** beschriebenen Techniken angelegt.

[0123] Bei mindestens einer Ausführungsform wird ein synthetisches Bild angelegt und ein System wird konfiguriert, um einen Diskriminator zu verwenden 1516, um einen Sichtpunkt, eine Bestimmung, ob ein eingegebenes Bild echt oder falsch ist, und einen Satz von Erscheinungsbildparametern vorherzusagen. Bei mindestens einer Ausführungsform wird ein Diskriminator gemäß Techniken, die in Verbindung mit **Fig. 2** beschrieben sind, umgesetzt, um Sichtpunkte und Erscheinungen vorherzusagen.

[0124] Bei mindestens einer Ausführungsform ist ein System konfiguriert, um einen Sichtpunktkonsistenzverlust mindestens zum Teil basierend auf einem vorhergesagten Sichtpunkt (z. B. von einem Diskriminator, der einen Sichtpunkt eines synthetischen Bildes vorhersagt, zu erhalten) und einem eingegebenen Sichtpunkt (z. B. Sichtpunkt, der von einem Generator verwendet wird, um ein synthetisches Bild anzulegen) zu berechnen 1518. Bei mindestens einer Ausführungsform wird der Sichtpunktkonsistenzverlust gemäß Techniken berechnet, die in Verbindung mit den **Fig. 5** und/oder **Fig. 7** beschrieben sind.

[0125] **Fig. 16A** zeigt ein veranschaulichendes Beispiel eines Prozesses 1600A zum Berechnen von Symmetrieverlust gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform wird ein Verfahren 1600A (oder ein anderer Prozess, der hierin beschrieben ist, oder Variationen und/oder Kombinationen davon) teilweise oder ganz unter der Steuerung eines oder mehrerer Computersysteme ausgeführt, die mit ausführbaren Anweisungen konfiguriert sind und als Code (z. B. ausführbare Anweisungen, ein oder mehrere Computerprogramme oder eine oder mehrere Anwendungen) umgesetzt werden, die gemeinsam auf einem oder mehreren Prozessoren, durch Hardware oder Kombinationen davon ausgeführt werden. Bei mindestens einer Ausführungsform ist der Code auf einem computerlesbaren Speichermedium gespeichert, zum Beispiel in Form eines Computerprogramms, das eine Vielzahl von computerlesbaren Anweisungen umfasst, die von einem oder mehreren Prozessoren ausgeführt werden können. Ein computerlesbares Speichermedium ist bei mindestens einer Ausführungsform ein nichtflüchtiges computerlesbares Medium. Bei mindestens einer Ausführungsform werden mindestens einige computerlesbare Anweisungen, die zum Ausführen des Prozesses 1600A verwendbar sind, nicht ausschließlich unter Verwendung von transienten Signale (z. B. einer sich ausbreitenden transienten elektrischen oder elektromagnetischen Übertragung) gespeichert. Ein nichtflüchtiges computerlesbares Medium beinhaltet nicht notwendigerweise nichtflüchtige Datenspeicherschaltungen (z. B. Puffer, Caches und Warteschlangen) innerhalb von Transceivern für flüchtige Signale. Bei mindestens einer Ausführungsform wird der Prozess 1600A mindestens zum Teil auf einem Computersystem ausgeführt, wie etwa den an anderer Stelle in dieser Offenbarung beschriebenen. Bei mindestens einer Ausführungsform berechnet ein erstes Computersystem Symmetrieverlust. Bei mindestens einer Ausführungsform können Techniken, die in Verbindung mit **Fig. 8** beschrieben sind, für den Prozess 1600A gelten.

[0126] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1600A ausführt, ausführbaren Code, um ein Eingabebild zu erhalten 1602. Bei mindestens einer Ausführungsform ist ein Eingabebild ein Bild einer Sammlung von einem oder mehreren Bildern eines Objekttyps. Bei mindestens einer Ausführungsform kann ein Objekttyp bedeuten, dass alle Bilder einer Sammlung von Bildern Bilder sind, die Autos beinhalten. Bei mindestens einer Ausführungsform erhält ein System ein Eingabebild, das ein Objekt in einer spezifischen Ausrichtung, die spezifische Erscheinungsbildmerkmale umfasst, abbildet.

[0127] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1600A ausführt, ausführbaren Code, um eine Transformation an einem Eingabebild auszuführen 1604, wodurch ein transformiertes Bild erzeugt wird. Bei mindestens einer Ausführungsform wird eine Transformation auf ein Eingabebild angewandt, in dem das Eingabebild horizontal gespiegelt wird, um ein transformiertes Bild zu erzeugen. Bei mindestens einer Ausführungsform wird eine Transformation von einem oder mehreren Systemen angewandt, die mit einem Diskriminator assoziiert sind. Bei mindestens einer Ausführungsform werden eine oder mehrere Bildverarbeitungstechniken auf ein Eingabebild angewandt, um ein transformiertes Bild zu erzeugen. Bei mindestens einer Ausführungsform werden Azimut- und Kippwinkel

eines Sichtpunkts eines Objekts innerhalb eines Eingabebildes umgekehrt, wenn das Eingabebild transformiert wird, um ein transformiertes Bild zu erzeugen, und Elevationswinkel bleiben über das Eingabebild und das transformierte Bild gleich.

[0128] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1600A ausführt, ausführbaren Code, um für ein transformiertes Bild mindestens einen Sichtpunkt vorherzusagen 1606. Bei mindestens einer Ausführungsform sagt ein Diskriminator für ein transformiertes Bild einen Sichtpunkt, einen Satz von Erscheinungsbildparametern, eine Echt/Falsch-Klassifizierung oder eine beliebige Kombination davon vorher. Bei mindestens einer Ausführungsform ist ein Diskriminator mit einem oder mehreren neuronalen Netzwerken assoziiert, die dazu trainiert sind, einen Sichtpunkt sowie andere Charakteristiken aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform empfängt ein Diskriminator ein transformiertes Bild und sagt einen Sichtpunkt vorher. Bei mindestens einer Ausführungsform entspricht ein Sichtpunkt einer spezifischen Ausrichtung eines Objekts, das in einem Bild abgebildet ist, und umfasst spezifische Werte für einen Satz von Parametern, der einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfasst.

[0129] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1600A ausführt, ausführbaren Code, um eine Transformation auf einen vorhergesagten Sichtpunkt anzuwenden 1608. Bei mindestens einer Ausführungsform wird eine Transformation oder Inverse davon auf einen vorhergesagten Sichtpunkt angewandt. Falls bei mindestens einer Ausführungsform ein Eingabebild um Winkel (φ,θ,ψ) gedreht wird, um ein transformiertes Bild zu erzeugen, kann vorhergesagter Sichtpunkt, der basierend auf dem transformierten Bild erzeugt wird, umgekehrt um Winkel $(-\varphi,-\theta,-\psi)$ gedreht werden.

[0130] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1600A ausführt, einen ausführbaren Code, um für ein Eingabebild mindestens einen Sichtpunkt vorherzusagen 1610. Bei mindestens einer Ausführungsform sagt ein Diskriminator für ein Eingabebild einen Sichtpunkt, einen Satz von Erscheinungsbildparametern, eine Echt/Falsch-Klassifizierung oder eine beliebige Kombination davon vorher. Bei mindestens einer Ausführungsform empfängt ein Diskriminator ein Eingabebild und sagt einen Sichtpunkt vorher. Bei mindestens einer Ausführungsform entspricht ein Sichtpunkt einer spezifischen Ausrichtung eines Objekts, das in einem Bild abgebildet ist, und umfasst spezifische Werte für einen Satz von Parametern, der einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfasst.

[0131] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1600A ausführt, ausführbaren Code zum Vergleichen 1612 von Sichtpunkten zum Berechnen von Symmetrieverlust. Bei mindestens einer Ausführungsform vergleicht ein System einen vorhergesagten Sichtpunkt für ein transformiertes Bild mit einem vorhergesagten Sichtpunkt eines Eingabebildes, wobei das transformierte Bild aus dem Eingabebild erzeugt wird. Bei mindestens einer Ausführungsform wird Symmetrieverlust berechnet, indem Größen von Azimut, Elevation und Kippen eines vorhergesagten Sichtpunkts eines Eingabebildes mit Größen von Azimut, Elevation und Kippen eines vorhergesagten Sichtpunkts eines transformierten Bildes verglichen werden, wobei sich Nullverlust ergibt, wenn die Größen der Parameter jedes Sichtpunkts gleich sind. Bei mindestens einer Ausführungsform wird Symmetrieverlust mindestens zum Teil darauf basierend berechnet, wie genau Erscheinungsbildparameter eines Bildes (z. B. Eingabebild) und einer transformierten Version dieses Bildes miteinander übereinstimmen.

[0132] **Fig. 16B** veranschaulicht einen Prozess 1600B zum Berechnen von Symmetrieverlust gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform wird **Fig. 16B** verwendet, um Parameter eines Generators als Teil des Trainierens eines neuronalen Netzwerkes zum Vorhersagen von Sichtpunkten zu aktualisieren. Bei mindestens einer Ausführungsform beinhaltet ein System, das den Prozess 1600B ausführt, ausführbaren Code, um einen Sichtpunkt und einen Satz von Erscheinungsbildattributen zu erhalten 1614. Bei mindestens einer Ausführungsform ist ein System konfiguriert, um eine Transformation auf einen erhaltenen Sichtpunkt anzuwenden 1616, um einen transformierten Sichtpunkt zu bestimmen. In mindestens einigen Ausführungsformen wird ein Sichtpunkt gespiegelt, um einen transformierten Sichtpunkt zu erhalten. Bei mindestens einer Ausführungsform wird eine Transformationsfunktion T() auf einen Satz von Parametern x_1, y_1 und z_1 angewandt, um einen transformierten Satz von Parametern x_2, y_2, z_2, die Azimut-, Kipp- und Elevationsparametern entsprechen können, zu erhalten.

[0133] Bei mindestens einer Ausführungsform wird ein Generator verwendet, um ein erstes synthetisches Bild zu, mindestens zum Teil basierend auf einem transformierten Sichtpunkt und einem Satz von Erscheinungsbildparametern, die berechnet oder anderweitig unter Verwendung von Techniken bestimmt werden

können, die in Verbindung mit anderen Schritten der **Fig. 16B** beschrieben sind, zu erzeugen 1618. Bei mindestens einer Ausführungsform ist ein System dazu konfiguriert, eine Transformation auf ein erstes synthetisches Bild anzuwenden 1620, das aus einem transformierten Sichtpunkt und einem Satz von Erscheinungsbildattributen erzeugt wird. Wenn bei mindestens einer Ausführungsform wird ein transformierter Sichtpunkt durch Ausführen einer Transformation T() erzeugt, dann wird eine inverse Transformation $T^{-1}()$ auf ein synthetisches Bild angewandt, wobei $T(T^{-1}(x,y,z))=(x,y,z)$. Wenn bei mindestens einer Ausführungsform eine Transformation ein Bild horizontal spiegelt, spiegelt eine Inverse ein Bild horizontal zurück in seine ursprüngliche Ausrichtung.

[0134] Bei mindestens einer Ausführungsform beinhaltet ein System ausführbare Anweisungen zum Erzeugen 1622 eines zweiten synthetischen Bildes mindestens zum Teil basierend auf einem Sichtpunkt und einem Satz von Erscheinungsbildparametern. Bei mindestens einer Ausführungsform wird derselbe Generator verwendet, um ein erstes synthetisches Bild basierend auf einem transformierten Sichtpunkt und ein zweites synthetisches Bild basierend auf einem ursprünglichen Sichtpunkt zu erzeugen. Bei mindestens einer Ausführungsform ist ein System dazu konfiguriert, ein erstes synthetisches Bild und ein zweites synthetisches Bild zu vergleichen 1624, um einen Symmetrieverlust zu berechnen, wobei sich ein Kosinus-Abstand von Null zwischen derartigen Bildern einen Nullverlust betrifft.

[0135] **Fig. 17** zeigt ein veranschaulichendes Beispiel eines Prozesses 1700 zum Berechnen des Verlustes des nächsten Nachbarn und des am weitesten entfernten Nachbarn gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform wird ein Prozess 1700 (oder ein anderer Prozess, der hierin beschrieben ist, oder Variationen und/oder Kombinationen davon) ganz oder teilweise unter der Steuerung eines oder mehrerer Computersysteme ausgeführt, die mit computerausführbaren Anweisungen konfiguriert sind, die als Code umgesetzt werden können (zum Beispiel computerausführbare Anweisungen, ein oder mehrere Computerprogramme oder eine oder mehrere Anwendungen), die gemeinsam auf einem oder mehreren Prozessoren, durch Hardware, Software oder Kombinationen davon ausführen. Bei mindestens einer Ausführungsform ist der Code auf einem computerlesbaren Speichermedium gespeichert, zum Beispiel in Form eines Computerprogramms, das eine Vielzahl von computerlesbaren Anweisungen umfasst, die von einem oder mehreren Prozessoren ausgeführt werden können. Ein computerlesbares Speichermedium ist bei mindestens einer Ausführungsform ein nichtflüchtiges computerlesbares Medium. Bei mindestens einer Ausführungsform werden mindestens einige computerlesbare Anweisungen, die zum Ausführen des Prozesses 1700 verwendbar sind, nicht ausschließlich unter Verwendung von transienten Signalen (z. B. einer sich ausbreitenden transienten elektrischen oder elektromagnetischen Übertragung) gespeichert. Ein nichtflüchtiges computerlesbares Medium beinhaltet nicht notwendigerweise nichtflüchtige Datenspeicherschaltungen (z. B. Puffer, Caches und Warteschlangen) innerhalb von Transceivern für flüchtige Signale. Bei mindestens einer Ausführungsform wird der Prozess 1700 mindestens zum Teil auf einem Computersystem wie zum Beispiel denjenigen, die an anderer Stelle in dieser Offenbarung beschrieben sind, ausgeführt. Bei mindestens einer Ausführungsform berechnet ein erstes Computersystem den Verlust des nächsten und des am weitesten entfernten Nachbarn. Bei mindestens einer Ausführungsform können Techniken, die in Verbindung mit **Fig. 9** beschrieben sind, für den Prozess 1700 gelten.

[0136] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1700 ausführt, ausführbaren Code, um eine Sammlung von Bildern zu erhalten 1702. Bei mindestens einer Ausführungsform umfasst eine Sammlung von Bildern ein oder mehrere Bilder eines Objekttyps. Bei mindestens einer Ausführungsform kann ein Objekttyp bedeuten, dass alle Bilder einer Sammlung von Bildern Bilder sind, die Autos beinhalten. Bei mindestens einer Ausführungsform erhält ein System eine Sammlung von einem oder mehreren Bildern gemäß Techniken, die an anderer Stelle in dieser Offenbarung, wie **Fig. 13**, beschrieben sind.

[0137] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1700 ausführt, ausführbaren Code, um für ein Paar von Bildern einer Sammlung einen Kosinus-Abstand zu berechnen 1704, der Merkmalsähnlichkeit vergleicht. Bei mindestens einer Ausführungsform ist der Kosinus-Abstand ein mathematisches Komplement der Kosinusähnlichkeit (z. B. Kosinus-Abstand = 1-Kosinusähnlichkeit). Bei mindestens einer Ausführungsform ist Kosinusähnlichkeit ein Maß für die Ähnlichkeit zwischen zwei Vektoren, die Bilder, Text, Daten und/oder Variationen, basierend auf einem Kosinus eines Winkels zwischen ihnen, davon darstellen können. Bei mindestens einer Ausführungsform ist, wenn zwei Bilder Merkmale umfassen, die Objekten entsprechen, die ähnliche Sichtpunkte aufweisen, ein Kosinus-Abstand, der für die Bilder berechnet wird, gering. Bei mindestens einer Ausführungsform ist, wenn zwei Bilder Merkmale umfassen, die Objekten mit unterschiedlichen Sichtpunkten entsprechen, ein Kosinus-

Abstand, der für die Bilder berechnet wird, hoch. Bei mindestens einer Ausführungsform wird für jedes Bild einer Sammlung von Bildern in Bezug auf jedes andere Bild der Sammlung ein Kosinus-Abstand berechnet.

[0138] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1700 ausführt, ausführbaren Code zum Erzeugen 1706 eines Sichtpunktgraphen, wobei Knoten des Graphen Bildern einer Sammlung entsprechen und Kanten ihren Kosinus-Abständen entsprechen. Bei mindestens einer Ausführungsform werden Kosinus-Abstände basierend auf Merkmalsähnlichkeiten von Bildpaaren unter Verwendung eines faltenden neuronalen Netzwerks berechnet. Bei mindestens einer Ausführungsform werden Kanten eines Sichtpunktgraphen so gewichtet, dass dickere Kanten zwischen zwei Bildern einem höheren Grad an Ähnlichkeit zwischen den zwei Bildern entsprechen und dünnere Kanten zwischen zwei Bildern einem geringeren Grad an Ähnlichkeit zwischen den zwei Bildern entsprechen.

[0139] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1700 ausführt, ausführbaren Code, um ein Ankerbild eines Graphen auszuwählen 1708 und einen ersten Sichtpunkt vorherzusagen. Bei mindestens einer Ausführungsform wird ein Ankerbild aus einer Sammlung von Trainingsbildern ausgewählt, die dazu verwendet werden, einen Sichtpunktgraphen zu erzeugen. Bei mindestens einer Ausführungsform ist ein Diskriminator mit einem oder mehreren neuronalen Netzwerken assoziiert, die dazu trainiert sind, einen Sichtpunkt sowie andere Charakteristiken aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform empfängt ein Diskriminator ein Ankerbild und sagt einen ersten Sichtpunkt vorher. Bei mindestens einer Ausführungsform entspricht ein erster Sichtpunkt einer vorhergesagten spezifischen Ausrichtung eines Objekts, das in einem Ankerbild abgebildet ist, und umfasst spezifische Werte für einen Satz von Parametern, der Azimutparameter, Elevationsparameter und Kippparameter, die der Ausrichtung des Objekts entsprechen.

[0140] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1700 ausführt, ausführbaren Code, um einen Graphen zu verwenden 1710, um einen nächsten Nachbarn eines Ankerbildes auszuwählen und einen zweiten Sichtpunkt vorherzusagen. Bei mindestens einer Ausführungsform wird ein nächster Nachbar basierend auf Kantengewichten eines Sichtpunktgraphen bestimmt. Bei mindestens einer Ausführungsform ist ein nächster Nachbar ein Bild, das eine kürzeste Kante, die mit einem Ankerbild eines Sichtpunktgraphen verbunden ist, aufweist. Bei mindestens einer Ausführungsform ist ein nächster Nachbar zu einem Ankerbild ein Bild, das dem Ankerbild innerhalb einer Sammlung von Bildern am ähnlichsten ist. Bei mindestens einer Ausführungsform empfängt ein Diskriminator ein Bild des nächsten Nachbarn und sagt einen zweiten Sichtpunkt vorher. Bei mindestens einer Ausführungsform entspricht ein zweiter Sichtpunkt einer vorhergesagten spezifischen Ausrichtung eines Objekts, das in einem Bild des nächsten Nachbarn abgebildet ist, und umfasst spezifische Werte für einen Satz von Parametern, der Azimutparameter, Elevationsparameter und Kippparameter, die der Ausrichtung des Objekts entsprechen Objekt, umfasst.

[0141] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1700 ausführt, ausführbaren Code, um einen Graphen zu verwenden 1712, um einen am weitesten entfernten Nachbarn eines Ankerbildes auszuwählen und einen dritten Sichtpunkt vorherzusagen. Bei mindestens einer Ausführungsform wird ein am weitesten entfernter Nachbar basierend auf Kantengewichten eines Sichtpunktgraphen bestimmt. Bei mindestens einer Ausführungsform ist ein am weitesten entfernter Nachbar ein Bild, das eine längste Kante aufweist, die mit einem Ankerbild eines Sichtpunktgraphen verbunden ist. Bei mindestens einer Ausführungsform ist ein am weitesten entfernter Nachbar zu einem Ankerbild ein Bild, das sich innerhalb einer Sammlung von Bildern am meisten von dem Ankerbild unterscheidet. Bei mindestens einer Ausführungsform empfängt ein Diskriminator ein Bild des am weitesten entfernten Nachbarn und sagt einen dritten Sichtpunkt vorher. Bei mindestens einer Ausführungsform entspricht ein dritter Sichtpunkt einer vorhergesagten spezifischen Ausrichtung eines Objekts, das in einem Bild des am weitesten entfernten Nachbarn abgebildet ist, und umfasst spezifische Werte für einen Satz von Parametern, der Azimutparameter, Elevationsparameter und Kippparameter, die der Ausrichtung des Objekts entsprechen, umfasst.

[0142] Bei mindestens einer Ausführungsform beinhaltet ein System, das mindestens einen Teil des Prozesses 1700 ausführt, ausführbaren Code zum Berechnen 1714 des Verlustes des nächsten und des am weitesten entfernten Nachbarn. Bei mindestens einer Ausführungsform wird ein Verlust des nächsten Nachbarn zwischen einem ersten Sichtpunkt, der von einem Ankerbild vorhergesagt wird, und einem zweiten Sichtpunkt, der von einem Bild des nächsten Nachbarn zu dem Ankerbild vorhergesagt wird, berechnet. Bei mindestens einer Ausführungsform wird ein Verlust des nächsten Nachbarn derart berechnet, dass eine höhere Ähnlichkeit zwischen einem ersten Sichtpunkt und einem zweiten Sichtpunkt einem geringeren Verlust entspricht. Bei mindestens einer Ausführungsform wird Verlust des am weitesten entfernten Nachbarn zwischen

einem ersten Sichtpunkt, der von einem Ankerbild vorhergesagt wird, und einem dritten Sichtpunkt berechnet, der von einem Bild des am weitesten entfernten Nachbarn zu dem Ankerbild vorhergesagt wird. Bei mindestens einer Ausführungsform wird ein Verlust des am weitesten Nachbarn derart berechnet, dass eine geringere Ähnlichkeit zwischen einem ersten Sichtpunkt und einem dritten Sichtpunkt einem geringeren Verlust entspricht. Bei mindestens einer Ausführungsform wird Verlust des nächsten und des am weitesten entfernten Nachbarn basierend auf einer Kombination aus dem Verlust des nächsten Nachbars und dem Verlust des am weitesten entfernten Nachbarn berechnet.

ABLEITUNGS- UND TRAININGSLOGIK

[0143] Fig. 18A veranschaulicht Ableitungs- und/oder Trainingslogik 1815, die verwendet wird, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden nachstehend in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt.

[0144] Bei mindestens einer Ausführungsform kann die Ableitungs- und/oder Trainingslogik 1815 ohne Einschränkung Folgendes beinhalten: Code- und/oder Datenspeicher 1801, um Gewichtung und/oder Eingabe-/Ausgabedaten und/oder andere Parameter vorwärtsgerichtet zu speichern und/oder auszugeben, um Neuronen oder Schichten eines neuronalen Netzwerks zu konfigurieren, das in Aspekten einer oder mehrerer Ausführungsformen trainiert und/oder zum Ableiten verwendet wird. Bei mindestens einer Ausführungsform kann die Trainingslogik 1815 einen Code- und/oder Datenspeicher 1801 beinhalten oder mit diesem gekoppelt sein, um Graphcode oder andere Software zu speichern, um die Zeitsteuerung und/oder die Reihenfolge zu steuern, in der Gewichtungsinformationen und/oder Informationen zu anderen Parametern gespeichert werden sollen, um Logik zu konfigurieren, die Ganzzahl- und/oder Gleitkommaeinheiten (gemeinsam arithmetische Logikeinheiten (Arithmetic Logic Units, ALUs)) beinhalten. Bei mindestens einer Ausführungsform lädt Code, wie etwa Graphcode, Gewichtungsinformationen oder Informationen zu anderen Parametern in Prozessor-ALUs basierend auf einer Architektur eines neuronalen Netzwerks, dem ein derartiger Code entspricht. Bei mindestens einer Ausführungsform speichert der Code- und/oder Datenspeicher 1801 Gewichtungsparameter und/oder Eingabe-/Ausgabedaten jeder Schicht eines neuronalen Netzwerks, das trainiert oder in Verbindung mit einer oder mehreren Ausführungsformen während der Rückwärtspropagation von Eingabe-/Ausgabedaten und/oder Gewichtsparametern während des Trainings und/oder dem Ableiten unter Verwendung von Aspekten einer oder mehrerer Ausführungsformen verwendet wird. Bei mindestens einer Ausführungsform kann ein beliebiger Teil des Code- und/oder Datenspeichers 1801 in einem anderen chipinternen oder chipexternen Datenspeicher beinhaltet sein, was einen L1-, L2- oder L3-Cache oder Systemspeicher eines Prozessors beinhaltet.

[0145] Bei mindestens einer Ausführungsform kann ein beliebiger Teil des Code- und/oder Datenspeichers 1801 intern oder extern von einem oder mehreren Prozessoren oder anderen Hardwarelogikvorrichtungen oder -schaltungen sein. Bei mindestens einer Ausführungsform kann Code und/oder Code- und/oder Datenspeicher 1801 ein Cache-Speicher, ein dynamischer zufällig adressierbarer Speicher („DRAM“), ein statischer zufällig adressierbarer Speicher („SRAM“), ein nichtflüchtiger Speicher (z. B. Flash-Speicher) oder ein anderer Speicher sein. Bei mindestens einer Ausführungsform kann eine Wahl, ob der Code und/oder Code- und/oder Datenspeicher 1801 zum Beispiel intern oder extern von einem Prozessor ist oder DRAM, SRAM, Flash oder eine andere Speicherart umfasst, von verfügbarem chipinternen im Vergleich zu chip externen Speicher, Latenzanforderungen von ausgeführten Trainings- und/oder Ableitungsfunktionen, Batchgröße von Daten, die beim Ableiten und/oder Trainieren eines neuronalen Netzwerks verwendet werden, oder einer Kombination dieser Faktoren abhängen.

[0146] Bei mindestens einer Ausführungsform kann die Ableitungs- und/oder Trainingslogik 1815 ohne Einschränkung Folgendes beinhalten: einen Code- und/oder Datenspeicher 1805, um Gewichtung und/oder Eingabe-/Ausgabedaten rückwärtsgerichtet zu speichern und/oder auszugeben, die Neuronen oder Schichten eines neuronalen Netzwerks entsprechen, das in Aspekten einer oder mehrerer Ausführungsformen trainiert und/oder zum Ableiten verwendet wird. Bei mindestens einer Ausführungsform speichert der Code- und/oder Datenspeicher 1805 Gewichtungsparameter und/oder Eingabe-/Ausgabedaten jeder Schicht eines neuronalen Netzwerks, das trainiert oder in Verbindung mit einer oder mehreren Ausführungsformen während der Rückwärtspropagation von Eingabe-/Ausgabedaten und/oder Gewichtsparametern während des Trainings und/oder dem Ableiten unter Verwendung von Aspekten einer oder mehrerer Ausführungsformen verwendet wird. Bei mindestens einer Ausführungsform kann die Trainingslogik 1815 einen Code- und/oder Datenspeicher 1805 beinhalten oder mit diesem gekoppelt sein, um Graphcode oder andere Software zu speichern, um die Zeitsteuerung und/oder die Reihenfolge zu steuern, in der Gewichtungsinformationen und/oder Infor-

mationen zu anderen Parametern gespeichert werden sollen, um Logik zu konfigurieren, die Ganzzahl- und/oder Gleitkommaeinheiten (gemeinsam arithmetische Logikeinheiten (ALUs)) beinhalten. Bei mindestens einer Ausführungsform lädt Code, wie etwa Graphcode, Gewichtungsinformationen oder Informationen zu anderen Parametern in Prozessor-ALUs basierend auf einer Architektur eines neuronalen Netzwerks, dem ein derartiger Code entspricht. Bei mindestens einer Ausführungsform kann ein beliebiger Teil des Code- und/oder Datenspeichers 1805 in einem anderen chipinternen oder chipexternen Datenspeicher beinhaltet sein, was einen L1-, L2- oder L3-Cache oder Systemspeicher eines Prozessors beinhaltet. Bei mindestens einer Ausführungsform kann ein beliebiger Teil des Code- und/oder Datenspeichers 1805 intern oder extern von einem oder mehreren Prozessoren oder anderen Hardwarelogikvorrichtungen oder -schaltungen sein. Bei mindestens einer Ausführungsform kann der Code- und/oder Datenspeicher 1805 ein Pufferspeicher, DRAM, SRAM, ein nichtflüchtiger Speicher (z. B. Flash-Speicher) oder ein anderer Speicher sein. Bei mindestens einer Ausführungsform kann eine Wahl, ob der Code- und/oder Datenspeicher 1805 zum Beispiel intern oder extern von einem Prozessor ist oder DRAM, SRAM, Flash-Speicher oder eine andere Speicherart umfasst, von verfügbarem chipinternen im Vergleich zu chip externen Speicher, Latenzanforderungen von ausgeführten Trainings- und/oder Ableitungsfunktionen, Batchgröße von Daten, die beim Ableiten und/oder Trainieren eines neuronalen Netzwerks verwendet werden, oder einer Kombination dieser Faktoren abhängen.

[0147] Bei mindestens einer Ausführungsform können der Code- und/oder Datenspeicher 1801 und der Code- und/oder Datenspeicher 1805 getrennte Speicherstrukturen sein. Bei mindestens einer Ausführungsform können der Code- und/oder Datenspeicher 1801 und der Code- und/oder Datenspeicher 1805 dieselbe Speicherstruktur sein. Bei mindestens einer Ausführungsform können der Code- und/oder Datenspeicher 1801 und der Code- und/oder Datenspeicher 1805 zum Teil dieselbe Speicherstruktur und zum Teil separate Speicherstrukturen sein. Bei mindestens einer Ausführungsform kann ein beliebiger Teil des Code- und/oder Datenspeichers 1801 und des Code- und/oder Datenspeichers 1805 in einem anderen chipinternen oder chip externen Datenspeicher beinhaltet sein, was einen L1-, L2- oder L3-Cache oder Systemspeicher eines Prozessors beinhaltet.

[0148] Bei mindestens einer Ausführungsform kann die Ableitungs- und/oder Trainingslogik 1815 ohne Einschränkung Folgendes beinhalten: eine oder mehrere arithmetische Logikeinheit(en) („ALU(s)“) 1810, die Ganzzahl- und/oder Gleitkommaeinheiten beinhalten, um logische und/oder mathematische Operationen auszuführen, die mindestens zum Teil auf Trainings- und/oder Ableitungscode (z. B. Graphcode) basieren oder durch diesen angezeigt werden, deren Ergebnis Aktivierungen (z. B. Ausgabewerte von Schichten oder Neuronen innerhalb eines neuronalen Netzwerks) erzeugen kann, die in einem Aktivierungsspeicher 1820 gespeichert sind, die Funktionen von Eingabe/Ausgabe- und/oder Gewichtsparameterdaten sind, die in dem Code- und/oder Datenspeicher 1801 und/oder Code- und/oder Datenspeicher 1805 gespeichert sind. Bei mindestens einer Ausführungsform werden in dem Aktivierungsspeicher 1820 gespeicherte Aktivierungen gemäß linearer algebraischer und/oder matrixbasierter Mathematik erzeugt, die von ALU(s) 1810 als Reaktion auf das Ausführen von Anweisungen oder anderem Code ausgeführt wird, wobei in dem Code und/oder Datenspeicher 1805 und/oder dem Datenspeicher 1801 gespeicherte Gewichtungswerte als Operanden gemeinsam mit anderen Werten verwendet werden, wie etwa Bias-Werten, Gradienteninformationen, Impulswerten oder anderen Parametern oder Hyperparametern, von denen einige oder alle in dem Code- und/oder Datenspeicher 1805 oder dem Code- und/oder Datenspeicher 1801 oder einem anderen Speicher chipintern oder -extern gespeichert werden können.

[0149] Bei mindestens einer Ausführungsform sind die ALU(s) 1810 in einem oder mehreren Prozessoren oder anderen Hardwarelogikvorrichtungen oder -schaltungen beinhaltet, wohingegen bei einer weiteren Ausführungsform die ALU(s) 1810 außerhalb eines Prozessors oder einer anderen Hardwarelogikvorrichtung oder -schaltung liegen können, die sie verwendet (z. B. ein Co-Prozessor). Bei mindestens einer Ausführungsform können die ALUs 1810 in den Ausführungseinheiten eines Prozessors oder anderweitig in einer Bank von ALUs beinhaltet sein, auf welche die Ausführungseinheiten eines Prozessors, entweder innerhalb desselben Prozessors oder zwischen unterschiedlichen Prozessoren unterschiedlicher Typen (z. B. Zentraleinheiten, Grafikverarbeitungseinheiten, Einheiten mit fester Funktion usw.) verteilt, zugreifen können. Bei mindestens einer Ausführungsform können sich der Datenspeicher 1801, der Code- und/oder Datenspeicher 1805 und der Aktivierungsspeicher 1820 auf einem selben Prozessor oder einer anderen Hardwarelogikvorrichtung oder -schaltung befinden, wohingegen sie sich bei einer weiteren Ausführungsform in unterschiedlichen Prozessoren oder anderen Hardwarelogikvorrichtungen oder -schaltungen oder einer Kombination von gleichen und unterschiedlichen Prozessoren oder anderen Hardwarelogikvorrichtungen oder -schaltungen befinden können. Bei mindestens einer Ausführungsform kann ein beliebiger Teil des Aktivierungsspeichers 1820 in einem anderen chipinternen oder chipexternen Datenspeicher beinhaltet sein, was einen L1-, L2-

oder L3-Cache oder Systemspeicher eines Prozessors beinhaltet. Darüber hinaus kann Ableitungs- und/oder Trainingscode mit anderem Code gespeichert werden, auf den ein Prozessor oder eine andere Hardwarelogik oder -schaltung zugreifen kann, und unter Verwendung der Abruf-, Decodier-, Planungs-, Ausführungs-, Stilllegungs- und/oder anderen logischen Schaltungen eines Prozessors abgerufen und/oder verarbeitet werden.

[0150] Bei mindestens einer Ausführungsform kann der Aktivierungsspeicher 1820 ein Cache-Speicher, DRAM, SRAM, ein nichtflüchtiger Speicher (z. B. Flash-Speicher) oder ein anderer Speicher sein. Bei mindestens einer Ausführungsform kann sich der Aktivierungsspeicher 1820 vollständig oder teilweise innerhalb oder außerhalb von einem oder mehreren Prozessoren oder anderen logischen Schaltungen befinden. Bei mindestens einer Ausführungsform kann eine Wahl, ob der Aktivierungsspeicher 1820 zum Beispiel intern oder extern von einem Prozessor ist oder DRAM, SRAM, Flash-Speicher oder eine andere Speicherart umfasst, von verfügbarem chipinternen im Vergleich zu chipexternen Speicher, Latenzanforderungen von ausgeführten Trainings- und/oder Ableitungsfunktionen, Batchgröße von Daten, die beim Ableiten und/oder Trainieren eines neuronalen Netzwerks verwendet werden, oder einer Kombination dieser Faktoren abhängen. Bei mindestens einer Ausführungsform kann die Ableitungs- und/oder Trainingslogik 1815, die in **Fig. 18A** veranschaulicht ist, in Verbindung mit einer anwendungsspezifischen integrierten Schaltung („ASIC“) verwendet werden, wie etwa einer Tensorflow®-Verarbeitungseinheit von Google, einer Ableitungsverarbeitungseinheit (IPU) von Graphcore™ oder einem Nervana®-Prozessor (z. B. „Lake Crest“-Prozessor) von Intel Corp. Bei mindestens einer Ausführungsform kann die Ableitungs- und/oder Trainingslogik 1815, die in **Fig. 18A** veranschaulicht ist, in Verbindung mit Hardware einer Zentralverarbeitungseinheit (CPU), Hardware einer Grafikverarbeitungseinheit (GPU) oder anderer Hardware, wie etwa feldprogrammierbaren Gate-Arrays (FPGAs) verwendet werden.

[0151] **Fig. 18B** veranschaulicht die Ableitungs- und/oder Trainingslogik 1815 gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform kann die Ableitungs- und/oder Trainingslogik 1815 ohne Einschränkung Folgendes beinhalten: Hardwarelogik, bei der Berechnungsressourcen dediziert oder anderweitig ausschließlich in Verbindung mit Gewichtungswerten oder anderen Informationen verwendet werden, die einer oder mehreren Schichten von Neuronen innerhalb eines neuronalen Netzwerks entsprechen. Bei mindestens einer Ausführungsform kann die Ableitungs- und/oder Trainingslogik 1815, die in **Fig. 18B** veranschaulicht ist, in Verbindung mit einer anwendungsspezifischen integrierten Schaltung (ASIC) verwendet werden, wie etwa einer Tensorflow®-Verarbeitungseinheit von Google, einer Ableitungsverarbeitungseinheit (IPU) von Graphcore™ oder einem Nervana®-(z. B. „Lake Crest“-)Prozessor von Intel Corp. Bei mindestens einer Ausführungsform kann die Ableitungs- und/oder Trainingslogik 1815, die in **Fig. 18B** veranschaulicht ist, in Verbindung mit Hardware einer Zentralverarbeitungseinheit (CPU), Hardware einer Grafikverarbeitungseinheit (GPU) oder anderer Hardware, wie etwa feldprogrammierbaren Gate-Arrays (FPGAs) verwendet werden. Bei mindestens einer Ausführungsform beinhaltet die Ableitungs- und/oder Trainingslogik 1815 ohne Einschränkung den Code- und/oder Datenspeicher 1801 und den Code- und/oder Datenspeicher 1805, die verwendet werden können, um Code (z. B. Graphcode), Gewichtungswerte und/oder andere Informationen zu speichern, was Bias-Werte, Gradienteninformationen, Impulswerte und/oder andere Parameter- oder Hyperparameterinformationen beinhaltet. Bei mindestens einer Ausführungsform, die in **Fig. 18B** veranschaulicht ist, ist jeder von dem Code- und/oder Datenspeicher 1801 und dem Code- und/oder Datenspeicher 1805 mit einer dedizierten Berechnungsressource verknüpft, wie etwa einer Berechnungshardware 1802 bzw. einer Berechnungshardware 1806. Bei mindestens einer Ausführungsform umfasst jede von der Berechnungshardware 1802 und der Berechnungshardware 1806 eine oder mehrere ALUs, die mathematische Funktionen, wie etwa lineare algebraische Funktionen, nur an Informationen ausführen, die in dem Code- und/oder Datenspeicher 1801 und Code- und/oder Datenspeicher 1805 gespeichert sind, deren Ergebnis in dem Aktivierungsspeicher 1820 gespeichert ist.

[0152] Bei mindestens einer Ausführungsform entspricht jeder von dem Code- und/oder Datenspeicher 1801 und 1805 und die entsprechende Berechnungshardware 1802 bzw. 1806 unterschiedlichen Schichten eines neuronalen Netzwerks derart, dass die sich ergebende Aktivierung von einem „Speicher-/Berechnungspaar 1801/1802“ des Code- und/oder Datenspeichers 1801 und der Berechnungshardware 1802 als eine Eingabe für ein nächstes „Speicher-/Berechnungspaar 1805/1806“ des Code- und/oder Datenspeichers 1805 und der Berechnungshardware 1806 bereitgestellt wird, um eine konzeptionelle Organisation eines neuronalen Netzwerks widerzuspiegeln. Bei mindestens einer Ausführungsform kann jedes der Speicher-/Berechnungspaare 1801/1802 und 1805/1806 mehr als einer Schicht des neuronalen Netzwerks entsprechen. Bei mindestens einer Ausführungsform können zusätzliche Speicher-/Berechnungspaare (nicht gezeigt) nach oder parallel zu den Speicher-/Berechnungspaaren 1801/1802 und 1805/1806 in der Ableitungs- und/oder Trainingslogik 1815 beinhaltet sein.

TRAINING UND EINSATZ DES NEURONALEN NETZWERKS

[0153] Fig. 19 veranschaulicht ein Training und einen Einsatz eines tiefen neuronalen Netzwerks gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform wird ein untrainiertes neuronales Netzwerk 1906 unter Verwendung eines Trainingsdatensatzes 1902 trainiert. Bei mindestens einer Ausführungsform ist das Trainings-Framework 1904 ein PyTorch-Framework, wohingehend bei anderen Ausführungsformen das Trainings-Framework 1904 ein TensorFlow, Boost, Caffe, Microsoft Cognitive Toolkit/CNTK, MXNet, Chainer, Keras, Deeplearning4j oder ein anderes Trainings-Framework ist. Bei mindestens einer Ausführungsform trainiert das Trainings-Framework 1904 ein untrainiertes neuronales Netzwerk 1906 und ermöglicht ihm unter Verwendung von hierin beschriebenen Verarbeitungsressourcen, trainiert zu werden, um ein trainiertes neuronales Netzwerk 1908 zu erzeugen. Bei mindestens einer Ausführungsform können Gewichte willkürlich oder durch Vortraining unter Verwendung eines Deep-Belief-Netzwerks ausgewählt werden. Bei mindestens einer Ausführungsform kann das Trainieren entweder auf eine beaufsichtigte, zum Teil beaufsichtigte oder unbeaufsichtigte Weise ausgeführt werden.

[0154] Bei mindestens einer Ausführungsform wird das untrainierte neuronale Netzwerk 1906 unter Verwendung von beaufsichtigtem Lernen trainiert, wobei der Trainingsdatensatz 1902 eine Eingabe gepaart mit einer gewünschten Ausgabe für eine Eingabe beinhaltet, oder wobei der Trainingsdatensatz 1902 Eingabe beinhaltet, die eine bekannte Ausgabe aufweist und eine Ausgabe des neuronalen Netzwerks 1906 manuell abgestuft wird. Bei mindestens einer Ausführungsform wird ein untrainiertes neuronales Netzwerk 1906 auf eine überwachte Weise trainiert und verarbeitet Eingaben von dem Trainingsdatensatz 1902 und vergleicht sich ergebende Ausgaben mit einem Satz von erwarteten oder gewünschten Ausgaben. Bei mindestens einer Ausführungsform werden Fehler dann zurück durch das untrainierte neuronale Netzwerk 1906 propagiert. Bei mindestens einer Ausführungsform stellt das Trainings-Framework 1904 Gewichte ein, die das untrainierte neuronale Netzwerk 1906 steuern. Bei mindestens einer Ausführungsform beinhaltet das Trainings-Framework 1904 Werkzeuge, um zu überwachen, wie gut das untrainierte neuronale Netzwerk 1906 in Richtung eines Modells, wie dem trainierten neuronalen Netzwerk 1908, konvergiert, das dazu geeignet ist, korrekte Antworten zu erzeugen, wie zum Beispiel in dem Ergebnis 1914, basierend auf bekannten Eingabedaten, wie einem neuen Datensatz 1912. Bei mindestens einer Ausführungsform trainiert das Trainings-Framework 1904 das untrainierte neuronale Netzwerk 1906 wiederholt, während Gewichte angepasst werden, um eine Ausgabe des untrainierten neuronalen Netzwerks 1906 unter Verwendung einer Verlustfunktion und eines Anpassungsalgorithmus, wie einem stochastischen Gradientenabstieg, zu verfeinern. Bei mindestens einer Ausführungsform trainiert das Trainings-Framework 1904 das untrainierte neuronale Netzwerk 1906, bis das untrainierte neuronale Netzwerk 1906 eine gewünschte Genauigkeit erreicht. Bei mindestens einer Ausführungsform kann das trainierte neuronale Netzwerk 1908 dann eingesetzt werden, um eine beliebige Anzahl von Maschinenlernoperationen umzusetzen.

[0155] Bei mindestens einer Ausführungsform wird das untrainierte neuronale Netzwerk 1906 unter Verwendung von unbeaufsichtigtem Lernen trainiert, wobei das untrainierte neuronale Netzwerk 1906 versucht, sich selbst unter Verwendung von nicht gekennzeichneten Daten zu trainieren. Bei mindestens einer Ausführungsform beinhaltet der unüberwachte Lerntrainingsdatensatz 1902 Eingabedaten ohne jegliche assoziierte Ausgabedaten oder „Ground-Truth-Daten“. Bei mindestens einer Ausführungsform kann das untrainierte neuronale Netzwerk 1906 Gruppierungen innerhalb des Trainingsdatensatzes 1902 lernen und kann bestimmen, wie sich individuelle Eingaben den untrainierten Datensatz 1902 betreffen. Bei mindestens einer Ausführungsform kann unüberwachtes Training dazu verwendet werden, eine sich selbst organisierende Karte zu erzeugen, die eine Art von trainiertem neuronalem Netzwerk 1908 ist, die dazu in der Lage ist, Operationen auszuführen, die beim Reduzieren der Dimensionalität des neuen Datensatzes 1912 nützlich sind. Bei mindestens einer Ausführungsform kann unüberwachtes Training auch dazu verwendet werden, Anomalieerfassung auszuführen, was die Identifikation von Datenpunkten in einem neuen Datensatz 1912, die von normalen Mustern des neuen Datensatzes 1912 abweichen, erlaubt.

[0156] Bei mindestens einer Ausführungsform kann halb überwachtes Lernen verwendet werden, wobei es sich um eine Technik handelt, bei der der Trainingsdatensatz 1902 eine Mischung aus gekennzeichneten und nicht gekennzeichneten Daten beinhaltet. Bei mindestens einer Ausführungsform kann das Trainings-Framework 1904 dazu verwendet werden, inkrementelles Lernen auszuführen, wie zum Beispiel durch transferierte Lerntechniken. Bei mindestens einer Ausführungsform ermöglicht es inkrementelles Lernen dem trainierten neuronalen Netzwerk 1908, sich an den neuen Datensatz 1912 anzupassen, ohne Wissen, das innerhalb des Netzwerks während des anfänglichen Trainings vermittelt wird, zu vergessen.

RECHENZENTRUM

[0157] **Fig. 20** zeigt ein Beispiel eines Rechenzentrums 2000, in dem mindestens eine Ausführungsform verwendet werden kann. Bei mindestens einer Ausführungsform beinhaltet das Rechenzentrum 2000 eine Rechenzentruminfrastrukturschicht 2010, eine Framework-Schicht 2020, eine Softwareschicht 2030 und eine Anwendungsschicht 2040.

[0158] Bei mindestens einer Ausführungsform, wie es in **Fig. 20** gezeigt ist, kann die Rechenzentrumsinfrastrukturschicht 2010 einen Ressourcen-Orchestrator 2012, gruppierte Rechenressourcen 2014 und Knoten-Rechenressourcen („Knoten-C.R.s“) 2016(1)-2016(N) aufweisen, wobei „N“ eine beliebige ganze, positive Zahl darstellt. Bei mindestens einer Ausführungsform können die Knoten-C.R.s 2016(1)-2016(N) eine beliebige Anzahl von Zentraleinheiten („CPUs“) oder anderen Prozessoren (einschließlich Beschleunigern, Field Programmable Gate Arrays (FPGAs), Grafikprozessoren usw.), Speichereinrichtungen (z. B., dynamischer Festwertspeicher), Speichereinrichtungen (z. B. Festkörper- oder Festplattenlaufwerke), Netzwerk-Eingabe-/Ausgabeeinrichtungen („NW E/A“), Netzwerk-Switches, virtuelle Maschinen („VMs“), Stromversorgungsmodule und Kühlmodule, usw einschließen. Bei mindestens einer Ausführungsform kann es sich bei einem oder mehreren Knoten-C.R.s unter den Knoten-C.R.s 2016(1)-2016(N) um einen Server handeln, der über eine oder mehrere der oben erwähnten Rechenressourcen verfügt.

[0159] Bei mindestens einer Ausführungsform können die gruppierten Rechenressourcen 2014 getrennte Gruppierungen von Knoten-C.R.s aufweisen, die in einem oder mehreren Racks (nicht gezeigt) untergebracht sind, oder viele Racks, die in Rechenzentren an verschiedenen geografischen Standorten untergebracht sind (ebenfalls nicht gezeigt). Getrennte Gruppierungen von Knoten-CRs innerhalb gruppierter Berechnungsressourcen 2014 können gruppierte Rechen-, Netzwerk-, Arbeitsspeicher- oder Datenspeicherressourcen beinhalten, die konfiguriert oder zugewiesen sein können, um eine oder mehrere Arbeitslasten zu unterstützen. Bei mindestens einer Ausführungsform können mehrere Knoten-C.R.s, die CPUs oder Prozessoren aufweisen, in einem oder mehreren Racks gruppiert sein, um Rechenressourcen zur Unterstützung einer oder mehrerer Arbeitslasten bereitzustellen. Bei mindestens einer Ausführungsform können ein oder mehrere Racks auch eine beliebige Anzahl von Stromversorgungsmodulen, Kühlmodulen und Netzwerk-Switches in beliebiger Kombination aufweisen.

[0160] Bei mindestens einer Ausführungsform kann der Ressourcen-Orchestrator 2012 einen oder mehrere Knoten C.R.s 2016(1)-2016(N) und/oder gruppierte Rechenressourcen 2014 ausgestalten oder anderweitig steuern. Bei mindestens einer Ausführungsform kann der Ressourcen-Orchestrator 2012 eine Software-Design-Infrastruktur („SDI“)-Verwaltungseinheit für das Rechenzentrum 2000 aufweisen. Bei mindestens einer Ausführungsform kann der Ressourcen-Orchestrator Hardware, Software oder eine Kombination davon aufweisen.

[0161] Bei mindestens einer Ausführungsform, wie es in **Fig. 20** gezeigt ist, weist die Framework-Schicht 2020 einen Job Scheduler 2032, einen Konfigurationsmanager 2034, einen Ressourcenmanager 2036 und ein verteiltes Dateisystem 2038 auf. Bei mindestens einer Ausführungsform kann die Framework-Schicht 2020 einen Rahmen bzw. Framework zur Unterstützung der Software 2032 der Softwareschicht 2030 und/oder einer oder mehrerer Anwendung(en) 2042 der Anwendungsschicht 2040 aufweisen. Bei mindestens einer Ausführungsform kann die Software 2032 oder die Anwendung(en) 2042 jeweils webbasierte Dienstsoftware oder Anwendungen aufweisen, wie sie beispielsweise von Amazon Web Services, Google Cloud und Microsoft Azure bereitgestellt werden. Bei mindestens einer Ausführungsform kann es sich bei der Framework-Schicht 2020 um eine Art freies und quelloffenes Software-Webanwendungs-Framework wie Apache Spark™ (im Folgenden „Spark“) handeln, das ein verteiltes Dateisystem 2038 für die Verarbeitung großer Datenmengen (z. B. „Big Data“) nutzen kann. Bei mindestens einer Ausführungsform kann der Job Scheduler 2032 einen Spark-Treiber aufweisen, um die Planung von Arbeitslasten zu erleichtern, die von verschiedenen Schichten des Rechenzentrums 2000 unterstützt werden. Bei mindestens einer Ausführungsform kann der Konfigurationsmanager 2034 in der Lage sein, um verschiedene Schichten wie die Softwareschicht 2030 und die Framework-Schicht 2020, die Spark und das verteilte Dateisystem 2038 aufweist, zur Unterstützung der Verarbeitung großer Datenmengen zu konfigurieren. Bei mindestens einer Ausführungsform kann der Ressourcenmanager 2036 in der Lage sein, geclusterte oder gruppierte Computerressourcen zu verwalten, die zur Unterstützung des verteilten Dateisystems 2038 und des Job Schedulers 2032 zugeordnet oder zugewiesen sind. Bei mindestens einer Ausführungsform können geclusterte oder gruppierte Rechenressourcen gruppierte Rechenressourcen 2014 in der Infrastrukturschicht 2010 des Rechenzentrums aufweisen. Bei mindestens einer Ausführungsform kann der Ressourcenmanager 2036 mit dem Ressourcenorchestrator 2012 koordiniert sein, um diese zugeordneten oder zugewiesenen Computerressourcen zu verwalten.

[0162] Bei mindestens einer Ausführungsform kann die in der Softwareschicht 2030 enthaltene Software 2032 Software beinhalten, die von mindestens Abschnitten der Knoten-C.R.s2016(1)-2016(N), der gruppierten Rechenressourcen 2014 und/oder des verteilten Dateisystems 2038 der Framework-Schicht 2020 verwendet wird. Eine oder mehrere Arten von Software können Internet-Webseiten-Suchsoftware, E-Mail-Virenscan-Software, Datenbanksoftware und Streaming-Video-Content-Software beinhalten, ohne darauf beschränkt zu sein.

[0163] Bei mindestens einer Ausführungsform kann (können) die in der Anwendungsschicht 2040 enthaltene (n) Anwendung(en) 2042 eine oder mehrere Arten von Anwendungen aufweisen, die von mindestens Abschnitten der Knoten C.R.s 2016(1)-2016(N), gruppierten Rechenressourcen 2014 und/oder dem verteilten Dateisystem 2038 der Framework-Schicht 2020 verwendet werden. Eine oder mehrere Arten von Anwendungen können eine beliebige Anzahl einer Genomikanwendung, einer kognitiven Rechenanwendung und einer maschinellen Lernanwendung beinhalten, die Trainings- oder Ableitungssoftware, Framework-Software des maschinellen Lernens (z. B. PyTorch, TensorFlow, Caffe usw.) oder andere maschinelle Lernanwendungen beinhaltet, ohne darauf beschränkt zu sein, die in Verbindung mit einer oder mehreren Ausführungsformen verwendet werden.

[0164] Bei mindestens einer Ausführungsform kann jeder von Konfigurationsmanager 2034, Ressourcenmanager 2036 und Ressourcen-Orchestrator 2012 eine beliebige Anzahl und Art von selbstmodifizierenden Aktionen basierend auf einer beliebigen Menge und Art von Daten, die auf jede technisch machbare Weise ermittelt werden, umsetzen. Bei mindestens einer Ausführungsform können selbstmodifizierende Aktionen einen Rechenzentrumsbetreiber des Rechenzentrums 2000 davon entlasten, möglicherweise schlechte Konfigurationsentscheidungen zu treffen und möglicherweise nicht ausgelastete und/oder schlecht funktionierende Abschnitte eines Rechenzentrums zu vermeiden.

[0165] Bei mindestens einer Ausführungsform kann das Rechenzentrum 2000 Werkzeuge, Dienste, Software oder andere Ressourcen aufweisen, um ein oder mehrere Modelle zum maschinellen Lernen zu trainieren oder Informationen unter Verwendung eines oder mehrerer Modelle zum maschinellen Lernen gemäß einer oder mehrerer hier beschriebener Ausführungsformen vorherzusagen oder abzuleiten. Beispielsweise kann bei mindestens einer Ausführungsform ein Modell zum maschinellen Lernen durch Berechnung von Gewichtungsparametern gemäß einer neuronalen Netzwerkarchitektur unter Verwendung von Software und Rechenressourcen trainiert werden, die oben in Bezug auf das Rechenzentrum 2000 beschrieben wurden. Bei mindestens einer Ausführungsform können trainierte Modelle zum maschinellen Lernen, die einem oder mehreren neuronalen Netzen entsprechen, verwendet werden, um Informationen abzuleiten oder vorherzusagen, wobei die oben beschriebenen Ressourcen in Bezug auf das Rechenzentrum 2000 verwendet werden, indem Gewichtungsparameter verwendet werden, die durch eine oder mehrere hier beschriebene Trainingstechniken berechnet werden.

[0166] Bei mindestens einer Ausführungsform kann das Rechenzentrum CPUs, anwendungsspezifische integrierte Schaltungen (ASICs), GPUs, FPGAs oder andere Hardware verwenden, um Training und/oder Inferencing unter Verwendung der oben beschriebenen Ressourcen durchzuführen. Darüber hinaus können eine oder mehrere der oben beschriebenen Software- und/oder Hardwareressourcen als Dienst ausgestaltet sein, um es Benutzern zu ermöglichen, Informationen zu trainieren oder Inferencing durchzuführen, wie etwa Bilderkennung, Spracherkennung oder andere Dienste der künstlichen Intelligenz.

[0167] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform kann Ableitungs- und/oder Trainingslogik 1815 in dem System **Fig. 20** zum Ableiten oder Vorhersagen von Operationen mindestens zum Teil basierend auf Gewichtungsparametern verwendet werden, die unter Verwendung von neuronalen Netzwerktrainingsoperationen, Funktionen und/oder Architekturen von neuronalen Netzwerken oder in dieser Schrift beschriebenen Verwendungsfällen von neuronalen Netzwerken berechnet wurden.

[0168] Bei mindestens einer Ausführungsform wird das System **Fig. 20** eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird das System **Fig. 20** eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird das System **Fig. 20** eingesetzt, um ein oder mehrere neuronale Netzwerke umzuset-

zen, die einen Diskriminator und einen Generator umfassen, und das System **Fig. 20** wird in Verbindung mit einem oder mehreren Prozessen verwendet, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teile des Trainings berechnet werden.

Autonomes Fahrzeug

[0169] Fig. 21A veranschaulicht ein Beispiel eines autonomen Fahrzeugs 2100 gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform kann das autonome Fahrzeug 2100 (hier alternativ als „Fahrzeug 2100“ bezeichnet) ohne Einschränkung ein Personenfahrzeug sein, wie z. B ein Pkw, ein Lkw, ein Bus und/oder eine andere Art von Fahrzeug, das einen oder mehrere Fahrgäste aufnimmt. Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 ein Sattelschlepper sein, der für den Transport von Gütern verwendet wird. Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 ein Flugzeug, ein Roboterfahrzeug oder eine andere Art von Fahrzeug sein.

[0170] Autonome Fahrzeuge können in Form von Automatisierungsstufen beschrieben werden, die von der National Highway Traffic Safety Administration („NHTSA“), einer Abteilung des US-Verkehrsministeriums, und der Society of Automotive Engineers („SAE“) „Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles“ (z. B. Standard Nr. J3016-201806, veröffentlicht am 15. Juni 2018, Standard Nr. J3016-201609, veröffentlicht am 30. September 2016, sowie frühere und zukünftige Versionen dieses Standards) definiert sind. In einer oder mehreren Ausführungsformen kann das Fahrzeug 2100 in der Lage sein, um eine Funktionalität gemäß einer oder mehrerer der Stufen 1 bis 5 der Stufen des autonomen Fahrens auszuführen. Beispielsweise kann das Fahrzeug 2100 bei mindestens einer Ausführungsform in der Lage sein, je nach Ausführungsform bedingt automatisiert (Stufe 3), hochautomatisiert (Stufe 4) und/oder vollständig automatisiert (Stufe 5) zu fahren.

[0171] Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 ohne Einschränkung Komponenten wie ein Fahrgestell, eine Fahrzeugkarosserie, Räder (z. B. 2, 4, 6, 8, 18 usw.), Reifen, Achsen und andere Komponenten eines Fahrzeugs aufweisen. Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 ohne Einschränkung ein Antriebssystem 2150 aufweisen, wie z. B. einen Verbrennungsmotor, ein Hybrid-Elektroantrieb, einen reinen Elektromotor und/oder einen anderen Antriebssystemtyp. Bei mindestens einer Ausführungsform kann das Antriebssystem 2150 mit einem Antriebsstrang des Fahrzeugs 2100 verbunden sein, der unter anderem ein Getriebe aufweisen kann, um den Antrieb des Fahrzeugs 2100 zu ermöglichen. Bei mindestens einer Ausführungsform kann das Antriebssystem 2150 in Reaktion auf den Empfang von Signalen von einer Drosselklappe/einem Gaspedal (bzw. mehreren Gaspedalen) 2152 gesteuert werden.

[0172] Bei mindestens einer Ausführungsform wird ein Lenksystem 2154, das ohne Einschränkung ein Lenkrad aufweisen kann, verwendet, um ein Fahrzeug 2100 zu lenken (z. B entlang eines gewünschten Weges oder einer Route), wenn ein Antriebssystem 2150 in Betrieb ist (z. B wenn das Fahrzeug in Bewegung ist). Bei mindestens einer Ausführungsform kann ein Lenksystem 2154 Signale von einem oder mehreren Lenkaktoren 2156 empfangen. Ein Lenkrad kann für eine Vollautomatisierungs-Funktionalität (Stufe-5-Funktionalität) optional sein. Bei mindestens einer Ausführungsform kann ein Bremssensorsystem 2146 verwendet werden, um die Fahrzeugbremsen in Reaktion auf den Empfang von Signalen von einem oder mehreren Bremsaktuatoren 2148 und/oder Bremssensoren zu betätigen.

[0173] Bei mindestens einer Ausführungsform stellen eine oder mehrere Steuervorrichtungen 2136, die ohne Einschränkung ein oder mehrere Systeme auf Chips (System on Chips-„SoCs“) (in **Fig. 21A** nicht gezeigt) und/oder eine oder mehrere Grafikverarbeitungseinheiten (Graphics Processing Units-„GPUs“) beinhalten können, Signale (die z. B. Befehle darstellen) zu einer oder mehreren Komponenten und/oder Systemen des Fahrzeugs 2100 bereit. Bei mindestens einer Ausführungsform kann (können) die Steuerung(en) 2136 beispielsweise Signale zur Betätigung der Fahrzeugbremsen über die Bremsaktuatoren 2148, zur Betätigung des Lenksystems 2154 über den/die Lenkaktuator(en) 2156 und zur Betätigung des Antriebssystems 2150 über eine Drosselklappe / (ein) Gaspedal(e) 2152 senden. Die Steuervorrichtung 2136 kann eine oder mehrere fahrzeuginterne (z. B. integrierte) Rechenvorrichtungen (z. B. Supercomputer) beinhalten, die Sensorsignale verarbeiten und Betriebsbefehle (z. B. Signale, die Befehle darstellen) ausgeben, um das autonome Fahren zu ermöglichen und/oder einen menschlichen Fahrer beim Fahren des Fahrzeugs 2100 zu unterstützen. Bei mindestens einer Ausführungsform kann (können) die Steuervorrichtung(en) 2136 eine erste Steuervorrichtung 2136 für autonome Fahrfunktionen, eine zweite Steuervorrichtung 2136 für funktionale Sicherheitsfunktionen, eine dritte Steuervorrichtung 2136 für Funktionen der künstlichen Intelligenz (z. B. Computer

Vision), eine vierte Steuervorrichtung 2136 für Infotainment-Funktionen, eine fünfte Steuervorrichtung 2136 für Redundanz in Notfällen und/oder andere Steuervorrichtungen beinhalten. Bei mindestens einer Ausführungsform kann eine einzige Steuerung 2136 zwei oder mehrere der oben genannten Funktionen übernehmen, zwei oder mehr Steuerungen 2136 können eine einzige Funktion übernehmen und/oder eine beliebige Kombination davon.

[0174] Bei mindestens einer Ausführungsform stellt/stellen die Steuerung(en) 2136 Signale zur Steuerung einer oder mehrerer Komponenten und/oder Systeme des Fahrzeugs 2100 als Reaktion auf Sensordaten bereit, die von einem oder mehreren Sensoren empfangen werden (z. B Sensoreingaben). Bei mindestens einer Ausführungsform können Sensordaten beispielsweise und ohne Einschränkung von (einem) Global Navigation Satellite Systems („GNSS")-Sensor(en) 2158 (z. B, Global Positioning System-Sensor(en)), RADAR-Sensor(en) 2160, Ultraschallsensor(en) 2162, LIDAR-Sensor(en) 2164, Trägheitsmesseinheit-Sensor(en) („IMU") 2166 (z. B. Beschleunigungsmesser, Gyroskop(e), Magnetkompass(e), Magnetometer usw.), Mikrofon(en) 2196, Stereokamera(s) 2168, Weitwinkelkamera(s) 2170 (z. B., Fischaugenkameras), Infrarotkamera(s) 2172, Umgebungskamera(s) 2174 (z. B 360-Grad-Kameras), Fernkameras (nicht in **Fig. 21A** gezeigt), Mittelbereichskamera(s) (nicht in **Fig. 21A** gezeigt), Geschwindigkeitssensor(en) 2144 (z. B zur Messung der Geschwindigkeit des Fahrzeugs 2100), Vibrationssensor(en) 2142, Lenksensor(en) 2140, Bremssensor(en) (z. B als Teil des Bremssensorsystems 2146) und/oder anderen Sensortypen empfangen werden.

[0175] Bei mindestens einer Ausführungsform kann eine oder können mehrere der Steuervorrichtungen 2136 Eingaben (z. B. in Form von Eingabedaten) von einem Kombiinstrument 2132 des Fahrzeugs 2100 empfangen und Ausgaben (z. B. in Form von Ausgabedaten, Anzeigedaten usw.) über eine Mensch-Maschine-Schnittstelle („HMI")-Anzeige 2134, einen akustischen Melder, einen Lautsprecher und/oder über andere Komponenten des Fahrzeugs 2100 bereitstellen. Bei mindestens einer Ausführungsform können die Ausgaben Informationen, wie Fahrzeuggeschwindigkeit, Drehzahl, Zeit, Kartendaten (z. B. eine hochauflösende Karte (in **Fig. 21** Anicht gezeigt)), Positionsdaten (z. B. die Position des Fahrzeugs 2100, wie auf einer Karte), Richtung, Position anderer Fahrzeuge (z. B. ein Belegungsraster), Informationen über Objekte und den Status von Objekten, wie es von der/den Steuervorrichtung(en) 2136 wahrgenommen wird, usw. aufweisen. Bei mindestens einer Ausführungsform kann die HMI-Anzeige 2134 beispielsweise Informationen über das Vorhandensein eines oder mehrerer Objekte (z. B. ein Straßenschild, ein Warnschild, eine sich ändernde Ampel usw.) und/oder Informationen über Fahrmanöver anzeigen, die das Fahrzeug ausgeführt hat, gerade ausführt oder ausführen wird (z. B. Spurwechsel jetzt, Ausfahrt 34B in zwei Meilen nehmen usw.).

[0176] Bei mindestens einer Ausführungsform weist das Fahrzeug 2100 darüber hinaus eine Netzwerkschnittstelle 2124 auf, die (eine) drahtlose Antenne(n) 2126 und/oder (ein) Modem(e) zur Kommunikation über ein oder mehrere Netzwerke verwenden kann. Zum Beispiel kann bei mindestens einer Ausführungsform die Netzwerkschnittstelle 2124 in der Lage sein, über Long-Term Evolution („LTE"), Wideband Code Division Multiple Access („WCDMA"), Universal Mobile Telecommunications System („UMTS"), Global System for Mobile communication („GSM"), IMT-CDMA Multi-Carrier („CDMA2000"), etc. zu kommunizieren. Bei mindestens einer Ausführungsform kann/können die drahtlose(n) Antenne(n) 2126 auch die Kommunikation zwischen Objekten in der Umgebung (z. B. Fahrzeuge, mobile Einrichtungen usw.) ermöglichen, wobei lokale Netzwerke wie Bluetooth, Bluetooth Low Energy („LE"), Z-Wave, ZigBee usw. und/oder Weitverkehrsnetzwerke mit geringer Leistung („LPWANs") wie LoRaWAN, SigFox usw. verwendet werden.

[0177] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform kann Ableitungs- und/oder Trainingslogik 1815 in dem System in **Fig. 21A** zum Ableiten oder Vorhersagen von Operationen mindestens zum Teil basierend auf Gewichtungsparametern verwendet werden, die unter Verwendung von neuronalen Netzwerktrainingsoperationen, Funktionen und/oder Architekturen von neuronalen Netzwerken oder in dieser Schrift beschriebenen Anwendungsfällen von neuronalen Netzwerken berechnet werden.

[0178] Bei mindestens einer Ausführungsform wird das System **Fig. 21A** eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird das System **Fig. 21A** eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird das System **Fig. 21A** eingesetzt, um ein oder mehrere neuronale Netzwerke umzuset-

zen, die einen Diskriminator und einen Generator umfassen, und das System **Fig. 21A** wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0179] **Fig. 21B** veranschaulicht ein Beispiel von Kameralagen und -sichtfeldern für das autonome Fahrzeug 2100 der **Fig. 21A** gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform stellen die Kameras und die jeweiligen Sichtfelder eine beispielhafte Ausführungsform dar und sind nicht als einschränkend zu betrachten. Beispielsweise können bei mindestens einer Ausführungsform zusätzliche und/oder alternative Kameras vorhanden sein und/oder die Kameras können an anderen Stellen des Fahrzeugs 2100 angeordnet sein.

[0180] Bei mindestens einer Ausführungsform können die Kameratypen für Kameras Digitalkameras aufweisen, die für die Verwendung mit Komponenten und/oder Systemen des Fahrzeugs 2100 angepasst sein können, sind aber nicht darauf beschränkt. Die Kamera(s) kann (können) mit einer Automobilsicherheitsintegritätsstufe (Automotive Safety Integrity Level - „ASIL“) B und/oder mit einer anderen ASIL arbeiten. Bei mindestens einer Ausführungsform können die Kameratypen je nach Ausführungsform eine beliebige Bildaufnahmerate, wie 60 Bilder pro Sekunde (fps), 1220 fps, 240 fps usw., erreichen. Bei mindestens einer Ausführungsform können die Kameras Rolling Shutter, Global Shutter, einen anderen Verschlusstyp oder eine Kombination davon verwenden. Bei mindestens einer Ausführungsform kann die Farbfilteranordnung eine Rot-Klar-Klar-Klar-Farbfilteranordnung („RCCC“), eine Rot-Klar-Klar-Blau-Farbfilteranordnung („RCCB“), eine Rot-Blau-Grün-Klar-Farbfilteranordnung („RBGC“), eine Foveon X3-Farbfilteranordnung, eine Bayer-Sensor-Farbfilteranordnung („RGGB“), eine Monochromsensor-Farbfilteranordnung und/oder eine andere Art von Farbfilteranordnung aufweisen. Bei mindestens einer Ausführungsform können zur Erhöhung der Lichtempfindlichkeit Klar-Pixel-Kameras, wie z. B. Kameras mit einer RCCC-, einer RCCB- und/oder einer RBGC-Farbfilteranordnung, verwendet werden.

[0181] Bei mindestens einer Ausführungsform kann eine oder können mehrere Kameras verwendet werden, um fortschrittliche Fahrerassistenzsysteme („ADAS“) auszuführen (z. B. als Teil eines redundanten oder ausfallsicheren Designs). So kann bei mindestens einer Ausführungsform eine Multifunktions-Monokamera installiert sein, die Funktionen wie Spurhalteassistent, Verkehrszeichenassistent und intelligente Scheinwerfersteuerung bietet. Bei mindestens einer Ausführungsform kann eine oder können mehrere der Kameras (z. B alle Kameras) gleichzeitig Bilddaten (z. B Video) aufzeichnen und bereitstellen.

[0182] Bei mindestens einer Ausführungsform kann eine oder können mehrere der Kameras in einer Montageanordnung, wie z. B einer kundenspezifisch entworfenen (dreidimensionalen („3D“) gedruckten) Anordnung, montiert sein, um Streulicht und Reflexionen aus dem Fahrzeuginneren (z. B Reflexionen vom Armaturenbrett, die in den Windschutzscheibenspiegeln reflektiert werden) auszuschalten, die die Fähigkeit der Kamera zur Bilddatenerfassung beeinträchtigen können. Unter Bezugnahme auf die Außenspiegel-Montageanordnungen können bei mindestens einer Ausführungsform die Anordnungen für die Außenspiegel individuell in 3D gedruckt werden, so dass die Kameramontageplatte der Form des Außenspiegels entspricht. Bei mindestens einer Ausführungsform kann (können) die Kamera(s) in dem Außenspiegel integriert sein. Für Seitensichtkameras kann eine Kamera oder können Kameras bei mindestens einer Ausführungsform auch innerhalb von vier Säulen in jeder Ecke des Passagierraums integriert sein.

[0183] Bei mindestens einer Ausführungsform können Kameras mit einem Sichtfeld, das Abschnitte der Umgebung vor dem Fahrzeug 2100 aufweist (z. B. nach vorne gerichtete Kameras), für die Rundumsicht verwendet werden, um bei der Erkennung von nach vorne gerichteten Wegen und Hindernissen zu helfen, sowie mit Hilfe einer oder mehrerer Steuerungen 2136 und/oder Steuer-SoCs Informationen bereitzustellen, die für die Erstellung eines Belegungsrasters und/oder die Bestimmung bevorzugter Fahrzeugwege entscheidend sind. Bei mindestens einer Ausführungsform können nach vorne gerichtete Kameras verwendet werden, um viele der gleichen ADAS-Funktionen wie LIDAR auszuführen, einschließlich, ohne Einschränkung, Notbremsung, Fußgängererkennung und Kollisionsvermeidung. Bei mindestens einer Ausführungsform können nach vorne gerichtete Kameras auch für ADAS-Funktionen und -Systeme verwendet werden, einschließlich, aber nicht beschränkt auf eine Warnung vor dem Verlassen der Fahrspur („LDW“), autonome Geschwindigkeitsregelung („ACC“) und/oder andere Funktionen wie Verkehrszeichenerkennung.

[0184] Bei mindestens einer Ausführungsform kann eine Vielzahl von Kameras in einer nach vorne gerichteten Konfiguration verwendet werden, einschließlich z. B. einer monokularen Kameraplattform, die einen

CMOS-Farbbildwandler („Complementary Metal Oxide Semiconductor“) aufweist. Bei mindestens einer Ausführungsform kann eine Weitsichtkamera 2170 dazu verwendet werden, Objekte wahrzunehmen, die von einem Außenrand in das Blickfeld gelangen (z. B. Fußgänger, Querverkehr oder Fahrräder). Obwohl lediglich eine Weitsichtkamera 2170 in **Fig. 21B** veranschaulicht ist, kann es bei anderen Ausführungsformen eine beliebige Anzahl (die Null beinhaltet) von Weitsichtkameras 2170 an dem Fahrzeug 2100 geben. Bei mindestens einer Ausführungsform kann eine beliebige Anzahl von Weitwinkelkamera(s) 2198 (z. B. ein Weitwinkel-Stereokamerapaar) zur tiefenbasierten Objekterkennung verwendet werden, insbesondere für Objekte, für die ein neuronales Netz noch nicht trainiert worden ist. Bei mindestens einer Ausführungsform können die Weitwinkelkamera(s) 2198 auch zur Objekterkennung und -klassifizierung sowie zur grundlegenden Objektverfolgung verwendet werden.

[0185] Bei mindestens einer Ausführungsform kann eine beliebige Anzahl von Stereokamera(s) 2168 auch in einer nach vorne gerichteten Konfiguration vorhanden sein. Bei mindestens einer Ausführungsform kann eine oder können mehrere der Stereokamera(s) 2168 eine integrierte Steuereinheit beinhalten, die eine skalierbare Verarbeitungseinheit umfasst, die eine programmierbare Logik („FPGA“) und einen Multi-Core-Mikroprozessor mit einer integrierten Controller Area Network- („CAN“) oder Ethernet-Schnittstelle auf einem einzigen Chip bereitstellen kann. Bei mindestens einer Ausführungsform kann eine solche Einheit verwendet werden, um eine 3D-Karte der Umgebung des Fahrzeugs 2100 zu erstellen, die eine Abstandsschätzung für alle Punkte im Bild aufweist. Bei mindestens einer Ausführungsform kann eine oder können mehrere der Stereokamera(s) 2168 ohne Einschränkung kompakte(n) Stereosicht-Sensor(en) aufweisen, die ohne Einschränkung zwei Kameralinsen (je eine auf der linken und rechten Seite) und einen Bildverarbeitungschip enthalten können, der den Abstand zwischen dem Fahrzeug 2100 und dem Zielobjekt messen und die erzeugten Informationen (z. B. Metadaten) verwenden kann, um autonome Notbrems- und Spurhaltewarnfunktionen zu aktivieren. Bei mindestens einer Ausführungsform können auch andere Typen von Stereokameras 2168 zusätzlich oder alternativ zu den hier beschriebenen verwendet werden.

[0186] Bei mindestens einer Ausführungsform können Kameras mit einem Sichtfeld, das Abschnitte der Umgebung seitlich des Fahrzeugs 2100 aufweist (z. B. Seitenkameras), für die Umgebungsansicht verwendet werden und Informationen liefern, die zur Erstellung und Aktualisierung des Belegungsgitters sowie zur Erzeugung von Seitenaufprallwarnungen verwendet werden. Beispielsweise könnten bei mindestens einer Ausführungsform eine oder mehrere Surround-Kameras 2174 (z. B. vier Surround-Kameras 2174, wie in **Fig. 21B** veranschaulicht) auf den Fahrzeug 2100 positioniert sein. Die Surround-Kamera(s) 2174 kann/können ohne Einschränkung eine beliebige Anzahl und Kombination von Weitsichtkamera(s) 2170, Fischaugenkamera(s), 360-Grad-Kamera(s) und/oder dergleichen beinhalten. Zum Beispiel können bei mindestens einer Ausführungsform vier Fischaugenkameras an der Vorderseite, der Rückseite und den Seiten des Fahrzeugs 2100 positioniert sein. Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 drei Surround-Kamera(s) 2174 (z. B links, rechts und hinten) verwenden und eine oder mehrere andere Kamera(s) (z. B eine nach vorne gerichtete Kamera) als vierte Surround-View-Kamera nutzen.

[0187] Bei mindestens einer Ausführungsform können Kameras mit einem Sichtfeld, das Abschnitte der Umgebung hinter dem Fahrzeug 2100 aufweist (z. B. Rückfahrkameras), für die Einparkhilfe, die Umgebungsansicht, die Heckkollisionswarnungen und die Erstellung und Aktualisierung des Belegungsgitters verwendet werden. Bei mindestens einer Ausführungsform kann eine Vielzahl von Kameras verwendet werden, einschließlich, aber nicht beschränkt auf, Kameras, die auch als nach vorne gerichtete Kamera(s) geeignet sind (z. B Weitbereichskameras 2198 und/oder Mittelbereichskamera(s) 2176, Stereokamera(s) 2168), Infrarotkamera(s) 2172, usw.), wie es hier beschrieben ist.

[0188] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform kann Ableitungs- und/oder Trainingslogik 1815 in dem System in **Fig. 21B** zum Ableiten oder Vorhersagen von Operationen mindestens zum Teil basierend auf Gewichtungsparametern verwendet werden, die unter Verwendung von neuronalen Netzwerktrainingsoperationen, Funktionen und/oder Architekturen von neuronalen Netzwerken oder in dieser Schrift beschriebenen Anwendungsfällen von neuronalen Netzwerken berechnet wurden.

[0189] Bei mindestens einer Ausführungsform wird das System **Fig. 21B** eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird das System **Fig. 21B** eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sicht-

punkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird das System **Fig. 21B** eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und das System **Fig. 21B** wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0190] **Fig. 21C** ist ein Blockdiagramm, das eine beispielhafte Systemarchitektur für das autonome Fahrzeug 2100 der **Fig. 21A** gemäß mindestens einer Ausführungsform veranschaulicht. In mindestens einer Ausführungsform ist jede Komponente, jedes Merkmal und jedes System des Fahrzeugs 2100 in **Fig. 21C** als über einen Bus 2102 verbunden dargestellt. In mindestens einer Ausführungsform kann der Bus 2102 ohne Einschränkung eine CAN-Datenschnittstelle aufweisen (hier alternativ als „CAN-Bus“ bezeichnet). Bei mindestens einer Ausführungsform kann ein CAN ein Netzwerk innerhalb des Fahrzeugs 2100 sein, das zur Unterstützung der Steuerung verschiedener Merkmale und Funktionen des Fahrzeugs 2100 verwendet wird, wie z.B. Betätigung der Bremsen, Beschleunigung, Bremsen, Lenkung, Scheibenwischer usw. Bei mindestens einer Ausführungsform kann der Bus 2102 so ausgestaltet sein, dass er Dutzende oder sogar Hunderte von Knoten aufweist, von denen jeder seine eigene eindeutige Kennung hat (z. B. eine CAN-ID). Bei mindestens einer Ausführungsform kann der Bus 2102 ausgelesen werden, um den Lenkradwinkel, die Fahrgeschwindigkeit, die Motordrehzahl pro Minute („RPMs“), die Tastenpositionen und/oder andere Fahrzeugstatusanzeigen zu ermitteln. Bei mindestens einer Ausführungsform kann der Bus 2102 ein CAN-Bus sein, der ASIL B-konform ist.

[0191] Bei mindestens einer Ausführungsform können zusätzlich zu oder alternativ zu CAN auch FlexRay und/oder Ethernet verwendet werden. Bei mindestens einer Ausführungsform kann eine beliebige Anzahl von Bussen 2102 vorhanden sein, die ohne Einschränkung null oder mehr CAN-Busse, null oder mehr FlexRay-Busse, null oder mehr Ethernet-Busse und/oder null oder mehr andere Arten von Bussen mit einem anderen Protokoll aufweisen können. Bei mindestens einer Ausführungsform können zwei oder mehr Busse 2102 verwendet werden, um unterschiedliche Funktionen auszuführen, und/oder sie können zur Redundanz verwendet werden. Zum Beispiel kann ein erster Bus 2102 für die Kollisionsvermeidungsfunktionalität und ein zweiter Bus 2102 für die Betätigungssteuerung verwendet werden. Bei mindestens einer Ausführungsform kann jeder Bus 2102 mit beliebigen Komponenten des Fahrzeugs 2100 kommunizieren, und zwei oder mehr Busse 2102 können mit denselben Komponenten kommunizieren. Bei mindestens einer Ausführungsform kann jede beliebige Anzahl von System(en) auf (einem) Chip(s) („SoC(s)“) 2104, jede Steuerung 2136 und/oder jeder Computer im Fahrzeug Zugriff auf dieselben Eingabedaten (z. B Eingaben von Sensoren des Fahrzeugs 2100) haben und mit einem gemeinsamen Bus, wie dem CAN-Bus, verbunden sein.

[0192] Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 eine oder mehrere Steuervorrichtungen 2136, wie etwa die in dieser Schrift in Bezug auf **Fig. 21A** beschriebenen, beinhalten. Die Steuervorrichtung 2136 kann für eine Vielfalt von Funktionen verwendet werden. Bei mindestens einer Ausführungsform kann (können) die Steuervorrichtung(en) 2136 mit verschiedenen anderen Komponenten und Systemen des Fahrzeugs 2100 gekoppelt sein und zur Steuerung des Fahrzeugs 2100, zur künstlichen Intelligenz des Fahrzeugs 2100, zum Infotainment für das Fahrzeug 2100 und/oder ähnlichem verwendet werden.

[0193] Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 eine beliebige Anzahl von SoCs 2104 aufweisen. Jedes der SoCs 2104 kann, ohne Einschränkung, zentrale Verarbeitungseinheiten („CPU(s)“) 2106, Grafikverarbeitungseinheiten („GPU(s)“) 2108, Prozessor(en) 2110, Cache(s) 2112, Beschleuniger 2114, Datenspeicher 2116 und/oder andere nicht veranschaulichte Komponenten und Merkmale beinhalten. Bei mindestens einer Ausführungsform können SoC(s) 2104 zur Steuerung des Fahrzeugs 2100 in einer Vielzahl von Plattformen und Systemen verwendet werden. Bei mindestens einer Ausführungsform kann (können) SoC(s) 2104 beispielsweise in einem System (z. B dem System des Fahrzeugs 2100) mit einer High-Definition-Karte („HD“-Karte) 2122 kombiniert sein, die über eine Netzwerkschnittstelle 2124 von einem oder mehreren Servern (in **Fig. 21C** nicht gezeigt) Kartenauffrischungen und/oder -aktualisierungen erhalten kann.

[0194] Bei mindestens einer Ausführungsform kann (können) die CPU(s) 2106 einen CPU-Cluster oder CPU-Komplex (hier alternativ als „CCPLEX“ bezeichnet) aufweisen. Bei mindestens einer Ausführungsform kann (können) die CPU(s) 2106 mehrere Kerne und/oder Level Two („L2“) Caches aufweisen. Bei mindestens einer Ausführungsform kann (können) die CPU(s) 2106 beispielsweise acht Kerne in einer kohärenten Multiprozessorkonfiguration aufweisen. Bei mindestens einer Ausführungsform können die eine oder die mehreren CPUs 2106 vier Dual-Core-Cluster beinhalten, bei denen jedes Cluster einen dedizierten L2-

Cache (z. B. einen L2-Cache mit 2 Megabyte (MB)) aufweist. Bei mindestens einer Ausführungsform kann (können) die CPU(s) 2106 (z. B CCPLEX) dazu konfiguriert sein, den gleichzeitigen Clusterbetrieb derart zu unterstützen, dass jede Kombination von Clustern der CPU(s) 2106 zu jedem Zeitpunkt aktiv sein kann.

[0195] Bei mindestens einer Ausführungsform kann eine oder können mehrere der CPU(s) 2106 Energieverwaltungsfunktionen implementieren, die ohne Einschränkung eines oder mehrere der folgenden Merkmale aufweisen: einzelne Hardwareblöcke können im Leerlauf automatisch getaktet sein, um dynamische Energie zu sparen; jeder Kerntakt kann getaktet sein, wenn der Kern aufgrund der Ausführung von Wait for Interrupt („WFI“)/Wait for Event („WFE“)-Befehlen nicht aktiv Befehle ausführt; jeder Kern kann unabhängig stromgesteuert sein; jeder Kerncluster kann unabhängig taktgesteuert sein, wenn alle Kerne taktgesteuert oder stromgesteuert sind; und/oder jeder Kerncluster kann unabhängig stromgesteuert sein, wenn alle Kerne stromgesteuert sind. Bei mindestens einer Ausführungsform kann/können die CPU(s) 2106 darüber hinaus einen erweiterten Algorithmus für die Verwaltung von Energiezuständen implementieren, bei dem zulässige Energiezustände und erwartete Aufwachzeiten festgelegt werden und die Hardware/der Mikrocode den besten Energiezustand bestimmt, der für Kern, Cluster und CCPLEX einzunehmen ist. Bei mindestens einer Ausführungsform können die Prozessorkerne vereinfachte Sequenzen zur Eingabe des Energiezustands in Software unterstützen, wobei die Arbeit an den Mikrocode ausgelagert ist.

[0196] Bei mindestens einer Ausführungsform kann (können) die GPU(s) 2108 eine integrierte GPU aufweisen (hier alternativ als „iGPU“ bezeichnet). Bei mindestens einer Ausführungsform kann (können) die GPU (s) 2108 programmierbar sein und für parallele Arbeitslasten effizient sein. Bei mindestens einer Ausführungsform kann/können die GPU(s) 2108 einen erweiterten Tensor-Anweisungssatz verwenden. Bei mindestens einer Ausführungsform kann (können) (die) GPU(s) 2108 einen oder mehrere Streaming-Mikroprozessoren beinhalten, wobei jeder Streaming-Mikroprozessor einen L1-Cache (z. B. einen L1-Cache mit einer Speicherkapazität von mindestens 96 KB) beinhalten kann und zwei oder mehr Streaming-Mikroprozessoren einen L2-Cache (z. B. einen L2-Cache mit einer Speicherkapazität von 512 KB) gemeinsam nutzen können. Bei mindestens einer Ausführungsform kann (können) die GPU(s) 2108 mindestens acht Streaming-Mikroprozessoren aufweisen. Bei mindestens einer Ausführungsform kann (können) die GPU(s) 2108 eine oder mehrere Programmierschnittstellen (API(s)) für Berechnungen verwenden. Bei mindestens einer Ausführungsform kann (können) die GPU(s) 2108 eine oder mehrere parallele Rechenplattformen und/oder Programmiermodelle (z. B. CUDA von NVIDIA) verwenden.

[0197] Bei mindestens einer Ausführungsform kann eine oder können mehrere der GPU(s) 2108 für die beste Leistung in automobilen und eingebetteten Anwendungsfällen energieoptimiert sein. In einer Ausführungsform kann (können) die GPU(s) 2108 beispielsweise mit Fin-Feldeffekttransistoren („FinFETs“) hergestellt sein. Bei mindestens einer Ausführungsform kann jeder Streaming-Mikroprozessor eine Anzahl von in mehrere Blöcke unterteilten Rechenkernen mit gemischter Präzision enthalten. Beispielsweise können 64 PF32-Kerne und 32 PF64-Kerne in vier Verarbeitungsblöcke unterteilt sein. Bei mindestens einer Ausführungsform könnten jedem Verarbeitungsblock 16 FP32-Kerne, 8 FP64-Kerne, 16 INT32-Kerne, zwei NVIDIA-TENSOR-KERNE mit unterschiedlicher Genauigkeit zur Deep-Learning-Matrix-Arithmetik, ein Anweisungs-Cache der Stufe null („LO“), ein Warp-Planer, eine Versandeinheit und/oder eine Registerdatei mit 64 KB zugeordnet sein. Bei mindestens einer Ausführungsform können Streaming-Mikroprozessoren unabhängige parallele Ganzzahl- und Gleitkomma-Datenpfade aufweisen, um eine effiziente Ausführung von Arbeitslasten mit einer Mischung aus Berechnungen und Adressierungsberechnungen zu ermöglichen. Bei mindestens einer Ausführungsform können Streaming-Mikroprozessoren eine unabhängige Thread-Planungsfunktion aufweisen, um eine feinkörnigere Synchronisierung und Zusammenarbeit zwischen parallelen Threads zu ermöglichen. Bei mindestens einer Ausführungsform können Streaming-Mikroprozessoren einen kombinierten LI-Datencache und eine gemeinsame Speichereinheit aufweisen, um die Leistung zu verbessern und gleichzeitig die Programmierung zu vereinfachen.

[0198] Bei mindestens einer Ausführungsform können eine oder mehrere der einen oder der mehreren GPUs 2108 einen Speicher mit hoher Bandbreite (High Bandwidth Memory-„HBM“) und/oder ein HBM2-Speicherteilsystem mit 16 GB beinhalten, um bei einigen Beispielen ungefähr 900 GB/Sekunde an Spitzenspeicherbandbreite bereitzustellen. Bei mindestens einer Ausführungsform kann zusätzlich oder alternativ zum HBM-Speicher ein synchroner Grafik-Direktzugriffsspeicher („SGRAM“) verwendet werden, wie z. B. ein synchroner Grafik-Doppeldatenraten-Direktzugriffsspeicher vom Typ 5 („GDDRS“).

[0199] Bei mindestens einer Ausführungsform kann (können) die GPU(s) 2108 eine Unified-Memory-Technologie aufweisen. Bei mindestens einer Ausführungsform kann die Unterstützung von Adressübersetzungsdiensten („ATS“) verwendet werden, damit die GPU(s) 2108 direkt auf Seitentabellen der CPU(s) 2106

zugreifen können. Bei mindestens einer Ausführungsform kann eine Adressübersetzungsanforderung an die CPU(s) 2106 übermittelt werden, wenn die Speicherverwaltungseinheit („MMU“) der GPU(s) 2108 einen Fehler feststellt. Als Antwort darauf kann (können) die CPU(s) 2106 in ihren Seitentabellen nach einer virtuell-physikalischen Zuordnung der Adresse suchen und bei mindestens einer Ausführungsform die Übersetzung zurück an die GPU(s) 2108 übertragen. Bei mindestens einer Ausführungsform kann die Unified-Memory-Technologie einen einzigen, einheitlichen virtuellen Adressraum für den Speicher sowohl der CPU(s) 2106 als auch der GPU(s) 2108 ermöglichen, wodurch die Programmierung der GPU(s) 2108 und der Anschluss von Anwendungen an die GPU(s) 2108 vereinfacht wird.

[0200] Bei mindestens einer Ausführungsform kann (können) die GPU(s) 2108 eine beliebige Anzahl von Zugriffszählern aufweisen, die die Häufigkeit des Zugriffs der GPU(s) 2108 auf den Speicher anderer Prozessoren verfolgen können. Bei mindestens einer Ausführungsform können Zugriffszähler dazu beitragen, dass Speicherseiten in den physischen Speicher desjenigen Prozessors verschoben werden, der am häufigsten auf Seiten zugreift, wodurch die Effizienz von Speicherbereichen verbessert wird, die von Prozessoren gemeinsam genutzt werden.

[0201] Bei mindestens einer Ausführungsform kann einer oder können mehrere der SoC(s) 2104 eine beliebige Anzahl von Cache(s) 2112 aufweisen, einschließlich der hier beschriebenen. Bei mindestens einer Ausführungsform kann (können) der/die Cache(s) 2112 beispielsweise einen Level-3-Cache („L3“) aufweisen, der sowohl der/den CPU(s) 2106 als auch der/den GPU(s) 2108 zur Verfügung steht (z. B. der sowohl mit der/den CPU(s) 2106 als auch der/den GPU(s) 2108 verbunden ist). Bei mindestens einer Ausführungsform kann (können) der (die) Cache(s) 2112 einen Write-Back-Cache aufweisen, der die Zustände der Zeilen verfolgen kann, z. B. durch Verwendung eines Cache-Kohärenzprotokolls (z. B. MEI, MESI, MSI usw.). Bei mindestens einer Ausführungsform kann ein L3-Cache 4 MB von Speicher oder mehr, abhängig von der Ausführungsform, beinhalten, obwohl kleinere Cache-Speichergrößen verwendet werden können.

[0202] Bei mindestens einer Ausführungsform kann einer oder können mehrere der SoC(s) 2104 einen oder mehrere Beschleuniger 2114 aufweisen (z. B. Hardware-Beschleuniger, Software-Beschleuniger oder eine Kombination davon). Bei mindestens einer Ausführungsform kann (können) der (die) SoC(s) 2104 einen Hardwarebeschleunigungscluster aufweisen, der optimierte Hardwarebeschleuniger und/oder einen großen On-Chip-Speicher aufweisen kann. Bei mindestens einer Ausführungsform kann ein großer Speicher auf dem Chip (z. B. 4 MB SRAM) es einem Hardwarebeschleunigungs-Cluster ermöglichen, neuronale Netzwerke und andere Berechnungen zu beschleunigen. Bei mindestens einer Ausführungsform kann der Hardware-Beschleunigungscluster zur Ergänzung der GPU(s) 2108 und zur Entlastung einiger Tasks der GPU(s) 2108 verwendet werden (z. B. um mehr Zyklen der GPU(s) 2108 für die Durchführung anderer Tasks freizugeben). Bei mindestens einer Ausführungsform kann (können) der/die Beschleuniger 2114 für gezielte Arbeitslasten verwendet werden (z. B. Wahrnehmung, faltende neuronale Netze („CNNs“), rückgekoppelte neuronale Netze („RNNs“) usw.), die stabil genug sind, um für eine Beschleunigung geeignet zu sein. Bei mindestens einer Ausführungsform kann ein CNN ein regionenbasiertes oder regionales faltendes neuronales Netz („RCNNs“) und ein schnelles RCNN (z. B. wie es für die Objekterkennung verwendet wird) oder eine andere Art von CNN aufweisen.

[0203] Bei mindestens einer Ausführungsform kann (können) der (die) Beschleuniger 2114 (z. B. Hardware-Beschleunigungscluster) einen Deep-Learning-Beschleuniger („DLA“) aufweisen. (Ein) DLA(s) kann (können) ohne Einschränkung eine oder mehrere Tensor Processing Units („TPUs“) aufweisen, die so ausgestaltet sein können, dass sie zusätzliche zehn Billionen Operationen pro Sekunde für Deep-Learning-Anwendungen und Inferencing bereitstellen. Bei mindestens einer Ausführungsform kann es sich bei den TPUs um Beschleuniger handeln, die für die Durchführung von Bildverarbeitungsfunktionen ausgestaltet und optimiert sind (z. B. für CNNs, RCNNs usw.). Der (die) DLA(s) kann (können) darüber hinaus für einen bestimmten Satz neuronaler Netzwerktypen und Gleitkommaoperationen sowie für Inferencing optimiert sein. Bei mindestens einer Ausführungsform kann das Design von DLA(s) mehr Leistung pro Millimeter bieten als eine typische Allzweck-GPU und übertrifft in der Regel die Leistung einer CPU bei weitem. Bei mindestens einer Ausführungsform kann (können) die TPU(s) mehrere Funktionen ausführen, einschließlich einer Einzelinstanz-Faltungsfunktion, die z. B. INT8-, INT16- und FP16-Datentypen sowohl für Merkmale als auch für Gewichte sowie Postprozessorfunktionen unterstützt. Bei mindestens einer Ausführungsform können DLA(s) schnell und effizient neuronale Netze, insbesondere CNNs, auf verarbeiteten oder unverarbeiteten Daten für eine Vielzahl von Funktionen ausführen, einschließlich, zum Beispiel und ohne Einschränkung: ein CNN für die Objektidentifizierung und -erkennung unter Verwendung von Daten von Kamerasensoren; ein CNN für die Abstandsschätzung unter Verwendung von Daten von Kamerasensoren; ein CNN für die Erkennung und Identifizierung von Einsatzfahrzeugen und die Erkennung unter Verwendung von Daten von Mikrofonen

2196; ein CNN für die Gesichtserkennung und die Identifizierung von Fahrzeugeigentümern unter Verwendung von Daten von Kamerasensoren; und/oder ein CNN für sicherheitsrelevante und/oder sicherheitsbezogene Ereignisse.

[0204] Bei mindestens einer Ausführungsform kann (können) DLA(s) jede Funktion der GPU(s) 2108 ausführen, und durch die Verwendung eines Inferenzbeschleunigers kann ein Entwickler beispielsweise entweder DLA(s) oder GPU(s) 2108 für eine beliebige Funktion vorsehen. Bei mindestens einer Ausführungsform kann der Entwickler beispielsweise die Verarbeitung von CNNs und Gleitkommaoperationen auf DLA(s) konzentrieren und andere Funktionen der GPU(s) 2108 und/oder einem oder mehreren anderen Beschleunigern 2114 überlassen.

[0205] Bei mindestens einer Ausführungsform kann (können) der (die) Beschleuniger 2114 (z. B. Hardware-Beschleunigungscluster) einen programmierbaren Bildverarbeitungsbeschleuniger („PVA“) aufweisen, der hier alternativ auch als Computer-Vision-Beschleuniger bezeichnet werden kann. Bei mindestens einer Ausführungsform kann (können) der (die) PVA(s) so gestaltet und ausgestaltet sein, dass er (sie) Computer-Vision-Algorithmen für fortschrittliche Fahrerassistenzsysteme („ADAS“) 2138, autonomes Fahren, Augmented-Reality-Anwendungen („AR“) und/oder Virtual-Reality-Anwendungen („VR“) beschleunigt. PVA(s) können ein Gleichgewicht zwischen Leistung und Flexibilität bieten. Bei mindestens einer Ausführungsform kann jeder PVA beispielsweise und ohne Einschränkung eine beliebige Anzahl von Rechenkernen mit reduziertem Befehlssatz („RISC“), direkten Speicherzugriff („DMA“) und/oder eine beliebige Anzahl von Vektorprozessoren aufweisen.

[0206] Bei mindestens einer Ausführungsform können die RISC-Kerne mit Bildsensoren (z. B Bildsensoren einer der hier beschriebenen Kameras), Bildsignalprozessoren und/oder ähnlichem interagieren. Bei mindestens einer Ausführungsform kann jeder der RISC-Kerne eine beliebige Menge an Speicher aufweisen. Bei mindestens einer Ausführungsform können die RISC-Kerne je nach Ausführungsform eines von mehreren Protokollen verwenden. Bei mindestens einer Ausführungsform können RISC-Kerne ein Echtzeitbetriebssystem („RTOS“) ausführen. Bei mindestens einer Ausführungsform können RISC-Kerne mit einer oder mehreren integrierten Schaltungseinrichtungen, anwendungsspezifischen integrierten Schaltungen („ASICs“) und/oder Speichereinrichtungen implementiert sein. Bei mindestens einer Ausführungsform können RISC-Kerne beispielsweise einen Befehls-Cache und/oder einen eng gekoppelten RAM aufweisen.

[0207] Bei mindestens einer Ausführungsform kann ein DMA es Komponenten der PVA(s) ermöglichen, unabhängig von der/den CPU(s) 2106 auf den Systemspeicher zuzugreifen. Bei mindestens einer Ausführungsform kann ein DMA eine beliebige Anzahl von Merkmalen unterstützen, die zur Optimierung des PVAs verwendet werden, einschließlich, aber nicht beschränkt auf die Unterstützung von mehrdimensionaler Adressierung und/oder zirkulärer Adressierung. Bei mindestens einer Ausführungsform kann ein DMAbis zu sechs oder mehr Dimensionen der Adressierung unterstützen, die ohne Einschränkung Blockbreite, Blockhöhe, Blocktiefe, horizontales Block-Stepping, vertikales Block-Stepping und/oder Tiefen-Stepping aufweisen können.

[0208] Bei mindestens einer Ausführungsform können Vektorprozessoren programmierbare Prozessoren sein, die für eine effiziente und flexible Ausführung der Programmierung für Computer-Vision-Algorithmen ausgelegt sein können und Signalverarbeitungsfunktionen bieten. Bei mindestens einer Ausführungsform kann der PVA einen PVA-Kern und zwei Vektorverarbeitungs-Subsystem-Partitionen aufweisen. Bei mindestens einer Ausführungsform kann der PVA-Kern ein Prozessor-Subsystem, DMA-Engine(s) (z. B. zwei DMA-Engines) und/oder andere Peripheriegeräte aufweisen. Bei mindestens einer Ausführungsform kann das Vektorverarbeitungs-Subsystem als primäre Verarbeitungseinheit des PVAs fungieren und eine Vektorverarbeitungseinheit („VPU“), einen Befehlscache und/oder einen Vektorspeicher (z. B. „VMEM“) aufweisen. Bei mindestens einer Ausführungsform kann der VPU-Kern einen digitalen Signalprozessor aufweisen, wie z. B. einen digitalen Signalprozessor mit mehreren Daten für eine Anweisung („SIMD“) und sehr langen Anweisungsworten („VLIW“). Bei mindestens einer Ausführungsform kann eine Kombination aus SIMD und VLIW den Durchsatz und die Geschwindigkeit erhöhen.

[0209] Bei mindestens einer Ausführungsform kann jeder der Vektorprozessoren einen Befehls-Cache aufweisen und mit einem dedizierten Speicher verbunden sein. Infolgedessen kann bei mindestens einer Ausführungsform jeder der Vektorprozessoren so konfiguriert sein, dass er unabhängig von anderen Vektorprozessoren arbeitet. Bei mindestens einer Ausführungsform können Vektorprozessoren, die in einem bestimmten PVA enthalten sind, so konfiguriert sein, dass sie Datenparallelität verwenden. Beispielsweise können bei mindestens einer Ausführungsform mehrere Vektorprozessoren, die in einem einzigen PVA ent-

halten sind, denselben Computer-Vision-Algorithmus ausführen, jedoch für unterschiedliche Bildbereiche. Bei mindestens einer Ausführungsform können Vektorprozessoren, die in einem bestimmten PVA enthalten sind, gleichzeitig verschiedene Bildverarbeitungsalgorithmen für dasselbe Bild oder sogar verschiedene Algorithmen für aufeinander folgende Bilder oder Abschnitte eines Bildes ausführen. Bei mindestens einer Ausführungsform kann unter anderem eine beliebige Anzahl von PVAs in einem Hardware-Beschleunigungscluster und eine beliebige Anzahl von Vektorprozessoren in jedem PVA vorhanden sein. Bei mindestens einer Ausführungsform kann/können (der) PVA(s) einen zusätzlichen Fehlerkorrekturcode-Speicher („ECC“) aufweisen, um die Gesamtsystemsicherheit zu erhöhen.

[0210] Bei mindestens einer Ausführungsform kann (können) der (die) Beschleuniger 2114 (z. B. ein Hardware-Beschleunigungscluster) ein Computer-Vision-Netzwerk auf dem Chip und einen statischen Direktzugriffsspeicher („SRAM“) aufweisen, um einen SRAM mit hoher Bandbreite und geringer Latenz für den (die) Beschleuniger 2114 bereitzustellen. Bei mindestens einer Ausführungsform kann der Speicher auf dem Chip mindestens SRAM mit 4 MB beinhalten, der zum Beispiel und ohne Einschränkung acht feldkonfigurierbare Speicherblöcke umfasst, auf die sowohl durch einen PVA als auch einen DLA zugegriffen werden kann. Bei mindestens einer Ausführungsform kann jedes Paar von Speicherblöcken eine erweiterte Peripheriebusschnittstelle („APB“), Konfigurationsschaltungen, eine Steuerung und einen Multiplexer aufweisen. Bei mindestens einer Ausführungsform kann jeder beliebige Speichertyp verwendet werden. Bei mindestens einer Ausführungsform können PVA und DLA über einen Backbone auf den Speicher zugreifen, der PVA und DLA einen Hochgeschwindigkeitszugriff auf den Speicher ermöglicht. Bei mindestens einer Ausführungsform kann der Backbone ein Computer-Vision-Netzwerk auf dem Chip aufweisen, das PVA und DLA mit dem Speicher verbindet (z. B unter Verwendung einer APB).

[0211] Bei mindestens einer Ausführungsform kann das Computer-Vision-Netz auf dem Chip eine Schnittstelle aufweisen, die vor der Übertragung von Steuersignalen/Adressen/Daten feststellt, dass sowohl der PVA als auch der DLA bereitstehende und gültige Signale liefern. Bei mindestens einer Ausführungsform kann eine Schnittstelle getrennte Phasen und getrennte Kanäle für die Übertragung von Steuersignalen/Adressen/Daten sowie eine Burst-Kommunikation für die kontinuierliche Datenübertragung vorsehen. Bei mindestens einer Ausführungsform kann eine Schnittstelle den Normen der Internationalen Organisation für Normung („ISO“) 26262 oder der Internationalen Elektrotechnischen Kommission („IEC“) 61508 entsprechen, obwohl auch andere Normen und Protokolle verwendet werden können.

[0212] Bei mindestens einer Ausführungsform können einer oder mehrere des einen oder der mehreren SOCs 2104 einen Echtzeit-Raytracing-Hardwarebeschleuniger beinhalten. Bei mindestens einer Ausführungsform kann der Echtzeit-Raytracing-Hardwarebeschleuniger verwendet werden, um schnell und effizient Positionen und Ausmaße von Objekten (z. B innerhalb eines Weltmodells) zu bestimmen, um Echtzeit-Visualisierungssimulationen zu erzeugen, für RADAR-Signalinterpretation, für Schallausbreitungssynthese und/oder -analyse, für die Simulation von SONAR-Systemen, für eine allgemeine Wellenausbreitungssimulation, für den Vergleich mit LIDAR-Daten zum Zwecke der Lokalisierung und/oder für andere Funktionen und/oder für andere Zwecke.

[0213] Bei mindestens einer Ausführungsform hat/haben der/die Beschleuniger 2114 (z. B. Hardware-Beschleuniger-Cluster) eine breite Palette von Anwendungen für das autonome Fahren. Bei mindestens einer Ausführungsform kann ein PVA ein programmierbarer Bildverarbeitungsbeschleuniger sein, der für wichtige Verarbeitungsschritte in ADAS und autonomen Fahrzeugen verwendet werden kann. Bei mindestens einer Ausführungsform eignen sich die Fähigkeiten eines PVAs gut für algorithmische Bereiche, die eine vorhersehbare Verarbeitung bei geringer Leistung und geringer Latenz benötigen. Mit anderen Worten: ein PVA eignet sich gut für halbdichte oder dichte reguläre Berechnungen, selbst bei kleinen Datensätzen, die vorhersehbare Laufzeiten mit geringer Latenz und geringem Stromverbrauch erfordern. Bei mindestens einer Ausführungsform sind für autonome Fahrzeuge, wie z. B Fahrzeug 2100, PVAs entwickelt, um klassische Computer-Vision-Algorithmen auszuführen, da sie effizient bei der Objekterkennung sind und mit ganzzahligen mathematischen Verfahren arbeiten.

[0214] Zum Beispiel wird bei mindestens einer Ausführungsform einer Technologie ein PVA verwendet, um Computer-Stereo-Vision durchzuführen. Bei mindestens einer Ausführungsform kann bei einigen Beispielen ein auf semiglobalem Matching basierender Algorithmus verwendet werden, obwohl dies nicht als Einschränkung gedacht ist. Bei mindestens einer Ausführungsform werden bei Anwendungen für das autonome Fahren der Stufen 3-5 Bewegungsschätzungen/Stereoabgleich während der Fahrt verwendet (z. B. Struktur aus Bewegung, Fußgängererkennung, Fahrspurerkennung usw.). Bei mindestens einer Ausführungsform kann der PVA eine Computer-Stereosichtfunktion auf Eingaben von zwei monokularen Kameras ausführen.

[0215] Bei mindestens einer Ausführungsform kann ein PVA verwendet werden, um einen dichten optischen Fluss auszuführen. Zum Beispiel kann ein PVA bei mindestens einer Ausführungsform RADAR-Rohdaten verarbeiten (z. B unter Verwendung einer 4D-Fast-Fourier-Transformation), um verarbeitete RADAR-Daten zu liefern. Bei mindestens einer Ausführungsform wird ein PVA für die Flugzeittiefenverarbeitung verwendet, indem Flugzeit-Rohdaten verarbeitet werden, um z. B verarbeitete Flugzeitdaten bereitzustellen.

[0216] Bei mindestens einer Ausführungsform kann ein DLA verwendet werden, um jede Art von Netzwerk zu betreiben, um die Steuerung und die Fahrsicherheit zu verbessern, einschließlich beispielsweise und ohne Einschränkung eines neuronalen Netzes, das für jede Objekterkennung ein Maß für das Vertrauen ausgibt. Bei mindestens einer Ausführungsform kann das Vertrauen als Wahrscheinlichkeit dargestellt oder interpretiert werden, oder als relative „Gewichtung“ jeder Erkennung im Vergleich zu anderen Erkennungen. Bei mindestens einer Ausführungsform ermöglicht es die Konfidenz dem System, darüber hinaus Entscheidungen darüber zu treffen, welche Erkennungen als echte positive Erkennungen und welche als falsch positive Erkennungen betrachtet werden sollen. Bei mindestens einer Ausführungsform kann zum Beispiel ein System einen Schwellenwert für Konfidenz festlegen und lediglich Erfassungen berücksichtigen, die den Schwellenwert als echte positive Erfassungen übersteigen. In einer Ausführungsform, in der ein automatisches Notbremssystem („AEB“) verwendet wird, würden falsch positive Erkennungen dazu führen, dass das Fahrzeug automatisch eine Notbremsung durchführt, was natürlich unerwünscht ist. Bei mindestens einer Ausführungsform können sehr sichere Erkennungen als Auslöser für ein AEB angesehen werden. Bei mindestens einer Ausführungsform kann ein DLA ein neuronales Netz zur Regression des Vertrauenswertes einsetzen. Bei mindestens einer Ausführungsform kann das neuronale Netzwerk als seine Eingabe mindestens eine Teilmenge von Parametern verwenden, wie die Abmessungen des Begrenzungsrahmens, die (z. B von einem anderen Teilsystem) erhaltene Schätzung der Grundfläche, die Ausgabe des/der IMU-Sensors/en 2166, die mit der Ausrichtung des Fahrzeugs 2100, Entfernung, die unter anderem Schätzungen der 3D-Position des Objekts, die von dem neuronalen Netzwerk und/oder anderen Sensoren (z. B LIDAR-Sensor(en) 2164 oder RADAR-Sensor(en) 2160) erhalten werden, korreliert.

[0217] Bei mindestens einer Ausführungsform kann ein oder können mehrere SoC(s) 2104 einen oder mehrere Datenspeicher 2116 (z. B einen Speicher) aufweisen. Bei mindestens einer Ausführungsform kann (können) der (die) Datenspeicher 2116 ein On-Chip-Speicher des (der) SoC(s) 2104 sein, der (die) neuronale Netze speichern kann (können), die auf GPU(s) 2108 und/oder einem DLA ausgeführt werden sollen. Bei mindestens einer Ausführungsform kann die Kapazität des/der Datenspeicher(s) 2116 groß genug sein, um mehrere Instanzen von neuronalen Netzen aus Gründen der Redundanz und Sicherheit zu speichern. Bei mindestens einer Ausführungsform kann/können der/die Datenspeicher 2112 L2 oder L3 Cache(s) umfassen.

[0218] Bei mindestens einer Ausführungsform kann einer oder können mehrere der SoC(s) 2104 eine beliebige Anzahl von Prozessoren 2110 (z. B eingebettete Prozessoren) aufweisen. Der Prozessor 2110 kann einen Hochfahr- und Leistungsverwaltungsprozessor beinhalten, der ein dedizierter Prozessor und ein Teilsystem sein kann, um die Hochfahrleistungs- und Verwaltungsfunktionen und die dazugehörige sicherheitsbezogene Durchsetzung zu handhaben. Bei mindestens einer Ausführungsform kann der Boot- und Energieverwaltungsprozessor ein Teil der Bootsequenz des/der SoC(s) 2104 sein und Laufzeit-Energieverwaltungsdienste bereitstellen. Bei mindestens einer Ausführungsform kann der Prozessor für die Boot-Energieversorgung und -Verwaltung Takt- und Spannungsprogrammierung, Unterstützung bei Systemübergängen mit niedrigem Energiebedarf, Verwaltung von SoC(s) 2104-Temperaturen und Temperatursensoren und/oder Verwaltung von SoC(s) 2104-Energieversorgungszuständen bereitstellen. Bei mindestens einer Ausführungsform kann jeder Temperatursensor als Ringoszillator implementiert sein, dessen Ausgangsfrequenz proportional zur Temperatur ist, und (ein) SoC(s) 2104 kann/können Ringoszillatoren verwenden, um Temperaturen von CPU(s) 2106, GPU(s) 2108 und/oder Beschleuniger(n) 2114 zu erfassen. Bei mindestens einer Ausführungsform kann der Boot- und Energieverwaltungsprozessor, wenn festgestellt wird, dass die Temperaturen einen Schwellenwert überschreiten, in eine Temperaturfehlerroutine eintreten und die SoC(s) 2104 in einen Zustand mit geringerer Leistung versetzen und/oder das Fahrzeug 2100 in einen Chauffeur-zu-sicherem-Halt-Modus versetzen (z. B. das Fahrzeug 2100 zu einem sicheren Halt bringen).

[0219] Bei mindestens einer Ausführungsform kann/können der/die Prozessor(en) 2110 darüber hinaus einen Satz eingebetteter Prozessoren aufweisen, die als Audioverarbeitungsmaschine dienen können. Bei mindestens einer Ausführungsform kann eine Audioverarbeitungs-Engine ein Audioteilsystem sein, das volle Hardwareunterstützung für Mehrkanalaudio über mehrere Schnittstellen und einen breiten und flexiblen Bereich von Audio-E/A-Schnittstellen ermöglicht. Bei mindestens einer Ausführungsform handelt es sich bei der Audioverarbeitungsmaschine um einen dedizierten Prozessorkern mit einem digitalen Signalprozessor mit dediziertem RAM.

[0220] Bei mindestens einer Ausführungsform kann/können der/die Prozessor(en) 2110 darüber hinaus eine „always on“-Prozessor-Maschine aufweisen, die die notwendigen Hardware-Funktionen zur Unterstützung von Sensor-Management mit geringem Stromverbrauch und Aufwach-Anwendungsfälle bereitstellen kann. Bei mindestens einer Ausführungsform kann die „always on“-Prozessor-Maschine ohne Einschränkung einen Prozessorkern, ein eng gekoppeltes RAM, unterstützende Peripheriegeräte (z. B Timer und Interrupt-Controller), verschiedene E/A-Controller-Peripheriegeräte und Routing-Logik aufweisen.

[0221] Bei mindestens einer Ausführungsform kann/können der/die Prozessor(en) 2110 darüber hinaus eine Sicherheits-Cluster-Maschine aufweisen, die ohne Einschränkung ein dediziertes Prozessor-Subsystem zur Handhabung des Sicherheitsmanagements für Automobilanwendungen aufweist. Bei mindestens einer Ausführungsform kann die Sicherheits-Cluster-Maschine ohne Einschränkung zwei oder mehr Prozessorkerne, ein eng gekoppeltes RAM, unterstützende Peripheriegeräte (z. B. Zeitgeber, eine Interrupt-Steuerung usw.) und/oder eine Routing-Logik aufweisen. In einem Sicherheitsmodus können bei mindestens einer Ausführungsform zwei oder mehr Kerne in einem Lockstep-Modus arbeiten und als ein einziger Kern mit einer Vergleichslogik funktionieren, um etwaige Unterschiede zwischen ihren Operationen zu erkennen. Bei mindestens einer Ausführungsform kann (können) der (die) Prozessor(en) 2110 darüber hinaus eine Echtzeit-Kamera-Maschine aufweisen, die ohne Einschränkung ein dediziertes Prozessor-Subsystem zur Handhabung des Echtzeit-Kameramanagements aufweisen kann. Bei mindestens einer Ausführungsform kann (können) der (die) Prozessor(en) 2110 darüber hinaus einen Signalprozessor mit hohem Dynamikbereich aufweisen, der ohne Einschränkung einen Bildsignalprozessor aufweisen kann, der eine Hardware-Maschine ist, die Teil der Kameraverarbeitungspipeline ist.

[0222] Bei mindestens einer Ausführungsform kann (können) der (die) Prozessor(en) 2110 einen Videobildkompositor aufweisen, der ein Verarbeitungsblock sein kann (z. B. auf einem Mikroprozessor implementiert), der Videonachverarbeitungsfunktionen implementiert, die von einer Videowiedergabeanwendung benötigt werden, um das endgültige Bild für das Spieler-Fenster zu erzeugen. Bei mindestens einer Ausführungsform kann der Videobildkompositor eine Linsenverzerrungskorrektur an der/den Weitwinkelkamera(s) 2170, der/den Surround-Kamera(s) 2174 und/oder an dem/den Sensor(en) der Überwachungskamera(s) in der Kabine vornehmen. Bei mindestens einer Ausführungsform wird/werden der/die Sensor(en) der Überwachungskamera(s) in der Kabine vorzugsweise von einem neuronalen Netz überwacht, das auf einer anderen Instanz des SoC 2104 läuft und so ausgestaltet ist, dass es Ereignisse in der Kabine erkennt und entsprechend reagiert. Bei mindestens einer Ausführungsform kann ein System im Fahrzeuginneren ohne Einschränkung Lippenlesen durchführen, um den Mobilfunkdienst zu aktivieren und einen Anruf zu tätigen, E-Mails zu diktieren, das Fahrtziel zu ändern, das Infotainmentsystem und die Einstellungen des Fahrzeugs zu aktivieren oder zu ändern oder sprachgesteuertes Surfen im Internet zu ermöglichen. Bei mindestens einer Ausführungsform stehen dem Fahrer bestimmte Funktionen zur Verfügung, wenn das Fahrzeug in einem autonomen Modus betrieben wird, und sind ansonsten deaktiviert.

[0223] Bei mindestens einer Ausführungsform kann der Videobildkompositor eine verbesserte zeitliche Rauschunterdrückung sowohl für eine räumliche als auch für eine zeitliche Rauschunterdrückung aufweisen. Zum Beispiel bei mindestens einer Ausführungsform, wenn Bewegung in einem Video auftritt, gewichtet die Rauschunterdrückung die räumliche Information angemessen und verringert das Gewicht der Information, die von benachbarten Bildern geliefert wird. Bei mindestens einer Ausführungsform, bei der ein Bild oder ein Abschnitt eines Bildes keine Bewegung aufweist, kann die vom Videobildkompositor durchgeführte zeitliche Rauschreduzierung Informationen aus dem vorherigen Bild verwenden, um das Rauschen im aktuellen Bild zu reduzieren.

[0224] Bei mindestens einer Ausführungsform kann der Videobildkompositor auch so ausgestaltet sein, dass er eine Stereorektifizierung an eingegebenen Stereolinsenrahmen durchführt. Bei mindestens einer Ausführungsform kann der Videobildkompositor darüber hinaus für die Gestaltung der Benutzeroberfläche verwendet werden, wenn der Desktop des Betriebssystems in Gebrauch ist und die GPU(s) 2108 nicht zum kontinuierlichen Rendern neuer Oberflächen benötigt werden. Bei mindestens einer Ausführungsform, wenn die GPU(s) 2108 eingeschaltet sind und aktiv 3D-Rendering durchführen, kann der Videobildkompositor verwendet werden, um die GPU(s) 2108 zu entlasten, um die Leistung und Reaktionsfähigkeit zu verbessern.

[0225] Bei mindestens einer Ausführungsform kann einer oder können mehrere der SoC(s) 2104 darüber hinaus eine serielle MIPI-Kameraschnittstelle zum Empfang von Video und Eingaben von Kameras, eine Hochgeschwindigkeitsschnittstelle und/oder einen Videoeingabeblock aufweisen, der für Kamera- und verwandte Pixeleingabefunktionen verwendet werden kann. Bei mindestens einer Ausführungsform kann einer oder können mehrere der SoC(s) 2104 darüber hinaus eine oder mehrere Eingabe-/Ausgabe-Steuerungen

aufweisen, die durch Software gesteuert werden können und für den Empfang von E/A-Signalen verwendet werden können, die keiner bestimmten Rolle zugeordnet sind.

[0226] Bei mindestens einer Ausführungsform können ein oder mehrere SoC(s) 2104 darüber hinaus eine breite Palette von Peripherieschnittstellen aufweisen, um die Kommunikation mit Peripheriegeräten, Audio-Encodern/Decodern („Codecs“), der Energieverwaltung und/oder anderen Einrichtungen zu ermöglichen. SoC(s) 2104 kann (können) verwendet werden, um Daten von Kameras (z. B. verbunden über Gigabit Multimedia Serial Link und Ethernet), Sensoren (z. B. LIDAR-Sensor(en) 2164, RADAR-Sensor(en) 2160 usw., die über Ethernet verbunden sein können), Daten von Bus 2102 (z. B. Geschwindigkeit des Fahrzeugs 2100, Lenkradposition usw.), Daten von GNSS-Sensor(en) 2158 (z. B. verbunden über Ethernet oder CAN-Bus) usw. zu verarbeiten. Bei mindestens einer Ausführungsform kann einer oder können mehrere der SoC(s) 2104 darüber hinaus dedizierte Hochleistungs-Massenspeichersteuerungen aufweisen, die ihre eigenen DMA-Maschinen aufweisen können und die verwendet werden können, um die CPU(s) 2106 von Routine-Datenverwaltungsaufgaben zu entlasten.

[0227] Bei mindestens einer Ausführungsform kann (können) der (die) SoC(s) 2104 eine End-to-End-Plattform mit einer flexiblen Architektur sein, die die Automatisierungsebenen 3 bis 5 umfasst und dadurch eine umfassende funktionale Sicherheitsarchitektur bereitstellt, die Computer-Vision- und ADAS-Techniken für Diversität und Redundanz nutzt und eine Plattform für einen flexiblen, zuverlässigen Fahrsoftware-Stack zusammen mit Deep-Learning-Werkzeugen bereitstellt. Bei mindestens einer Ausführungsform können die SoC(s) 2104 schneller, zuverlässiger und sogar energie- und platzsparender sein als herkömmliche Systeme. Zum Beispiel können bei mindestens einer Ausführungsform der/die Beschleuniger 2114 in Kombination mit der/den CPU(s) 2106, der/den GPU(s) 2108 und dem/den Datenspeicher(n) 2116 eine schnelle, effiziente Plattform für autonome Fahrzeuge der Stufe 3-5 bilden.

[0228] Bei mindestens einer Ausführungsform können Computer-Vision-Algorithmen auf CPUs ausgeführt werden, die unter Verwendung einer Hochsprachen-Programmierung, wie z. B C, ausgestaltet sein können, um eine Vielzahl von Verarbeitungsalgorithmen für eine Vielzahl von visuellen Daten auszuführen. Bei mindestens einer Ausführungsform sind CPUs jedoch oft nicht in der Lage, die Leistungsanforderungen vieler Bildverarbeitungsanwendungen zu erfüllen, wie z. B. die Anforderungen an die Ausführungszeit und den Stromverbrauch. Bei mindestens einer Ausführungsform sind viele CPUs nicht in der Lage, komplexe Objekterkennungsalgorithmen in Echtzeit auszuführen, die in fahrzeuginternen ADAS-Anwendungen und in praktischen autonomen Fahrzeugen der Stufe 3-5 verwendet werden.

[0229] Ausführungsformen, wie sie hier beschrieben sind, ermöglichen die gleichzeitige und/oder sequentielle Ausführung mehrerer neuronaler Netze und die Kombination der Ergebnisse, um autonome Fahrfunktionen der Stufe 3-5 zu ermöglichen. Zum Beispiel kann bei mindestens einer Ausführungsform ein CNN, das auf einem DLA oder einer diskreten GPU (z. B GPU(s) 2120) ausgeführt wird, eine Text- und Worterkennung aufweisen, die es dem Supercomputer ermöglicht, Verkehrsschilder zu lesen und zu verstehen, einschließlich Schildern, für die das neuronale Netz nicht speziell trainiert wurde. Bei mindestens einer Ausführungsform kann ein DLA darüber hinaus ein neuronales Netz aufweisen, das in der Lage ist, Verkehrszeichen zu identifizieren, zu interpretieren und semantisch zu verstehen, und dieses semantische Verständnis an die auf einem CPU-Komplex laufenden Wegplanungsmodule weiterzugeben.

[0230] Bei mindestens einer Ausführungsform können mehrere neuronale Netze gleichzeitig ausgeführt werden, wie beim Fahren der Stufe 3, 4 oder 5. Zum Beispiel kann bei mindestens einer Ausführungsform ein Warnschild mit der Aufschrift „Vorsicht: Blinkende Lichter deuten auf Vereisung hin“ zusammen mit einem elektrischen Licht unabhängig oder gemeinsam von mehreren neuronalen Netzen interpretiert werden. Bei mindestens einer Ausführungsform kann das Schild selbst von einem ersten eingesetzten neuronalen Netz (z. B. einem trainierten neuronalen Netz) als Verkehrsschild identifiziert werden, und der Text „Blinkende Lichter deuten auf Glatteis hin“ kann von einem zweiten eingesetzten neuronalen Netz interpretiert werden, das die (vorzugsweise auf einem CPU-Komplex ausgeführte) Wegplanungssoftware des Fahrzeugs darüber informiert, dass, wenn blinkende Lichter erkannt werden, Glatteis vorliegt. Bei mindestens einer Ausführungsform kann das Blinklicht durch den Betrieb eines dritten neuronalen Netzes über mehrere Bilder identifiziert werden, das die Wegplanungssoftware des Fahrzeugs über das Vorhandensein (oder Fehlen) von Blinklichtern informiert. Bei mindestens einer Ausführungsform können alle drei neuronalen Netze gleichzeitig laufen, beispielsweise innerhalb eines DLAs und/oder auf GPU(s) 2108.

[0231] Bei mindestens einer Ausführungsform kann ein CNN zur Gesichtserkennung und zur Identifizierung des Fahrzeugbesitzers Daten von Kamerasensoren verwenden, um die Anwesenheit eines autorisierten Fah-

rers und/oder Besitzers des Fahrzeugs 2100 zu identifizieren. Bei mindestens einer Ausführungsform kann eine immer aktive Sensorverarbeitungs-Maschine verwendet werden, um das Fahrzeug zu entriegeln, wenn sich der Besitzer der Fahrertür nähert, und um die Lichter einzuschalten, und, im Sicherheitsmodus, um das Fahrzeug zu deaktivieren, wenn der Besitzer das Fahrzeug verlässt. Auf diese Weise sorgen die SoC(s) 2104 für Sicherheit gegen Diebstahl und/oder Carjacking.

[0232] Bei mindestens einer Ausführungsform kann ein CNN zur Erkennung und Identifizierung von Einsatzfahrzeugen Daten von Mikrofonen 2196 verwenden, um Sirenen von Einsatzfahrzeugen zu erkennen und zu identifizieren. Bei mindestens einer Ausführungsform verwenden die SoC(s) 2104 ein CNN zur Klassifizierung von Umwelt- und Stadtgeräuschen sowie zur Klassifizierung visueller Daten. Bei mindestens einer Ausführungsform wird ein CNN, das auf einem DLA läuft, darauf trainiert, die relative Annäherungsgeschwindigkeit von Einsatzfahrzeugen zu erkennen (z. B. unter Verwendung des Dopplereffekts). Bei mindestens einer Ausführungsform kann ein CNN auch so trainiert werden, dass es Einsatzfahrzeuge identifiziert, die spezifisch für das lokale Gebiet sind, in dem das Fahrzeug unterwegs ist, wie es von GNSS-Sensor(en) 2158 identifiziert wird. Bei mindestens einer Ausführungsform wird ein CNN bei einem Einsatz in Europa versuchen, europäische Sirenen zu erkennen, und bei einem Einsatz in den Vereinigten Staaten wird das CNN versuchen, nur nordamerikanische Sirenen zu identifizieren. Bei mindestens einer Ausführungsform kann, sobald ein Einsatzfahrzeug erkannt wird, ein Steuerprogramm verwendet werden, um eine Sicherheitsroutine für Einsatzfahrzeuge auszuführen, das Fahrzeug zu verlangsamen, an den Straßenrand zu fahren, das Fahrzeug zu parken und/oder das Fahrzeug im Leerlauf laufen zu lassen, mit Hilfe des/der Ultraschallsensors/en 2162, bis das/die Einsatzfahrzeug(e) vorbeifahren.

[0233] Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 eine oder mehrere CPU(s) 2118 (z. B diskrete CPU(s) oder dCPU(s)) aufweisen, die über eine Hochgeschwindigkeitsverbindung (z. B PCIe) mit dem/den SoC(s) 2104 verbunden sein können. Bei mindestens einer Ausführungsform kann (können) die CPU(s) 2118 beispielsweise einen X86-Prozessor aufweisen. (Eine) CPU(s) 2118 kann/können verwendet werden, um eine Vielzahl von Funktionen auszuführen, einschließlich der Schlichtung potenziell inkonsistenter Ergebnisse zwischen ADAS-Sensoren und SoC(s) 2104 und/oder der Überwachung des Status und des Zustands der Steuerung(en) 2136 und/oder eines Infotainment-Systems auf einem Chip („Infotainment-SoC“) 2130, zum Beispiel.

[0234] Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 GPU(s) 2120 (z. B diskrete GPU(s) oder dGPU(s)) aufweisen, die mit dem/den SoC(s) 2104 über eine Hochgeschwindigkeitsverbindung (z. B NVIDIAs NVLINK) gekoppelt sein können. Bei mindestens einer Ausführungsform kann/können GPU(s) 2120 zusätzliche künstliche Intelligenzfunktionalität bereitstellen, beispielsweise durch Ausführen redundanter und/oder unterschiedlicher neuronaler Netze, und kann/können verwendet werden, um neuronale Netze zu trainieren und/oder zu aktualisieren, was zumindest teilweise auf Eingaben (z. B Sensordaten) von Sensoren des Fahrzeugs 2100 basiert.

[0235] Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 darüber hinaus eine Netzwerkschnittstelle 2124 aufweisen, die ohne Einschränkung eine oder mehrere drahtlose Antennen 2126 aufweisen kann (z. B eine oder mehrere drahtlose Antennen 2126 für verschiedene Kommunikationsprotokolle, wie z. B eine Mobilfunkantenne, eine Bluetooth-Antenne, usw.). Bei mindestens einer Ausführungsform kann die Netzwerkschnittstelle 2124 verwendet werden, um eine drahtlose Verbindung über das Internet mit einer Cloud (z. B. mit einem oder mehreren Servern und/oder anderen Netzwerkeinrichtungen), mit anderen Fahrzeugen und/oder mit Recheneinrichtungen (z. B. Clienteinrichtungen von Fahrgästen) zu ermöglichen. Bei mindestens einer Ausführungsform kann zur Kommunikation mit anderen Fahrzeugen eine direkte Verbindung zwischen dem Fahrzeug 210 und einem anderen Fahrzeug und/oder eine indirekte Verbindung (z. B. über Netzwerke und das Internet) hergestellt werden. Bei mindestens einer Ausführungsform können direkte Verbindungen unter Verwendung einer Fahrzeug-zu-Fahrzeug-Kommunikationsverbindung bereitgestellt werden. Eine Fahrzeug-zu-Fahrzeug-Kommunikationsverbindung kann dem Fahrzeug 2100 Informationen über Fahrzeuge in der Nähe zu dem Fahrzeug 2100 bereitstellen (z. B. Fahrzeuge vor, seitlich von und/oder hinter dem Fahrzeug 2100). Bei mindestens einer Ausführungsform kann die vorgenannte Funktionalität Teil einer kooperativen adaptiven Geschwindigkeitsregelungsfunktion des Fahrzeugs 2100 sein.

[0236] Bei mindestens einer Ausführungsform kann die Netzwerkschnittstelle 2124 ein SoC aufweisen, das Modulations- und Demodulationsfunktionen bereitstellt und die Steuerung(en) 2136 in die Lage versetzt, über drahtlose Netzwerke zu kommunizieren. Bei mindestens einer Ausführungsform kann die Netzwerkschnittstelle 2124 ein Hochfrequenz-Frontend für die Aufwärtskonvertierung von einem Basisband auf eine Hochfrequenz und die Abwärtskonvertierung von einer Hochfrequenz auf ein Basisband aufweisen. Bei min-

destens einer Ausführungsform können die Frequenzumwandlungen auf jede technisch mögliche Weise durchgeführt werden. Beispielsweise können Frequenzumwandlungen durch bekannte Verfahren und/oder unter Verwendung von Superheterodyn-Verfahren durchgeführt werden. Bei mindestens einer Ausführungsform kann die Hochfrequenz-Front-End-Funktionalität durch einen separaten Chip bereitgestellt werden. Bei mindestens einer Ausführungsform können Netzwerkschnittstellen drahtlose Funktionen für das Kommunizieren über LTE, WCDMA, UMTS, GSM, CDMA2000, Bluetooth, Bluetooth LE, Wi-Fi, Z-Wave, ZigBee, LoRaWAN und/oder andere drahtlose Protokolle beinhalten.

[0237] Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 darüber hinaus einen oder mehrere Datenspeicher 2128 aufweisen, die ohne Einschränkung einen Off-Chip-Speicher (z. B Off-SoC(s) 2104) aufweisen können. Bei mindestens einer Ausführungsform kann (können) der (die) Datenspeicher 2128 ohne Einschränkung ein oder mehrere Speicherelemente aufweisen, darunter RAM, SRAM, dynamischer Direktzugriffsspeicher („DRAM“), Video-Direktzugriffsspeicher („VRAM“), Flash, Festplatten und/oder andere Komponenten und/oder Einrichtungen, die mindestens ein Bit an Daten speichern können.

[0238] Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 darüber hinaus GNSS-Sensor(en) 2158 (z. B GPS- und/oder unterstützte GPS-Sensoren) aufweisen, um bei der Kartierung, der Wahrnehmung, der Erstellung von Belegungsrastern und/oder der Pfadplanung zu helfen. Bei mindestens einer Ausführungsform kann eine beliebige Anzahl von GNSS-Sensor(en) 2158 verwendet werden, die beispielsweise und ohne Einschränkung ein GPS aufweisen, das einen USB-Anschluss mit einer Ethernetzu-Seriell-Brücke (z. B. RS-232) verwendet.

[0239] Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 darüber hinaus RADAR-Sensor(en) 2160 aufweisen. Der/die RADAR-Sensor(en) 2160 kann/können von einem Fahrzeug 2100 für die Fahrzeugerkennung über große Entfernungen verwendet werden, selbst bei Dunkelheit und/oder schlechten Wetterbedingungen. Bei mindestens einer Ausführungsform können die RADAR-Funktionssicherheitsstufen ASIL B sein. Der/die RADAR-Sensor(en) 2160 kann/können CAN und/oder den Bus 2102 (z. B. zur Übertragung der von dem/den RADAR-Sensor(en) 2160 erzeugten Daten) zur Steuerung und zum Zugriff auf Objektverfolgungsdaten verwenden, wobei bei einigen Beispielen der Zugriff auf Rohdaten über ein Ethernet erfolgt. Bei mindestens einer Ausführungsform kann eine breite Palette von RADAR-Sensortypen verwendet werden. Zum Beispiel und ohne Einschränkung können RADAR-Sensor(en) 2160 für die Verwendung von Front-, Heck- und Seiten-RADAR geeignet sein. Bei mindestens einer Ausführungsform handelt es sich bei einem oder mehreren der RADAR-Sensoren 2160 um Puls-Doppler-RADAR-Sensor(en).

[0240] Bei mindestens einer Ausführungsform kann (können) der (die) RADAR-Sensor(en) 2160 verschiedene Konfigurationen aufweisen, wie z. B. große Reichweite mit engem Sichtfeld, kurze Reichweite mit breitem Sichtfeld, seitliche Abdeckung mit kurzer Reichweite usw. Bei mindestens einer Ausführungsform kann das RADAR mit großer Reichweite für die adaptive Geschwindigkeitsregelung verwendet werden. Bei mindestens einer Ausführungsform können RADAR-Systeme mit großer Reichweite ein breites Sichtfeld bieten, was durch zwei oder mehr unabhängige Abtastungen, z. B. innerhalb eines Bereichs von 250 m, realisiert wird. Bei mindestens einer Ausführungsform kann/können der/die RADAR-Sensor(en) 2160 dabei helfen, zwischen stationären und sich bewegenden Objekten zu unterscheiden, und kann/können vom ADAS-System 2138 zur Notbremsunterstützung und zur Vorwärtskollisionswarnung verwendet werden. Der eine oder die mehreren Sensoren 2160, die in einem Langstrecken-RADAR-System beinhaltet sind, können ohne Einschränkung monostatisches multimodales RADAR mit mehreren (z. B. sechs oder mehr) feststehenden RADAR-Antennen und eine Hochgeschwindigkeits-CAN und FlexRay-Schnittstelle beinhalten. Bei mindestens einer Ausführungsform mit sechs Antennen können vier mittige Antennen ein konzentriertes Strahlmuster anlegen, das ausgestaltet ist, um die Umgebung des Fahrzeugs 2100 bei höheren Geschwindigkeiten mit minimaler Störung durch Verkehr in angrenzenden Spuren aufzuzeichnen. Bei mindestens einer Ausführungsform können die beiden anderen Antennen das Sichtfeld erweitern, so dass Fahrzeuge, die in die Fahrspur des Fahrzeugs 2100 einfahren oder diese verlassen, schnell erfasst werden können.

[0241] Bei mindestens einer Ausführungsform können RADAR-Systeme mittlerer Reichweite beispielsweise eine Reichweite von bis zu 160 m (vorn) oder 80 m (hinten) und ein Sichtfeld von bis zu 42 Grad (vorn) oder 150 Grad (hinten) aufweisen. Bei mindestens einer Ausführungsform können Kurzstrecken-RADAR-Systeme ohne Einschränkung eine beliebige Anzahl von RADAR-Sensoren 2160 aufweisen, die an beiden Enden des hinteren Stoßfängers installiert sein können. Bei mindestens einer Ausführungsform kann ein RADAR-Sensorsystem, wenn es an beiden Enden des hinteren Stoßfängers installiert ist, zwei Strahlen erzeugen, die den toten Winkel im hinteren Bereich und neben dem Fahrzeug ständig überwachen. Bei mindestens einer

Ausführungsform können RADAR-Systeme mit kurzer Reichweite in dem ADAS-System 2138 zur Erfassung des toten Winkels und/oder zur Unterstützung beim Spurwechsel verwendet werden.

[0242] Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 darüber hinaus Ultraschallsensor(en) 2162 aufweisen. Der eine oder die mehreren Ultraschallsensoren 2162, die an einer Vorderseite, einer Rückseite und/oder seitlichen Stellen des Fahrzeugs 2100 positioniert sein können, können für die Einparkunterstützung und/oder zum Anlegen und Aktualisieren eines Belegungsrasters verwendet werden. Bei mindestens einer Ausführungsform kann eine Vielzahl von Ultraschallsensoren 2162 verwendet werden, und unterschiedliche Ultraschallsensoren 2162 können für unterschiedliche Erfassungsbereiche (z. B. 2,5 m, 4 m) verwendet werden. Bei mindestens einer Ausführungsform kann/können der/die Ultraschallsensor(en) 2162 bei funktionalen Sicherheitsstufen von ASIL B arbeiten.

[0243] Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 LIDAR-Sensor(en) 2164 aufweisen. Der/die LIDAR-Sensor(en) 2164 kann/können zur Objekt- und Fußgängererkennung, Notbremsung, Kollisionsvermeidung und/oder anderen Funktionen verwendet werden. Bei mindestens einer Ausführungsform kann/können der/die LIDAR-Sensor(en) 2164 die funktionale Sicherheitsstufe ASIL B aufweisen. Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 mehrere LIDAR-Sensoren 2164 (z. B zwei, vier, sechs usw.) aufweisen, die Ethernet verwenden können (z. B um Daten an einen Gigabit-Ethernet-Switch zu liefern).

[0244] Bei mindestens einer Ausführungsform kann/können der/die LIDAR-Sensor(en) 2164 in der Lage sein, eine Liste von Objekten und deren Entfernungen für ein 360-Grad-Sichtfeld zu liefern. Bei mindestens einer Ausführungsform kann/können der/die handelsübliche(n) LIDAR-Sensor(en) 2164 eine angezeigte Reichweite von etwa 100 m haben, mit einer Genauigkeit von 2 cm bis 3 cm und mit Unterstützung für eine 100-Mbps-Ethernet-Verbindung, zum Beispiel. Bei mindestens einer Ausführungsform können ein oder mehrere nicht vorstehende LIDAR-Sensoren 2164 verwendet werden. Bei einer solchen Ausführungsform kann/-können der/die LIDAR-Sensor(en) 2164 als eine kleine Einrichtung implementiert sein, die in die Front, das Heck, die Seiten und/oder die Ecken des Fahrzeugs 2100 eingebettet sein kann. Bei mindestens einer Ausführungsform kann/können der/die LIDAR-Sensor(en) 2164 in einer solchen Ausführungsform ein horizontales Sichtfeld von bis zu 120 Grad und ein vertikales Sichtfeld von bis zu 35 Grad mit einer Reichweite von 200 m selbst für Objekte mit geringem Reflexionsvermögen bieten. Bei mindestens einer Ausführungsform kann/können der/die frontmontierte(n) LIDAR-Sensor(en) 2164 für ein horizontales Sichtfeld zwischen 45 Grad und 135 Grad ausgestaltet sein.

[0245] Bei mindestens einer Ausführungsform können auch LIDAR-Technologien, wie z. B. 3D Flash LIDAR, verwendet werden. 3D Flash LIDAR verwendet einen Blitz eines Lasers als Sendequelle, um die Umgebung des Fahrzeugs 2100 bis zu einer Entfernung von etwa 200 m zu beleuchten. Bei mindestens einer Ausführungsform weist eine Flash-LIDAR-Einheit ohne Einschränkung einen Rezeptor auf, der die Laufzeit des Laserpulses und das reflektierte Licht auf jedem Pixel aufzeichnet, was wiederum der Entfernung des Fahrzeugs 2100 zu Objekten entspricht. Bei mindestens einer Ausführungsform kann es der Flash-LIDAR ermöglichen, mit jedem Laserblitz hochgenaue und verzerrungsfreie Bilder der Umgebung zu erzeugen. Bei mindestens einer Ausführungsform können vier Flash-LIDAR-Sensoren eingesetzt werden, einer auf jeder Seite des Fahrzeugs 2100. Bei mindestens einer Ausführungsform weisen 3D-Blitz-LIDAR-Systeme ohne Einschränkung eine Festkörper-3D-Star-Array-LIDAR-Kamera auf, die außer einem Gebläse keine beweglichen Teile aufweist (z. B. eine nicht scannende LIDAR-Einrichtung). Bei mindestens einer Ausführungsform kann die Flash-LIDAR-Einrichtung einen 5-Nanosekunden-Laserimpuls der Klasse I (augensicher) pro Bild verwenden und das reflektierte Laserlicht in Form von 3D-Entfernungspunktwolken und koregistrierten Intensitätsdaten erfassen.

[0246] Bei mindestens einer Ausführungsform kann das Fahrzeug darüber hinaus einen oder mehrere IMU-Sensoren 2166 aufweisen. Bei mindestens einer Ausführungsform kann/können der/die IMU-Sensor(en) 2166 in der Mitte der Hinterachse des Fahrzeugs 2100 angeordnet sein. Bei mindestens einer Ausführungsform kann/können der/die IMU-Sensor(en) 2166 beispielsweise und ohne Einschränkung einen oder mehrere Beschleunigungsmesser, Magnetometer, Gyroskop(e), Magnetkompass(e) und/oder andere Sensortypen aufweisen. Bei mindestens einer Ausführungsform, wie z. B. bei sechsachsigen Anwendungen, kann/können der/die IMU-Sensor(en) 2166 ohne Einschränkung Beschleunigungsmesser und Gyroskope aufweisen. Bei mindestens einer Ausführungsform, wie z. B bei neunachsigen Anwendungen, kann/können der/die IMU-Sensor(en) 2166 ohne Einschränkung Beschleunigungsmesser, Gyroskope und Magnetometer aufweisen.

[0247] Bei mindestens einer Ausführungsform können der eine oder die mehreren IMU-Sensoren 2166 als ein GPS-gestütztes Miniaturhochleistungsträgheits-Navigationssystem (GPS-Aided Inertial Navigation System-„GPS/INS“) umgesetzt sein, das Trägheitssensoren von mikroelektromechanischen Systemen („MEMS“), einen hochempfindlichen GPS-Empfänger und erweiterte Kalman-Filteralgorithmen kombiniert, um Schätzungen der Position, Geschwindigkeit und des Verhaltens bereitzustellen. Bei mindestens einer Ausführungsform kann/können der/die IMU-Sensor(en) 2166 das Fahrzeug 2100 in die Lage versetzen, den Kurs zu schätzen, ohne dass Eingaben von einem Magnetsensor erforderlich sind, indem Änderungen der Geschwindigkeit vom GPS direkt mit dem/den IMU-Sensor(en) 2166 beobachtet und korreliert werden. Bei mindestens einer Ausführungsform können IMU-Sensor(en) 2166 und GNSS-Sensor(en) 2158 in einer einzigen integrierten Einheit kombiniert sein.

[0248] Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 ein oder mehrere Mikrofone 2196 aufweisen, die im und/oder um das Fahrzeug 2100 herum angeordnet sind. Bei mindestens einer Ausführungsform kann (können) das (die) Mikrofon(e) 2196 u.a. zur Erkennung und Identifizierung von Einsatzfahrzeugen verwendet werden.

[0249] Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 darüber hinaus eine beliebige Anzahl von Kameratypen aufweisen, einschließlich Stereokamera(s) 2168, Weitwinkelkamera(s) 2170, Infrarotkamera(s) 2172, Umgebungskamera(s) 2174, Weitbereichskamera(s) 2198, Mittelbereichskamera(s) 2176 und/oder anderer Kameratypen. Bei mindestens einer Ausführungsform können Kameras verwendet werden, um Bilddaten rund um den gesamten Umfang des Fahrzeugs 2100 zu erfassen. Bei mindestens einer Ausführungsform hängen die Typen der verwendeten Kameras vom Fahrzeug 2100 ab. Bei mindestens einer Ausführungsform kann eine beliebige Kombination von Kameratypen verwendet werden, um die erforderliche Abdeckung um das Fahrzeug 2100 herum zu gewährleisten. Bei mindestens einer Ausführungsform kann die Anzahl der Kameras je nach Ausführungsform unterschiedlich sein. Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 beispielsweise sechs, sieben, zehn, zwölf oder eine andere Anzahl von Kameras aufweisen. Kameras können zum Beispiel ohne Einschränkung Gigabit Multimedia Serial Link („GMSE“) und/oder Gigabit Ethernet unterstützen. Bei mindestens einer Ausführungsform wird jede der Kameras hierin zuvor unter Bezugnahme auf die **Fig. 21A** und **Fig. 21B** näher beschrieben.

[0250] Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 darüber hinaus einen oder mehrere Schwingungssensoren 2142 aufweisen. Ein Vibrationssensor oder Vibrationssensoren 2142 können Vibrationen von Komponenten des Fahrzeugs 2100, wie etwa einer Achse oder von Achsen, messen. Zum Beispiel können bei mindestens einer Ausführungsform Änderungen der Schwingungen eine Änderung der Straßenoberfläche anzeigen. Bei mindestens einer Ausführungsform, wenn zwei oder mehr Schwingungssensoren 2142 verwendet werden, können Unterschiede zwischen den Schwingungen verwendet werden, um die Reibung oder den Schlupf der Straßenoberfläche zu bestimmen (z. B wenn der Unterschied in den Schwingungen zwischen einer angetriebenen Achse und einer frei drehenden Achse besteht).

[0251] Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 ein ADAS-System 2138 aufweisen. Das ADAS-System 2138 kann bei einigen Beispielen ohne Einschränkung ein SoC aufweisen. Bei mindestens einer Ausführungsform kann das ADAS-System 2138 ohne Einschränkung eine beliebige Anzahl und Kombination eines autonomen/adaptiven/automatischen Geschwindigkeitsregelsystems („ACC“), eines kooperativen adaptiven Geschwindigkeitsregelsystems („CACC“), eines Vorwärtscrashwarnsystems („FCW“), eines automatischen Notbremssystems („AEB“) aufweisen, ein System zur Warnung vor dem Verlassen der Fahrspur („LDW“), ein Spurhalteassistent („LKA“), ein System zur Warnung vor dem toten Winkel („BSW“), ein System zur Warnung vor rückwärtigem Querverkehr („RCTW“), ein System zur Kollisionswarnung („CW“), ein System zur Zentrierung der Fahrspur („LC“) und/oder andere Systeme, Merkmale und/oder Funktionen.

[0252] Bei mindestens einer Ausführungsform kann das ACC-System RADAR-Sensor(en) 2160, LIDAR-Sensor(en) 2164 und/oder eine beliebige Anzahl von Kameras verwenden. Bei mindestens einer Ausführungsform kann das ACC-System ein ACC-System in Längsrichtung und/oder ein ACC-System in Querrichtung aufweisen. Bei mindestens einer Ausführungsform überwacht und steuert das ACC-System in Längsrichtung den Abstand zum unmittelbar vor dem Fahrzeug 2100 befindlichen Fahrzeug und passt die Geschwindigkeit des Fahrzeugs 2100 automatisch an, um einen sicheren Abstand zu vorausfahrenden Fahrzeugen einzuhalten. Bei mindestens einer Ausführungsform übernimmt das seitliche ACC-System die Abstandshaltung und rät dem Fahrzeug 2100, bei Bedarf die Fahrspur zu wechseln. Bei mindestens einer Ausführungsform ist das seitliche ACC-System mit anderen ADAS-Anwendungen wie LC und CW verbunden.

[0253] Bei mindestens einer Ausführungsform verwendet das CACC-System Informationen von anderen Fahrzeugen, die über die Netzwerkschnittstelle 2124 und/oder die Funkantenne(n) 2126 von anderen Fahrzeugen über eine drahtlose Verbindung oder indirekt über eine Netzwerkverbindung (z. B. über das Internet) empfangen werden können. Bei mindestens einer Ausführungsform können direkte Verbindungen durch eine Fahrzeug-zu-Fahrzeug-Kommunikationsverbindung („V2V“) bereitgestellt werden, während indirekte Verbindungen durch eine Infrastruktur-zu-Fahrzeug-Kommunikationsverbindung („I2V“) bereitgestellt werden können. Im Allgemeinen liefert das V2V-Kommunikationskonzept Informationen über unmittelbar vorausfahrende Fahrzeuge (z. B. Fahrzeuge, die sich unmittelbar vor und auf derselben Spur wie Fahrzeug 2100 befinden), während das 12V-Kommunikationskonzept Informationen über den weiter vorausfahrenden Verkehr liefert. Bei mindestens einer Ausführungsform kann das CACC-System entweder eine oder beide 12V- und V2V-Informationsquellen aufweisen. Bei mindestens einer Ausführungsform kann das CACC-System angesichts der Informationen über vorausfahrende Fahrzeuge vor Fahrzeug 2100 zuverlässiger sein und es hat das Potenzial, den Verkehrsfluss zu verbessern und Staus auf der Straße zu reduzieren.

[0254] Bei mindestens einer Ausführungsform ist das FCW-System so konzipiert, dass es den Fahrer vor einer Gefahr warnt, so dass er korrigierend eingreifen kann. Bei mindestens einer Ausführungsform verwendet das FCW-System eine nach vorn gerichtete Kamera und/oder RADAR-Sensor(en) 2160, die mit einem dedizierten Prozessor, DSP, FPGA und/oder ASIC gekoppelt sind, der elektrisch mit der Rückmeldung an den Fahrer gekoppelt ist, z. B. mit einer Anzeige, einem Lautsprecher und/oder einer vibrierenden Komponente. Bei mindestens einer Ausführungsform kann das FCW-System eine Warnung bereitstellen, z. B. in Form eines Tons, einer visuellen Warnung, einer Vibration und/oder eines schnellen Bremsimpulses.

[0255] Bei mindestens einer Ausführungsform erkennt das AEB-System eine drohende Vorwärtskollision mit einem anderen Fahrzeug oder einem anderen Objekt und kann automatisch die Bremsen betätigen, wenn der Fahrer nicht innerhalb eines bestimmten Zeit- oder Entfernungsparameters korrigierend eingreift. Bei mindestens einer Ausführungsform kann das AEB-System (eine) nach vorne gerichtete Kamera(s) und/oder RADAR-Sensor(en) 2160 verwenden, die mit einem speziellen Prozessor, DSP, FPGA und/oder ASIC gekoppelt sind. Bei mindestens einer Ausführungsform warnt das AEB-System, wenn es eine Gefahr erkennt, in der Regel zunächst den Fahrer, damit er korrigierende Maßnahmen ergreift, um eine Kollision zu vermeiden, und wenn der Fahrer keine korrigierenden Maßnahmen ergreift, kann das AEB-System automatisch die Bremsen betätigen, um die Auswirkungen der vorhergesagten Kollision zu verhindern oder zumindest abzumildern. Bei mindestens einer Ausführungsform kann das AEB-System Techniken wie eine dynamische Bremsunterstützung und/oder eine Crash-Imminent-Bremsung bzw. Bremsung bei bevorstehendem Zusammenstoß aufweisen.

[0256] Bei mindestens einer Ausführungsform bietet das LDW-System optische, akustische und/oder taktile Warnungen, wie z. B. Lenkrad- oder Sitzvibrationen, um den Fahrer zu warnen, wenn das Fahrzeug 2100 die Fahrbahnmarkierungen überquert. Bei mindestens einer Ausführungsform wird das LDW-System nicht aktiviert, wenn der Fahrer ein absichtliches Verlassen der Fahrspur anzeigt, indem er einen Blinker betätigt. Bei mindestens einer Ausführungsform kann das LDW-System nach vorne gerichtete Kameras verwenden, die mit einem speziellen Prozessor, DSP, FPGA und/oder ASIC gekoppelt sind, der elektrisch mit der Rückmeldung an den Fahrer gekoppelt ist, z. B. mit einer Anzeige, einem Lautsprecher und/oder einer vibrierenden Komponente. Bei mindestens einer Ausführungsform ist das LKA-System eine Variante des LDW-Systems. Das LKA-System sorgt für einen Lenkeingriff oder ein Bremsen, um das Fahrzeug 2100 zu korrigieren, wenn das Fahrzeug 2100 beginnt, die Fahrspur zu verlassen.

[0257] Bei mindestens einer Ausführungsform erkennt und warnt das BSW-System den Fahrer vor Fahrzeugen, die sich im toten Winkel des Fahrzeugs befinden. Bei mindestens einer Ausführungsform kann das BSW-System eine optische, akustische und/oder taktile Warnung ausgeben, um darauf hinzuweisen, dass das Zusammenführen oder Wechseln der Fahrspur unsicher ist. Bei mindestens einer Ausführungsform kann das BSW-System eine zusätzliche Warnung ausgeben, wenn der Fahrer einen Blinker betätigt. Bei mindestens einer Ausführungsform kann das BSW-System (eine) nach hinten gerichtete Kamera(s) und/oder (einen) RADAR-Sensor(s) 2160 verwenden, der/die mit einem speziellen Prozessor, DSP, FPGA und/oder ASIC gekoppelt ist/sind, der/die elektrisch mit dem Fahrerfeedback gekoppelt ist/sind, wie z. B eine Anzeige, ein Lautsprecher und/oder eine vibrierende Komponente.

[0258] Bei mindestens einer Ausführungsform kann das RCTW-System eine visuelle, akustische und/oder taktile Benachrichtigung liefern, wenn ein Objekt außerhalb des Bereichs der Rückfahrkamera erkannt wird, wenn das Fahrzeug 2100 rückwärts fährt. Bei mindestens einer Ausführungsform weist das RCTW-System ein AEB-System auf, um sicherzustellen, dass die Fahrzeugbremsen betätigt werden, um einen Unfall zu ver-

meiden. Bei mindestens einer Ausführungsform kann das RCTW-System einen oder mehrere nach hinten gerichtete(n) RADAR-Sensor(en) 2160 verwenden, der/die mit einem dedizierten Prozessor, DSP, FPGA und/oder ASIC gekoppelt ist/sind, der/die elektrisch mit einer Fahrerrückkopplung gekoppelt ist/sind, wie z. B. eine Anzeige, ein Lautsprecher und/oder eine vibrierende Komponente.

[0259] Bei mindestens einer Ausführungsform können herkömmliche ADAS-Systeme zu falschpositiven Ergebnissen neigen, die für den Fahrer ärgerlich und ablenkend sein können, aber typischerweise nicht katastrophal sind, weil herkömmliche ADAS-Systeme den Fahrer warnen und ihm die Möglichkeit geben, zu entscheiden, ob eine Sicherheitsbedingung wirklich vorliegt und entsprechend zu handeln. Bei mindestens einer Ausführungsform entscheidet das Fahrzeug 2100 bei widersprüchlichen Ergebnissen selbst, ob das Ergebnis eines Primärrechners oder eines Sekundärrechners (z. B. der ersten Steuerung 2136 oder der zweiten Steuerung 2136) beachtet werden soll. Bei mindestens einer Ausführungsform kann das ADAS-System 2138 beispielsweise ein Backup- und/oder Sekundärcomputer sein, der Wahrnehmungsinformationen an ein Rationalitätsmodul des Backup-Computers liefert. Bei mindestens einer Ausführungsform kann ein Rationalitätsmonitor des Backup-Rechners eine redundante, diverse Software auf Hardwarekomponenten ausführen, um Fehler bei der Wahrnehmung und bei dynamischen Fahraufgaben zu erkennen. Bei mindestens einer Ausführungsform können die Ausgaben des ADAS-Systems 2138 an eine übergeordnete MCU weitergeleitet werden. Bei mindestens einer Ausführungsform bestimmt die überwachende MCU bei Konflikten zwischen den Ausgaben des Primärrechners und des Sekundärrechners, wie der Konflikt beigelegt werden kann, um einen sicheren Betrieb zu gewährleisten.

[0260] Bei mindestens einer Ausführungsform kann der Primärcomputer so ausgestaltet sein, dass er der übergeordneten MCU einen Vertrauenswert liefert, der das Vertrauen des Primärcomputers in das gewählte Ergebnis angibt. Bei mindestens einer Ausführungsform kann die überwachende MCU der Anweisung des Primärcomputers folgen, wenn der Vertrauenswert einen Schwellenwert überschreitet, unabhängig davon, ob der Sekundärcomputer ein widersprüchliches oder inkonsistentes Ergebnis liefert. Bei mindestens einer Ausführungsform, bei der der Vertrauenswert den Schwellenwert nicht erreicht und der primäre und der sekundäre Computer unterschiedliche Ergebnisse (z. B. einen Konflikt) anzeigen, kann die überwachende MCU zwischen den Computern vermitteln, um das geeignete Ergebnis zu bestimmen.

[0261] Bei mindestens einer Ausführungsform kann die überwachende MCU so ausgestaltet sein, dass sie ein neuronales Netz bzw. neuronale Netze ausführt, das bzw. die trainiert und so ausgestaltet ist bzw. sind, dass es bzw. sie zumindest teilweise auf der Grundlage der Ausgaben des Primärcomputers und des Sekundärcomputers die Bedingungen bestimmt bzw. bestimmen, unter denen der Sekundärcomputer Fehlalarme auslöst. Bei mindestens einer Ausführungsform kann (können) das (die) neuronale(n) Netz(e) in der überwachenden MCU lernen, wann der Ausgabe des Sekundärcomputers vertraut werden kann und wann nicht. Zum Beispiel kann bei mindestens einer Ausführungsform, wenn der sekundäre Computer ein RADAR-basiertes FCW-System ist, ein neuronales Netz in der überwachenden MCU lernen, wenn das FCW-System metallische Objekte identifiziert, die in Wirklichkeit keine Gefahren sind, wie z. B. ein Abflussgitter oder ein Schachtdeckel, der einen Alarm auslöst. Bei mindestens einer Ausführungsform, wenn der Sekundärcomputer ein kamerabasiertes LDW-System ist, kann ein neuronales Netz in der überwachenden MCU lernen, das LDW-System außer Kraft zu setzen, wenn Radfahrer oder Fußgänger vorhanden sind und ein Verlassen der Fahrspur tatsächlich das sicherste Manöver ist. Bei mindestens einer Ausführungsform kann die überwachende MCU einen DLA oder eine GPU aufweisen, die für die Ausführung von neuronalen Netzen mit zugehörigem Speicher geeignet sind. Bei mindestens einer Ausführungsform kann die überwachende MCU eine Komponente des/der SoC(s) 2104 umfassen und/oder in einer solchen enthalten sein.

[0262] Bei mindestens einer Ausführungsform kann das ADAS-System 2138 einen sekundären Computer aufweisen, der die ADAS-Funktionalität unter Verwendung herkömmlicher Regeln der Computer Vision ausführt. Bei mindestens einer Ausführungsform kann der sekundäre Computer klassische Computer-Vision-Regeln (wenn-dann) verwenden, und das Vorhandensein eines neuronalen Netzes (von neuronalen Netzen) in der übergeordneten MCU kann die Zuverlässigkeit, Sicherheit und Leistung verbessern. Beispielsweise macht bei mindestens einer Ausführungsform eine vielfältige Umsetzung und absichtliche Nichtidentität ein Gesamtsystem fehlertoleranter, insbesondere für Fehler, die durch eine Funktion von Software (oder einer Software-Hardware-Schnittstelle) verursacht werden. Zum Beispiel, bei mindestens einer Ausführungsform, wenn es einen Software-Bug oder Fehler in der Software gibt, die auf dem primären Computer läuft, und wenn ein nicht-identischer Software-Code, der auf dem sekundären Computer läuft, dasselbe Gesamtergebnis liefert, dann kann die überwachende MCU ein größeres Vertrauen haben, dass das Gesamtergebnis korrekt ist und der Bug in der Software oder Hardware auf dem primären Computer keinen wesentlichen Fehler verursacht.

[0263] Bei mindestens einer Ausführungsform kann die Ausgabe des ADAS-Systems 2138 in den Wahrnehmungsblock des Primärrechners und/oder den Block für dynamische Fahraufgaben des Primärrechners eingespeist werden. Wenn beispielsweise bei mindestens einer Ausführungsform das ADAS-System 2138 eine Vorwärtscrash-Warnung aufgrund eines unmittelbar vorausliegenden Objekts anzeigt, kann der Wahrnehmungsblock diese Information bei der Identifizierung von Objekten verwenden. Bei mindestens einer Ausführungsform kann der sekundäre Computer über ein eigenes neuronales Netz verfügen, das trainiert ist und so das Risiko von Fehlalarmen reduziert, wie es hier beschrieben ist.

[0264] Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 darüber hinaus ein Infotainment-SoC 2130 aufweisen (z. B. ein bordeigenes Infotainment-System (IVI)). Obwohl es als SoC dargestellt und beschrieben ist, kann das Infotainment-System 2130 bei mindestens einer Ausführungsform kein SoC sein und kann ohne Einschränkung zwei oder mehr diskrete Komponenten aufweisen. Bei mindestens einer Ausführungsform kann das Infotainment-SoC 2130 ohne Einschränkung eine Kombination aus Hardware und Software, die verwendet werden kann, um dem Fahrzeug 2100 Audio (z. B. Musik, einen persönlichen digitalen Assistenten, Navigationsanweisungen, Nachrichten, Funk usw.) bereitzustellen, Video (z. B. TV, Filme, Streaming usw.), Telefon (z. B. Freisprechen), Netzwerkanbindungstechnik (z. B. LTE, WiFi usw.) und/oder Informationsdienste (z. B. Navigationssysteme, Rückwärtseinparkunterstützung, ein Funkdatensystem, fahrzeugbezogene Informationen, wie etwa Kraftstoffpegel, insgesamt zurückgelegte Entfernung, Bremsflüssigkeitspegel, Ölpegel, Tür offen/geschlossen, Luftfilterinformationen usw.) beinhalten. Das Infotainment-SoC 2130 kann beispielsweise Radios, Plattenspieler, Navigationssysteme, Videoplayer, eine USB- und Bluetooth-Konnektivität, Carputer, In-Car-Entertainment, WiFi, Audiobedienelemente am Lenkrad, eine Freisprecheinrichtung, ein Heads-up-Display („HUD“), eine HMI-Anzeige 2134, eine Telematikeinrichtung, ein Bedienfeld (z. B. zur Steuerung und/oder Interaktion mit verschiedenen Komponenten, Funktionen und/oder Systemen) und/oder andere Komponenten aufweisen. Bei mindestens einer Ausführungsform kann das Infotainment-SoC 2130 darüber hinaus verwendet werden, um dem/den Benutzer(n) des Fahrzeugs Informationen (z. B visuell und/oder akustisch) bereitzustellen, wie z. B Informationen vom ADAS-System 2138, Informationen zum autonomen Fahren, wie z. B geplante Fahrzeugmanöver, Trajektorien, Umgebungsinformationen (z. B Kreuzungsinformationen, Fahrzeuginformationen, Straßeninformationen, usw.), und/oder andere Informationen.

[0265] Bei mindestens einer Ausführungsform kann das Infotainment-SoC 2130 eine beliebige Menge und Art von GPU-Funktionalität aufweisen. Bei mindestens einer Ausführungsform kann das Infotainment-SoC 2130 über den Bus 2102 (z. B CAN-Bus, Ethernet, etc.) mit anderen Einrichtungen, Systemen und/oder Komponenten des Fahrzeugs 2100 kommunizieren. Bei mindestens einer Ausführungsform kann das Infotainment-SoC 2130 mit einer Überwachungs-MCU gekoppelt sein, so dass die GPU des Infotainment-Systems einige Selbstfahrfunktionen ausführen kann, falls die primäre(n) Steuerung(en) 2136 (z. B. Primär- und/oder Backup-Computer des Fahrzeugs 2100) ausfallen. Bei mindestens einer Ausführungsform kann das Infotainment-SoC 2130 das Fahrzeug 2100 in einen Chauffeur-zu-sicherem-Halt-Modus versetzen, wie es hier beschrieben ist.

[0266] Bei mindestens einer Ausführungsform kann das Fahrzeug 2100 darüber hinaus ein Kombiinstrument 2132 aufweisen (z. B. ein digitales Armaturenbrett, ein elektronisches Kombiinstrument, eine digitale Instrumententafel usw.). Das Instrumenten-Cluster 2132 kann ohne Einschränkung eine Steuervorrichtung und/oder einen Supercomputer (z. B. eine separate Steuervorrichtung oder einen separaten Supercomputer) beinhalten. Bei mindestens einer Ausführungsform kann das Kombiinstrument 2132 ohne Einschränkung eine beliebige Anzahl und Kombination von Instrumenten aufweisen, wie z. B. Tachometer, Kraftstoffstand, Öldruck, Drehzahlmesser, Kilometerzähler, Blinker, Schaltstellungsanzeige, Sicherheitsgurtwarnleuchte(n), Parkbremswarnleuchte(n), Motorstörungsleuchte(n), Informationen über zusätzliche Rückhaltesysteme (z. B. Airbags), Beleuchtungssteuerungen, Sicherheitssystemsteuerungen, Navigationsinformationen usw. Bei einigen Beispielen können die Informationen auf dem Infotainment-SoC 2130 und dem Kombiinstrument 2132 angezeigt und/oder gemeinsam genutzt werden. Bei mindestens einer Ausführungsform kann das Kombiinstrument 2132 einen Teil des Infotainment-SoC 2130 aufweisen, oder umgekehrt.

[0267] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform kann Ableitungs- und/oder Trainingslogik 1815 in dem System **Fig. 21C** zum Ableiten oder Vorhersagen von Operationen mindestens zum Teil basierend auf Gewichtungsparametern verwendet werden, die unter Verwendung von neuronalen Netzwerktrainingsopera-

tionen, Funktionen und/oder Architekturen von neuronalen Netzwerken oder in dieser Schrift beschriebenen Anwendungsfällen von neuronalen Netzwerken berechnet wurden.

[0268] Bei mindestens einer Ausführungsform wird das System **Fig. 21C** eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird das System **Fig. 21C** eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird das System **Fig. 21C** eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und das System **Fig. 21C** wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0269] **Fig. 21D** ist ein Diagramm eines Systems 2176 zur Kommunikation zwischen einem oder mehreren cloudbasierten Servern und dem autonomen Fahrzeug 2100 der **Fig. 21A** gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform kann das System 2176 ohne Einschränkung den/die Server 2178, das/die Netzwerk(e) 2190 und eine beliebige Anzahl und Art von Fahrzeugen, einschließlich des Fahrzeugs 2100, beinhalten. Der/die Server 2178 kann/können ohne Einschränkung eine Vielzahl von GPUs 2184(A)-2184(H) (hierin kollektiv als GPUs 2184 bezeichnet), PCIe-Switches 2182(A)-2182(H) (hierin kollektiv als PCIe-Switches 2182 bezeichnet), und/oder CPUs 2180(A)-2180(B) (hierin kollektiv als CPUs 2180 bezeichnet) aufweisen. GPUs 2184, CPUs 2180 und PCIe-Switches 2182 können über Hochgeschwindigkeitsverbindungen miteinander verbunden sein, wie z. B. und ohne Einschränkung über die von NVIDIA entwickelten NVLink-Schnittstellen 2188 und/oder PCIe-Verbindungen 2186. Bei mindestens einer Ausführungsform sind die GPUs 2184 über ein NVLink- und/oder NVSwitch-SoC und die GPUs 2184 und PCIe-Switches 2182 über PCIe-Verbindungen verbunden. Bei mindestens einer Ausführungsform sind zwar acht GPUs 2184, zwei CPUs 2180 und vier PCIe-Switches 2182 dargestellt, dies ist jedoch nicht als Einschränkung zu verstehen. Bei mindestens einer Ausführungsform kann jeder der Server 2178 ohne Einschränkung eine beliebige Anzahl von GPUs 2184, CPUs 2180 und/oder PCIe-Switches 2182 in beliebiger Kombination aufweisen. Bei mindestens einer Ausführungsform kann/können der/die Server 2178 beispielsweise jeweils acht, sechzehn, zweiunddreißig und/oder mehr GPUs 2184 beinhalten.

[0270] Bei mindestens einer Ausführungsform kann (können) der (die) Server 2178 über das (die) Netzwerk (e) 2190 und von Fahrzeugen Bilddaten empfangen, die für Bilder repräsentativ sind, die unerwartete oder veränderte Straßenzustände zeigen, wie beispielsweise kürzlich begonnene Straßenarbeiten. Bei mindestens einer Ausführungsform kann/können der/die Server 2178 über das/die Netzwerk(e) 2190 und an Fahrzeuge neuronale Netzwerke 2192, aktualisierte neuronale Netzwerke 2192 und/oder Karteninformationen 2194 übertragen, die ohne Einschränkung Informationen über den Verkehr und die Straßenbedingungen beinhalten. Bei mindestens einer Ausführungsform können die Aktualisierungen der Karteninformationen 2194 ohne Einschränkung Aktualisierungen für die HD-Karte 2122, z.B. Informationen zu Baustellen, Schlaglöchern, Umleitungen, Überschwemmungen und/oder anderen Hindernissen beinhalten. Bei mindestens einer Ausführungsform können neuronale Netze 2192, aktualisierte neuronale Netze 2192 und/oder Karteninformationen 2194 aus neuem Training und/oder Erfahrungen resultieren, die in Daten repräsentiert sind, die von einer beliebigen Anzahl von Fahrzeugen in der Umgebung empfangen wurden, und/oder zumindest teilweise auf einem Training basieren, das in einem Rechenzentrum durchgeführt wurde (z. B. unter Verwendung von Server(n) 2178 und/oder anderen Servern).

[0271] Bei mindestens einer Ausführungsform kann/können der/die Server 2178 verwendet werden, um Modelle zum maschinellen Lernen (z. B neuronale Netze) zumindest teilweise auf der Grundlage von Trainingsdaten zu trainieren. Die Trainingsdaten können von Fahrzeugen erzeugt werden und/oder können in einer Simulation (z. B. unter Verwendung einer Spiel-Engine) erzeugt werden. Bei mindestens einer Ausführungsform ist eine beliebige Menge an Trainingsdaten markiert (z. B. bei denen das assoziierte neuronale Netzwerk von dem überwachten Lernen profitiert) und/oder anderes Vorverarbeiten erfährt. Bei mindestens einer Ausführungsform ist eine beliebige Menge an Trainingsdaten nicht markiert und/oder vorverarbeitet (z. B. bei denen das assoziierte neuronale Netzwerk kein überwachtes Lernen erfordert). Bei mindestens einer Ausführungsform können, sobald Modelle zum maschinellen Lernen trainiert sind, Modelle zum maschinellen Lernen von Fahrzeugen verwendet werden (z. B Übertragung an Fahrzeuge über Netzwerk(e) 2190, und/oder Modelle zum maschinellen Lernen können von Server(n) 2178 zur Fernüberwachung von Fahrzeugen verwendet werden.

[0272] Bei mindestens einer Ausführungsform kann (können) der (die) Server 2178 Daten von Fahrzeugen empfangen und Daten auf aktuelle neuronale Echtzeit-Netze für intelligentes Inferencing in Echtzeit anwenden. Bei mindestens einer Ausführungsform kann/können der/die Server 2178 Deep-Learning-Supercomputer und/oder dedizierte KI-Computer aufweisen, die von GPU(s) 2184 angetrieben werden, wie z. B. die von NVIDIA entwickelten DGX- und DGX-Station-Maschinen. Bei mindestens einer Ausführungsform kann/können der/die Server 2178 jedoch eine Deep-Learning-Infrastruktur aufweisen, die CPU-betriebene Rechenzentren verwendet.

[0273] Bei mindestens einer Ausführungsform kann die Deep-Learning-Infrastruktur von Server(n) 2178 zu schnellem Inferencing in Echtzeit fähig sein und diese Fähigkeit nutzen, um den Zustand von Prozessoren, Software und/oder zugehöriger Hardware im Fahrzeug 2100 zu bewerten und zu überprüfen. Bei mindestens einer Ausführungsform kann die Deep-Learning-Infrastruktur beispielsweise periodische Aktualisierungen vom Fahrzeug 2100 erhalten, wie etwa eine Bildsequenz und/oder Objekte, die das Fahrzeug 2100 in dieser Bildsequenz lokalisiert hat (z. B. über Computer Vision und/oder andere maschinelle Objektklassifizierungstechniken). Bei mindestens einer Ausführungsform kann die Deep-Learning-Infrastruktur ihr eigenes neuronales Netzwerk laufen lassen, um Objekte zu identifizieren und sie mit den Objekten, die von dem Fahrzeug 2100 identifiziert werden, zu vergleichen, und falls die Ergebnisse nicht übereinstimmen und die Deep-Learning-Infrastruktur zu dem Schluss kommt, dass die KI in dem Fahrzeug 2100 eine Fehlfunktion aufweist, kann/können der/die Server 2178 ein Signal an das Fahrzeug 2100 senden, das einen ausfallsicheren Computer des Fahrzeugs 2100 anweist, die Steuerung zu übernehmen, die Fahrgäste zu benachrichtigen und ein sicheres Parkmanöver auszuführen.

[0274] Bei mindestens einer Ausführungsform kann/können der/die Server 2178 GPU(s) 2184 und einen oder mehrere programmierbare Inferenzbeschleuniger (z. B NVIDIAs TensorRT 3) aufweisen. Bei mindestens einer Ausführungsform kann die Kombination von GPU-gesteuerten Servern und Inferenzbeschleunigung eine Reaktionsfähigkeit in Echtzeit ermöglichen. Bei mindestens einer Ausführungsform, z. B. wenn die Leistung weniger kritisch ist, können für das Inferencing auch Server mit CPUs, FPGAs und anderen Prozessoren verwendet werden. Bei mindestens einer Ausführungsform wird (werden) die Hardwarestruktur(en) 1815 zur Ausführung einer oder mehrerer Ausführungsformen verwendet. Details bezüglich der (den) Hardware-Struktur(en) 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt.

COMPUTERSYSTEME

[0275] **Fig. 22** ist ein Blockdiagramm, das ein beispielhaftes Computersystem veranschaulicht, das ein System mit miteinander verbundenen Einrichtungen und Komponenten, ein System-on-a-Chip (SOC) oder eine Kombination davon 2200 sein kann, das gemäß mindestens einer Ausführungsform einen Prozessor aufweist, der Ausführungseinheiten zur Ausführung eines Befehls enthält. Bei mindestens einer Ausführungsform kann das Computersystem 2200 ohne Einschränkung eine Komponente, wie z. B. einen Prozessor 2202, aufweisen, um Ausführungseinheiten einschließlich Logik zur Durchführung von Algorithmen zur Verarbeitung von Daten gemäß der vorliegenden Offenbarung einzusetzen, wie z. B. bei der hier beschriebenen Ausführungsform. Bei mindestens einer Ausführungsform kann das Computersystem 2200 Prozessoren aufweisen, wie z. B. die PENTIUM®-Prozessorfamilie, Xeon™-, Itanium®-, XScale™- und/oder StrongARM™-, Intel® Core™- oder Intel® Nervana™-Mikroprozessoren, die von der Intel Corporation in Santa Clara, Kalifornien, erhältlich sind, obwohl auch andere Systeme (einschließlich PCs mit anderen Mikroprozessoren, technische Workstations, Set-Top-Boxen und dergleichen) verwendet werden können. Bei mindestens einer Ausführungsform kann das Computersystem 2200 eine Version des Betriebssystems WINDOWS ausführen, das von der Microsoft Corporation in Redmond, Washington, erhältlich ist, obwohl auch andere Betriebssysteme (z. B UNIX und Linux), eingebettete Software und/oder grafische Benutzeroberflächen verwendet werden können.

[0276] Ausführungsformen können auch bei anderen Ausführungen wie Handheld-Geräten und eingebetteten Anwendungen verwendet werden. Einige Beispiele für tragbare Einrichtungen weisen Mobiltelefone, Internetprotokollgeräte, Digitalkameras, persönliche digitale Assistenten („PDAs“) und Handheld-PCs auf. Bei mindestens einer Ausführungsform können eingebettete Anwendungen einen Mikrocontroller, einen digitalen Signalprozessor („DSP“), ein System auf einem Chip, Netzwerkcomputer („NetPCs“), Set-Top-Boxen, Netzwerk-Hubs, Wide-Area-Network-Switches („WAN“) oder jedes andere System aufweisen, das eine oder mehrere Anweisungen gemäß mindestens einer Ausführungsform ausführen kann.

[0277] Bei mindestens einer Ausführungsform kann das Computersystem 2200 ohne Einschränkung einen Prozessor 2202 aufweisen, der ohne Einschränkung eine oder mehrere Ausführungseinheiten 2208 aufwei-

sen kann, um das Training eines Modells zum maschinellen Lernen und/oder Inferencing gemäß den hier beschriebenen Techniken durchzuführen. Bei mindestens einer Ausführungsform ist das System 22 ein Einzelprozessor-Desktop- oder -Serversystem, aber in einer anderen Ausführungsform kann das System 22 ein Multiprozessorsystem sein. Bei mindestens einer Ausführungsform kann der Prozessor 2202 ohne Einschränkung einen CISC-Mikroprozessor (Complex Instruction Set Computer), einen RISC-Mikroprozessor (Reduced Instruction Set Computing), einen VLIW-Mikroprozessor (Very Long Instruction Word), einen Prozessor, der eine Kombination von Befehlssätzen implementiert, oder eine beliebige andere Einrichtung, wie z. B. einen digitalen Signalprozessor, aufweisen. Bei mindestens einer Ausführungsform kann der Prozessor 2202 mit einem Prozessorbus 2210 verbunden sein, der Datensignale zwischen dem Prozessor 2202 und anderen Komponenten im Computersystem 2200 übertragen kann.

[0278] Bei mindestens einer Ausführungsform kann der Prozessor 2202 ohne Einschränkung einen internen Level 1 („L1“) Cache-Speicher („Cache“) 2204 aufweisen. Bei mindestens einer Ausführungsform kann der Prozessor 2202 einen einzigen internen Cache oder mehrere Ebenen eines internen Caches aufweisen. Bei mindestens einer Ausführungsform kann sich der Cache-Speicher außerhalb des Prozessors 2202 befinden. Andere Ausführungsformen können auch eine Kombination aus internen und externen Caches aufweisen, abhängig von der jeweiligen Implementierung und den Bedürfnissen. Bei mindestens einer Ausführungsform kann die Registerdatei 2206 verschiedene Datentypen in verschiedenen Registern speichern, einschließlich, ohne Einschränkung, Ganzzahlregister, Gleitkommaregister, Statusregister und Befehlszeigerregister.

[0279] Bei mindestens einer Ausführungsform befindet sich die Ausführungseinheit 2208, die ohne Einschränkung eine Logik zur Durchführung von Ganzzahl- und Gleitkommaoperationen aufweist, ebenfalls im Prozessor 2202. Der Prozessor 2202 kann auch einen Mikrocode- („ucode“-) Nur-LeseSpeicher („ROM“) beinhalten, der Mikrocode für bestimmte Makroanweisungen speichert. Bei mindestens einer Ausführungsform kann die Ausführungseinheit 2208 eine Logik zur Handhabung eines gepackten Befehlssatzes 2209 aufweisen. Bei mindestens einer Ausführungsform können durch das Vorweisen eines gepackten Befehlssatzes 2209 in einem Befehlssatz eines Mehrzweckprozessors 2202 zusammen mit einer zugehörigen Schaltung zur Ausführung von Befehlen die von vielen Multimedia-Anwendungen verwendeten Operationen unter Verwendung gepackter Daten in einem Mehrzweckprozessor 2202 durchgeführt werden. In einer oder mehreren Ausführungsformen können viele Multimedia-Anwendungen beschleunigt und effizienter ausgeführt werden, indem die volle Breite des Datenbusses eines Prozessors für die Durchführung von Operationen mit gepackten Daten genutzt wird, wodurch die Notwendigkeit entfällt, kleinere Dateneinheiten über den Datenbus des Prozessors zu übertragen, um eine oder mehrere Operationen mit einem Datenelement nach dem anderen durchzuführen.

[0280] Bei mindestens einer Ausführungsform kann die Ausführungseinheit 2208 auch in Mikrocontrollern, eingebetteten Prozessoren, Grafikeinrichtungen, DSPs und anderen Arten von Logikschaltungen verwendet werden. Bei mindestens einer Ausführungsform kann das Computersystem 2200, ohne Einschränkung, einen Speicher 2220 aufweisen. Bei mindestens einer Ausführungsform kann der Speicher 2220 als dynamische Random-Access-Memory- („DRAM“) Einrichtung, statische Random-Access-Memory- („SRAM“) Einrichtung, Flash-Speichereinrichtung oder andere Speichereinrichtung implementiert sein. Der Speicher 2220 kann (eine) Anweisung(en) 2219 und/oder Daten 2221 speichern, die von Datensignalen, die von dem Prozessor 2202 ausgeführt werden können, dargestellt sind.

[0281] Bei mindestens einer Ausführungsform kann der Systemlogikchip mit dem Prozessorbus 2210 und dem Speicher 2220 verbunden sein. Bei mindestens einer Ausführungsform kann der Systemlogikchip ohne Einschränkung einen Speichersteuerungs-Hub („MCH“) 2216 aufweisen, und der Prozessor 2202 kann mit dem MCH 2216 über den Prozessorbus 2210 kommunizieren. Bei mindestens einer Ausführungsform kann der MCH 2216 einen Speicherpfad 2218 mit hoher Bandbreite zum Speicher 2220 für die Befehls- und Datenspeicherung sowie für die Speicherung von Grafikbefehlen, Daten und Texturen bereitstellen. Bei mindestens einer Ausführungsform kann der MCH 2216 Datensignale zwischen dem Prozessor 2202, dem Speicher 2220 und anderen Komponenten im Computersystem 2200 leiten und Datensignale zwischen dem Prozessorbus 2210, dem Speicher 2220 und einem System-E/A 2222 überbrücken. Bei mindestens einer Ausführungsform kann der Systemlogikchip einen Grafikanschluss zur Verbindung mit einer Grafiksteuerung bereitstellen. Bei mindestens einer Ausführungsform kann der MCH 2216 über einen Speicherpfad 2218 mit hoher Bandbreite mit dem Speicher 2220 gekoppelt sein, und die Grafik-/Videokarte 2212 kann über eine AGP-Verbindung 2214 mit dem MCH 2216 gekoppelt sein.

[0282] Bei mindestens einer Ausführungsform kann das Computersystem 2200 einen System-E/A-Bus 2222 verwenden, bei dem es sich um einen proprietären Hub-Interface-Bus handelt, um den MCH 2216 mit dem

E/A-Controller-Hub („ICH“) 2230 zu verbinden. Bei mindestens einer Ausführungsform kann der ICH 2230 direkte Verbindungen zu einigen E/A-Einrichtungen über einen lokalen E/A-Bus bereitstellen. Bei mindestens einer Ausführungsform kann der lokale E/A-Bus ohne Einschränkung einen Hochgeschwindigkeits-E/A-Bus zur Verbindung von Peripheriegeräten mit dem Speicher 2220, dem Chipsatz und dem Prozessor 2202 aufweisen. Beispiele können unter anderem einen Audiocontroller 2229, einen Firmware-Hub („Flash-BIOS“) 2228, einen drahtlosen Transceiver 2226, einen Datenspeicher 2224, einen Legacy-E/A-Controller 2223 mit Benutzereingabe- und Tastaturschnittstellen, einen seriellen Erweiterungsanschluss 2227, wie Universal Serial Bus („USB“), und eine Netzwerksteuerung 2234 aufweisen. Der Datenspeicher 2224 kann ein Festplattenlaufwerk, ein Diskettenlaufwerk, eine CD-ROM-Vorrichtung, ein Flash-Speichervorrichtung oder eine andere Massenspeichervorrichtung umfassen.

[0283] Bei mindestens einer Ausführungsform zeigt **Fig. 22** ein System, das miteinander verbundene Hardware-Einrichtungen oder „Chips“ aufweist, während bei anderen Ausführungen **Fig. 22** ein beispielhaftes System on a Chip („SoC“) zeigen kann. Bei mindestens einer Ausführungsform können die in **Fig. cc** dargestellten Einrichtungen mit proprietären Verbindungen, standardisierten Verbindungen (z. B. PCIe) oder einer Kombination davon miteinander verbunden sein. Bei mindestens einer Ausführungsform sind eine oder mehrere Komponenten des Systems 2200 über Compute-Express-Link (CXL)-Verbindungen miteinander verbunden.

[0284] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform kann Ableitungs- und/oder Trainingslogik 1815 in dem System **Fig. 22** zum Ableiten oder Vorhersagen von Operationen mindestens zum Teil basierend auf Gewichtungsparametern verwendet werden, die unter Verwendung von neuronalen Netzwerktrainingsoperationen, Funktionen und/oder Architekturen von neuronalen Netzwerken oder in dieser Schrift beschriebenen Verwendungsfällen von neuronalen Netzwerken berechnet wurden.

[0285] Bei mindestens einer Ausführungsform wird das System **Fig. 22** eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird das System **Fig. 22** eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird das System **Fig. 22** eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und das System **Fig. 22** wird in Verbindung mit einem oder mehreren Prozessen verwendet, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts in einem Bild auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen berechnet werden, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0286] **Fig. 23** ist ein Blockdiagramm, das eine elektronische Einrichtung 2300 zur Verwendung eines Prozessors 2310 gemäß mindestens einer Ausführungsform zeigt. Bei mindestens einer Ausführungsform kann die elektronische Einrichtung 2300 beispielsweise und ohne Einschränkung ein Notebook, ein Tower-Server, ein Rack-Server, ein Blade-Server, ein Laptop, ein Desktop-Computer, ein Tablet, eine mobile Einrichtung, ein Telefon, ein eingebetteter Computer oder jede andere geeignete elektronische Einrichtung sein.

[0287] Bei mindestens einer Ausführungsform kann das System 2300 ohne Einschränkung einen Prozessor 2310 aufweisen, der kommunikativ mit einer beliebigen Anzahl oder Art von Komponenten, Peripheriegeräten, Modulen oder Einrichtungen gekoppelt ist. Bei mindestens einer Ausführungsform ist der Prozessor 2310 unter Verwendung eines Busses oder einer Schnittstelle gekoppelt, wie etwa eines 1°C-Busses, eines Systemverwaltungsbusses (System Management Bus-„SMBus“), eines Busses mit geringer Pin-Anzahl (Low Pin Count-LPC), einer seriellen Peripherieschnittstelle (Serial Peripheral Interface-„SPI“), eines High-Definition-Audio-Busses („HDA“-Busses), eines Serial-Advance-Technology-Attachment-Busses („SATA“-Busses), eines universellen seriellen Busses („USB“) (Versionen 1, 2, 3) oder eines Busses eines universellen asynchronen Empfänger/Senders (Universal Asynchronous Receiver/Transmitter-„UART“). Bei mindestens einer Ausführungsform veranschaulicht **Fig. 23** ein System, das miteinander verbundene Hardwarevorrichtungen oder „Chips“ beinhaltet, wohingegen bei anderen Ausführungsformen **Fig. 23** ein beispielhaftes System auf einem Chip („SoC“) veranschaulichen kann. Bei mindestens einer Ausführungsform können die in **Fig. 23** dargestellten Einrichtungen mit proprietären Verbindungen, standardisierten Verbindungen (z. B. PCIe) oder einer Kombination davon miteinander verbunden sein. Bei mindestens einer Ausführungsform sind eine oder

mehrere Komponenten von **Fig. 23** über Compute-Express-Link (CXL)-Verbindungen miteinander verbunden.

[0288] Bei mindestens einer Ausführungsform kann **Fig. 23** eine Anzeige 2324, einen Touchscreen 2325, ein Touchpad 2330, eine Near-Field-Communications-Einheit („NFC“) 2345, einen Sensor-Hub 2340, einen Wärmesensor 2346, einen Express-Chipsatz („EC“) 2335, ein Trusted Platform Module („TPM“) 2338, BIOS/Firmware/Flash-Speicher („BIOS, FW Flash“) 2322, ein DSP 2360, ein Laufwerk („SSD oder HDD“) 2320, wie eine Solid State Disk („SSD“) oder eine Festplatte („HDD“), eine drahtlose lokale Netzwerkeinheit („WLAN“) 2350, eine Bluetooth-Einheit 2352, eine drahtlose Wide Area Network-Einheit („WWAN“) 2356, ein Global Positioning System (GPS) 2355, eine Kamera („USB 3. 0-Kamera“) 2354, wie z. B. eine USB 3.0-Kamera, oder eine Low Power Double Data Rate („EPDDR“)-Speichereinheit („EPDDR3“) 2315, die z. B. in dem LPDDR3-Standard umgesetzt ist, beinhalten. Diese Komponenten können in jeder geeigneten Weise implementiert sein.

[0289] Bei mindestens einer Ausführungsform können andere Komponenten mit dem Prozessor 2310 über die oben beschriebenen Komponenten kommunikativ verbunden sein. Bei mindestens einer Ausführungsform können ein Beschleunigungsmesser 2341, ein Umgebungslichtsensor („ALS“) 2342, ein Kompass 2343 und ein Gyroskop 2344 kommunikativ mit dem Sensor-Hub 2340 verbunden sein. Bei mindestens einer Ausführungsform können ein Wärmesensor 2339, ein Lüfter 2337, eine Tastatur 2346 und ein Touchpad 2330 kommunikativ mit dem EC 2335 verbunden sein. Bei mindestens einer Ausführungsform können der Lautsprecher 2363, ein Kopfhörer 2364 und ein Mikrofon („mic“) 2365 kommunikativ mit einer Audioeinheit („audio codec and class d amp“) 2364 gekoppelt sein, die ihrerseits kommunikativ mit dem DSP 2360 gekoppelt sein kann. Bei mindestens einer Ausführungsform kann die Audioeinheit 2364 beispielsweise und ohne Einschränkung einen Audiocodierer/-decoder („Codec“) und einen Verstärker der Klasse D aufweisen. Bei mindestens einer Ausführungsform kann die SIM-Karte („SIM“) 2357 mit der WWAN-Einheit 2356 kommunikativ gekoppelt sein. Bei mindestens einer Ausführungsform können Komponenten wie die WLAN-Einheit 2350 und die Bluetooth-Einheit 2352 sowie die WWAN-Einheit 2356 in einem Next Generation Form Factor („NGFF“) implementiert sein.

[0290] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform kann Ableitungs- und/oder Trainingslogik 1815 in dem System **Fig. 23** zum Ableiten oder Vorhersagen von Operationen mindestens zum Teil basierend auf Gewichtungsparametern verwendet werden, die unter Verwendung von neuronalen Netzwerktrainingsoperationen, Funktionen und/oder Architekturen von neuronalen Netzwerken oder in dieser Schrift beschriebenen Verwendungsfällen von neuronalen Netzwerken berechnet wurden.

[0291] Bei mindestens einer Ausführungsform wird das System **Fig. 23** eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird das System **Fig. 23** eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird das System **Fig. 23** eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und das System **Fig. 23** wird in Verbindung mit einem oder mehreren Prozessen verwendet, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts in einem Bild auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen berechnet werden, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0292] **Fig. 24** veranschaulicht ein Computersystem 2400 gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform ist das Computersystem 2400 ausgestaltet, um verschiedene in dieser Offenbarung beschriebene Prozesse und Verfahren zu implementieren.

[0293] Bei mindestens einer Ausführungsform umfasst das Computersystem 2400 ohne Einschränkung mindestens eine Zentraleinheit („CPU“) 2402, die an einen Kommunikationsbus 2410 angeschlossen ist, der unter Verwendung eines beliebigen geeigneten Protokolls implementiert ist, wie PCI („Peripheral Component Interconnect“), Peripheral Component Interconnect Express („PCI-Express“), AGP („Accelerated Graphics Port“), HyperTransport oder ein anderes Bus- oder Punkt-zu-Punkt-Kommunikationsprotokoll. Bei mindestens einer Ausführungsform weist das Computersystem 2400 ohne Einschränkung einen Hauptspeicher

2404 und eine Steuerlogik auf (z. B implementiert als Hardware, Software oder eine Kombination davon), und die Daten werden im Hauptspeicher 2404 gespeichert, der die Form eines Direktzugriffsspeichers („RAM“) annehmen kann. Bei mindestens einer Ausführungsform stellt ein Netzwerkschnittstellen-Subsystem („Netzwerkschnittstelle“) 2422 eine Schnittstelle zu anderen Recheneinrichtungen und Netzwerken bereit, um Daten von dem Computersystem 2400 zu empfangen und an andere Systeme zu übermitteln.

[0294] In mindestens einer Ausführungsform weist das Computersystem 2400 ohne Einschränkung Eingabeeinrichtungen 2408, ein Parallelverarbeitungssystem 2412 und Anzeigeeinrichtungen 2406 auf, die unter Verwendung einer herkömmlichen Kathodenstrahlröhre („CRT“), einer Flüssigkristallanzeige („LCD“), einer lichtemittierenden Diode („LED“), einer Plasmaanzeige oder anderer geeigneter Anzeigetechnologien implementiert sein können. Bei mindestens einer Ausführungsform werden Benutzereingaben von Eingabeeinrichtungen 2408 wie Tastatur, Maus, Touchpad, Mikrofon und anderen empfangen. Bei mindestens einer Ausführungsform kann jedes der vorgenannten Module auf einer einzigen Halbleiterplattform angeordnet sein, um ein Verarbeitungssystem zu bilden.

[0295] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform kann Ableitungs- und/oder Trainingslogik 1815 in dem System **Fig. 24** zum Ableiten oder Vorhersagen von Operationen mindestens zum Teil basierend auf Gewichtungsparametern verwendet werden, die unter Verwendung von neuronalen Netzwerktrainingsoperationen, Funktionen und/oder Architekturen von neuronalen Netzwerken oder in dieser Schrift beschriebenen Verwendungsfällen von neuronalen Netzwerken berechnet wurden.

[0296] Bei mindestens einer Ausführungsform wird das System **Fig. 24** eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird das System **Fig. 24** eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird das System **Fig. 24** eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und das System **Fig. 24** wird in Verbindung mit einem oder mehreren Prozessen verwendet, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts in einem Bild auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen berechnet werden, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0297] **Fig. 25** veranschaulicht ein Computersystem 2500 gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform weist das Computersystem 2500, ohne Einschränkung, einen Computer 2510 und einen USB-Stick 2520 auf. Bei mindestens einer Ausführungsform kann der Computer 2510 ohne Einschränkung eine beliebige Anzahl und Art von Prozessor(en) (nicht dargestellt) und einen Speicher (nicht dargestellt) aufweisen. Bei mindestens einer Ausführungsform weist der Computer 2510, ohne Einschränkung, einen Server, eine Cloud-Instanz, einen Laptop und einen Desktop-Computer auf.

[0298] Bei mindestens einer Ausführungsform weist der USB-Stick 2520, ohne Einschränkung, eine Verarbeitungseinheit 2530, eine USB-Schnittstelle 2540 und eine USB-Schnittstellenlogik 2550 auf. Bei mindestens einer Ausführungsform kann die Verarbeitungseinheit 2530 ein beliebiges Befehlsausführungssystem, ein Gerät oder eine Einrichtung sein, die in der Lage ist, Befehle auszuführen. Bei mindestens einer Ausführungsform kann die Verarbeitungseinheit 2530 ohne Einschränkung eine beliebige Anzahl und Art von Verarbeitungskernen (nicht dargestellt) aufweisen. Bei mindestens einer Ausführungsform umfasst der Verarbeitungskern 2530 eine anwendungsspezifische integrierte Schaltung („ASIC“), die für die Durchführung beliebiger Mengen und Arten von Operationen im Zusammenhang mit maschinellem Lernen optimiert ist. Bei mindestens einer Ausführungsform ist der Verarbeitungskern 2530 beispielsweise eine Tensor Processing Unit („TPC“), die für die Durchführung von Inferenzoperationen des maschinellen Lernens optimiert ist. Bei mindestens einer Ausführungsform ist der Verarbeitungskern 2530 eine Bildverarbeitungseinheit („VPU“), die für die Durchführung von Bildverarbeitungs- und maschinellen Lemoperationen optimiert ist.

[0299] Bei mindestens einer Ausführungsform kann die USB-Schnittstelle 2540 eine beliebige Art von USB-Stecker oder USB-Buchse sein. Bei mindestens einer Ausführungsform ist die USB-Schnittstelle 2540 beispielsweise eine USB 3.0 Typ-C-Buchse für Daten und Strom. Bei mindestens einer Ausführungsform ist die USB-Schnittstelle 2540 ein USB-3.0-Typ-A-Stecker. Bei mindestens einer Ausführungsform kann die USB-

Schnittstellenlogik 2550 eine beliebige Menge und Art von Logik aufweisen, die es der Verarbeitungseinheit 2530 ermöglicht, sich über den USB-Anschluss 2540 mit einer Einrichtung (z. B. einem Computer 2510) zu verbinden.

[0300] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform kann Ableitungs- und/oder Trainingslogik 1815 in dem System **Fig. 25** zum Ableiten oder Vorhersagen von Operationen mindestens zum Teil basierend auf Gewichtungsparametern verwendet werden, die unter Verwendung von neuronalen Netzwerktrainingsoperationen, Funktionen und/oder Architekturen von neuronalen Netzwerken oder in dieser Schrift beschriebenen Verwendungsfällen von neuronalen Netzwerken berechnet wurden.

[0301] Bei mindestens einer Ausführungsform wird das System **Fig. 25** eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird das System **Fig. 25** eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird das System **Fig. 25** eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und das System **Fig. 25** wird in Verbindung mit einem oder mehreren Prozessen verwendet, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teile des Trainings berechnet werden.

[0302] **Fig. 26A** veranschaulicht eine beispielhafte Architektur, in der eine Vielzahl von GPUs 2610-2613 mit einer Vielzahl von Mehrkern-Prozessoren 2605-2606 über Hochgeschwindigkeitsverbindungen 2640-2643 (z. B Busse, Punkt-zu-Punkt-Verbindungen usw.) kommunikativ gekoppelt ist. Bei einer Ausführungsform unterstützen die Hochgeschwindigkeitsverbindungen 2640-2643 einen Kommunikationsdurchsatz von 4 GB/s, 30 GB/s, 80 GB/s oder mehr. Es können verschiedene Verbindungsprotokolle verwendet werden, die PCIe 4.0 oder 5.0 und NVLink 2.0 einschließen, aber nicht darauf beschränkt sind.

[0303] Zusätzlich und in einer Ausführungsform sind zwei oder mehr GPUs 2610-2613 über Hochgeschwindigkeitsverbindungen 2629-2630 miteinander verbunden, die mit denselben oder anderen Protokollen/Verbindungen implementiert sein können als die für Hochgeschwindigkeitsverbindungen 2640-2643 verwendeten. Gleichermaßen können zwei oder mehr Mehrkernprozessoren 2605-2606 über eine Hochgeschwindigkeitsverbindung 2628 verbunden sein, die symmetrische Mehrprozessor-Busse (Symmetrie Multi-Processor-„SMP“-Busse) sein können, die mit 20 GB/s, 30 GB/s, 120 GB/s oder mehr betrieben werden. Alternativ kann die gesamte Kommunikation zwischen den verschiedenen in **Fig. 26A** gezeigten Systemkomponenten über dieselben Protokolle/Leitungen erfolgen (z. B. über eine gemeinsame Verbindungs-Fabric).

[0304] In einer Ausführungsform ist jeder Mehrkern-Prozessor 2605-2606 kommunikativ mit einem Prozessorspeicher 2601-2602 über Speicherverbindungen 2626-2627 verbunden, und jeder Grafikprozessor 2610-2613 ist kommunikativ mit dem Grafikprozessorspeicher 2620-2623 über Grafikprozessorspeicherverbindungen 2650-2653 verbunden. Die Speicherverbindungen 2626-2627 und 2650-2653 können gleiche oder unterschiedliche Speicherzugriffstechnologien einsetzen. Beispielsweise können die Prozessorspeicher 2601-2602 und die GPU-Speicher 2620-2623 flüchtige Speicher wie dynamische Direktzugriffsspeicher (DRAMs) (einschließlich gestapelter DRAMs), Grafik-DDR-SDRAM (GDDR) (z. B. GDDR5, GDDR6) oder High Bandwidth Memory (HBM) aufweisen und/oder nichtflüchtige Speicher wie 3D XPoint oder Nano-Ram sein. In einer Ausführungsform kann ein Abschnitt der Prozessorspeicher 2601-2602 ein flüchtiger Speicher und ein anderer Abschnitt ein nichtflüchtiger Speicher sein (z. B. unter Verwendung einer zweistufigen Speicherhierarchie (2LM)).

[0305] Wie es hier beschrieben ist, können zwar verschiedene Prozessoren 2605-2606 und GPUs 2610-2613 physisch mit einem bestimmten Speicher 2601-2602 bzw. 2620-2623 verbunden sein, doch kann eine einheitliche Speicherarchitektur implementiert sein, bei der ein und derselbe virtuelle Systemadressraum (auch als „effektiver Adressraum“ bezeichnet) auf verschiedene physische Speicher verteilt ist. Beispielsweise können die Prozessorspeicher 2601-2602 jeweils 64 GB des Systemspeicheradressbereichs umfassen und können die GPU-Speicher 2620-2623 jeweils 32 GB des Systemspeicheradressbereichs umfassen (was insgesamt bei diesem Beispiel zu 256 GB adressierbarem Speicher ergibt).

[0306] **Fig. 26B** veranschaulicht zusätzliche Details für eine Verbindung zwischen einem Mehrkernprozessor 2607 und einem Grafikbeschleunigungsmodul 2646 gemäß einer beispielhaften Ausführungsform. Das Grafikbeschleunigungsmodul 2646 kann einen oder mehrere GPU-Chips aufweisen, die auf einer Linecard integriert sind, die über eine Hochgeschwindigkeitsverbindung 2640 mit dem Prozessor 2607 verbunden ist. Alternativ kann das Grafikbeschleunigungsmodul 2646 auf einem gleichen Package oder Chip wie der Prozessor 2607 integriert sein.

[0307] Bei mindestens einer Ausführungsform weist der veranschaulichte Prozessor 2607 eine Vielzahl von Kernen 2660A-2660D jeweils mit einem Translations-Lookaside-Puffer 2661A-2661D und einem oder mehreren Caches 2662A-2662D auf. Bei mindestens einer Ausführungsform können die Kerne 2660A-2660D verschiedene andere Komponenten zur Ausführung von Befehlen und Verarbeitung von Daten aufweisen, die nicht dargestellt sind. Die Caches 2662A-2662D können Level-1- (LI) und Level-2- (L2) Caches umfassen. Zusätzlich können ein oder mehrere gemeinsam genutzte Caches 2656 in den Caches 2662A-2662D beinhaltet sein und von den Sätzen von Kernen 2660A-2660D gemeinsam genutzt werden. Eine Ausführungsform des Prozessors 2607 weist beispielsweise 24 Kerne auf, jeder mit seinem eigenen L1-Cache, zwölf gemeinsam genutzten L2-Caches und zwölf gemeinsam genutzten L3-Caches. In dieser Ausführungsform werden ein oder mehrere L2 und L3 Caches von zwei benachbarten Kernen gemeinsam genutzt. Der Prozessor 2607 und das Grafikbeschleunigungsmodul 2646 sind mit dem Systemspeicher 2614 verbunden, der die Prozessorspeicher 2601-2602 der **Fig. 26A** beinhalten kann.

[0308] Die Kohärenz von Daten und Anweisungen, die in verschiedenen Caches 2662A-2662D, 2656 und in dem Systemspeicher 2614 gespeichert sind, wird durch Kommunikation zwischen den Kernen über einen Kohärenzbus 2664 aufrechterhalten. Beispielsweise kann jeder Cache über eine Cache-Kohärenzlogik/-schaltung verfügen, die mit ihm verbunden ist, um als Reaktion auf erkannte Lese- oder Schreibvorgänge in bestimmten Cache-Zeilen über den Kohärenzbus 2664 zu kommunizieren. In einer Implementierung wird ein Cache-Snooping-Protokoll über den Kohärenzbus 2664 implementiert, um Cache-Zugriffe mitzulesen.

[0309] In einer Ausführungsform koppelt eine Proxy-Schaltung 2625 das Grafikbeschleunigungsmodul 2646 kommunikativ an den Kohärenzbus 2664, so dass das Grafikbeschleunigungsmodul 2646 an einem Cache-Kohärenzprotokoll als Peer der Kerne 2660A-2660D teilnehmen kann. Insbesondere sorgt eine Schnittstelle 2635 für die Konnektivität mit der Proxy-Schaltung 2625 über die Hochgeschwindigkeitsverbindung 2640 (z. B. ein PCIe-Bus, NVLink usw.), und eine Schnittstelle 2637 verbindet das Grafikbeschleunigungsmodul 2646 mit der Verbindung 2640.

[0310] In einer Implementierung bietet eine Beschleuniger-Integrationsschaltung 2636 Cache-Verwaltungs-, Speicherzugriffs-, Kontextverwaltungs- und Unterbrechungs-Verwaltungsdienste im Auftrag einer Vielzahl von Grafikverarbeitungsmaschinen 2631, 2632, N des Grafikbeschleunigungsmoduls 2646. Die Grafikverarbeitungsmaschinen 2631, 2632, N können jeweils eine separate Grafikverarbeitungseinheit (GPU) umfassen. Alternativ können die Grafikverarbeitungsmaschinen 2631, 2632, N verschiedene Arten von Grafikverarbeitungsmaschinen innerhalb eines Grafikprozessors umfassen, wie z. B. Grafikausführungseinheiten, Medienverarbeitungsmaschinen (z. B. Video-Encoder/Decoder), Sampler und Blit-Module. Bei mindestens einer Ausführungsform kann das Grafikbeschleunigungsmodul 2646 eine GPU mit einer Vielzahl von Grafikverarbeitungseinheiten 2631-2632, N sein, oder die Grafikverarbeitungseinheiten 2631-2632, N können einzelne GPUs sein, die in einem gemeinsamen Gehäuse, einer Linecard oder einem Chip integriert sind.

[0311] In einer Ausführungsform weist die Beschleuniger-Integrationsschaltung 2636 eine Speicherverwaltungseinheit (MMU) 2639 auf, um verschiedene Speicherverwaltungsfunktionen wie Übersetzungen von virtuellem zu physischem Speicher (auch als Übersetzungen von effektivem zu realem Speicher bezeichnet) und Speicherzugriffsprotokolle für den Zugriff auf den Systemspeicher 2614 durchzuführen. Die MMU 2639 kann auch einen Translations-Lookaside-Buffer (TLB) (nicht gezeigt) aufweisen, um Übersetzungen von virtuellen/effektiven in physische/reale Adressen zwischenzuspeichern. Bei einer Ausführungsform speichert ein Cache 2638 Befehle und Daten für den effizienten Zugriff durch die Grafikprozessoren 2631-2632, N. Bei einer Ausführungsform werden die in dem Cache 2638 und in den Grafikspeichern 2633-2634, M gespeicherten Daten mit den Kern-Caches 2662A-2662D, 2656 und dem Systemspeicher 2614 kohärent gehalten. Wie erwähnt, kann dies über eine Proxy-Schaltung 2625 im Namen des Cache 2638 und der Speicher 2633-2634, M erfolgen (z. B Senden von Aktualisierungen an den Cache 2638 im Zusammenhang mit Änderungen/Zugriffen auf Cache-Zeilen in den Prozessor-Caches 2662A-2662D, 2656 und Empfangen von Aktualisierungen von dem Cache 2638).

[0312] Ein Satz von Registern 2645 speichert Kontextdaten für Threads, die von Grafikverarbeitungsmaschinen 2631-2632, N ausgeführt werden, und eine Kontextverwaltungsschaltung 2648 verwaltet Thread-Kontexte. Beispielsweise kann die Kontextverwaltungsschaltung 2648 Speicher- und Wiederherstellungsoperationen ausführen, um Kontexte verschiedener Threads während Kontextumschaltungen zu speichern und wiederherzustellen (z. B. wenn ein erster Thread gesichert und ein zweiter Thread gespeichert wird, damit ein zweiter Thread von einer Grafikverarbeitungsmaschine ausgeführt werden kann). Bei einer Kontextumschaltung kann die Kontextverwaltungsschaltung 2648 beispielsweise aktuelle Registerwerte in einem bestimmten Bereich im Speicher speichern (z. B. durch einen Kontextzeiger identifiziert). Die Registerwerte können dann bei der Rückkehr zu einem Kontext wiederhergestellt werden. In einer Ausführungsform empfängt und verarbeitet eine Unterbrechungsverwaltungsschaltung 2647 Unterbrechungen, die von Systemeinrichtungen empfangen werden.

[0313] Bei einer Umsetzung werden virtuelle/effektive Adressen von einer Grafikverarbeitungsmaschine 2631 von der MMU 2639 in echte/physische Adressen in dem Systemspeicher 2614 übersetzt. Eine Ausführungsform der Beschleuniger-Integrationsschaltung 2636 unterstützt mehrere (z. B. 4, 8, 16) Grafikbeschleunigermodule 2646 und/oder andere Beschleunigereinrichtungen. Das Grafikbeschleunigermodul 2646 kann für eine einzelne Anwendung bestimmt sein, die auf dem Prozessor 2607 ausgeführt wird, oder es kann von mehreren Anwendungen gemeinsam genutzt werden. In einer Ausführungsform wird eine virtualisierte Grafikausführungsumgebung vorgestellt, in der die Ressourcen der Grafikprozessoren 2631-2632, N von mehreren Anwendungen oder virtuellen Maschinen (VMs) gemeinsam genutzt werden. Bei mindestens einer Ausführungsform können die Ressourcen in „Slices“ unterteilt sein, die verschiedenen VMs und/oder Anwendungen auf der Grundlage von Verarbeitungsanforderungen und Prioritäten, die mit VMs und/oder Anwendungen verbunden sind, zugewiesen werden.

[0314] Bei mindestens einer Ausführungsform fungiert eine Beschleuniger-Integrationsschaltung 2636 als Brücke zu einem System für das Grafikbeschleunigungsmodul 2646 und bietet Adressübersetzung und Systemspeicher-Cache-Dienste. Darüber hinaus kann die Beschleuniger-Integrationsschaltung 2636 Virtualisierungsfunktionen für einen Host-Prozessor bereitstellen, um die Virtualisierung der Grafikverarbeitungsmodule 2631-2632, Interrupts und die Speicherverwaltung zu verwalten.

[0315] Da die Hardwareressourcen der Grafikprozessoren 2631-2632, N explizit auf einen realen Adressraum abgebildet werden, den der Host-Prozessor 2607 sieht, kann jeder Host-Prozessor diese Ressourcen direkt mit einem effektiven Adresswert adressieren. Eine Funktion der Beschleuniger-Integrationsschaltung 2636 ist in einer Ausführungsform die physische Trennung der Grafikverarbeitungsmaschinen 2631-2632, N, so dass sie für ein System als unabhängige Einheiten erscheinen.

[0316] Bei mindestens einer Ausführungsform sind ein oder mehrere Grafikspeicher 2633-2634, M mit jeder der Grafikverarbeitungsmaschinen 2631-2632, N verbunden. Die Grafikspeicher 2633-2634, M speichern Anweisungen und Daten, die von jeder der Grafikverarbeitungsmaschinen 2631-2632, N verarbeitet werden. Die Grafikspeicher 2633-2634, M können flüchtige Speicher, wie DRAMs (einschließlich gestapelter DRAMs), GDDR-Speicher (z. B GDDR5, GDDR6) oder HBM, aufweisen und/oder können nichtflüchtige Speicher, wie 3D XPoint oder Nano-Ram, sein.

[0317] Bei einer Ausführungsform werden zur Verringerung des Datenverkehrs über die Verbindung 2640 Zuordnungs-Techniken verwendet, um sicherzustellen, dass die in den Grafikspeichern 2633-2634, M gespeicherten Daten Daten sind, die am häufigsten von den Grafikverarbeitungsmaschinen 2631-2632, N verwendet werden und vorzugsweise nicht von den Kernen 2660A-2660D (mindestens nicht häufig) verwendet werden. In ähnlicher Weise versucht ein Zuordnungs-Mechanismus bzw. Biasing-Mechanismus, Daten, die von Kernen (und vorzugsweise nicht von den Grafikverarbeitungsmaschinen 2631-2632, N) benötigt werden, in den Caches 2662A-2662D, 2656 der Kerne und im Systemspeicher 2614 zu halten.

[0318] **Fig. 26C** zeigt eine weitere beispielhafte Ausführungsform, bei der die Beschleuniger-Integrationsschaltung 2636 in den Prozessor 2607 integriert ist. In dieser Ausführungsform kommunizieren die Grafikprozessoren 2631-2632, N direkt über die Hochgeschwindigkeitsverbindung 2640 mit der Beschleuniger-Integrationsschaltung 2636 über die Schnittstelle 2637 und die Schnittstelle 2635 (die wiederum jede Form von Bus oder Schnittstellenprotokoll verwenden kann). Die Beschleunigerintegrationsschaltung 2636 kann dieselben Operationen wie diejenigen ausführen, die in Bezug auf **Fig. 26B** beschrieben sind, aber möglicherweise mit einem höheren Durchsatz aufgrund ihrer Nähe zu dem Kohärenzbus 2664 und den Caches 2662A-2662D, 2656. Eine Ausführungsform unterstützt verschiedene Programmiermodelle, einschließlich eines Programmiermodells für dedizierte Prozesse (ohne Virtualisierung des Grafikbeschleunigungsmoduls) und gemein-

sam genutzter Programmiermodelle (mit Virtualisierung), die Programmiermodelle aufweisen können, die von der Beschleuniger-Integrationsschaltung 2636 gesteuert werden, und Programmiermodelle, die vom Grafikbeschleunigungsmodul 2646 gesteuert werden.

[0319] Bei mindestens einer Ausführungsform sind die Grafikverarbeitungsmaschinen 2631-2632, N für eine einzige Anwendung oder einen einzigen Prozess unter einem einzigen Betriebssystem bestimmt. Bei mindestens einer Ausführungsform kann eine einzelne Anwendung andere Anwendungsanforderungen an die Grafikverarbeitungs-Engines 2631-2632, N weiterleiten, wodurch eine Virtualisierung innerhalb einer VM/Partition bereitgestellt wird.

[0320] Bei mindestens einer Ausführungsform können die Grafikverarbeitungsmaschinen 2631-2632, N, von mehreren VM-/Anwendungspartitionen gemeinsam genutzt werden. Bei mindestens einer Ausführungsform können gemeinsam genutzte Modelle einen Systemhypervisor verwenden, um die Grafikverarbeitungsmaschinen 2631-2632, N zu virtualisieren und den Zugriff durch jedes Betriebssystem zu ermöglichen. Bei Systemen mit einer einzigen Partition ohne Hypervisor gehören die Grafikprozessoren 2631-2632, N zu einem Betriebssystem. Bei mindestens einer Ausführungsform kann ein Betriebssystem die Grafikverarbeitungsmaschinen 2631-2632, N virtualisieren, um jedem Prozess oder jeder Anwendung Zugriff zu gewähren.

[0321] Bei mindestens einer Ausführungsform wählt das Grafikbeschleunigungsmodul 2646 oder eine einzelne Grafikverarbeitungsmaschine 2631-2632, N ein Prozesselement mithilfe eines Prozesshandles aus. In einer Ausführungsform werden Prozesselemente im Systemspeicher 2614 gespeichert und sind unter Verwendung einer Übersetzungstechnik von effektiver Adresse zu realer Adresse adressierbar, was hier beschrieben ist. Bei mindestens einer Ausführungsform kann ein Prozesshandle ein implementierungsspezifischer Wert sein, der einem Host-Prozess zur Verfügung gestellt wird, wenn er seinen Kontext bei der Grafikverarbeitungsmaschine 2631-2632, N registriert (d. h. wenn er die Systemsoftware aufruft, um ein Prozesselement zu einer verknüpften Prozesselementliste hinzuzufügen). Bei mindestens einer Ausführungsform können die unteren 16 Bits eines Prozesshandles ein Offset des Prozesselements innerhalb einer verknüpften Prozesselementliste sein.

[0322] Fig. 26D zeigt ein beispielhaftes Beschleuniger-Integrations-Slice 2690. Wie hierin verwendet, umfasst ein „Slice“ einen spezifizierten Abschnitt der Verarbeitungsressourcen der Beschleuniger-Integrationsschaltung 2636. Der effektive Anwendungsadressraum 2682 im Systemspeicher 2614 speichert Prozesselemente 2683. In einer Ausführungsform werden die Prozesselemente 2683 als Reaktion auf GPU-Aufrufe 2681 von Anwendungen 2680, die auf dem Prozessor 2607 ausgeführt werden, gespeichert. Ein Prozesselement 2683 enthält den Prozessstatus für die entsprechende Anwendung 2680. Ein Arbeitsdeskriptor (Work Descriptor (WD)) 2684, der in dem Prozesselement 2683 enthalten ist, kann ein einziger Auftrag sein, der von einer Anwendung angeforderter wird oder kann einen Zeiger auf eine Warteschlange von Aufträgen enthalten. Bei mindestens einer Ausführungsform ist der WD 2684 ein Zeiger auf eine Auftragsanforderungs-Warteschlange im Adressraum 2682 einer Anwendung.

[0323] Das Grafikbeschleunigungsmodul 2646 und/oder die einzelnen Grafikverarbeitungsmaschinen 2631-2632, N können von allen oder von einer Teilmenge der Prozesse in einem System gemeinsam genutzt werden. Bei mindestens einer Ausführungsform kann eine Infrastruktur zum Einrichten des Prozessstatus und zum Senden eines WD 2684 an ein Grafikbeschleunigungsmodul 2646 zum Starten eines Auftrags in einer virtualisierten Umgebung vorhanden sein.

[0324] Bei mindestens einer Ausführungsform ist ein Programmiermodell für dedizierte Prozesse implementierungsspezifisch. In diesem Modell besitzt ein einzelner Prozess das Grafikbeschleunigungsmodul 2646 oder eine einzelne Grafikverarbeitungsmaschine 2631. Da das Grafikbeschleunigungsmodul 2646 einem einzelnen Prozess gehört, initialisiert ein Hypervisor die Beschleuniger-Integrationsschaltung 2636 für eine besitzende Partition, und ein Betriebssystem initialisiert die Beschleuniger-Integrationsschaltung 2636 für einen besitzenden Prozess, wenn das Grafikbeschleunigungsmodul 2646 zugewiesen wird.

[0325] Im Betrieb holt eine WD-Abrufeinheit 2691 in dem Beschleuniger-Integrations-Slice 2690 den nächsten WD 2684 ab, der eine Angabe der Arbeit aufweist, die von einer oder mehreren Grafikverarbeitungsmaschinen des Grafikbeschleunigungsmoduls 2646 zu erledigen ist. Die Daten aus dem WD 2684 können, wie veranschaulicht, in Registern 2645 gespeichert und von der MMU 2639, der Unterbrechungsverwaltungsschaltung 2647 und/oder der Kontextverwaltungsschaltung 2648 verwendet werden. Eine Ausführungsform der MMU 2639 weist beispielsweise eine Segment-/Seitenlaufschaltung für den Zugriff auf Segment-/Seitentabellen 2686 im virtuellen Adressraum 2685 des Betriebssystems auf. Die Unterbrechungsverwaltungs-

schaltung 2647 kann vom Grafikbeschleunigungsmodul 2646 empfangene Unterbrechungsereignisse 2692 verarbeiten. Bei der Ausführung von Grafikoperationen wird eine effektive Adresse 2693, die von einer Grafikverarbeitungsmaschine 2631-2632, N erzeugt wird, von der MMU 2639 in eine echte Adresse übersetzt.

[0326] In einer Ausführungsform wird für jede Grafikverarbeitungsmaschine 2631-2632, N und/oder jedes Grafikbeschleunigungsmodul 2646 ein und derselbe Satz von Registern 2645 dupliziert und kann von einem Hypervisor oder Betriebssystem initialisiert werden. Jedes dieser duplizierten Register kann in einem Beschleuniger-Integrations-Slice 2690 vorhanden sein. Beispielhafte Register, die von einem Hypervisor initialisiert werden können, sind in Tabelle 1 aufgeführt.

Tabelle 1 - Von dem Hypervisor initialisierte Register

1	Slice-Steuerungsregister
2	Reale Adresse (RA) Bereichszeiger geplanter Prozesse
3	Autoritätsmasken-Überschreibungsregister
4	Unterbrechungsvektor-Tabelleneintrags-Offset
5	Unterbrechungsvektor-Tabelleneintragsgrenze
6	Statusregister
7	Logische Partitions-ID
8	Reale Adresse (RA) Hypervisor-Beschleuniger-Nutzungsdatensatzzeiger
9	Speicherbeschreibungsregister

[0327] Beispielhafte Register, die von einem Betriebssystem initialisiert werden können, sind in Tabelle 2 gezeigt.

Tabelle 2 - Vom Betriebssystem initialisierte Register

1	Prozess- und Thread-Identifikation
2	Effektive Adresse (EA) Kontext-Speicher/Wiederherstellungs-Zeiger
3	Virtuelle Adresse (VA) Beschleuniger-Nutzungsdatensatzzeiger
4	Virtuelle Adresse (VA) Zeiger auf die Speichersegmenttabelle
5	Autoritätsmaske
6	Arbeitsdeskriptor

[0328] In einer Ausführungsform ist jeder WD 2684 spezifisch für ein bestimmtes Grafikbeschleunigungsmodul 2646 und/oder die Grafikverarbeitungsmaschinen 2631-2632, N. Er enthält alle Informationen, die von einer Grafikverarbeitungsmaschine 2631-2632, N benötigt werden, um Arbeit zu verrichten, oder er kann ein Zeiger auf einen Speicherplatz sein, an dem eine Anwendung eine Befehlswarteschlange von zu verrichtender Arbeit eingerichtet hat.

[0329] **Fig. 26E** veranschaulicht zusätzliche Details für eine beispielhafte Ausführungsform eines gemeinsamen Modells. Diese Ausführungsform weist einen realen Hypervisor-Adressraum 2698 auf, in dem eine Prozesselementliste 2699 gespeichert ist. Auf den realen Hypervisor-Adressraum 2698 kann über einen Hypervisor 2696, der Grafikbeschleunigungsmodul-Maschinen für das Betriebssystem 2695 virtualisiert, zugegriffen werden.

[0330] Bei mindestens einer Ausführungsform erlauben gemeinsame Programmiermodelle allen oder einer Teilmenge von Prozessen aus allen oder einer Teilmenge von Partitionen in einem System, ein Grafikbeschleunigungsmodul 2646 zu verwenden. Es gibt zwei Programmiermodelle, bei denen das Grafikbeschleunigungsmodul 2646 von mehreren Prozessen und Partitionen gemeinsam genutzt wird: zeitlich geteilte und grafisch gerichtete gemeinsame Nutzung.

[0331] Bei diesem Modell ist der System-Hypervisor 2696 Besitzer des Grafikbeschleunigungsmoduls 2646 und stellt seine Funktion allen Betriebssystemen 2695 zur Verfügung. Damit ein Grafikbeschleunigungsmodul

2646 die Virtualisierung durch den System-Hypervisor 2696 unterstützen kann, kann das Grafikbeschleunigungsmodul 2646 folgende Bedingungen erfüllen: 1) Eine Auftragsanforderung einer Anwendung muss autonom sein (d. h., der Zustand braucht zwischen den Aufträgen nicht aufrechterhalten zu werden), oder das Grafikbeschleunigungsmodul 2646 muss einen Mechanismus zur Kontextsicherung und -wiederherstellung bereitstellen. 2) Das Grafikbeschleunigungsmodul 2646 garantiert, dass die Auftragsanforderung einer Anwendung in einer bestimmten Zeitspanne abgeschlossen wird, einschließlich etwaiger Übersetzungsfehler, oder das Grafikbeschleunigungsmodul 2646 bietet die Möglichkeit, die Verarbeitung eines Auftrags zu unterbrechen. 3) Dem Grafikbeschleunigungsmodul 2646 muss Fairness zwischen den Prozessen garantiert werden, wenn es in einem gerichteten gemeinsamen Programmiermodell arbeitet.

[0332] Bei mindestens einer Ausführungsform muss die Anwendung 2680 einen Systemaufruf des Betriebssystems 2695 mit einem Grafikbeschleunigungsmodul 2646-Typ, einem Arbeitsdeskriptor (WD), einem AMR-Wert (Authority Mask Register) und einem CSRP-Zeiger (Context Save/Restore Area Pointer) ausführen. Bei mindestens einer Ausführungsform beschreibt der Typ des Grafikbeschleunigungsmoduls 2646 eine gezielte Beschleunigungsfunktion für einen Systemaufruf. Bei mindestens einer Ausführungsform kann der Typ des Grafikbeschleunigungsmoduls 2646 ein systemspezifischer Wert sein. Bei mindestens einer Ausführungsform ist der WD speziell für das Grafikbeschleunigungsmodul 2646 formatiert und kann in Form eines Grafikbeschleunigungsmodul 2646-Befehls, eines effektiven Adresszeigers auf eine benutzerdefinierte Struktur, eines effektiven Adresszeigers auf eine Befehlswarteschlange oder einer anderen Datenstruktur vorliegen, die die vom Grafikbeschleunigungsmodul 2646 zu verrichtende Arbeit beschreibt. In einer Ausführungsform ist ein AMR-Wert ein AMR-Zustand, der für einen aktuellen Prozess zu verwenden ist. Bei mindestens einer Ausführungsform ähnelt ein an ein Betriebssystem übergebener Wert der Einstellung eines AMR durch eine Anwendung. Wenn die Implementierungen der Beschleuniger-Integrationsschaltung 2636 und des Grafikbeschleunigungsmoduls 2646 kein Benutzer-Autoritätsmasken-Überschreibungsregister (User Authority Mask Override Register (UAMOR)) unterstützen, kann ein Betriebssystem einen aktuellen UAMOR-Wert auf einen AMR-Wert anwenden, bevor ein AMR in einem Hypervisor-Aufruf übergeben wird. Der Hypervisor 2696 kann optional einen aktuellen AMOR-Wert (Authority Mask Override Register) anwenden, bevor ein AMR in einem Prozesselement 2683 angeordnet wird. Bei mindestens einer Ausführungsform ist CSRP eines der Register 2645, die eine effektive Adresse eines Bereichs im Adressraum 2682 einer Anwendung für das Grafikbeschleunigungsmodul 2646 zur Speicherung und Wiederherstellung des Kontextstatus enthalten. Dieser Zeiger ist optional, wenn kein Zustand zwischen Aufträgen gespeichert werden muss oder wenn ein Auftrag vorzeitig beendet wird. Bei mindestens einer Ausführungsform kann der Kontextspeicher-/Wiederherstellungsbereich im Systemspeicher verankert sein.

[0333] Beim Empfang eines Systemaufrufs kann das Betriebssystem 2695 überprüfen, ob die Anwendung 2680 registriert ist und die Berechtigung zur Verwendung des Grafikbeschleunigungsmoduls 2646 erhalten hat. Das Betriebssystem 2695 ruft dann den Hypervisor 2696 mit den in Tabelle 3 dargestellten Informationen auf.

Tabelle 3 - Hypervisor-Aufrufparameter von dem Betriebssystem

1	Ein Arbeitsdeskriptor (WD)
2	Ein Autoritätsmaskenregister- (AMR)-Wert (möglicherweise maskiert)
3	Eine effektive Adresse (EA) Kontext-Sicherungs-/Wiederherstellungs-Bereichszeigers (CSRP)
4	Eine Prozess-ID (PID) und optional eine Thread-ID (TID)
5	Eine virtuelle Adresse (VA) Beschleuniger-Nutzungsdatensatzzeiger (AURP)
6	Virtuelle Adresse eines Speichersegmenttabellenzeigers (SSTP)
7	Eine logische Interrupt-Service-Nummer (LISN)

[0334] Beim Empfang eines Hypervisor-Aufrufs überprüft Hypervisor 2696, ob das Betriebssystem 2695 registriert ist und die Berechtigung zur Verwendung des Grafikbeschleunigungsmoduls 2646 erhalten hat. Der Hypervisor 2696 setzt dann das Prozesselement 2683 in eine verknüpfte Prozesselementliste für einen entsprechenden Grafikbeschleunigungsmodultyp 2646. Ein Prozesselement kann die in Tabelle 4 dargestellten Informationen aufweisen.

Tabelle 4 - Prozesselementinformationen

1	Ein Arbeitsdeskriptor (WD)

2	Ein Autoritätsmaskenregister- (AMR)-Wert (möglicherweise maskiert)
3	Eine effektive Adresse (EA) Kontext-Sicherungs-/Wiederherstellungs-Bereichszeigers (CSRP)
4	Eine Prozess-ID (PID) und optional eine Thread-ID (TID)
5	Eine virtuelle Adresse (VA) Beschleuniger-Nutzungsdatensatzzeiger (AURP)
6	Virtuelle Adresse eines Speichersegmenttabellenzeigers (SSTP)
7	Eine logische Interrupt-Service-Nummer (LISN)
8	Unterbrechungsvektortabelle, abgeleitet von Hypervisor-Aufrufparametern
9	Ein Statusregister- (SR-) Wert
10	Eine logische Partitions-ID (LPID)
11	Reale Adresse (RA) Hypervisor-Beschleuniger-Nutzungsdatensatzzeiger
12	Speicherbeschreibungsregister (SDR)

[0335] Bei mindestens einer Ausführungsform initialisiert der Hypervisor eine Vielzahl von Registern 2645 für Beschleuniger-Integrations-Slices 2690.

[0336] Wie es in **Fig. 26F** dargestellt ist, wird bei mindestens einer Ausführungsform ein einheitlicher Speicher verwendet, der über einen gemeinsamen virtuellen Speicheradressraum adressierbar ist, der für den Zugriff auf physische Prozessorspeicher 2601-2602 und GPU-Speicher 2620-2623 verwendet wird. Bei dieser Implementierung verwenden die auf den GPUs 2610-2613 ausgeführten Operationen denselben virtuellen/effektiven Speicheradressraum für den Zugriff auf die Prozessorspeicher 2601-2602 und umgekehrt, was die Programmierbarkeit vereinfacht. Bei einer Ausführungsform wird ein erster Abschnitt eines virtuellen/effektiven Adressraums dem Prozessorspeicher 2601 zugeordnet, ein zweiter Abschnitt dem zweiten Prozessorspeicher 2602, ein dritter Abschnitt dem GPU-Speicher 2620 usw. Bei mindestens einer Ausführungsform wird dadurch ein gesamter virtueller/effektiver Speicherraum (manchmal auch als effektiver Adressraum bezeichnet) über jeden der Prozessorspeicher 2601-2602 und GPU-Speicher 2620-2623 verteilt, wodurch jeder Prozessor oder jede GPU auf jeden physischen Speicher mit einer diesem Speicher zugeordneten virtuellen Adresse zugreifen kann.

[0337] In einer Ausführungsform stellt die Bias/Kohärenz-Management-Schaltung 2694A-2694E innerhalb einer oder mehrerer MMUs 2639A-2639E die Cache-Kohärenz zwischen den Caches eines oder mehrerer Host-Prozessoren (z. B. 2605) und GPUs 2610-2613 sicher und implementiert Biasing-Verfahren, die angeben, in welchen physischen Speichern bestimmte Datentypen gespeichert werden sollen. Während mehrere Instanzen der Bias/Kohärenz-Management-Schaltung 2694A-2694E in **Fig. 26F** dargestellt sind, kann die Bias/Kohärenz-Schaltung innerhalb einer MMU eines oder mehrerer Host-Prozessoren 2605 und/oder innerhalb der Beschleuniger-Integrationsschaltung 2636 implementiert sein.

[0338] Eine Ausführungsform ermöglicht es, dass GPU-angeschlossener Speicher 2620-2623 als Teil des Systemspeichers abgebildet ist und dass auf ihn unter Verwendung der SVM-Technologie (Shared Virtual Memory) zugegriffen wird, ohne jedoch Leistungsnachteile zu erleiden, die mit einer vollständigen System-Cache-Kohärenz verbunden sind. Bei mindestens einer Ausführungsform bietet die Möglichkeit des Zugriffs auf GPU-angeschlossenen Speicher 2620-2623 als Systemspeicher ohne lästigen Cache-Kohärenz-Overhead eine vorteilhafte Betriebsumgebung für GPU-Offload. Diese Anordnung ermöglicht es der Software des Host-Prozessors 2605, Operanden einzustellen und auf Berechnungsergebnisse zuzugreifen, ohne den Overhead herkömmlicher E/A-DMA-Datenkopien. Solche herkömmlichen Kopien beinhalten Treiberaufrufe, Unterbrechungen und speicherabbildende E/A- (MMIO-) Zugriffe, die alle im Vergleich zu einfachen Speicherzugriffen ineffizient sind. Bei mindestens einer Ausführungsform kann die Fähigkeit, ohne Cache-Kohärenz-Overheads auf den GPU-verbundenen Speicher 2620-2623 zuzugreifen, für die Ausführungszeit einer ausgelagerten Berechnung entscheidend sein. In Fällen mit erheblichem Streaming-Schreibspeicherverkehr kann der Cache-Kohärenz-Overhead beispielsweise die effektive Schreibbandbreite einer GPU 2610-2613 erheblich reduzieren. Bei mindestens einer Ausführungsform können die Effizienz des Operanden-Setups, die Effizienz des Ergebniszugriffs und die Effizienz der GPU-Berechnung eine Rolle bei der Bestimmung der Effektivität eines GPU-Offloads spielen.

[0339] Bei mindestens einer Ausführungsform wird die Auswahl eines GPU-Bias und eines Host-Prozessor-Bias durch eine Bias-Tracker-Datenstruktur gesteuert. Es kann z. B eine Bias-Tabelle verwendet werden, die eine seitengranulare Struktur sein kann (d.h. mit der Granularität einer Speicherseite gesteuert), die 1 oder 2 Bits pro GPU-angeschlossener Speicherseite beinhaltet. Bei mindestens einer Ausführungsform kann eine Bias-Tabelle in einem gestohlenen Speicherbereich eines oder mehrerer GPU-angeschlossener Speicher 2620-2623 implementiert sein, mit oder ohne Bias-Cache in GPU 2610-2613 (z. B. um häufig/kürzlich verwendete Einträge einer Bias-Tabelle zu cachen). Alternativ dazu kann eine gesamte Bias-Tabelle in einer GPU verwaltet werden.

[0340] Bei mindestens einer Ausführungsform wird vor dem tatsächlichen Zugriff auf einen GPU-Speicher auf einen Bias-Tabelleneintrag zugegriffen, der mit jedem Zugriff auf den GPU-angeschlossenen Speicher 2620-2623 assoziiert ist, was die folgenden Operationen bewirkt. Zunächst werden lokale Anfragen von GPU 2610-2613, die ihre Seite im GPU-Bias finden, direkt an einen entsprechenden GPU-Speicher 2620-2623 weitergeleitet. Lokale Anforderungen von einer GPU, die ihre Seite in dem Host-Bias finden, werden an den Prozessor 2605 weitergeleitet (z. B., wie oben erörtert, über eine Hochgeschwindigkeitsverbindung). In einer Ausführungsform werden Anfragen vom Prozessor 2605, die eine angeforderte Seite im Host-Prozessor-Bias finden, wie ein normaler Speicherlesezugriff abgeschlossen. Alternativ können Anforderungen, die an eine GPU-biased bzw. GPU-gebundene Seite gerichtet sind, an die GPU 2610-2613 weitergeleitet werden. Bei mindestens einer Ausführungsform kann eine GPU dann eine Seite in einen Host-Prozessor-Bias überführen, wenn sie die Seite gerade nicht verwendet. Bei mindestens einer Ausführungsform kann ein Bias-Zustand einer Seite entweder durch einen softwarebasierten Mechanismus, einen hardwareunterstützten softwarebasierten Mechanismus oder für eine begrenzte Anzahl von Fällen durch einen rein hardwarebasierten Mechanismus geändert werden.

[0341] Ein Mechanismus zum Ändern des Bias-Zustands verwendet einen API-Aufruf (z. B. OpenCL), der wiederum den Vorrichtungstreiber einer GPU aufruft, der wiederum eine Mitteilung an eine GPU sendet (oder einen Befehlsdeskriptor einreiht), die sie anweist, einen VBias-Zustand zu ändern und bei einigen Übergängen eine Cache-Flushing-Operation in einem Host auszuführen. Bei mindestens einer Ausführungsform wird die Cache-Flushing-Operation für einen Übergang von dem Bias des Host-Prozessors 2605 zum Bias der GPU verwendet, aber nicht für einen entgegengesetzten Übergang.

[0342] Bei einer Ausführungsform wird die Cache-Kohärenz aufrechterhalten, indem bewirkt wird, dass GPU-gebundene Seiten vorübergehend von dem Host-Prozessor 2605 nicht zwischengespeichert werden können. Um auf diese Seiten zuzugreifen, kann der Prozessor 2605 den Zugriff von der GPU 2610 anfordern, die den Zugriff möglicherweise nicht sofort gewährt. Um die Kommunikation zwischen dem Prozessor 2605 und der GPU 2610 zu verringern, ist es daher vorteilhaft sicherzustellen, dass GPU-gebundene Seiten diejenigen sind, die von einer GPU aber nicht von dem Hostprozessor 2605 benötigt werden und umgekehrt.

[0343] Hardware-Struktur(en) 1815 werden verwendet, um eine oder mehrere Ausführungsformen auszuführen. Details bezüglich der einen oder den mehreren Hardware-Strukturen 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt.

[0344] **Fig. 27** zeigt beispielhafte integrierte Schaltungen und zugehörige Grafikprozessoren, die unter Verwendung eines oder mehrerer IP-Kerne gemäß verschiedenen hier beschriebenen Ausführungsformen hergestellt werden können. Zusätzlich zu den dargestellten Schaltungen können bei mindestens einer Ausführungsform weitere Logik und Schaltungen vorhanden sein, einschließlich zusätzlicher Grafikprozessoren/-kerne, Steuervorrichtungen für periphere Schnittstellen oder Allzweck-Prozessorkerne.

[0345] **Fig. 27** ist ein Blockdiagramm, das eine beispielhafte integrierte Schaltung 2700 mit einem System auf einem Chip veranschaulicht, die unter Verwendung eines oder mehrerer IP-Kerne gemäß mindestens einer Ausführungsform hergestellt werden kann. Bei mindestens einer Ausführungsform weist die integrierte Schaltung 2700 einen oder mehrere Anwendungsprozessor(en) 2705 (z. B. CPUs), mindestens einen Grafikprozessor 2710 auf und kann zusätzlich einen Bildprozessor 2715 und/oder einen Videoprozessor 2720 aufweisen, von denen jeder ein modularer IP-Kern sein kann. Bei mindestens einer Ausführungsform weist die integrierte Schaltung 2700 eine Peripherie- oder Buslogik auf, darunter eine USB-Steuerung 2725, eine UART-Steuerung 2730, eine SPI/SDIO-Steuerung 2735 und eine I.sup.2S/I.sup.2C-Steuerung 2740. Bei mindestens einer Ausführungsform kann die integrierte Schaltung 2700 eine Anzeigeeinrichtung 2745 aufweisen, die mit einer oder mehreren HDMI- (High-Definition Multimedia Interface-) Steuerungen 2750 und einer MIPI- (Mobile Industry Processor Interface-) Anzeigenschnittstelle 2755 verbunden ist. Bei mindestens einer Ausführungsform kann der Speicher durch ein Flash-Speicher-Subsystem 2760 bereitgestellt sein, das einen

Flash-Speicher und eine Flash-SpeicherSteuerung aufweist. Bei mindestens einer Ausführungsform kann die Speicherschnittstelle über eine Speichersteuerung 2765 für den Zugriff auf SDRAM- oder SRAM-Speichereinrichtungen bereitgestellt sein. Bei mindestens einer Ausführungsform weisen einige integrierte Schaltungen zusätzlich eine eingebettete Sicherheits-Maschine 2770 auf.

[0346] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform kann eine Ableitungs- und/oder Trainingslogik 1815 in der integrierten Schaltung 2700 zum Ableiten oder Vorhersagen von Operationen mindestens zum Teil basierend auf Gewichtungsparametern verwendet werden, die unter Verwendung von neuronalen Netzwerktrainingsoperationen, Funktionen und/oder Architekturen von neuronalen Netzwerken oder in dieser Schrift beschriebenen Anwendungsfällen von neuronalen Netzwerken berechnet wurden.

[0347] Bei mindestens einer Ausführungsform wird die integrierte Schaltung 2700 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird die integrierte Schaltung 2700 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird die integrierte Schaltung 2700 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und die integrierte Schaltung 2700 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0348] Die **Fig. 28A** und **Fig. 28B** veranschaulichen beispielhafte integrierte Schaltungen und assoziierte Grafikprozessoren, die unter Verwendung eines oder mehrerer IP-Kerne gemäß verschiedenen Ausführungsformen, wie sie hier beschrieben sind, gefertigt werden können. Zusätzlich zu den dargestellten Schaltungen können bei mindestens einer Ausführungsform weitere Logik und Schaltungen vorhanden sein, einschließlich zusätzlicher Grafikprozessoren/-kerne, Steuervorrichtungen für periphere Schnittstellen oder Allzweck-Prozessorkerne.

[0349] Die **Fig. 28A** und **Fig. 28B** sind Blockdiagramme, die beispielhafte Grafikprozessoren zur Verwendung in einem SoC gemäß den hierin beschriebenen Ausführungsformen veranschaulichen. **Fig. 28A** veranschaulicht einen beispielhaften Grafikprozessor 2810 eines Systems auf einer integrierten Chipschaltung, der unter Verwendung eines oder mehrerer IP-Kerne gefertigt werden kann, gemäß mindestens einer Ausführungsform. **Fig. 28B** veranschaulicht einen zusätzlichen beispielhaften Grafikprozessor 2840 eines Systems auf einer integrierten Chipschaltung, der unter Verwendung eines oder mehrerer IP-Kerne gefertigt werden kann, gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform ist der Grafikprozessor 2810 der **Fig. 28A** ein stromsparender Grafikprozessorkern. Bei mindestens einer Ausführungsform ist der Grafikprozessor 2840 der **Fig. 28B** ein Grafikprozessorkern mit höherer Leistung. Bei mindestens einer Ausführungsform kann jeder der Grafikprozessoren 2810, 2840 eine Variante des Grafikprozessors 2710 von **Fig. 27** sein.

[0350] Bei mindestens einer Ausführungsform beinhaltet der Grafikprozessor 2810 einen Vertex-Prozessor 2805 und einen oder mehrere Fragmentprozessoren 2815A-2815N (z. B. 2815A, 2815B, 2815C, 2815D bis 2815N-1 und 2815N). Bei mindestens einer Ausführungsform kann der Grafikprozessor 2810 derartig unterschiedliche Shader-Programme über getrennte Logik derart ausführen, dass der Vertex-Prozessor 2805 optimiert wird, um Operationen für Vertex-Shader-Programme auszuführen, während ein oder mehrere Fragmentprozessoren 2815A-2815N Shading-Operationen für Fragmente (z. B. Bildpunkt) für Fragment- oder Bildpunkt-Shader-Programme ausführen. Bei mindestens einer Ausführungsform führt der Vertex-Prozessor 2805 eine Vertex-Verarbeitungsstufe einer 3D-Grafikpipeline durch und erzeugt Primitives und Vertex-Daten. Bei mindestens einer Ausführungsform verwenden die Fragmentprozessoren 2815A-2815N die vom Vertex-Prozessor 2805 erzeugten Primitiv- und Vertex-Daten, um einen Bildpuffer zu erzeugen, der auf einer Anzeigeeinrichtung angezeigt wird. Bei mindestens einer Ausführungsform ist/sind der/die Fragmentprozessor(en) 2815A-2815N für die Ausführung von Fragment-Shader-Programmen optimiert, wie sie in einer OpenGL-API vorgesehen sind, die verwendet werden können, um ähnliche Operationen wie ein Pixel-Shader-Programm durchzuführen, wie sie in einer Direct 3D-API vorgesehen sind.

[0351] Bei mindestens einer Ausführungsform weist der Grafikprozessor 2810 zusätzlich eine oder mehrere Speicherverwaltungseinheiten (MMUs) 2820A-2820B, einen oder mehrere Cache(s) 2825A-2825B und eine oder mehrere Schaltungsverbindungen 2830A-2830B auf. Bei mindestens einer Ausführungsform sorgen eine oder mehrere MMU(s) 2820A-2820B für die Zuordnung von virtuellen zu physischen Adressen für den Grafikprozessor 2810, einschließlich für den Vertex-Prozessor 2805 und/oder den/die Fragmentprozessor(en) 2815A-2815N, der/die zusätzlich zu den in einem oder mehreren Cache(s) 2825A-2825B gespeicherten Vertex- oder Bild-/Texturdaten auf im Speicher gespeicherte Vertex- oder Bild-/Texturdaten verweisen kann/-können. Bei mindestens einer Ausführungsform kann eine oder können mehrere MMU(s) 2820A-2820B mit anderen MMUs innerhalb des Systems synchronisiert werden, einschließlich einer oder mehrerer MMUs, die mit einem oder mehreren Anwendungsprozessoren 2705, Bildprozessoren 2715 und/oder Videoprozessoren 2720 der **Fig. 27** assoziiert sind, so dass sich jeder Prozessor 2705-2720 an einem gemeinsamen genutzten oder vereinheitlichten virtuellen Speichersystem beteiligen kann. Bei mindestens einer Ausführungsform ermöglichen es eine oder mehrere Schaltungsverbindung(en) 2830A-2830B dem Grafikprozessor 2810, eine Schnittstelle mit anderen IP-Kernen innerhalb des SoC, entweder über einen internen Bus des SoC oder über eine direkte Verbindung herzustellen.

[0352] Bei mindestens einer Ausführungsform beinhaltet der Grafikprozessor 2840 eine oder mehrere MMU(s) 2820A-2820B, Caches 2825A-2825B und Schaltungs-Interconnects 2830A-2830B des Grafikprozessors 2810 der **Fig. 28A**. Bei mindestens einer Ausführungsform beinhaltet der Grafikprozessor 2840 einen oder mehrere Shader-Kern(e) 2855A-2855N (z. B. 2855A, 2855B, 2855C, 2855D, 2855E, 2855F bis 2855N-1 und 2855N), was eine einheitliche Shader-Kern-Architektur ermöglicht, bei der ein einziger Kern oder Typ oder Kern alle Arten von programmierbarem Shader-Code ausführen kann, einschließlich Shader-Programmcode zur Implementierung von Vertex-Shadern, Fragment-Shadern und/oder Compute-Shadern. Bei mindestens einer Ausführungsform kann die Anzahl der Shader-Kerne variieren. Bei mindestens einer Ausführungsform weist der Grafikprozessor 2840 einen Inter-Core-Task-Manager 2845 auf, der als Thread-Dispatcher fungiert, um Ausführungs-Threads an einen oder mehrere Shader-Kerne 2855A-2855N und eine Tiling-Einheit 2858 zu verteilen, um Tiling-Operationen für kachelbasiertes Rendering zu beschleunigen, bei denen Rendering-Operationen für eine Szene im Bildraum unterteilt sind, um beispielsweise eine lokale räumliche Kohärenz innerhalb einer Szene auszunutzen oder die Nutzung interner Caches zu optimieren.

[0353] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform kann eine Ableitungs- und/oder Trainingslogik 1815 in der integrierten Schaltung **Fig. 28A** und/oder **Fig. 28B** zum Ableiten oder Vorhersagen von Operationen mindestens zum Teil basierend auf Gewichtungsparametern verwendet werden, die unter Verwendung von neuronalen Netzwerktrainingsoperationen, Funktionen und/oder Architekturen von neuronalen Netzwerken oder in dieser Schrift beschriebenen Anwendungsfällen von neuronalen Netzwerken berechnet wurden.

[0354] Bei mindestens einer Ausführungsform wird die integrierte Schaltung **Fig. 28A** und/oder **Fig. 28B** eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird die integrierte Schaltung **Fig. 28A** und/oder **Fig. 28B** eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird die integrierte Schaltung **Fig. 28A** und/oder **Fig. 28B** eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umzusetzen, und eine integrierte Schaltung der **Fig. 28A** und/oder **Fig. 28B** wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die ein oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0355] Die **Fig. 29A** und **Fig. 29B** veranschaulichen eine zusätzliche beispielhafte Grafikprozessorlogik gemäß den hierin beschriebenen Ausführungsformen. **Fig. 29A** veranschaulicht einen Grafikkern 2900, der in dem Grafikprozessor 2710 der **Fig. 27** bei mindestens einer Ausführungsform beinhaltet sein kann, und ein einheitlicher Shader-Kern 2855A-2855N bei mindestens einer Ausführungsform, wie in **Fig. 28B**, sein kann. **Fig. 29B** veranschaulicht eine hochparallele Allzweck-Grafikverarbeitungseinheit 2930, die bei mindestens einer Ausführungsform zum Einsatz auf einem Mehrchipmodul geeignet ist.

[0356] Bei mindestens einer Ausführungsform beinhaltet der Grafikkern 2900 einen gemeinsam genutzten Anweisungs-Cache 2902, eine Textureinheit 2918 und einen Cache/gemeinsamen Speicher 2920, die den Ausführungsressourcen innerhalb des Grafikkerns 2900 gemeinsam sind. Bei mindestens einer Ausführungsform kann der Grafikkern 2900 mehrere Slices 2901A-2901N oder Partitionen für jeden Kern aufweisen, und ein Grafikprozessor kann mehrere Instanzen des Grafikkerns 2900 aufweisen. Die Slices 2901A-2901N können eine Unterstützungslogik aufweisen, die einen lokalen Befehlscache 2904A-2904N, einen Thread-Scheduler 2906A-2906N, einen Thread-Dispatcher 2908A-2908N und einen Satz von Registern 2910A-2910N umfasst. Bei mindestens einer Ausführungsform können Slices 2901A-2901N einen Satz zusätzlicher Funktionseinheiten (AFUs 2912A-2912N), Gleitkommaeinheiten (FPU 2914A-2914N), ganzzahlige arithmetische Logikeinheiten (ALUs 2916-2916N), Adressberechnungseinheiten (ACU 2913A-2913N), doppeltgenaue Gleitkommaeinheiten (DPFPU 2915A-2915N) und Matrixverarbeitungseinheiten (MPU 2917A-2917N) beinhalten.

[0357] Bei mindestens einer Ausführungsform können die FPUs 2914A-2914N Gleitkommaoperationen mit einfacher Genauigkeit (32 Bit) und halber Genauigkeit (16 Bit) ausführen, während die DPFPUs 2915A-2915N Gleitkommaoperationen mit doppelter Genauigkeit (64 Bit) ausführen. Bei mindestens einer Ausführungsform können die ALUs 2916A-2916N ganzzahlige Operationen mit variabler Genauigkeit mit einer Genauigkeit von 8-Bit, 16-Bit und 32-Bit ausführen und können für Operationen mit gemischter Präzision konfiguriert sein. Bei mindestens einer Ausführungsform können die MPUs 2917A-2917N auch für Matrixoperationen mit gemischter Genauigkeit, die Gleitkomma- und 8-Bit-Ganzzahloperationen mit halber Genauigkeit beinhalten, konfiguriert sein. Bei mindestens einer Ausführungsform können die MPUs 2917-2917N eine Vielfalt von Matrixoperationen ausführen, um Anwendungs-Frameworks des maschinellen Lernens zu beschleunigen, die das Ermöglichen von Unterstützung für beschleunigte allgemeine Matrix-zu-Matrix-Multiplikation (general matrix to matrix multiplication-GEMM) beinhalten. Bei mindestens einer Ausführungsform können die AFUs 2912A-2912N zusätzliche logische Operationen durchführen, die von Gleitkomma- oder Ganzzahl-Einheiten nicht unterstützt werden, einschließlich trigonometrischer Operationen (z. B. Sinus, Cosinus usw.).

[0358] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform kann eine Ableitungs- und/oder Trainingslogik 1815 in dem Grafikkern 2900 zum Ableiten oder Vorhersagen von Operationen mindestens zum Teil basierend auf Gewichtungsparametern verwendet werden, die unter Verwendung von neuronalen Netzwerktrainingsoperationen, Funktionen und/oder Architekturen von neuronalen Netzwerken oder in dieser Schrift beschriebenen Anwendungsfällen von neuronalen Netzwerken berechnet wurden.

[0359] Bei mindestens einer Ausführungsform wird der Grafikkern 2900 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird der Grafikkern 2900 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird der Grafikkern 2900 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und der Grafikkern 2900 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0360] **Fig. 29B** veranschaulicht eine Allzweckverarbeitungseinheit (GPGPU) 2930, die konfiguriert werden kann, um hochparallele Rechenoperationen zu ermöglichen, die von einem Array von Grafikverarbeitungseinheiten ausgeführt werden sollen, bei mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform kann die GPGPU 2930 direkt mit anderen Instanzen der GPGPU 2930 verbunden sein, um einen Multi-GPU-Cluster zu bilden und die Trainingsgeschwindigkeit für tiefe neuronale Netze zu verbessern. Bei mindestens einer Ausführungsform weist die GPGPU 2930 eine Host-Schnittstelle 2932 auf, um eine Verbindung mit einem Host-Prozessor zu ermöglichen. Bei mindestens einer Ausführungsform handelt es sich bei der Host-Schnittstelle 2932 um eine PCI-Express-Schnittstelle. Bei mindestens einer Ausführungsform kann es sich bei der Host-Schnittstelle 2932 um eine herstellerspezifische Kommunikationsschnittstelle oder Kommunikationsstruktur handeln. Bei mindestens einer Ausführungsform empfängt die GPGPU 2930 Befehle von einem Host-Prozessor und verwendet einen globalen Scheduler 2934, um die mit diesen Befehlen verbunde-

nen Ausführungsthreads auf eine Reihe von Compute-Clustern 2936A-2936H zu verteilen. Bei mindestens einer Ausführungsform teilen sich die Compute-Cluster 2936A-2936H einen Cache-Speicher 2938. Bei mindestens einer Ausführungsform kann der Cache-Speicher 2938 als übergeordneter Cache für Cache-Speicher innerhalb von Compute-Clustern 2936A-2936H dienen.

[0361] Bei mindestens einer Ausführungsform weist die GPGPU 2930 einen Speicher 2944A-2944B auf, der über eine Reihe von Speichersteuerungen 2942A-2942B mit Compute-Clustern 2936A-2936H gekoppelt ist. Bei mindestens einer Ausführungsform kann der Speicher 2944A-2944B verschiedene Arten von Speichereinrichtungen aufweisen, einschließlich dynamischer Direktzugriffsspeicher (DRAM) oder Grafik-Direktzugriffsspeicher, wie synchroner Grafik-Direktzugriffsspeicher (SGRAM), einschließlich Grafik-Doppeldatenraten-Speicher (GDDR).

[0362] Bei mindestens einer Ausführungsform beinhalten die Rechen-Cluster 2936A-2936H jeweils einen Satz von Grafikkernen, wie den Grafikkern 2900 der **Fig. 29A**, der mehrere Arten von Ganzzahl- und Gleitkomma-Logikeinheiten beinhalten kann, die Rechenoperationen mit einer Reihe von Genauigkeiten, einschließlich für maschinelle Lernberechnungen geeignete, ausführen können. Zum Beispiel kann bei mindestens einer Ausführungsform mindestens eine Teilmenge von Gleitkommaeinheiten in jedem der Compute-Cluster 2936A-2936H so ausgestaltet sein, dass sie 16-Bit- oder 32-Bit-Gleitkommaoperationen durchführen, während eine andere Teilmenge von Gleitkommaeinheiten so ausgestaltet sein kann, dass sie 64-Bit-Gleitkommaoperationen durchführen kann.

[0363] Bei mindestens einer Ausführungsform können mehrere Instanzen der GPGPU 2930 für den Betrieb als ein Compute-Cluster ausgestaltet sein. Bei mindestens einer Ausführungsform variiert die von den Compute-Clustern 2936A-2936H für die Synchronisation und den Datenaustausch verwendete Kommunikation zwischen den Ausführungsformen. Bei mindestens einer Ausführungsform kommunizieren mehrere Instanzen der GPGPU 2930 über die Host-Schnittstelle 2932. Bei mindestens einer Ausführungsform weist die GPGPU 2930 einen E/A-Hub 2939 auf, der die GPGPU 2930 mit einem GPU-Link 2940 koppelt, der eine direkte Verbindung zu anderen Instanzen der GPGPU 2930 ermöglicht. Bei mindestens einer Ausführungsform ist die GPU-Verbindung 2940 mit einer dedizierten GPU-zu-GPU-Brücke gekoppelt, die die Kommunikation und Synchronisation zwischen mehreren Instanzen der GPGPU 2930 ermöglicht. Bei mindestens einer Ausführungsform ist der GPU-Link 2940 mit einer Hochgeschwindigkeits-Verbindung gekoppelt, um Daten an andere GPGPUs oder Parallelprozessoren zu senden und zu empfangen. Bei mindestens einer Ausführungsform befinden sich mehrere Instanzen der GPGPU 2930 in getrennten Datenverarbeitungssystemen und kommunizieren über eine Netzwerkeinrichtung, die über die Host-Schnittstelle 2932 zugänglich ist. Bei mindestens einer Ausführungsform kann die GPU-Verbindung 2940 so ausgestaltet sein, dass sie zusätzlich oder alternativ zur Hostschnittstelle 2932 eine Verbindung zu einem Hostprozessor ermöglicht.

[0364] Bei mindestens einer Ausführungsform kann die GPGPU 2930 so ausgestaltet sein, dass sie neuronale Netze trainiert. Bei mindestens einer Ausführungsform kann die GPGPU 2930 innerhalb einer Ableitungs-Plattform verwendet werden. Bei mindestens einer Ausführungsform, bei der die GPGPU 2930 für Inferencing verwendet wird, kann die GPGPU weniger Compute-Cluster 2936A-2936H aufweisen, als wenn die GPGPU zum Training eines neuronalen Netzes verwendet wird. Bei mindestens einer Ausführungsform kann sich die mit dem Speicher 2944A-2944B verbundene Speichertechnologie zwischen Inferencing- und Trainingskonfigurationen unterscheiden, wobei den Trainingskonfigurationen Speichertechnologien mit höherer Bandbreite zugewiesen werden. Bei mindestens einer Ausführungsform kann die Inferencing-Konfiguration der GPGPU 2930 Inferencing-spezifische Anweisungen unterstützen. Zum Beispiel kann bei mindestens einer Ausführungsform eine Inferencing-Konfiguration Unterstützung für eine oder mehrere 8-Bit-Ganzzahl-Punktprodukt-Anweisungen bieten, die während Inferencing-Operationen für eingesetzte neuronale Netze verwendet werden können.

[0365] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform kann eine Ableitungs- und/oder Trainingslogik 1815 in der GPGPU 2930 zum Ableiten oder Vorhersagen von Operationen mindestens zum Teil basierend auf Gewichtungsparametern, die unter Verwendung von neuronalen Netzwerktrainingsoperationen, Funktionen und/oder Architekturen von neuronalen Netzwerken oder in dieser Schrift beschriebenen Anwendungsfällen von neuronalen Netzwerken berechnet wurden, verwendet werden.

[0366] Bei mindestens einer Ausführungsform wird die GPGPU 2930 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird die GPGPU 2930 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird die GPGPU 2930 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und die GPGPU 2930 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes selbstüberwacht zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0367] **Fig. 30** ist ein Blockdiagramm, das ein Rechensystem 3000 gemäß mindestens einer Ausführungsform zeigt. Bei mindestens einer Ausführungsform weist das Rechensystem 3000 ein Verarbeitungsteilsystem 3001 mit einem oder mehreren Prozessor(en) 3002 und einem Systemspeicher 3004 auf, die über einen Verbindungspfad kommunizieren, der einen Speicher-Hub 3005 aufweisen kann. Bei mindestens einer Ausführungsform kann der Speicher-Hub 3005 eine separate Komponente innerhalb einer Chipsatzkomponente sein oder in einen oder mehrere Prozessor(en) 3002 integriert sein. Bei mindestens einer Ausführungsform ist der Speicher-Hub 3005 über eine Kommunikationsverbindung 3006 mit einem E/A-Subsystem 3011 verbunden. Bei mindestens einer Ausführungsform beinhaltet das E/A-Subsystem 3011 einen E/A-Hub 3007, der es dem Rechensystem 3000 ermöglicht, Eingaben von einer oder mehreren Eingabeeinrichtung(en) 3008 zu empfangen. Bei mindestens einer Ausführungsform kann der E/A-Hub 3007 eine Anzeigesteuerung, die in einem oder mehreren Prozessor(en) 3002 enthalten sein kann, in die Lage versetzen, Ausgaben an eine oder mehrere Anzeigevorrichtung(en) 3010A zu liefern. Bei mindestens einer Ausführungsform können eine oder mehrere Anzeigevorrichtungen 3010A, die mit dem E/A-Hub 3007 gekoppelt sind, eine lokale, interne oder eingebettete Anzeigevorrichtung beinhalten.

[0368] Bei mindestens einer Ausführungsform weist das Verarbeitungssubsystem 3001 einen oder mehrere parallele(n) Prozessor(en) 3012 auf, die über einen Bus oder eine andere Kommunikationsverbindung 3013 mit dem Speicher-Hub 3005 verbunden sind. Bei mindestens einer Ausführungsform kann es sich bei der Kommunikationsverbindung 3013 um eine beliebige Anzahl von standardbasierten Kommunikationsverbindungstechnologien oder -protokollen handeln, wie z. B. PCI Express, ist aber nicht darauf beschränkt, oder um eine herstellerspezifische Kommunikationsschnittstelle oder Kommunikationsstruktur. Bei mindestens einer Ausführungsform bilden ein oder mehrere parallele(r) Prozessor(en) 3012 ein rechnerisch fokussiertes Parallel- oder Vektorverarbeitungssystem, das eine große Anzahl von Verarbeitungskernen und/oder Verarbeitungsclustern aufweisen kann, wie z. B. einen MIC-Prozessor (Many Integrated Core). Bei mindestens einer Ausführungsform bilden ein oder mehrere Parallelprozessoren 3012 ein Grafikverarbeitungsteilsystem, das Pixel an eine von einer oder mehreren Anzeigevorrichtungen 3010A, die über den E/A-Hub 3007 gekoppelt sind, ausgeben kann. Bei mindestens einer Ausführungsform kann ein oder können mehrere Parallelprozessor(en) 3012 auch eine Anzeigesteuervorrichtung und eine Anzeigeschnittstelle (nicht gezeigt) beinhalten, um eine direkte Verbindung mit einer oder mehreren Anzeigevorrichtung(en) 3010B zu ermöglichen.

[0369] Bei mindestens einer Ausführungsform kann eine Systemspeichereinheit 3014 mit dem E/AE/A-Hub 3007 verbunden sein, um einen Speichermechanismus für das Computersystem 3000 bereitzustellen. Bei mindestens einer Ausführungsform kann ein E/A-Schalter 3016 verwendet werden, um einen Schnittstellenmechanismus bereitzustellen, um Verbindungen zwischen dem E/A-Hub 3007 und anderen Komponenten zu ermöglichen, wie z. B. einem Netzwerkadapter 3018 und/oder einem drahtlosen Netzwerkadapter 3019, der in die Plattform integriert sein kann, und verschiedenen anderen Vorrichtungen, die über eine oder mehrere Add-in-Vorrichtung(en) 3020 hinzugefügt werden können. Bei mindestens einer Ausführungsform kann der Netzwerkadapter 3018 ein Ethernet-Adapter oder ein anderer kabelgebundener Netzwerkadapter sein. Bei mindestens einer Ausführungsform kann der drahtlose Netzwerkadapter 3019 eine oder mehrere Wi-Fi-, Bluetooth-, Near Field Communication (NFC)- oder andere Netzwerkeinrichtungen aufweisen, die ein oder mehrere drahtlose Funkgeräte enthalten.

[0370] Bei mindestens einer Ausführungsform kann das Rechensystem 3000 andere, nicht explizit dargestellte Komponenten aufweisen, einschließlich USB- oder andere Anschlüsse, optische Speicherlaufwerke, Videoaufnahmegeräte und dergleichen, die ebenfalls mit dem E/AE/A-Hub 3007 verbunden sein können. Bei mindestens einer Ausführungsform können Kommunikationspfade, die verschiedene Komponenten in **Fig. 30** miteinander verbinden, unter Verwendung beliebiger geeigneter Protokolle implementiert sein, wie z. B PCI (Peripheral Component Interconnect)-basierte Protokolle (z. B PCI-Express) oder andere Bus- oder Punkt-

zu-Punkt-Kommunikationsschnittstellen und/oder Protokolle, wie z. B NV-Link High-Speed-Interconnect oder Interconnect-Protokolle.

[0371] Bei mindestens einer Ausführungsform weisen ein oder mehrere Parallelprozessor(en) 3012 eine für die Grafik- und Videoverarbeitung optimierte Schaltung auf, die beispielsweise eine Videoausgangsschaltung umfasst und eine Grafikverarbeitungseinheit (GPU) darstellt. Bei mindestens einer Ausführungsform enthalten ein oder mehrere Parallelprozessor(en) 3012 Schaltkreise, die für die allgemeine Verarbeitung optimiert sind. Bei mindestens einer Ausführungsform können Komponenten des Rechensystems 3000 mit einem oder mehreren anderen Systemelementen auf einem einzigen integrierten Schaltkreis integriert sein. Beispielsweise können bei mindestens einer Ausführungsform ein oder mehrere parallele(r) Prozessor(en) 3012, ein Speicher-Hub 3005, ein Prozessor(en) 3002 und ein E/AE/A-Hub 3007 in einer integrierten Schaltung mit einem System mit einem System auf einem Chip (SoC) integriert sein. Bei mindestens einer Ausführungsform können die Komponenten des Rechnersystems 3000 in einem einzigen Gehäuse integriert sein, um eine System-in-Package-Konfiguration (SIP) auszugestalten. Bei mindestens einer Ausführungsform kann mindestens ein Abschnitt der Komponenten des Rechensystems 3000 in ein Multi-Chip-Modul (MCM) integriert sein, das mit anderen Multi-Chip-Modulen zu einem modularen Rechensystem zusammengeschaltet sein kann.

[0372] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform kann Ableitungs- und/oder Trainingslogik 1815 in dem System FIG. 3000 zum Ableiten oder Vorhersagen von Operationen mindestens zum Teil basierend auf Gewichtungsparametern verwendet werden, die unter Verwendung von neuronalen Netzwerktrainingsoperationen, Funktionen und/oder Architekturen von neuronalen Netzwerken oder in dieser Schrift beschriebenen Anwendungsfällen von neuronalen Netzwerken berechnet wurden.

[0373] Bei mindestens einer Ausführungsform wird das System 3000 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird das System 3000 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird das System 3000 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und das System 3000 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes selbstüberwacht zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

PROZESSOREN

[0374] **Fig. 31A** veranschaulicht einen Parallelprozessor 3100 gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform können verschiedene Komponenten des Parallelprozessors 3100 unter Verwendung einer oder mehrerer integrierter Schaltungseinrichtungen, wie z. B programmierbare Prozessoren, anwendungsspezifische integrierte Schaltungen (ASICs) oder feldprogrammierbare Gate-Arrays (FPGAs), implementiert werden. Bei mindestens einer Ausführungsform ist der veranschaulichte Parallelprozessor 3100 eine Variante eines oder mehrerer Parallelprozessoren 3012, die in **Fig. 30** gemäß einer beispielhaften Ausführungsform gezeigt sind.

[0375] Bei mindestens einer Ausführungsform beinhaltet der Parallelprozessor 3100 eine Parallelverarbeitungseinheit 3102. Bei mindestens einer Ausführungsform weist die Parallelverarbeitungseinheit 3102 eine E/AE/A-Einheit 3104 auf, die die Kommunikation mit anderen Einrichtungen, einschließlich anderer Instanzen der Parallelverarbeitungseinheit 3102, ermöglicht. Bei mindestens einer Ausführungsform kann die E/AE/A-Einheit 3104 direkt mit anderen Einrichtungen verbunden sein. Bei mindestens einer Ausführungsform ist die E/AE/A-Einheit 3104 über eine Hub- oder Switch-Schnittstelle, wie z. B. den Speicher-Hub 3005, mit anderen Einrichtungen verbunden. Bei mindestens einer Ausführungsform bilden die Verbindungen zwischen Speicher-Hub 3005 und E/AE/A-Einheit 3104 eine Kommunikationsverbindung 3013. Bei mindestens einer Ausführungsform ist die E/A-Einheit 3104 mit einer Host-Schnittstelle 3106 und einem Speicher-Kreuzschiene 3116 verbunden, wobei die Host-Schnittstelle 3106 Befehle zur Ausführung von Verarbeitungsvorgängen und die Speicher-Kreuzschiene 3116 Befehle zur Ausführung von Speicheroperationen empfängt.

[0376] Bei mindestens einer Ausführungsform, wenn die Host-Schnittstelle 3106 einen Befehlspuffer über die E/AE/A-Einheit 3104 empfängt, kann die Host-Schnittstelle 3106 Arbeitsoperationen zur Ausführung dieser Befehle an ein Frontend 3108 leiten. Bei mindestens einer Ausführungsform ist das vordere Ende 3108 mit einem Scheduler 3110 gekoppelt, der so ausgestaltet ist, dass er Befehle oder andere Arbeitselemente an eine Verarbeitungsclusteranordnung 3112 verteilt. Bei mindestens einer Ausführungsform stellt der Planer 3110 sicher, dass die Verarbeitungs-Cluster-Anordnung 3112 ordnungsgemäß konfiguriert ist und sich in einem gültigen Zustand befindet, bevor Aufgaben an die Verarbeitungs-Cluster-Anordnung 3112 von Verarbeitungs-Cluster-Anordnung 3112 verteilt werden. Bei mindestens einer Ausführungsform ist der Scheduler 3110 über Firmware-Logik implementiert, die auf einem Mikrocontroller ausgeführt wird. Bei mindestens einer Ausführungsform ist der Mikrocontrollerimplementierte Scheduler 3110 so ausgestaltet, dass er komplexe Ablaufsteuerungs- und Arbeitsverteilungsoperationen mit grober und feiner Granularität durchführen kann, was eine schnelle Unterbrechung und Kontextumschaltung von Threads ermöglicht, die auf der Verarbeitungsanordnung 3112 ausgeführt werden. Bei mindestens einer Ausführungsform kann die Host-Software Arbeitslasten für die Planung auf der Verarbeitungsanordnung 3112 über eine von mehreren Grafikverarbeitungs-Doorbells nachweisen. Bei mindestens einer Ausführungsform können die Arbeitslasten dann automatisch durch die Logik des Schedulers 3110 innerhalb eines Mikrocontrollers, der den Scheduler 3110 aufweist, auf der Verarbeitungsanordnung 3112 verteilt werden.

[0377] Bei mindestens einer Ausführungsform kann die Verarbeitungs-Cluster-Anordnung 3112 bis zu „N“ Verarbeitungs-Cluster (z. B. Cluster 3114A, Cluster 3114B bis Cluster 3114N) beinhalten. Bei mindestens einer Ausführungsform kann jeder Cluster 3114A-3114N der Verarbeitungsclusteranordnung 3112 eine große Anzahl von gleichzeitigen Threads ausführen. Bei mindestens einer Ausführungsform kann der Scheduler 3110 den Clustern 3114A-3114N der Verarbeitungsclusteranordnung 3112 Arbeit zuweisen, indem er verschiedene Ablaufsteuerungs- und/oder Arbeitsverteilungsalgorithmen verwendet, die je nach der Arbeitslast variieren können, die für jede Art von Programm oder Berechnung entsteht. Bei mindestens einer Ausführungsform kann die Planung dynamisch durch den Scheduler 3110 erfolgen oder teilweise durch eine Compilerlogik während der Kompilierung der Programmlogik unterstützt werden, die für die Ausführung durch die Verarbeitungsclusteranordnung 3112 ausgestaltet ist. Bei mindestens einer Ausführungsform können verschiedene Cluster 3114A-3114N der Verarbeitungscusteranordnung 3112 für die Verarbeitung verschiedener Arten von Programmen oder für die Durchführung verschiedener Arten von Berechnungen zugewiesen werden.

[0378] Bei mindestens einer Ausführungsform kann die Verarbeitungsclusteranordnung 3112 so ausgestaltet sein, dass sie verschiedene Arten von Parallelverarbeitungsoperationen durchführt. Bei mindestens einer Ausführungsform ist die Verarbeitungsclusteranordnung 3112 so ausgestaltet, dass sie parallele Allzweck-Rechenoperationen durchführt. Zum Beispiel kann bei mindestens einer Ausführungsform die Verarbeitungsclusteranordnung 3112 eine Logik aufweisen, um Verarbeitungsaufgaben auszuführen, einschließlich der Filterung von Video- und/oder Audiodaten, der Durchführung von Modellierungsoperationen, einschließlich physikalischer Operationen, und der Durchführung von Datentransformationen.

[0379] Bei mindestens einer Ausführungsform ist die Verarbeitungsclusteranordnung 3112 so ausgestaltet, dass sie parallele Grafikverarbeitungsoperationen durchführt. Bei mindestens einer Ausführungsform kann die Verarbeitungs-Cluster-Anordnung 3112 zusätzliche Logik beinhalten, um die Ausführung derartiger Grafikverarbeitungsoperationen zu unterstützen, einschließlich, aber nicht beschränkt auf Texturabtastlogik, um Texturoperationen, sowie Tesselationslogik und andere Vertexverarbeitungslogik auszuführen. Bei mindestens einer Ausführungsform kann die Verarbeitungsclusteranordnung 3112 so ausgestaltet sein, dass sie grafikverarbeitungsbezogene Shader-Programme ausführt, wie z. B. Vertex-Shader, Tesselation-Shader, Geometrie-Shader und Pixel-Shader. Bei mindestens einer Ausführungsform kann die Parallelverarbeitungseinheit 3102 Daten aus dem Systemspeicher über die E/AE/A-Einheit 3104 zur Verarbeitung übertragen. Bei mindestens einer Ausführungsform können die übertragenen Daten während der Verarbeitung im On-Chip-Speicher (z. B. im Parallelprozessorspeicher 3122) gespeichert und dann in den Systemspeicher zurückgeschrieben werden.

[0380] Bei mindestens einer Ausführungsform, wenn die Parallelverarbeitungseinheit 3102 zur Durchführung der Grafikverarbeitung verwendet wird, kann der Scheduler 3110 so ausgestaltet sein, dass er eine Verarbeitungslast in ungefähr gleich große Tasks aufteilt, um eine bessere Verteilung der Grafikverarbeitungsoperationen auf mehrere Cluster 3114A-3114N der Verarbeitungsclusteranordnung 3112 zu ermöglichen. Bei mindestens einer Ausführungsform können Abschnitte der Verarbeitungsclusteranordnung 3112 so ausgestaltet sein, dass sie verschiedene Arten der Verarbeitung durchführen. Zum Beispiel kann bei mindestens einer Ausführungsform ein erster Abschnitt so ausgestaltet sein, dass er Vertex-Shading und Topologieerzeugung

durchführt, ein zweiter Abschnitt kann so ausgestaltet sein, dass er Tesselations- und Geometrie-Shading durchführt, und ein dritter Abschnitt kann so ausgestaltet sein, dass er Pixel-Shading oder andere Screenspace-Operationen durchführt, um ein gerendertes Bild für die Anzeige zu erzeugen. Bei mindestens einer Ausführungsform können Zwischendaten, die von einem oder mehreren Clustern 3114A-3114N erzeugt werden, in Puffern gespeichert werden, damit Zwischendaten zwischen den Clustern 3114A-3114N zur weiteren Verarbeitung übertragen werden können.

[0381] Bei mindestens einer Ausführungsform kann die Verarbeitungsclusteranordnung 3112 über den Scheduler 3110, der Befehle zur Definition von VerarbeitungsTasks vom Frontend 3108 erhält, auszuführende Verarbeitungs-Tasks empfangen. Bei mindestens einer Ausführungsform können Verarbeitungsaufgaben Indizes von zu verarbeitenden Daten beinhalten, z. B. Oberflächendaten (Patch-Daten), Primitivdaten, Vertex-Daten und/oder Bildpunktdaten sowie Zustandsparameter und Befehle, die definieren, wie Daten verarbeitet werden sollen (z. B. welches Programm ausgeführt werden soll). Bei mindestens einer Ausführungsform kann der Scheduler 3110 so ausgestaltet sein, dass er den Tasks entsprechende Indizes abruft oder Indizes vom Frontend 3108 empfängt. Bei mindestens einer Ausführungsform kann das Frontend 3108 so ausgestaltet sein, dass es sicherstellt, dass die Verarbeitungsclusteranordnung 3112 in einem gültigen Zustand konfiguriert ist, bevor eine durch eingehende Befehlspuffer (z. B. Batch-Puffer, Push-Puffer usw.) spezifizierte Arbeitslast eingeleitet wird.

[0382] Bei mindestens einer Ausführungsform kann jede von einer oder mehreren Instanzen der Parallelverarbeitungseinheit 3102 mit dem Parallelprozessorspeicher 3122 koppeln. Bei mindestens einer Ausführungsform kann auf den Parallelprozessorspeicher 3122 über das Speicherkoppelfeld 3116 zugegriffen werden, die Speicheranforderungen von der Verarbeitungsclusteranordnung 3112 sowie der E/AE/A-Einheit 3104 empfangen kann. Bei mindestens einer Ausführungsform kann das Speicherkoppelfeld 3116 über eine Speicherschnittstelle 3118 auf den Parallelprozessorspeicher 3122 zugreifen. Bei mindestens einer Ausführungsform kann die Speicherschnittstelle 3118 mehrere Partitionseinheiten (z. B. Partitionseinheit 3120A, Partitionseinheit 3120B bis Partitionseinheit 3120N), die jeweils an einen Teil (z. B. Speichereinheit) des Parallelprozessorspeichers 3122 koppeln können, beinhalten. Bei mindestens einer Ausführungsform ist eine Anzahl von Partitionseinheiten 3120A-3120N dazu konfiguriert, gleich einer Anzahl von Speichereinheiten zu sein, dass eine erste Partitionseinheit 3120A eine entsprechende erste Speichereinheit 3124A aufweist, eine zweite Partitionseinheit 3120B eine entsprechende zweite Speichereinheit 3124B aufweist, und eine N-te Partitionseinheit 3120N eine entsprechende N-te Speichereinheit 3124N aufweist. Bei mindestens einer Ausführungsform kann eine Anzahl von Partitionseinheiten 3120A-3120N nicht gleich einer Anzahl von Speichereinrichtungen sein.

[0383] Bei mindestens einer Ausführungsform können die Speichereinheiten 3124A-3124N verschiedene Arten von Speichereinrichtungen aufweisen, einschließlich dynamischer Direktzugriffsspeicher (DRAM) oder Grafik-Direktzugriffsspeicher, wie synchroner Grafik-Direktzugriffsspeicher (SGRAM), einschließlich Grafik-Doppeldatenraten-Speicher (GDDR). Bei mindestens einer Ausführungsform können die Speichereinheiten 3124A-3124N auch 3D-Stapelspeicher aufweisen, einschließlich, aber nicht beschränkt auf Speicher mit hoher Bandbreite (High Bandwidth Memory (HBM)). Bei mindestens einer Ausführungsform können Rendering-Ziele, wie z. B. Frame-Puffer oder Textur-Maps, über die Speichereinheiten 3124A-3124N hinweg gespeichert werden, so dass die Partitionseinheiten 3120A-3120N Abschnitte jedes Rendering-Ziels parallel schreiben können, um die verfügbare Bandbreite des Parallelprozessorspeichers 3122 effizient zu nutzen. Bei mindestens einer Ausführungsform kann eine lokale Instanz des Parallelprozessorspeichers 3122 zugunsten eines vereinheitlichten Speicherentwurfs ausgeschlossen werden, der den Systemspeicher in Verbindung mit dem lokalen Cache-Speicher nutzt.

[0384] Bei mindestens einer Ausführungsform kann jeder der Cluster 3114A-3114N der Verarbeitungsclusteranordnung 3112 Daten verarbeiten, die in jede der Speichereinheiten 3124A-3124N im Parallelprozessorspeicher 3122 geschrieben werden. Bei mindestens einer Ausführungsform kann das Speicherkoppelfeld 3116 so ausgestaltet sein, dass es eine Ausgabe jedes Clusters 3114A-3114N an eine beliebige Partitionseinheit 3120A-3120N oder an einen anderen Cluster 3114A-3114N überträgt, der zusätzliche Verarbeitungsoperationen an einer Ausgabe durchführen kann. Bei mindestens einer Ausführungsform kann jeder Cluster 3114A-3114N mit der Speicherschnittstelle 3118 über das Speicherkoppelfeld 3116 kommunizieren, um von verschiedenen externen Einrichtungen zu lesen oder in diese zu schreiben. Bei mindestens einer Ausführungsform weist die Speicherkreuzschiene 3116 eine Verbindung zur Speicherschnittstelle 3118 auf, um mit der E/A-Einheit 3104 zu kommunizieren, sowie eine Verbindung zu einer lokalen Instanz des Parallelprozessorspeichers 3122, so dass die Verarbeitungseinheiten in den verschiedenen Verarbeitungs-Clustern 3114A-3114N mit dem Systemspeicher oder einem anderen Speicher, der nicht lokal zur Parallelverarbei-

tungseinheit 3102 gehört, kommunizieren können. Bei mindestens einer Ausführungsform kann die Speicherkreuzschiene 3116 virtuelle Kanäle verwenden, um Verkehrsströme zwischen den Clustern 3114A-3114N und den Partitionseinheiten 3120A-3120N zu trennen.

[0385] Bei mindestens einer Ausführungsform können mehrere Instanzen der Parallelverarbeitungseinheit 3102 auf einer einzigen Add-in-Karte bereitgestellt sein, oder mehrere Add-in-Karten können miteinander verbunden sein. Bei mindestens einer Ausführungsform können unterschiedliche Instanzen der Parallelverarbeitungseinheit 3102 dazu konfiguriert sein zusammenzuarbeiten, selbst wenn unterschiedliche Instanzen unterschiedliche Anzahlen von Verarbeitungskernen, unterschiedliche Mengen an lokalem Parallelprozessorspeicher und/oder andere Konfigurationsunterschiede aufweisen. Zum Beispiel können bei mindestens einer Ausführungsform einige Instanzen der Parallelverarbeitungseinheit 3102 im Vergleich zu anderen Ausführungen Gleitkommaeinheiten mit höherer Präzision aufweisen. Bei mindestens einer Ausführungsform können Systeme, die eine oder mehrere Instanzen der Parallelverarbeitungseinheit 3102 oder des Parallelprozessors 3100 enthalten, in einer Vielzahl von Ausführungsformen und Formfaktoren implementiert sein, einschließlich, aber nicht beschränkt auf Desktop-, Laptop- oder Handheld-Personalcomputer, Server, Workstations, Spielkonsolen und/oder eingebettete Systeme.

[0386] **Fig. 31B** ist ein Blockdiagramm einer Partitionseinheit 3120 gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform ist die Partitionseinheit 3120 eine Instanz einer von Partitionseinheiten 3120A-3120N der **Fig. 31A**. Bei mindestens einer Ausführungsform weist die Partitionseinheit 3120 einen L2-Cache 3121, eine Rahmenpufferschnittstelle 3125 und eine ROP 3126 (Rasteroperationseinheit) auf. Der L2-Cache 3121 ist ein Lese-/Schreib-Cache, der dazu konfiguriert ist, von der Speicherkreuzschiene 3116 und der ROP 3126 empfangene Lade- und Speicheroperationen auszuführen. Bei mindestens einer Ausführungsform werden Lesefehler und dringende Rückschreibanforderungen von dem L2-Cache 3121 zur Verarbeitung an die Frame-Speicherschnittstelle 3125 ausgegeben. Bei mindestens einer Ausführungsform können Aktualisierungen auch über die Frame-Pufferschnittstelle 3125 zur Verarbeitung an einen Frame-Puffer gesendet werden. Bei mindestens einer Ausführungsform ist die Rahmenpufferschnittstelle 3125 mit einer der Speichereinheiten im Parallelprozessorspeicher verbunden, wie den Speichereinheiten 3124A-3124N von **Fig. 31** (z. B. innerhalb des Parallelprozessorspeichers 3122).

[0387] Bei mindestens einer Ausführungsform ist die ROP 3126 eine Verarbeitungseinheit, die Rasteroperationen wie Schablonieren, Z-Test, Überblendung und ähnliches durchführt. Bei mindestens einer Ausführungsform gibt die ROP 3126 dann verarbeitete Grafikdaten aus, die im Grafikspeicher abgelegt werden. Bei mindestens einer Ausführungsform weist die ROP 3126 eine Komprimierungslogik auf, um Tiefen- oder Farbdaten zu komprimieren, die in den Speicher geschrieben werden, und Tiefen- oder Farbdaten zu dekomprimieren, die aus dem Speicher gelesen werden. Bei mindestens einer Ausführungsform kann die Komprimierungslogik eine verlustfreie Komprimierungslogik sein, die einen oder mehrere von mehreren Komprimierungsalgorithmen verwendet. Bei mindestens einer Ausführungsform kann die Art der von der ROP 3126 durchgeführten Komprimierung auf der Grundlage statistischer Merkmale der zu komprimierenden Daten variieren. Beispielsweise wird bei mindestens einer Ausführungsform eine Delta-Farbkompression auf Tiefen- und Farbdaten auf einer Pro-Kachel-Basis ausgeführt.

[0388] Bei mindestens einer Ausführungsform ist die ROP 3126 in jedem Verarbeitungscluster (z. B. Cluster 3114A-3114N von **Fig. 31**) und nicht in der Partitionseinheit 3120 vorhanden. Bei mindestens einer Ausführungsform werden Lese- und Schreibanforderungen für Pixeldaten über das Speicherkoppelfeld 3116 anstelle von Pixelfragmentdaten übertragen. Bei mindestens einer Ausführungsform können verarbeitete Grafikdaten auf einer Anzeigevorrichtung angezeigt werden, wie etwa auf einer oder mehreren Anzeigevorrichtungen 3010 der **Fig. 30**, die zur Weiterverarbeitung durch den/die Prozessor(en) 3002 geleitet oder zur weiteren Verarbeitung durch eine der Verarbeitungseinheiten innerhalb des Parallelprozessors 3100 der **Fig. 31A** geleitet werden.

[0389] **Fig. 31C** ist ein Blockdiagramm eines Verarbeitungsclusters 3114 innerhalb einer Parallelverarbeitungseinheit gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform ist ein Verarbeitungscluster eine Instanz von einem der Verarbeitungscluster 3114A-3114N von **Fig. 31**. Bei mindestens einer Ausführungsform kann der Verarbeitungscluster 3114 so ausgestaltet sein, dass er viele Threads parallel ausführt, wobei sich der Begriff „Thread“ auf eine Instanz eines bestimmten Programms bezieht, das auf einem bestimmten Satz von Eingabedaten ausgeführt wird. Bei mindestens einer Ausführungsform werden SIMD-Befehlsausgabetechniken (Single-Instruction, Multiple-Data) verwendet, um die parallele Ausführung einer großen Anzahl von Threads zu unterstützen, ohne mehrere unabhängige Befehlseinheiten bereitzustellen. Bei mindestens einer Ausführungsform werden Single-Instruction-Multiple-Thread (SIMT)-Techniken ver-

wendet, um die parallele Ausführung einer großen Anzahl von im Allgemeinen synchronisierten Threads zu unterstützen, wobei eine gemeinsame Befehlseinheit ausgestaltet ist, um Befehle an einen Satz von Verarbeitungsmaschinen innerhalb jedes der Verarbeitungscluster auszugeben.

[0390] Bei mindestens einer Ausführungsform kann der Betrieb des Verarbeitungsclusters 3114 über einen Pipeline-Manager 3132 gesteuert werden, der die Verarbeitungs-Tasks an parallele SIMT-Prozessoren verteilt. Bei mindestens einer Ausführungsform empfängt der Pipeline-Manager 3132 Anweisungen vom Scheduler 3110 der **Fig. 31** und verwaltet die Ausführung dieser Anweisungen über einen Grafik-Multiprozessor 3134 und/oder eine Textureinheit 3136. Bei mindestens einer Ausführungsform ist der Grafikmultiprozessor 3134 eine beispielhafte Instanz eines SIMT-Parallelprozessors. Bei mindestens einer Ausführungsform können jedoch verschiedene Typen von SIMT-Parallelprozessoren mit unterschiedlichen Architekturen im Verarbeitungscluster 3114 vorhanden sein. Bei mindestens einer Ausführungsform können ein oder mehrere Instanzen des Grafik-Multiprozessors 3134 in einem Verarbeitungscluster 3114 vorhanden sein. Bei mindestens einer Ausführungsform kann der Grafikmultiprozessor 3134 Daten verarbeiten, und ein Datenkoppelfeld 3140 kann verwendet werden, um die verarbeiteten Daten an eines von mehreren möglichen Zielen zu verteilen, einschließlich anderer Shader-Einheiten. Bei mindestens einer Ausführungsform kann der Pipeline-Manager 3132 die Verteilung der verarbeiteten Daten erleichtern, indem er Ziele für die verarbeiteten Daten angibt, die über das Datenkoppelfeld 3140 verteilt werden sollen.

[0391] Bei mindestens einer Ausführungsform kann jeder Grafik-Multiprozessor 3134 innerhalb des Verarbeitungs-Clusters 3114 einen identischen Satz funktionaler Ausführungslogik beinhalten (z. B. arithmetische Logikeinheiten, Ladespeichereinheiten usw.). Bei mindestens einer Ausführungsform kann die funktionale Ausführungslogik in einer Pipeline ausgestaltet sein, so dass neue Befehle ausgegeben werden können, bevor vorherige Befehle abgeschlossen sind. Bei mindestens einer Ausführungsform unterstützt die funktionale Ausführungslogik eine Vielzahl von Operationen, darunter Ganzzahl- und Gleitkommaarithmetik, Vergleichsoperationen, boolesche Operationen, Bitverschiebung und die Berechnung verschiedener algebraischer Funktionen. Bei mindestens einer Ausführungsform kann dieselbe Hardware mit Funktionseinheiten genutzt werden, um verschiedene Operationen auszuführen, und es kann eine beliebige Kombination von Funktionseinheiten vorhanden sein.

[0392] Bei mindestens einer Ausführungsform bilden die an den Verarbeitungscluster 3114 übertragenen Anweisungen einen Thread. Bei mindestens einer Ausführungsform ist ein Satz von Threads, die über einen Satz von Parallelverarbeitungsmaschinen ausgeführt werden, eine Thread-Gruppe. Bei mindestens einer Ausführungsform führt die Thread-Gruppe ein Programm mit unterschiedlichen Eingabedaten aus. Bei mindestens einer Ausführungsform kann jeder Thread innerhalb einer Thread-Gruppe einer anderen Verarbeitungsmaschine innerhalb eines Grafik-Multiprozessors 3134 zugewiesen sein. Bei mindestens einer Ausführungsform kann eine Thread-Gruppe weniger Threads aufweisen als die Anzahl der Verarbeitungseinheiten im Grafik-Multiprozessor 3134. Bei mindestens einer Ausführungsform kann, wenn eine Thread-Gruppe weniger Threads als eine Anzahl von Verarbeitungsmaschinen aufweist, eine oder mehrere der Verarbeitungsmaschinen während der Zyklen, in denen diese Thread-Gruppe verarbeitet wird, im Leerlauf sein. Bei mindestens einer Ausführungsform kann eine Thread-Gruppe auch mehr Threads aufweisen als eine Anzahl von Verarbeitungsmaschinen im Grafik-Multiprozessor 3134. Bei mindestens einer Ausführungsform kann, wenn eine Thread-Gruppe mehr Threads aufweist als die Anzahl der Verarbeitungsmaschinen im Grafik-Multiprozessor 3134, die Verarbeitung über aufeinander folgende Taktzyklen erfolgen. Bei mindestens einer Ausführungsform können mehrere Thread-Gruppen gleichzeitig auf einem Grafik-Multiprozessor 3134 ausgeführt werden.

[0393] Bei mindestens einer Ausführungsform weist der Grafik-Multiprozessor 3134 einen internen Cache-Speicher auf, um Lade- und Speicheroperationen durchzuführen. Bei mindestens einer Ausführungsform kann der Grafik-Multiprozessor 3134 auf einen internen Cache verzichten und einen Cache-Speicher (z. B L1-Cache 3148) innerhalb des Verarbeitungsclusters 3114 verwenden. Bei mindestens einer Ausführungsform hat jeder Grafik-Multiprozessor 3134 auch Zugriff auf L2-Caches innerhalb von Partitionseinheiten (z. B. die Partitionseinheiten 3120A-3120N der **Fig. 31**), die von allen Verarbeitungs-Clustern 3114 gemeinsam genutzt werden und zur Datenübertragung zwischen Threads verwendet werden können. Bei mindestens einer Ausführungsform kann der Grafik-Multiprozessor 3134 auch auf den globalen Speicher außerhalb des Chips zugreifen, der einen oder mehrere lokale Parallelprozessorspeicher und/oder Systemspeicher aufweisen kann. Bei mindestens einer Ausführungsform kann jeder Speicher außerhalb der Parallelverarbeitungseinheit 3102 als globaler Speicher verwendet werden. Bei mindestens einer Ausführungsform weist der Verarbeitungscluster 3114 mehrere Instanzen des Grafik-Multiprozessors 3134 auf, die sich gemeinsame Anweisungen und Daten teilen können, die im L1 -Cache 3148 gespeichert sein können.

[0394] Bei mindestens einer Ausführungsform kann jeder Verarbeitungscluster 3114 eine MMU 3145 (Memory Management Unit) aufweisen, die so ausgestaltet ist, dass sie virtuelle Adressen in physische Adressen umsetzt. Bei mindestens einer Ausführungsform können sich eine oder mehrere Instanzen der MMU 3145 innerhalb der Speicherschnittstelle 3118 von **Fig. 31** befinden. Bei mindestens einer Ausführungsform weist die MMU 3145 einen Satz von Seitentabelleneinträgen (PTEs) auf, die dazu dienen, eine virtuelle Adresse auf eine physische Adresse einer Kachel abzubilden (weitere Informationen über Kacheln), sowie optional einen Cache-Zeilenindex. Bei mindestens einer Ausführungsform kann die MMU 3145 Adressübersetzungs-Lookaside-Puffer (TLB) oder Caches aufweisen, die sich im Grafik-Multiprozessor 3134 oder im L1-Cache oder im Verarbeitungscluster 3114 befinden können. Bei mindestens einer Ausführungsform wird die physikalische Adresse verarbeitet, um die Zugriffslokalität auf die Oberflächendaten zu verteilen, um eine effiziente Anforderungsverschachtelung zwischen den Partitionseinheiten zu ermöglichen. Bei mindestens einer Ausführungsform kann der Cache-Zeilen-Index verwendet werden, um zu bestimmen, ob eine Anforderung für eine Cache-Zeile ein Treffer (Hit) oder Fehlzugriff (Miss) ist.

[0395] Bei mindestens einer Ausführungsform kann ein Verarbeitungs-Cluster 3114 derartig konfiguriert sein, dass jeder Grafikmultiprozessor 3134 an eine Textureinheit 3136 zum Ausführen von Texturzuordnungsoperationen, z. B. zum Bestimmen von Texturabtastpositionen, Lesen von Texturdaten und Filtern von Texturdaten, gekoppelt ist. Bei mindestens einer Ausführungsform werden die Texturdaten aus einem internen Textur-L1-Cache (nicht gezeigt) oder aus einem L1-Cache innerhalb des Grafik-Multiprozessors 3134 gelesen und je nach Bedarf aus einem L2-Cache, einem lokalen Parallelprozessorspeicher oder dem Systemspeicher abgerufen. Bei mindestens einer Ausführungsform gibt jeder Grafikmultiprozessor 3134 verarbeitete Tasks an das Datenkoppelfeld 3140 aus, um die verarbeitete Task einem anderen Verarbeitungscluster 3114 zur weiteren Verarbeitung zur Verfügung zu stellen oder um die verarbeitete Task über das Speicherkoppelfeld 3116 in einem L2-Cache, im lokalen Parallelprozessorspeicher oder im Systemspeicher zu speichern. Bei mindestens einer Ausführungsform ist eine preROP 3142 (Pre-Raster Operations Unit) derart konzipiert, dass sie Daten von dem Grafik-Multiprozessor 3134 empfängt und Daten an ROP-Einheiten, die sich in den hierin beschriebenen Partitionseinheiten befinden können (z. B. die Partitionseinheiten 3120A-3120N der **Fig. 31**), weiterleitet. Bei mindestens einer Ausführungsform kann die PreROP-Einheit 3142 Optimierungen für die Farbmischung durchführen, Pixelfarbdaten organisieren und Adressübersetzungen vornehmen.

[0396] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform kann eine Ableitungs- und/oder Trainingslogik 1815 in dem Grafikverarbeitungs-Cluster 3114 zum Ableiten oder Vorhersagen von Operationen mindestens zum Teil basierend auf Gewichtungsparametern verwendet werden, die unter Verwendung von neuronalen Netzwerktrainingsoperationen, Funktionen und/oder Architekturen von neuronalen Netzwerken oder in dieser Schrift beschriebenen Anwendungsfällen von neuronalen Netzwerken berechnet wurden.

[0397] Bei mindestens einer Ausführungsform wird das Grafikverarbeitungs-Cluster 3114 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird das Grafikverarbeitungs-Cluster 3114 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird das Grafikverarbeitungs-Cluster 3114 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und das Grafikverarbeitungs-Cluster 3114 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um ein auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0398] **Fig. 31D** zeigt einen Grafik-Multiprozessor 3134 gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform ist der Grafik-Multiprozessor 3134 mit dem Pipeline-Manager 3132 des Verarbeitungsclusters 3114 gekoppelt. Bei mindestens einer Ausführungsform weist der Grafikmultiprozessor 3134 eine Ausführungspipeline auf, die unter anderem einen Anweisungs-Cache 3152, eine Befehlseinheit 3154, eine Adresszuordnungseinheit 3156, eine Registerdatei 3158, einen oder mehrere GPGPU-Kerne 3162 und eine oder mehrere Lade-/Speichereinheiten 3166 beinhaltet. Die GPGPU-Kerne 3162 und die Lade-/Speichereinheiten 3166 sind über eine Speicher- und ein Cache-Interconnect 3168 mit dem Cache-Speicher 3172 und dem gemeinsam genutzten Speicher 3170 gekoppelt.

[0399] Bei mindestens einer Ausführungsform empfängt der Befehlscache 3152 einen Strom von auszuführenden Befehlen vom Pipeline-Manager 3132. Bei mindestens einer Ausführungsform werden die Befehle im Befehlscache 3152 zwischengespeichert und von der Befehlseinheit 3154 zur Ausführung weitergeleitet. Bei mindestens einer Ausführungsform kann die Anweisungseinheit 3154 Anweisungen als Thread-Gruppen (z. B. Warps) versenden, wobei jeder Thread einer Thread-Gruppe einer anderen Ausführungseinheit innerhalb der GPGPU-Kerne 3162 zugewiesen ist. Bei mindestens einer Ausführungsform kann ein Befehl auf einen lokalen, gemeinsam genutzten oder globalen Adressraum zugreifen, indem er eine Adresse innerhalb eines einheitlichen Adressraums angibt. Bei mindestens einer Ausführungsform kann die Adressabbildungseinheit 3156 verwendet werden, um Adressen in einem vereinheitlichten Adressraum in eine eindeutige Speicheradresse zu übersetzen, auf die die Lade-/Speichereinheiten 3166 zugreifen können.

[0400] Bei mindestens einer Ausführungsform stellt die Registerdatei 3158 einen Satz von Registern für Funktionseinheiten des Grafik-Multiprozessors 3134 bereit. Bei mindestens einer Ausführungsform stellt die Registerdatei 3158 einen temporären Speicher für Operanden bereit, die mit Datenpfaden von funktionellen Einheiten (z. B. GPGPU-Kernen 3162, Lade-/Speichereinheit(en) 3166) des Grafikmultiprozessors 3134 verbunden sind. Bei mindestens einer Ausführungsform ist die Registerdatei 3158 zwischen den einzelnen Funktionseinheiten aufgeteilt, so dass jeder Funktionseinheit ein eigener Abschnitt der Registerdatei 3158 zugewiesen ist. Bei mindestens einer Ausführungsform ist die Registerdatei 3158 auf verschiedene Warps aufgeteilt, die vom Grafikmultiprozessor 3134 ausgeführt werden.

[0401] Bei mindestens einer Ausführungsform können die GPGPU-Kerne 3162 jeweils Gleitkommaeinheiten (FPUs) und/oder ganzzahlige arithmetische Logikeinheiten (ALUs) aufweisen, die zur Ausführung von Anweisungen des Grafikmultiprozessors 3134 verwendet werden. Die GPGPU-Kerne 3162 können sich in ihrer Architektur ähneln oder unterscheiden. Bei mindestens einer Ausführungsform weist ein erster Abschnitt der GPGPU-Kerne 3162 eine FPU mit einfacher Genauigkeit und eine Ganzzahl-ALU auf, während ein zweiter Abschnitt der GPGPU-Kerne eine FPU mit doppelter Genauigkeit aufweist. Bei mindestens einer Ausführungsform können die FPUs den IEEE 754-2008-Standard für Gleitkommaarithmetik implementieren oder Gleitkommaarithmetik mit variabler Genauigkeit ermöglichen. Bei mindestens einer Ausführungsform kann der Grafikmultiprozessor 3134 zusätzlich eine oder mehrere Festfunktions- oder Sonderfunktionseinheiten aufweisen, um spezifische Funktionen wie das Kopieren von Rechtecken oder Pixel-Blending-Operationen durchzuführen. Bei mindestens einer Ausführungsform kann einer oder können mehrere der GPGPU-Kerne auch eine feste oder spezielle Funktionslogik aufweisen.

[0402] Bei mindestens einer Ausführungsform weisen die GPGPU-Kerne 3162 eine SIMD-Logik auf, die in der Lage ist, einen einzigen Befehl für mehrere Datensätze auszuführen. Bei mindestens einer Ausführungsform können GPGPU-Kerne 3162 physikalisch SIMD4-, SIMD8- und SIMD16-Befehle und logisch SIMD1-, SIMD2- und SIMD32-Befehle ausführen. Bei mindestens einer Ausführungsform können SIMD-Befehle für GPGPU-Kerne zur Kompilierzeit von einem Shader-Compiler oder automatisch bei der Ausführung von Programmen erzeugt werden, die für SPMD- oder SIMT-Architekturen (Single Program Multiple Data) geschrieben und kompiliert wurden. Bei mindestens einer Ausführungsform können mehrere Threads eines Programms, das für ein SIMT-Ausführungsmodell ausgestaltet ist, über einen einzigen SIMD-Befehl ausgeführt werden. Beispielsweise können bei mindestens einer Ausführungsform acht SIMT-Threads, die gleiche oder ähnliche Operationen durchführen, über eine einzige SIMD8-Logikeinheit parallel ausgeführt werden.

[0403] Bei mindestens einer Ausführungsform ist die Speicher- und Cache-Verbindung 3168 ein Verbindungsnetzwerk, das jede Funktionseinheit des Grafik-Multiprozessors 3134 mit der Registerdatei 3158 und dem gemeinsamen Speicher 3170 verbindet. Bei mindestens einer Ausführungsform ist die Speicher- und Cache-Verbindung 3168 eine Koppelfeldverbindung, die es der Lade-/Speichereinheit 3166 ermöglicht, Lade- und Speicheroperationen zwischen dem gemeinsamen Speicher 3170 und der Registerdatei 3158 durchzuführen. Bei mindestens einer Ausführungsform kann die Registerdatei 3158 mit derselben Frequenz wie die GPGPU-Kerne 3162 arbeiten, so dass die Datenübertragung zwischen den GPGPU-Kernen 3162 und der Registerdatei 3158 eine sehr geringe Latenzzeit aufweist. Bei mindestens einer Ausführungsform kann der gemeinsam genutzte Speicher 3170 verwendet werden, um die Kommunikation zwischen Threads zu ermöglichen, die auf Funktionseinheiten innerhalb des Grafik-Multiprozessors 3134 ausgeführt werden. Bei mindestens einer Ausführungsform kann der Cache-Speicher 3172 beispielsweise als Daten-Cache verwendet werden, um Texturdaten, die zwischen Funktionseinheiten und der Textureinheit 3136 übertragen werden, zwischenzuspeichern. Bei mindestens einer Ausführungsform kann der gemeinsame Speicher 3170 auch als programmgesteuerter Cache verwendet werden. Bei mindestens einer Ausführungsform können Threads, die auf GPGPU-Kernen 3162 ausgeführt werden, zusätzlich zu den automatisch zwischengespei-

cherten Daten, die in den Cache-Speicher 3172 gespeichert sind, programmatisch Daten in dem gemeinsam genutzten Speicher speichern.

[0404] Bei mindestens einer Ausführungsform ist ein Parallelprozessor oder eine GPGPU, wie hierin beschrieben, kommunikativ mit Host-/Prozessorkernen gekoppelt, um Grafikoperationen, Operationen des maschinellen Lernens, Musteranalyseoperationen und verschiedene allgemeine GPU-Funktionen (GPGPU-Funktionen) zu beschleunigen. Bei mindestens einer Ausführungsform kann die GPU über einen Bus oder eine andere Verbindung (z. B. eine Hochgeschwindigkeitsverbindung wie PCIe oder NVLink) mit dem Host-Prozessor (den Prozessorkernen) kommunikativ verbunden sein. Bei mindestens einer Ausführungsform kann die GPU in demselben Gehäuse oder Chip wie die Kerne integriert sein und über einen internen Prozessorbus bzw. eine interne Verbindung (d. h. innerhalb des Gehäuses oder Chips) mit den Kernen kommunizieren. Bei mindestens einer Ausführungsform können die Prozessorkerne unabhängig von der Art des Anschlusses der GPU der GPU Arbeit in Form von Befehlsfolgen/Befehlen zuweisen, die in einem Arbeitsdeskriptor enthalten sind. Bei mindestens einer Ausführungsform verwendet die GPU dann dedizierte Schaltkreise/Logiken zur effizienten Verarbeitung dieser Befehle/Anweisungen.

[0405] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform kann eine Ableitungs- und/oder Trainingslogik 1815 in dem Grafikmultiprozessor 3134 zum Ableiten oder Vorhersagen von Operationen mindestens zum Teil basierend auf Gewichtungsparametern verwendet werden, die unter Verwendung von neuronalen Netzwerktrainingsoperationen, Funktionen und/oder Architekturen von neuronalen Netzwerken oder in dieser Schrift beschriebenen Anwendungsfällen von neuronalen Netzwerken berechnet werden.

[0406] Bei mindestens einer Ausführungsform wird der Grafik-Multiprozessor 3134 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird der Grafik-Multiprozessor 3134 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird der Grafik-Multiprozessor 3134 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und der Grafik-Multiprozessor 3134 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung zu identifizieren eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0407] Fig. 32 zeigt ein Multi-GPU-Rechnersystem 3200 gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform kann das Multi-GPU-Computersystem 3200 einen Prozessor 3202 beinhalten, der über einen Hostschnittstellen-Schalter 3204 an mehrere Allzweck-Grafikverarbeitungseinheiten (GPGPUs) 3206A-D gekoppelt ist. Bei mindestens einer Ausführungsform ist der Host-Schnittstellen-Switch 3204 eine PCI-Express-Switch-Einrichtung, die den Prozessor 3202 mit einem PCI-Express-Bus verbindet, über den der Prozessor 3202 mit den GPGPUs 3206A-D kommunizieren kann. Die GPGPUs 3206A-D können über eine Reihe von Hochgeschwindigkeits-Punkt-zu-Punkt-GPU-zu-GPU-Verbindungen 3216 miteinander verbunden sein. Bei mindestens einer Ausführungsform sind die GPU-zu-GPU-Verbindungen 3216 mit jeder der GPGPUs 3206A-D über eine eigene GPU-Verbindung verbunden. Bei mindestens einer Ausführungsform ermöglichen die P2P-GPU-Verbindungen 3216 direkte Kommunikation zwischen den einzelnen GPGPUs 3206A-D, ohne dass eine Kommunikation über den Host-Schnittstellenbus 3204, an den der Prozessor 3202 angeschlossen ist, erforderlich ist. Bei mindestens einer Ausführungsform, bei der der GPU-zu-GPU-Verkehr auf P2P-GPU-Verbindungen 3216 geleitet wird, bleibt der Host-Schnittstellenbus 3204 für den Systemspeicherzugriff oder für die Kommunikation mit anderen Instanzen des Multi-GPU-Computersystems 3200 verfügbar, zum Beispiel über eine oder mehrere Netzwerkeinrichtungen. Während bei mindestens einer Ausführungsform die GPGPUs 3206A-D mit dem Prozessor 3202 über den Host-Schnittstellen-Switch 3204 verbunden sind, weist der Prozessor 3202 bei mindestens einer Ausführungsform eine direkte Unterstützung für P2P-GPU-Verbindungen 3216 auf und kann direkt mit den GPGPUs 3206A-D verbunden sein.

[0408] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B**

bereitgestellt. Bei mindestens einer Ausführungsform kann eine Ableitungs- und/oder Trainingslogik 1815 in dem Multi-GPU-Computersystem 3200 zum Ableiten oder Vorhersagen von Operationen mindestens zum Teil basierend auf Gewichtungsparametern verwendet werden, die unter Verwendung von neuronalen Netzwerktrainingsoperationen, Funktionen und/oder Architekturen von neuronalen Netzwerken oder in dieser Schrift beschriebenen Anwendungsfällen von neuronalen Netzwerken berechnet wurden.

[0409] Bei mindestens einer Ausführungsform wird das Multi-GPU-Rechensystem 3200 verwendet, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird das Multi-GPU-Rechensystem 3200 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird das Multi-GPU-Rechensystem 3200 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und das Multi-GPU-Rechensystem 3200 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere Neuronen trainieren Netzwerke, um eine Ausrichtung eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0410] Fig. 33 ist ein Blockdiagramm eines Grafikprozessors 3300 gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform beinhaltet der Grafikprozessor 3300 eine Ring-Interconnect 3302, ein Pipeline-Frontend 3304, eine Media-Engine 3337 und Grafikkerne 3380A-3380N. Bei mindestens einer Ausführungsform verbindet die Ringverbindung 3302 den Grafikprozessor 3300 mit anderen Verarbeitungseinheiten, einschließlich anderer Grafikprozessoren oder eines oder mehrerer Mehrzweckprozessorkerne. Bei mindestens einer Ausführungsform ist der Grafikprozessor 3300 einer von vielen Prozessoren, die in ein Mehrkern-Verarbeitungssystem integriert sind.

[0411] Bei mindestens einer Ausführungsform empfängt der Grafikprozessor 3300 Stapel von Befehlen über die Ringverbindung 3302. Bei mindestens einer Ausführungsform werden die eingehenden Befehle von einem Befehls-Streamer 3303 im Pipeline-Frontend 3304 interpretiert. Bei mindestens einer Ausführungsform weist der Grafikprozessor 3300 eine skalierbare Ausführungslogik auf, um die 3D-Geometrieverarbeitung und die Medienverarbeitung über den/die Grafikkern(e) 3380A-3380N durchzuführen. Bei mindestens einer Ausführungsform liefert der Befehls-Streamer 3303 für 3D-Geometrieverarbeitungsbefehle Befehle an die Geometrie-Pipeline 3336. Bei mindestens einer Ausführungsform liefert der Befehls-Streamer 3303 für mindestens einige Medienverarbeitungsbefehle Befehle an ein Video-Frontend 3334, das mit einer Medien-Maschine 3337 koppelt. Bei mindestens einer Ausführungsform beinhaltet die Media-Engine 3337 eine Video-Qualitäts-Engine(VQE) 3330 für die Video- und Bildnachbearbeitung und eine Multi-Format-Encoder/-Decoder-Engine(MFX) 3333, um hardwarebeschleunigte Codierung und Decodierung von Mediendaten zu ermöglichen. Bei mindestens einer Ausführungsform erzeugen die Geometriepipeline 3336 und die Medien-Engine 3337 jeweils Ausführungs-Threads für Thread-Ausführungsressourcen, die von mindestens einem Grafikkern 3380A bereitgestellt werden.

[0412] Bei mindestens einer Ausführungsform weist der Grafikprozessor 3300 skalierbare Thread-Ausführungsressourcen auf, die modulare Kerne 3380A-3380N (manchmal als Kern-Slices bezeichnet) aufweisen, von denen jeder mehrere Sub-Kerne 3350A-550N, 3360A-3360N (manchmal als Kern-Sub-Slices bezeichnet) hat. Bei mindestens einer Ausführungsform kann der Grafikprozessor 3300 eine beliebige Anzahl von Grafikkernen 3380A bis 3380N aufweisen. Bei mindestens einer Ausführungsform beinhaltet der Grafikprozessor 3300 einen Grafikkern 3380A, der mindestens einen ersten Teilkern 3350A und einen zweiten Teilkern 3360A aufweist. Bei mindestens einer Ausführungsform ist der Grafikprozessor 3300 ein stromsparender Prozessor mit einem einzigen Teilkern (z. B. 3350A). Bei mindestens einer Ausführungsform weist der Grafikprozessor 3300 mehrere Grafikkerne 3380A-3380N auf, von denen jeder einen Satz von ersten Sub-Kernen 3350A-3350N und einen Satz von zweiten Sub-Kernen 3360A-3360N aufweist. Bei mindestens einer Ausführungsform beinhaltet jeder Teilkern in den ersten Teilkernen 3350A-3350N mindestens einen ersten Satz von Ausführungseinheiten 3352A-3352N und Medien-/Textur-Samplern 3354A-3354N. Bei mindestens einer Ausführungsform beinhaltet jeder Teilkern in den zweiten Teilkernen 3360A-3360N mindestens einen zweiten Satz von Ausführungseinheiten 3362A-3362N und Samplern 3364A-3364N. Bei mindestens einer Ausführungsform teilt sich jeder Sub-Kern 3350A-3350N, 3360A-3360N einen Satz gemeinsam genutzter Ressourcen 3370A-3370N. Bei mindestens einer Ausführungsform weisen die gemeinsam genutzten Ressourcen einen gemeinsamen Cache-Speicher und eine Pixeloperationslogik auf.

[0413] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform kann eine Ableitungs- und/oder Trainingslogik 1815 in dem Prozessor 3300 zum Ableiten oder Vorhersagen von Operationen mindestens zum Teil basierend auf Gewichtungsparametern verwendet werden, die unter Verwendung von neuronalen Netzwerktrainingsoperationen, Funktionen und/oder Architekturen von neuronalen Netzwerken oder in dieser Schrift beschriebenen Anwendungsfällen von neuronalen Netzwerken berechnet wurden.

[0414] Bei mindestens einer Ausführungsform wird der Grafikprozessor 3300 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird der Grafikprozessor 3300 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird der Grafikprozessor 3300 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und der Grafikprozessor 3300 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0415] **Fig. 34** ist ein Blockdiagramm, das die Mikroarchitektur eines Prozessors 3400 veranschaulicht, der logische Schaltungen zur Ausführung von Befehlen gemäß mindestens einer Ausführungsform aufweisen kann. Bei mindestens einer Ausführungsform kann der Prozessor 3400 Befehle ausführen, die x86-Befehle, ARM-Befehle, spezielle Befehle für anwendungsspezifische integrierte Schaltungen (ASICs) usw. aufweisen. Bei mindestens einer Ausführungsform kann der Prozessor 3410 Register zum Speichern gepackter Daten aufweisen, wie z. B. 64 Bit breite MMXTM-Register in Mikroprozessoren, die mit der MMX-Technologie der Intel Corporation aus Santa Clara, Kalifornien, ausgestattet sind. Bei mindestens einer Ausführungsform können MMX-Register, die sowohl als Ganzzahl- als auch als Gleitkommaregister verfügbar sind, mit gepackten Datenelementen arbeiten, die mit SIMD- (Single Instruction, Multiple Data) und SSE- (Streaming SIMD Extensions) Anweisungen einhergehen. Bei mindestens einer Ausführungsform können 128 Bit breite XMM-Register, die sich auf SSE2-, SSE3-, SSE4-, AVX- oder darüber hinausgehende Technologien beziehen (allgemein als „SSEx“ bezeichnet), solche gepackten Datenoperanden enthalten. Bei mindestens einer Ausführungsform können die Prozessoren 3410 Anweisungen zur Beschleunigung von Algorithmen für maschinelles Lernen oder Deep Learning, Training oder Inferencing ausführen.

[0416] Bei mindestens einer Ausführungsform weist der Prozessor 3400 ein In-Order-Front-End („Front-End“) 3401 auf, um auszuführende Befehle abzurufen und Befehle vorzubereiten, die später in der Prozessor-Pipeline verwendet werden sollen. Bei mindestens einer Ausführungsform kann das Frontend 3401 mehrere Einheiten aufweisen. Bei mindestens einer Ausführungsform holt ein Befehls-Prefetcher 3426 Befehle aus dem Speicher und leitet sie an einen Befehlsdecodierer 3428 weiter, der wiederum Befehle decodiert oder interpretiert. Bei mindestens einer Ausführungsform decodiert der Befehlsdecodierer 3428 beispielsweise einen empfangenen Befehl in eine oder mehrere Operationen, die als „Mikrobefehle“ oder „Mikrooperationen“ (auch „Mikro-Ops“ oder „Uops“ genannt) bezeichnet werden und von der Maschine ausgeführt werden können. Bei mindestens einer Ausführungsform zerlegt der Befehlsdecodierer 3428 den Befehl in einen Op-Code und entsprechende Daten- und Steuerfelder, die von der Mikroarchitektur zur Durchführung von Operationen gemäß mindestens einer Ausführungsform verwendet werden können. Bei mindestens einer Ausführungsform kann ein Trace-Cache 3430 decodierte uops zu programmgeordneten Sequenzen oder Traces in einer uop-Warteschlange 3434 zur Ausführung zusammenstellen. Bei mindestens einer Ausführungsform, wenn der Trace-Cache 3430 auf eine komplexe Anweisung stößt, stellt ein Mikrocode-ROM 3432 die für den Abschluss der Operation erforderlichen uops bereit.

[0417] Bei mindestens einer Ausführungsform können einige Anweisungen in eine einzige Mikro-OP umgewandelt werden, während andere mehrere Mikro-Ops benötigen, um den Betrieb vollständig abzuschließen. Bei mindestens einer Ausführungsform kann der Befehlsdecodierer 3428 auf den Mikrocode-ROM 3432 zugreifen, um den Befehl auszuführen, wenn für die Ausführung eines Befehls mehr als vier Mikro-Ops erforderlich sind. Bei mindestens einer Ausführungsform kann ein Befehl in eine kleine Anzahl von Mikro-Ops zur Verarbeitung im Befehlsdecodierer 3428 decodiert werden. Bei mindestens einer Ausführungsform kann ein Befehl im Mikrocode-ROM 3432 gespeichert sein, falls eine Anzahl von Mikro-OPs zur Ausführung des Vorgangs erforderlich ist. Bei mindestens einer Ausführungsform bezieht sich der Trace-Cache 3430 auf ein pro-

grammierbare Logik-Anordnung („PLA“) als Einstiegspunkt, um einen korrekten Mikrobefehlszeiger für das Lesen von Mikrocode-Sequenzen zur Vervollständigung eines oder mehrerer Befehle aus dem Mikrocode-ROM 3432 zu bestimmen. Bei mindestens einer Ausführungsform kann das Front-End 3401 einer Maschine, nachdem der Mikrocode-ROM 3432 die Sequenzierung von Mikro-Ops für eine Anweisung beendet hat, das Abrufen von Mikro-Ops aus dem Ablaufverfolgungszwischenspeicher 3430 wieder aufnehmen.

[0418] Bei mindestens einer Ausführungsform kann die Out-of-Order-Ausführungs-Maschine (Out-of-Order-Engine”) 3403 Befehle für die Ausführung vorbereiten. Bei mindestens einer Ausführungsform beinhaltet die Out-of-order-Ausführungslogik eine Anzahl von Puffern zum Glätten und Neuordnen des Flusses von Anweisungen zum Optimieren der Leistung, wenn sie die Pipeline herunter gehen und zur Ausführung geplant werden, beinhaltet die Out-of-order-Ausführungs-Engine 3403 ohne Einschränkung einen Zuordner/Registerumbenenner 3440, eine Speicher-uop-Warteschlange 3442, eine Ganzzahl/Gleitkomma-uop-Warteschlange 3444, einen Speicherplaner 3446, einen schnellen Planer 3402, einen langsamen/allgemeinen Gleitkomma-Planer („slow/general FP scheduler“) 3404 und einen einfachen Gleitkomma-Planer („simple FP scheduler“) 3406. Bei mindestens einer Ausführungsform werden der schnelle Scheduler 3402, der langsame/allgemeine Gleitkomma-Scheduler 3404 und der einfache Gleitkomma-Scheduler 3406 hier auch gemeinsam als „uop-Scheduler 3402, 3404, 3406“ bezeichnet. Der Zuordner/Registerumbenenner 3440 ordnet Maschinenpuffer und Ressourcen zu, die jede uop benötigt, um ausgeführt zu werden. Bei mindestens einer Ausführungsform benennt der Allokator/Register-Renamer 3440 logische Register auf Einträge in einer Registerdatei um. Bei mindestens einer Ausführungsform weist der Allokator/Register-Renamer 3440 außerdem jedem uop einen Eintrag in einer von zwei uop-Warteschlangen zu, der Speicher-uop-Warteschlange 3442 für Speicheroperationen und der Ganzzahl-/Gleitkomma-uop-Warteschlange 3444 für Nicht-Speicheroperationen, und zwar vor dem Speicher-Scheduler 3446 und den uop-Schedulern 3402, 3404, 3406. Bei mindestens einer Ausführungsform bestimmen die uop-Scheduler 3402, 3404, 3406 auf der Grundlage der Bereitschaft ihrer abhängigen Eingangsregister-Operandenquellen und der Verfügbarkeit der Ausführungsressourcen, die die uops für den Abschluss ihrer Operation benötigen, wann ein uop zur Ausführung bereit ist. Bei mindestens einer Ausführungsform kann der schnelle Scheduler 3402 bei jeder Hälfte des Haupttaktzyklus einplanen, während der langsame/allgemeine Gleitkomma-Scheduler 3404 und der einfache Gleitkomma-Scheduler 3406 einmal pro Hauptprozessortaktzyklus einplanen können. Bei mindestens einer Ausführungsform vermitteln die uop-Scheduler 3402, 3404, 3406 für Dispatch-Anschlüsse, um uops für die Ausführung zu planen.

[0419] Bei mindestens einer Ausführungsform beinhaltet der Ausführungsblock b11 ohne Einschränkung ein Ganzzahl-Registerdatei/Umgehungs-Netzwerk 3408, ein(e) Gleitkommaregisterdatei/Umgehungsnetzwerk („eine FP-Registerdatei/Umgehungs-Netzwerk“) 3410, Adresserzeugungseinheiten („AGUs“) 3412 und 3414, schnelle arithmetische Logikeinheiten (ALUs) („schnelle ALUs“) 3416 und 3418, eine langsame arithmetische Logikeinheit („langsame ALU“) 3420, eine Gleitkomma-ALU („FP“) 3422 und eine Gleitkomma-Bewegungseinheit („FP-Bewegung“) 3424. Bei mindestens einer Ausführungsform werden ein Ganzzahl-Registerdatei/Umgehungs-Netzwerk 3408 und ein Gleitkomma-Registerdatei/Umgehungs-Netzwerk 3410 hier auch als „Registerdateien 3408, 3410“ bezeichnet. Bei mindestens einer Ausführungsform werden die AGUSs 3412 und 3414, die schnellen ALUs 3416 und 3418, die langsame ALU 3420, die Gleitkomma-ALU 3422 und die Gleitkomma-Bewegungseinheit 3424 hier auch als „Ausführungseinheiten 3412, 3414, 3416, 3418, 3420, 3422 und 3424“ bezeichnet. Bei mindestens einer Ausführungsform kann der Ausführungsblock b11 ohne Einschränkung eine beliebige Anzahl (einschließlich Null) und Art von Registerdateien, Umgehungsnetzwerken, Adresserzeugungseinheiten und Ausführungseinheiten in beliebiger Kombination aufweisen.

[0420] Bei mindestens einer Ausführungsform können die Registerdateien 3408, 3410 zwischen den uop-Schedulern 3402, 3404, 3406 und den Ausführungseinheiten 3412, 3414, 3416, 3418, 3420, 3422 und 3424 angeordnet sein. Bei mindestens einer Ausführungsform führt das Integer-Registerdatei/Umgehungs-Netzwerk 3408 Integer-Operationen durch. Bei mindestens einer Ausführungsform führt das Gleitkommaregisterdatei/ Umgehungs-Netzwerk 3410 Gleitkommaoperationen durch. Bei mindestens einer Ausführungsform kann jede der Registerdateien 3408, 3410 ohne Einschränkung ein Umgehungsnetzwerk aufweisen, das gerade abgeschlossene Ergebnisse, die noch nicht in die Registerdatei geschrieben wurden, umleiten oder an neue abhängige Uops weiterleiten kann. Bei mindestens einer Ausführungsform können die Registerdateien 3408, 3410 Daten miteinander austauschen. Bei mindestens einer Ausführungsform kann das Integer-Registerdatei/Umgehungs-Netzwerk 3408 ohne Einschränkung zwei separate Registerdateien aufweisen, eine Registerdatei für zweiunddreißig Bits von Daten niedriger Ordnung und eine zweite Registerdatei für zweiunddreißig Bits von Daten hoher Ordnung. Bei mindestens einer Ausführungsform kann das Gleitkomma-Registerdatei/Umgehungs-Netzwerk 3410 ohne Einschränkung 128 Bit breite Einträge aufweisen, da Gleitkomma-Befehle typischerweise Operanden mit einer Breite von 64 bis 128 Bit aufweisen.

[0421] Bei mindestens einer Ausführungsform können die Ausführungseinheiten 3412, 3414, 3416, 3418, 3420, 3422, 3424 Befehle ausführen. Bei mindestens einer Ausführungsform speichern Registerdateien 3408, 3410 Ganzzahl- und Gleitkommadaten-Operandenwerte, die für die Ausführung von Mikrobefehlen erforderlich sind. Bei mindestens einer Ausführungsform kann der Prozessor 3400 ohne Einschränkung eine beliebige Anzahl und Kombination von Ausführungseinheiten 3412, 3414, 3416, 3418, 3420, 3422, 3424 aufweisen. Bei mindestens einer Ausführungsform können die Gleitkomma-ALU 3422 und die Gleitkomma-Bewegungseinheit 3424 Gleitkomma-, MMX-, SIMD-, AVX- und SSE- oder andere Operationen ausführen, einschließlich spezieller maschineller Lernbefehle. Bei mindestens einer Ausführungsform kann die Gleitkomma-ALU 3422 ohne Einschränkung einen 64-Bit-durch-64-Bit-Gleitkomma-Teiler aufweisen, um Divisions-, Quadratwurzel- und Restwert-Mikrooperationen auszuführen. Bei mindestens einer Ausführungsform können Befehle, die einen Gleitkommawert beinhalten, mit Gleitkomma-Hardware verarbeitet werden. Bei mindestens einer Ausführungsform können ALU-Operationen an schnelle ALUs 3416, 3418 weitergeleitet werden. Bei mindestens einer Ausführungsform können die schnellen ALUS 3416, 3418 schnelle Operationen mit einer effektiven Latenzzeit von einem halben Taktzyklus ausführen. Bei mindestens einer Ausführungsform gehen die meisten komplexen ganzzahligen Operationen an die langsame ALU 3420, da die langsame ALU 3420 ohne Einschränkung ganzzahlige Ausführungshardware für Operationen mit langer Latenzzeit aufweisen kann, wie z. B. einen Multiplizierer, Schiebeeinheiten, eine Flag-Logik und eine Verzweigungsverarbeitung. Bei mindestens einer Ausführungsform können Speicherlade-/Speicheroperationen von AGUS 3412, 3414 ausgeführt werden. Bei mindestens einer Ausführungsform können die schnelle ALU 3416, die schnelle ALU 3418 und die langsame ALU 3420 Ganzzahloperationen mit 64-Bit-Datenoperanden durchführen. Bei mindestens einer Ausführungsform können die schnelle ALU 3416, die schnelle ALU 3418 und die langsame ALU 3420 umgesetzt werden, um eine Vielfalt von Datenbitgrößen zu unterstützen, darunter sechzehn, zweiunddreißig, 128, 256, usw. Bei mindestens einer Ausführungsform können die Gleitkomma-ALU 3422 und die Gleitkomma-Bewegungseinheit 3424 umgesetzt sein, um einen Bereich von Operanden mit Bits verschiedener Breiten zu unterstützen. Bei mindestens einer Ausführungsform können die Gleitkomma-ALU 3422 und die Gleitkomma-Bewegungseinheit 3424 mit 128 Bit breiten gepackten Datenoperanden in Verbindung mit SIMD- und Multimedia-Anweisungen arbeiten.

[0422] Bei mindestens einer Ausführungsform leiten die uop-Scheduler 3402, 3404, 3406 abhängige Operationen ein, bevor die Ausführung der übergeordneten Last beendet ist. Bei mindestens einer Ausführungsform kann der Prozessor 3400, da uops spekulativ geplant und im Prozessor 3400 ausgeführt werden können, auch eine Logik zur Behandlung von Speicherfehlern aufweisen. Bei mindestens einer Ausführungsform kann es, wenn ein Datenladen in den Datencache fehlerhaft ist, abhängige Operationen in der Pipeline geben, die den Scheduler mit vorübergehend falschen Daten verlassen haben. Bei mindestens einer Ausführungsform verfolgt ein Wiederholungsmechanismus die Anweisungen, die falsche Daten verwenden, und führt sie erneut aus. Bei mindestens einer Ausführungsform kann es erforderlich sein, abhängige Operationen erneut auszuführen, während unabhängige Operationen zu Ende geführt werden können. Bei mindestens einer Ausführungsform können Scheduler und ein Wiederholungsmechanismus von mindestens einer Ausführungsform eines Prozessors auch so ausgelegt sein, dass sie Befehlssequenzen für Textstring-Vergleichsoperationen abfangen.

[0423] Bei mindestens einer Ausführungsform kann sich der Begriff „Register“ auf prozessorinterne Speicherplätze beziehen, die als Teil von Befehlen verwendet werden können, um Operanden zu identifizieren. Bei mindestens einer Ausführungsform kann es sich bei den Registern um solche handeln, die von außerhalb des Prozessors (aus der Sicht eines Programmierers) verwendet werden können. Bei mindestens einer Ausführungsform können die Register nicht auf einen bestimmten Schaltungstyp beschränkt sein. Vielmehr kann ein Register bei mindestens einer Ausführungsform Daten speichern, Daten bereitstellen und hier beschriebene Funktionen ausführen. Bei mindestens einer Ausführungsform können die hier beschriebenen Register durch Schaltkreise innerhalb eines Prozessors unter Verwendung einer beliebigen Anzahl verschiedener Techniken implementiert sein, wie z. B. dedizierte physische Register, dynamisch zugewiesene physische Register unter Verwendung von Registerumbenennung, Kombinationen aus dedizierten und dynamisch zugewiesenen physischen Registern usw. Bei mindestens einer Ausführungsform werden in Ganzzahlregistern 32-Bit-Ganzzahldaten gespeichert. In mindestens einer Ausführungsform enthält eine Registerdatei auch acht Multimedia-SIMD-Register für gepackte Daten.

[0424] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform können Teile der oder die gesamte Ableitungs- und/oder Trainingslogik 1815 in den Ausführungsblock 3411 und andere gezeigte oder nicht gezeigte Speicher

oder Register integriert sein. Beispielsweise können bei mindestens einer Ausführungsform in dieser Schrift beschriebene Trainings- und/oder Ableitungstechniken eine oder mehrere der in dem Ausführungsblock 3411 veranschaulichten ALUs verwenden. Darüber hinaus können Gewichtungsparameter in chipinternen oder chip externen Speicher und/oder Registern (gezeigt oder nicht gezeigt) gespeichert werden, die ALUs des Ausführungsblocks 3411 konfigurieren, um einen oder mehrere Algorithmen maschinellen Lernens, neuronale Netzwerkarchitekturen, Anwendungsfälle oder in dieser Schrift beschriebene Trainingsmethoden auszuführen.

[0425] Bei mindestens einer Ausführungsform werden ALUs des EXE-Blocks 3411 verwendet, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform werden ALUs des EXE-Blocks 3411 verwendet, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform werden die ALUs des EXE-Blocks 3411 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und die ALUs des EXE-Blocks 3411 werden in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke dazu trainieren, eine Ausrichtung eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0426] Fig. 35 veranschaulicht einen Deep-Learning-Anwendungsprozessor 3500 gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform verwendet der Deep-Learning-Anwendungsprozessor 3500 Anweisungen, die bei Ausführung durch den Deep-Learning-Anwendungsprozessor 3500 bewirken, dass der Deep-Learning-Anwendungsprozessor 3500 einige oder alle der in dieser Offenbarung beschriebenen Prozesse und Methoden ausführt. Bei mindestens einer Ausführungsform ist der Deep-Learning-Anwendungsprozessor 3500 eine anwendungsspezifische integrierte Schaltung (ASIC). Bei mindestens einer Ausführungsform führt der Anwendungsprozessor 3500 Matrixmultiplikationsoperationen aus, die entweder als Ergebnis der Ausführung einer oder mehrerer Anweisungen oder beider in Hardware „fest verdrahtet“ sind. Bei mindestens einer Ausführungsform beinhaltet der Deep-Learning-Anwendungsprozessor 3500 ohne Einschränkung Verarbeitungs-Cluster 3510(1)-3510(12), Inter-Chip-Links („ICLs“) 3520(1)-3520(12), Inter-Chip-Steuerungen (Inter-Chip Controllers-„ICCs“) 3530(1)-3530(2), Speicher mit hoher Bandbreite der zweiten Generation („HBM2“) 3540(1)-3540(4), Speichersteuervorrichtungen („Mem Ctrlrs“) 3542(1)-3542(4), eine physische Speicherschicht mit hoher Bandbreite („HBM PHY“) 3544(1)-3544(4), eine Verwaltungssteuervorrichtungszentralverarbeitungseinheit („Verwaltungssteuervorrichtungs-CPU“) 3550, einen seriellen peripheren Schnittstellen-, Inter-integrierten Schaltungs- und Allzweck-Eingabe-/Ausgabe-Block („SPI, I2C, GPIO“) 3560, eine Steuervorrichtung zur schnellen Verbindung von Peripheriegerätekomponenten und einen Block zum direkten Speicherzugriff („PCIe Controller und DMA“) 3570 sowie einen sechzehnspurigen Anschluss zur schnellen Verbindung von Peripheriegerätekomponenten („PCI Express x 16“) 3580.

[0427] Bei mindestens einer Ausführungsform können Verarbeitungs-Cluster 3510 Deep-Learning-Operationen, die Ableitungs- oder Vorhersageoperationen beinhalten, basierend auf Gewichtungsparametern ausführen, die eine oder mehrere Trainingsmethoden berechnet haben, einschließlich der hierin beschriebenen. Bei mindestens einer Ausführungsform kann jeder Verarbeitungs-Cluster 3510 ohne Einschränkung eine beliebige Anzahl und Art von Prozessoren beinhalten. Bei mindestens einer Ausführungsform kann der Deep-Learning-Anwendungsprozessor 3500 eine beliebige Anzahl und Art von Verarbeitungs-Clustern 3500 beinhalten. Bei mindestens einer Ausführungsform sind Inter-Chip-Links 3520 bidirektional. Bei mindestens einer Ausführungsform ermöglichen die Inter-Chip-Links 3520 und die Inter-Chip-Steuervorrichtung 3530 mehreren Deep-Learning-Anwendungsprozessoren 3500 den Austausch von Informationen, die Aktivierungsinformationen beinhalten, die sich aus der Ausführung eines oder mehrerer Algorithmen des maschinellen Lernens ergeben, die in einem oder mehreren neuronalen Netzwerken verkörpert sind. Bei mindestens einer Ausführungsform kann der Deep-Learning-Anwendungsprozessor 3500 eine beliebige Anzahl (die Null beinhaltet) und Art von ICLs 3520 und ICCs 3530 beinhalten.

[0428] Bei mindestens einer Ausführungsform stellen HBM2s 3540 insgesamt 32 Gigabyte (GB) Speicher bereit. Der HBM2 3540(i) ist sowohl mit einer Speichersteuervorrichtung 3542(i) als auch einer HBM PHY 3544(i) verknüpft. Bei mindestens einer Ausführungsform kann eine beliebige Anzahl von HBM2s 3540 eine beliebige Art und Gesamtmenge von Speicher mit hoher Bandbreite bereitstellen und kann mit einer beliebigen Anzahl (die Null beinhaltet) und Art von Speichersteuerungen 3542 und HBM-PHYs 3544 assoziiert sein. Bei mindestens einer Ausführungsform können SPI, I2C, GPIO 3560, PCIe-Steuervorrichtung und

DMA 3570 und/oder PCIe 3580 durch eine beliebige Anzahl und Art von Blöcken ersetzt werden, die eine beliebige Anzahl und Art von Kommunikationsstandards auf eine beliebige technisch machbare Weise ermöglichen.

[0429] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform wird der Deep-Learning-Anwendungsprozessor verwendet, um ein Modell maschinellen Lernens, wie etwa ein neuronales Netzwerk, zu trainieren, um dem Deep-Learning-Anwendungsprozessor 3500 bereitgestellte Informationen vorherzusagen oder abzuleiten. Bei mindestens einer Ausführungsform wird der Deep-Learning-Anwendungsprozessor 3500 verwendet, um Informationen basierend auf eines trainierten Modells maschinellen Lernens (z. B. neuronales Netzwerk) abzuleiten oder vorherzusagen, das von einem anderen Prozessor oder System oder von dem Deep-Learning-Anwendungsprozessor 3500 trainiert wurde. Bei mindestens einer Ausführungsform kann der Prozessor 3500 verwendet werden, um einen oder mehrere in dieser Schrift beschriebene Anwendungsfälle für neuronale Netzwerke auszuführen.

[0430] Bei mindestens einer Ausführungsform wird der Prozessor 3500 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird der Prozessor 3500 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird der Prozessor 3500 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und der Prozessor 3500 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes selbstüberwacht zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0431] **Fig. 36** ist ein Blockdiagramm eines neuromorphen Prozessors 3600 gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform kann der neuromorphe Prozessor 3600 eine oder mehrere Eingaben von Quellen außerhalb des neuromorphen Prozessors 3600 empfangen. Bei mindestens einer Ausführungsform können diese Eingaben an ein oder mehrere Neuronen 3602 innerhalb des neuromorphen Prozessors 3600 übermittelt werden. Bei mindestens einer Ausführungsform können Neuronen 3602 und Komponenten davon unter Verwendung von Schaltungen oder Logik umgesetzt sein, die eine oder mehrere arithmetische Logikeinheiten (ALUs) beinhalten. Bei mindestens einer Ausführungsform kann der neuromorphe Prozessor 3600 ohne Einschränkung Tausende oder Millionen von Instanzen von Neuronen 3602 beinhalten, aber es kann eine beliebige geeignete Anzahl von Neuronen 3602 verwendet werden. Bei mindestens einer Ausführungsform kann jede Instanz des Neurons 3602 einen Neuroneneingang 3604 und einen Neuronenausgang 3606 beinhalten. Bei mindestens einer Ausführungsform können die Neuronen 3602 Ausgaben erzeugen, die an Eingänge anderer Instanzen von Neuronen 3602 übermittelt werden können. Beispielsweise können bei mindestens einer Ausführungsform Neuroneneingänge 3604 und Neuronenausgänge 3606 über Synapsen 3608 miteinander verbunden sein.

[0432] Bei mindestens einer Ausführungsform können die Neuronen 3602 und die Synapsen 3608 derartig miteinander verbunden sein, dass der neuromorphe Prozessor 3600 arbeitet, um von dem neuromorphen Prozessor 3600 empfangene Informationen zu verarbeiten oder zu analysieren. Bei mindestens einer Ausführungsform können die Neuronen 3602 einen Ausgangsimpuls (oder „Feuer“ oder „Spitze“) übermitteln, wenn durch den Neuroneneingang 3604 empfangene Eingaben einen Schwellenwert überschreiten. Bei mindestens einer Ausführungsform können die Neuronen 3602 an den Neuroneneingängen 3604 empfangene Signale summieren oder integrieren. Beispielsweise können die Neuronen 3602 bei mindestens einer Ausführungsform als undichte Integrate-and-Fire-Neuronen umgesetzt sein, wobei, falls eine Summe (als ein „Membranpotenzial“ bezeichnet) einen Schwellenwert überschreitet, das Neuron 3602 eine Ausgabe (oder „Feuer“) unter Verwendung einer Übertragungsfunktion, wie etwa einer Sigmoid- oder Schwellenfunktion, überschreitet. Bei mindestens einer Ausführungsform kann ein undichtes Integrate-and-Fire-Neuron Signale, die an Neuroneneingängen 3604 empfangen werden, zu einem Membranpotenzial summieren und kann auch einen Abklingfaktor (oder Leck) anwenden, um ein Membranpotenzial zu verringern. Bei mindestens einer Ausführungsform kann ein undichtes Integrate-and-Fire-Neuron feuern, falls mehrere Eingangssignale an den Neuroneneingängen 3604 schnell genug empfangen werden, um einen Schwellenwert zu überschreiten (d. h. bevor ein Membranpotenzial zu weit abfällt, um zu feuern). Bei mindestens einer Ausführungsform

können die Neuronen 3602 unter Verwendung von Schaltungen oder Logik umgesetzt sein, die Eingaben empfangen, Eingaben in ein Membranpotenzial integrieren und ein Membranpotenzial abklingen lassen. Bei mindestens einer Ausführungsform können Eingaben gemittelt werden, oder es kann eine beliebige andere geeignete Übertragungsfunktion verwendet werden. Darüber hinaus können die Neuronen 3602 bei mindestens einer Ausführungsform ohne Einschränkung Vergleicherschaltungen oder Logik beinhalten, die eine Ausgangsspitze am Neuronenausgang 3606 erzeugen, wenn das Ergebnis des Anwendens einer Übertragungsfunktion auf den Neuroneneingang 3604 einen Schwellenwert überschreitet. Bei mindestens einer Ausführungsform kann das Neuron 3602, sobald es feuert, zuvor empfangene Eingabeinformationen ignorieren, indem es zum Beispiel ein Membranpotenzial auf 0 oder einen anderen geeigneten Standardwert zurücksetzt. Bei mindestens einer Ausführungsform kann das Neuron 3602 nach einem geeigneten Zeitraum (oder Refraktärzeitraum) den normalen Betrieb wieder aufnehmen, sobald das Membranpotenzial auf 0 zurückgesetzt ist.

[0433] Bei mindestens einer Ausführungsform können die Neuronen 3602 durch Synapsen 3608 miteinander verbunden sein. Bei mindestens einer Ausführungsform können die Synapsen 3608 arbeiten, um Signale von einem Ausgang eines ersten Neurons 3602 an einen Eingang eines zweiten Neurons 3602 zu übermitteln. Bei mindestens einer Ausführungsform können die Neuronen 3602 Informationen über mehr als eine Instanz der Synapse 3608 übermitteln. Bei mindestens einer Ausführungsform können eine oder mehrere Instanzen des Neuronenausgangs 3606 über eine Instanz der Synapse 3608 mit einer Instanz des Neuroneneingangs 3604 in demselben Neuron 3602 verbunden sein. Bei mindestens einer Ausführungsform kann eine Instanz des Neurons 3602, die eine über eine Instanz der Synapse 3608 zu übermittelnde Ausgabe erzeugt, in Bezug auf diese Instanz der Synapse 3608 als ein „präsynaptisches Neuron“ bezeichnet werden. Bei mindestens einer Ausführungsform kann eine Instanz des Neurons 3602, die eine über eine Instanz der Synapse 3608 übermittelte Eingabe empfängt, in Bezug auf diese Instanz der Synapse 3608 als ein „postsynaptisches Neuron“ bezeichnet werden. Da eine Instanz des Neurons 3602 Eingaben von einer oder mehreren Instanzen der Synapse 3608 empfangen kann und auch Ausgaben über eine oder mehrere Instanzen der Synapse 3608 übermitteln kann, kann eine einzelne Instanz des Neurons 3602 daher sowohl ein „präsynaptisches Neuron“ als auch ein „postsynaptisches Neuron“ in Bezug auf verschiedene Instanzen von Synapsen 3608 bei mindestens einer Ausführungsform sein.

[0434] Bei mindestens einer Ausführungsform können die Neuronen 3602 in einer oder mehreren Schichten organisiert sein. Jede Instanz des Neurons 3602 kann einen Neuronenausgang 3606 aufweisen, der durch eine oder mehrere Synapsen 3608 zu einem oder mehreren Neuroneneingängen 3604 auffächern kann. Bei mindestens einer Ausführungsform können Neuronenausgänge 3606 von Neuronen 3602 in einer ersten Schicht 3610 mit Neuroneneingängen 3604 von Neuronen 3602 in einer zweiten Schicht 3612 verbunden sein. Bei mindestens einer Ausführungsform kann die Schicht 3610 als eine „vorwärtsgekoppelte Schicht“ bezeichnet werden. Bei mindestens einer Ausführungsform kann jede Instanz des Neurons 3602 in einer Instanz der ersten Schicht 3610 zu jeder Instanz des Neurons 3602 in der zweiten Schicht 3612 auffächern. Bei mindestens einer Ausführungsform kann die erste Schicht 3610 als eine „vollständig verbundene vorwärtsgekoppelte Schicht“ bezeichnet werden. Bei mindestens einer Ausführungsform kann jede Instanz des Neurons 3602 in einer Instanz der zweiten Schicht 3612 auf weniger als alle Instanzen des Neurons 3602 in einer dritten Schicht 3614 auffächern. Bei mindestens einer Ausführungsform kann die zweite Schicht 3612 als „kaum verbundene vorwärtsgekoppelte Schicht“ bezeichnet sein. Bei mindestens einer Ausführungsform können sich Neuronen 3602 in der zweiten Schicht 3612 zu Neuronen 3602 in mehreren anderen Schichten auffächern, was auch zu Neuronen 3602 in der (gleichen) zweiten Schicht 3612 beinhaltet. Bei mindestens einer Ausführungsform kann die zweite Schicht 3612 als „wiederkehrende Schicht“ bezeichnet werden. Der neuromorphe Prozessor 3600 kann ohne Einschränkung eine beliebige geeignete Kombination von wiederkehrenden Schichten und vorwärtsgekoppelten Schichten beinhalten, die ohne Einschränkung sowohl kaum verbundene vorwärtsgekoppelte Schichten als auch vollständig verbundene vorwärtsgekoppelte Schichten beinhalten.

[0435] Bei mindestens einer Ausführungsform kann der neuromorphe Prozessor 3600 ohne Einschränkung eine umkonfigurierbare Interconnect-Architektur oder dedizierte fest verdrahtete Interconnects beinhalten, um die Synapse 3608 mit den Neuronen 3602 zu verbinden. Bei mindestens einer Ausführungsform kann der neuromorphe Prozessor 3600 ohne Einschränkung eine Schaltung oder Logik beinhalten, die es Synapsen ermöglicht, unterschiedlichen Neuronen 3602 nach Bedarf basierend auf der Topologie des neuronalen Netzwerks und dem Eingangs-/Ausgangslastfaktor von Neuronen zugeordnet zu werden. Beispielsweise können die Synapsen 3608 bei mindestens einer Ausführungsform mit Neuronen 3602 unter Verwendung eine Interconnect-Fabric, wie etwa Network-on-Chip, oder mit dedizierten Verbindungen verbunden sein. Bei

mindestens einer Ausführungsform können Synapsenverbindungen und Komponenten davon unter Verwendung von Schaltungen oder Logik umgesetzt sein.

[0436] Bei mindestens einer Ausführungsform wird der neuromorphe Prozessor 3600 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird der neuromorphe Prozessor 3600 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird der neuromorphe Prozessor 3600 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und der neuromorphe Prozessor 3600 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0437] Fig. 37 ist ein Blockdiagramm eines Verarbeitungssystems gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform weist das System 3700 einen oder mehrere Prozessoren 3702 und einen oder mehrere Grafikprozessoren 3708 auf und kann ein Einzelprozessor-Desktop-System, ein Multiprozessor-Workstation-System oder ein Server-System mit einer großen Anzahl von Prozessoren 3702 oder Prozessorkernen 3707 sein. Bei mindestens einer Ausführungsform ist das System 3700 eine Verarbeitungsplattform, die in eine integrierte System-on-a-Chip (SoC)-Schaltung zur Verwendung in mobilen, tragbaren oder eingebetteten Einrichtungen integriert ist.

[0438] Bei mindestens einer Ausführungsform kann das System 3700 eine serverbasierte Spielplattform, eine Spielkonsole, einschließlich einer Spiel- und Medienkonsole, eine mobile Spielkonsole, eine Handheld-Spielkonsole oder eine Online-Spielkonsole aufweisen oder darin integriert sein. Bei mindestens einer Ausführungsform ist das System 3700 ein Mobiltelefon, ein Smartphone, eine Recheneinrichtung für Tablets oder eine mobile Interneteinrichtung. Bei mindestens einer Ausführungsform kann das Verarbeitungssystem 3700 auch eine tragbare Einrichtung aufweisen, mit dieser gekoppelt oder in diese integriert sein, wie z. B. eine tragbare Einrichtung für eine intelligente Uhr, eine intelligente Brille, eine Augmented-Reality-Einrichtung oder eine Virtual-Reality-Einrichtung. Bei mindestens einer Ausführungsform ist das Verarbeitungssystem 3700 eine Fernseh- oder Set-Top-Box-Einrichtung mit einem oder mehreren Prozessoren 3702 und einer grafischen Schnittstelle, die von einem oder mehreren Grafikprozessoren 3708 erzeugt ist.

[0439] Bei mindestens einer Ausführungsform weisen ein oder mehrere Prozessoren 3702 jeweils einen oder mehrere Prozessorkerne 3707 auf, um Befehle zu verarbeiten, die, wenn sie ausgeführt werden, Operationen für System- und Benutzersoftware durchführen. Bei mindestens einer Ausführungsform ist jeder von einem oder mehreren Prozessorkernen 3707 so ausgestaltet, dass er einen bestimmten Befehlssatz 3709 verarbeitet. Bei mindestens einer Ausführungsform kann der Befehlssatz 3709 das Complex Instruction Set Computing (CISC), das Reduced Instruction Set Computing (RISC) oder das Rechnen über ein Very Long Instruction Word (VLIW) ermöglichen. Bei mindestens einer Ausführungsform können die Prozessorkerne 3707 jeweils einen anderen Befehlssatz 3709 verarbeiten, der Befehle aufweisen kann, um die Emulation anderer Befehlssätze zu erleichtern. Bei mindestens einer Ausführungsform kann der Prozessorkern 3707 auch andere verarbeitende Einrichtungen aufweisen, wie etwa einen digitalen Signalprozessor (DSP).

[0440] Bei mindestens einer Ausführungsform weist der Prozessor 3702 einen Cache-Speicher 3704 auf. Bei mindestens einer Ausführungsform kann der Prozessor 3702 einen einzigen internen Cache oder mehrere Ebenen eines internen Caches aufweisen. Bei mindestens einer Ausführungsform wird der Cache-Speicher von verschiedenen Komponenten des Prozessors 3702 gemeinsam genutzt. Bei mindestens einer Ausführungsform verwendet der Prozessor 3702 auch einen externen Cache (z. B einen Level-3 (L3) Cache oder Last Level Cache (LLC)) (nicht dargestellt), der unter Verwendung bekannter Cache-Kohärenztechniken von den Prozessorkernen 3707 gemeinsam genutzt werden kann. Bei mindestens einer Ausführungsform ist zusätzlich eine Registerdatei 3706 in dem Prozessor 3702 vorhanden, die verschiedene Arten von Registern zur Speicherung unterschiedlicher Datentypen beinhalten kann (z. B. Ganzzahlregister, Gleitkommaregister, Statusregister und ein Anweisungszeigerregister). Bei mindestens einer Ausführungsform kann die Registerdatei 3706 Allzweckregister oder andere Register aufweisen.

[0441] Bei mindestens einer Ausführungsform sind ein oder mehrere Prozessoren 3702 mit einem oder mehreren Schnittstellenbussen 3710 gekoppelt, um Kommunikationssignale, wie etwa Adress-, Daten- oder

Steuersignale, zwischen dem Prozessor 3702 und anderen Komponenten im System 3700 zu übermitteln. Bei mindestens einer Ausführungsform kann der Schnittstellenbus 3710 ein Prozessorbus sein, beispielsweise eine Version eines Direct-Media-Interface-Busses (DMI-Busses). Bei mindestens einer Ausführungsform ist die Schnittstelle 3710 nicht auf einen DMI-Bus beschränkt und kann einen oder mehrere Peripheral Component Interconnect-Busse (z. B PCI, PCI Express), Speicherbusse oder andere Arten von Schnittstellenbussen beinhalten. Bei mindestens einer Ausführungsform beinhalten Prozessor(en) 3702 eine integrierte Speichersteuerung 3716 und einen Plattformsteuervorrichtungs-Hub 3730. Bei mindestens einer Ausführungsform erleichtert die Speichersteuerung 3716 die Kommunikation zwischen einer Speichereinrichtung und anderen Komponenten des Systems 3700, während der Plattform-Controller-Hub (PCH) 3730 Verbindungen zu E/A-Einrichtungen über einen lokalen E/A-Bus bereitstellt.

[0442] Bei mindestens einer Ausführungsform kann die Speichereinrichtung 3720 eine dynamische Direktzugriffsspeichereinrichtung (DRAM), eine statische Direktzugriffsspeichereinrichtung (SRAM), eine Flash-Speichereinrichtung, eine Phasenwechsel-Speichereinrichtung oder eine andere Speichereinrichtung mit geeigneter Leistung sein, um als Prozessspeicher zu dienen. Bei mindestens einer Ausführungsform kann die Speichereinrichtung 3720 als Systemspeicher für das System 3700 arbeiten, um Daten 3722 und Befehle 3721 zur Verwendung zu speichern, wenn ein oder mehrere Prozessoren 3702 eine Anwendung oder einen Prozess ausführen. Bei mindestens einer Ausführungsform ist die Speichersteuerung 3716 auch mit einem optionalen externen Grafikprozessor 3712 gekoppelt, der mit einem oder mehreren Grafikprozessoren 3708 in den Prozessoren 3702 kommunizieren kann, um Grafik- und Medienoperationen durchzuführen. Bei mindestens einer Ausführungsform kann eine Anzeigeeinrichtung 3711 an den (die) Prozessor(en) 3702 angeschlossen sein. Bei mindestens einer Ausführungsform kann die Anzeigevorrichtung 3711 eine oder mehrere interne Anzeigevorrichtungen, wie z. B. in einer mobilen elektronischen Einrichtung oder einem Laptop, oder eine externe Anzeigevorrichtung, die über eine Anzeigeschnittstelle (z. B. DisplayPort usw.) angeschlossen ist, beinhalten. Bei mindestens einer Ausführungsform kann die Anzeigevorrichtung 3711 eine am Kopf montierte Anzeige (HMD) wie eine stereoskopische Anzeigevorrichtung zur Verwendung bei Virtual-Reality-Anwendungen (VR) oder Augmented-Reality-Anwendungen (AR) aufweisen.

[0443] Bei mindestens einer Ausführungsform ermöglicht der Plattformsteuervorrichtungs-Hub 3730 den Anschluss von Peripheriegeräten an die Speichereinrichtung 3720 und dem Prozessor 3702 über einen Hochgeschwindigkeits-E/A-Bus. Bei mindestens einer Ausführungsform beinhalten die E/A-Peripheriegeräte unter anderem eine Audio-Steuervorrichtung 3746, eine Netzwerk-Steuervorrichtung 3734, eine Firmware-Schnittstelle 3728, einen drahtlosen Transceiver 3726, Berührungssensoren 3725 und eine Einrichtung zur Datenspeicherung 3724 (z. B. Festplattenlaufwerk, Flash-Speicher usw.). Bei mindestens einer Ausführungsform kann die Datenspeichereinrichtung 3724 über eine Speicherschnittstelle (z. B. SATA) oder über einen Peripheriebus, wie einen Peripheral Component Interconnect Bus (z. B. PCI, PCI Express), angeschlossen sein. Bei mindestens einer Ausführungsform können die Berührungssensoren 3725 Touchscreen-Sensoren, Drucksensoren oder Fingerabdrucksensoren aufweisen. Bei mindestens einer Ausführungsform kann der drahtlose Transceiver 3726 ein Wi-Fi-Transceiver, ein Bluetooth-Transceiver oder ein Mobilfunk-Transceiver wie ein 3G-, 4G- oder Long Term Evolution (LTE)-Transceiver sein. Bei mindestens einer Ausführungsform ermöglicht die Firmware-Schnittstelle 3728 die Kommunikation mit der System-Firmware und kann z. B. eine einheitliche erweiterbare Firmware-Schnittstelle (UEFI) sein. Bei mindestens einer Ausführungsform kann die Netzwerksteuerung 3734 eine Netzwerkverbindung mit einem kabelgebundenen Netzwerk ermöglichen. Bei mindestens einer Ausführungsform ist eine Hochleistungs-Netzwerksteuerung (nicht dargestellt) mit dem Schnittstellenbus 3710 gekoppelt. Bei mindestens einer Ausführungsform handelt es sich bei der Audio-Steuerung 3746 um eine mehrkanalige High-Definition-Audio-Steuerung. Bei mindestens einer Ausführungsform weist das System 3700 eine optionale Legacy-E/A-Steuerung 3740 zur Kopplung von Legacy-Einrichtungen (z. B. Personal System 2 (PS/2)) mit dem System auf. Bei mindestens einer Ausführungsform kann der Plattformsteuerungs-Hub 3730 auch an eine oder mehrere Universal Serial Bus (USB)-Steuerungen 3742 angeschlossen sein, die Eingabeeinrichtungen wie Tastatur- und Mauskombinationen 3743, eine Kamera 3744 oder andere USB-Eingabeeinrichtungen anschließen.

[0444] Bei mindestens einer Ausführungsform kann eine Instanz der Speichersteuerung 3716 und des Plattformsteuerungs-Hubs 3730 in einen diskreten externen Grafikprozessor, wie den externen Grafikprozessor 3712, integriert sein. Bei mindestens einer Ausführungsform kann der Plattformsteuerungs-Hub 3730 und/oder die Speichersteuerung 3716 extern bezüglich eines oder mehrerer Prozessor(en) 3702 sein. Beispielsweise kann das System 3700 bei mindestens einer Ausführungsform eine externe Speichersteuerung 3716 und einen Plattformsteuervorrichtungs-Hub 3730 beinhalten, der als Speichersteuervorrichtungs-Hub und Peripherie-Steuervorrichtungs-Hub innerhalb eines System-Chipsatzes ausgestaltet sein kann, der mit dem (den) Prozessor(en) 3702 in Verbindung steht.

[0445] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform können Teile der oder die gesamte Ableitungs- und/oder Trainingslogik 1815 in den Grafikprozessor 3700 integriert sein. Beispielsweise können bei mindestens einer Ausführungsform in dieser Schrift beschriebene Trainings- und/oder Ableitungsmethoden eine oder mehrere der in einer 3D-Pipeline 3712 realisierten ALUs verwenden. Darüber hinaus können bei mindestens einer Ausführungsform in dieser Schrift beschriebene Ableitungs- und/oder Trainingsoperationen unter Verwendung einer anderen Logik als der in den **Fig. 18A** oder **Fig. 18B** veranschaulichten Logik ausgeführt werden. Bei mindestens einer Ausführungsform können Gewichtungsparameter in chipinternen oder chipexternen Speicher und/oder Registern (gezeigt oder nicht gezeigt) gespeichert werden, die ALUs des Grafikprozessors 3700 konfigurieren, um einen oder mehrere Algorithmen maschinellen Lernens, neuronale Netzwerkarchitekturen, Anwendungsfälle oder in dieser Schrift beschriebene Trainingsmethoden auszuführen.

[0446] Bei mindestens einer Ausführungsform wird der Grafikprozessor 3700 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird der Grafikprozessor 3700 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird der Grafikprozessor 3700 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und der Grafikprozessor 3700 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts in einem Bild auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0447] Fig. 38 ist ein Blockdiagramm eines Prozessors 3800 mit einem oder mehreren Prozessorkernen 3802A-3802N, einer integrierten Speichersteuerung 3814 und einem integrierten Grafikprozessor 3808, gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform kann der Prozessor 3800 zusätzliche Kerne beinhalten und den zusätzlichen Kern 3802N, der durch gestrichelte Kästen dargestellt ist, beinhalten. Bei mindestens einer Ausführungsform weist jeder der Prozessorkerne 3802A-3802N eine oder mehrere interne Cache-Einheiten 3804A-3804N auf. Bei mindestens einer Ausführungsform hat jeder Prozessorkern auch Zugriff auf eine oder mehrere gemeinsam genutzte Cache-Einheiten 3806.

[0448] Bei mindestens einer Ausführungsform stellen die internen Cache-Einheiten 3804A-3804N und die gemeinsam genutzten Cache-Einheiten 3806 eine Cache-Speicherhierarchie innerhalb des Prozessors 3800 dar. Bei mindestens einer Ausführungsform können die Cache-Speichereinheiten 3804A-3804N mindestens eine Ebene eines Befehls- und Daten-Caches innerhalb jedes Prozessorkerns und eine oder mehrere Ebenen eines gemeinsam genutzten Mid-Level-Caches, wie z. B eine Ebene 2 (L2), Ebene 3 (L3), Ebene 4 (L4) oder andere Cache-Ebenen, aufweisen, wobei eine höchste Cache-Ebene vor einem externen Speicher als LLC klassifiziert ist. Bei mindestens einer Ausführungsform hält die Cache-Kohärenzlogik die Kohärenz zwischen verschiedenen Cache-Einheiten 3806 und 3804A-3804N aufrecht.

[0449] Bei mindestens einer Ausführungsform kann der Prozessor 3800 auch einen Satz von einer oder mehreren Bussteuerungseinheiten 3816 und einen Systemagentenkern 3810 aufweisen. Bei mindestens einer Ausführungsform verwalten eine oder mehrere Bussteuerungseinheiten 3816 einen Satz von Peripheriebussen, wie einen oder mehrere PCI- oder PCI-Express-Busse. Bei mindestens einer Ausführungsform bietet der Systemagenten-Kern 3810 Verwaltungsfunktionen für verschiedene Prozessorkomponenten. Bei mindestens einer Ausführungsform weist der Systemagenten-Kern 3810 eine oder mehrere integrierte Speichersteuerungen 3814 auf, um den Zugriff auf verschiedene externe Speichereinrichtungen (nicht dargestellt) zu verwalten.

[0450] Bei mindestens einer Ausführungsform weist einer oder weisen mehrere der Prozessorkerne 3802A-3802N Unterstützung für gleichzeitiges Multithreading auf. Bei mindestens einer Ausführungsform weist der Systemagentenkern 3810 Komponenten zur Koordinierung und zum Betrieb der Kerne 3802A-3802N während der Multithreading-Verarbeitung auf. Bei mindestens einer Ausführungsform kann der Systemagentenkern 3810 zusätzlich eine Leistungssteuerungseinheit (PCU) aufweisen, die Logik und Komponenten zur Regelung eines oder mehrerer Leistungszustände der Prozessorkerne 3802A-3802N und des Grafikprozessors 3808 aufweist.

[0451] Bei mindestens einer Ausführungsform beinhaltet der Prozessor 3800 zusätzlich einen Grafikprozessor 3808 zur Ausführung von Grafikverarbeitungsoperationen. Bei mindestens einer Ausführungsform ist der Grafikprozessor 3808 mit gemeinsamen Cache-Einheiten 3806 und dem Systemagenten-Kern 3810 gekoppelt, der eine oder mehrere integrierte Speichersteuerungen 3814 aufweist. Bei mindestens einer Ausführungsform weist der Systemagenten-Kern 3810 auch eine Anzeigesteuerung 3811 auf, um die Ausgabe des Grafikprozessors an eine oder mehrere gekoppelte Anzeigen zu steuern. Bei mindestens einer Ausführungsform kann die Anzeigesteuerung 3811 auch ein separates Modul sein, das über mindestens eine Zwischenverbindung mit dem Grafikprozessor 3808 verbunden ist, oder sie kann in den Grafikprozessor 3808 integriert sein.

[0452] Bei mindestens einer Ausführungsform wird eine ringbasierte Verbindungseinheit 3812 verwendet, um interne Komponenten des Prozessors 3800 zu verbinden. Bei mindestens einer Ausführungsform kann eine alternative Verbindungseinheit verwendet werden, wie z. B. eine Punkt-zu-Punkt-Verbindung, eine geschaltete Verbindung oder andere Techniken. Bei mindestens einer Ausführungsform ist der Grafikprozessor 3808 über eine E/A-Verbindung 3813 mit dem Ring-Interconnect 3812 verbunden.

[0453] Bei mindestens einer Ausführungsform stellt die E/A-Verbindung 3813 mindestens eine von mehreren Arten von E/A-Interconnects dar, die ein On-Package-E/A-Interconnect beinhalten, das die Kommunikation zwischen verschiedenen Prozessorkomponenten und einem eingebetteten Hochleistungsspeichermodul 3818, wie z. B. einem eDRAM-Modul, erleichtert. Bei mindestens einer Ausführungsform verwenden jeder der Prozessorkerne 3802A-3802N und der Grafikprozessor 3808 eingebettete Speichermodule 3818 als gemeinsam genutzten Cache der letzten Ebene.

[0454] Bei mindestens einer Ausführungsform sind die Prozessorkerne 3802A-3802N homogene Kerne, die eine gemeinsame Befehlssatzarchitektur ausführen. Bei mindestens einer Ausführungsform sind die Prozessorkerne 3802A-3802N in Bezug auf die Befehlssatzarchitektur (ISA) heterogen, wobei ein oder mehrere Prozessorkerne 3802A-3802N einen gemeinsamen Befehlssatz ausführen, während ein oder mehrere andere Kerne der Prozessorkerne 3802A-38-02N eine Teilmenge eines gemeinsamen Befehlssatzes oder einen anderen Befehlssatz ausführen. Bei mindestens einer Ausführungsform sind die Prozessorkerne 3802A-3802N in Bezug auf die Mikroarchitektur heterogen, wobei ein oder mehrere Kerne mit einem relativ höheren Energieverbrauch mit einem oder mehreren Kernen mit einem niedrigeren Energieverbrauch gekoppelt sind. Bei mindestens einer Ausführungsform kann der Prozessor 3800 auf einem oder mehreren Chips oder als integrierte SoC-Schaltung implementiert sein.

[0455] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform können Teile der oder die gesamte Ableitungs- und/oder Trainingslogik 1815 in den Grafikprozessor 3810 integriert sein. Beispielsweise können bei mindestens einer Ausführungsform in dieser Schrift beschriebene Trainings- und/oder Ableitungstechniken eine oder mehrere der in einer 3D-Pipeline 3712 realisierten ALUs, Grafikkern(e) 3815A, gemeinsam genutzte Funktionslogik 3816, Grafikkern(e) 3815B, gemeinsam genutzte Funktionslogik 3820 oder andere Logik in **Fig. 38** verwenden. Darüber hinaus können bei mindestens einer Ausführungsform in dieser Schrift beschriebene Ableitungs- und/oder Trainingsoperationen unter Verwendung einer anderen Logik als der in den **Fig. 18A** oder **Fig. 18B** veranschaulichten Logik ausgeführt werden. Bei mindestens einer Ausführungsform können Gewichtungsparameter in chipinternen oder chipexternen Speicher und/oder Registern (gezeigt oder nicht gezeigt) gespeichert werden, die ALUs des Grafikprozessors 3810 konfigurieren, um einen oder mehrere Algorithmen maschinellen Lernens, neuronale Netzwerkarchitekturen, Anwendungsfälle oder in dieser Schrift beschriebene Trainingsmethoden auszuführen.

[0456] Bei mindestens einer Ausführungsform wird der Grafikprozessor 3810 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird der Grafikprozessor 3810 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird der Grafikprozessor 3810 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und der Grafikprozessor 3810 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts in einem Bild auf selbstüberwachte Weise zu identifizieren, indem

mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0457] **Fig. 39** ist ein Blockdiagramm eines Grafikprozessors 3900, bei dem es sich um eine diskrete Grafikverarbeitungseinheit oder um einen mit einer Vielzahl von Prozessorkernen integrierten Grafikprozessor handeln kann. Bei mindestens einer Ausführungsform kommuniziert der Grafikprozessor 3900 über eine einem Speicher zugeordnete E/A-Schnittstelle mit Registern auf dem Grafikprozessor 3900 und mit Befehlen, die im Speicher abgelegt sind. Bei mindestens einer Ausführungsform weist der Grafikprozessor 3900 eine Speicherschnittstelle 3914 für den Zugriff auf den Speicher auf. Bei mindestens einer Ausführungsform ist die Speicherschnittstelle 3914 eine Schnittstelle zum lokalen Speicher, einem oder mehreren internen Caches, einem oder mehreren gemeinsam genutzten externen Caches und/oder zum Systemspeicher.

[0458] Bei mindestens einer Ausführungsform beinhaltet der Grafikprozessor 3900 auch eine Anzeigesteuervorrichtung 3902, um Anzeigeausgangsdaten an eine Anzeigevorrichtung 3920 zu steuern. Bei mindestens einer Ausführungsform beinhaltet die Anzeigesteuervorrichtung 3902 Hardware für eine oder mehrere Überlagerungsebenen für die Anzeigevorrichtung 3920 und die Zusammensetzung mehrerer Schichten von Video- oder Benutzerschnittstellenelementen. Bei mindestens einer Ausführungsform kann die Anzeigeeinrichtung 3920 eine interne oder externe Anzeigeeinrichtung sein. Bei mindestens einer Ausführungsform handelt es sich bei der Anzeigeeinrichtung 3920 um eine am Kopf getragene Anzeigeeinrichtung, wie z. B. eine Virtual-Reality- (VR-) Anzeigeeinrichtung oder eine Augmented-Reality- (AR-) Anzeigeeinrichtung. Bei mindestens einer Ausführungsform weist der Grafikprozessor 3900 eine Videocodec-Maschine 3906 auf, um Medien in, aus oder zwischen einem oder mehreren Mediencodierformaten zu codieren, zu decodieren oder zu transcodieren, einschließlich, aber nicht beschränkt auf Moving Picture Experts Group (MPEG)-Formate wie MPEG-2, Advanced Video Coding (AVC)-Formate wie H.264 /MPEG-4 AVC, sowie die Society of Motion Picture & Television Engineers (SMPTE) 421M/VC-1 und Joint Photographic Experts Group (JPEG) Formate wie JPEG und Motion JPEG (MJPEG) Formate.

[0459] Bei mindestens einer Ausführungsform weist der Grafikprozessor 3900 eine BLIT-Maschine (Block Image Transfer) 3904 auf, um zweidimensionale (2D) Rasterisierungsoperationen durchzuführen, einschließlich z. B. Bit-Boundary Block Transfers. Bei mindestens einer Ausführungsform werden 2D-Grafikoperationen jedoch mit einer oder mehreren Komponenten der Grafikverarbeitungs-Engine (GPE) 3910 ausgeführt. Bei mindestens einer Ausführungsform ist die GPE 3910 eine Rechenmaschine zur Durchführung von Grafikoperationen, die dreidimensionale (3D) Grafikoperationen und Medienoperationen einschließen.

[0460] Bei mindestens einer Ausführungsform weist die GPE 3910 eine 3D-Pipeline 3912 zur Durchführung von 3D-Operationen auf, wie z. B. das Rendern dreidimensionaler Bilder und Szenen unter Verwendung von Verarbeitungsfunktionen, die auf 3D-Primitivformen (z. B. Rechteck, Dreieck usw.) wirken. Die 3D-Pipeline 3912 weist programmierbare und feste Funktionselemente auf, die verschiedene Aufgaben ausführen und/oder Ausführungs-Threads zu einem 3D/Media-Subsystem 3915 erzeugen. Während die 3D-Pipeline 3912 zur Durchführung von Medienoperationen verwendet werden kann, weist die GPE 3910 bei mindestens einer Ausführungsform auch eine Medien-Pipeline 3916 auf, die zur Durchführung von Medienoperationen, wie Videonachbearbeitung und Bildverbesserung, verwendet wird.

[0461] Bei mindestens einer Ausführungsform weist die Medienpipeline 3916 feste Funktions- oder programmierbare Logikeinheiten auf, um eine oder mehrere spezialisierte Medienoperationen wie Videodecodierbeschleunigung, Videoentflechtung und Videocodierbeschleunigung anstelle von oder im Auftrag der Videocodec-Maschine 3906 durchzuführen. Bei mindestens einer Ausführungsform weist die Medien-Pipeline 3916 zusätzlich eine Thread-Spawning-Einheit auf, um Threads zur Ausführung im 3D/Media-Subsystem 3915 zu erzeugen. Bei mindestens einer Ausführungsform führen die erzeugten Threads Berechnungen für Medienoperationen auf einer oder mehreren Grafikausführungseinheiten durch, die im 3D/Media-Subsystem 3915 vorhanden sind.

[0462] Bei mindestens einer Ausführungsform weist das 3D/Media-Subsystem 3915 eine Logik zur Ausführung von Threads auf, die von der 3D-Pipeline 3912 und der Media-Pipeline 3916 erzeugt werden. Bei mindestens einer Ausführungsform senden die 3D-Pipeline 3912 und die Medien-Pipeline 3916 Thread-Ausführungsanforderungen an das 3D/Media-Subsystem 3915, das eine Thread-Verteilungslogik beinhaltet, um verschiedene Anforderungen an verfügbare Thread-Ausführungsressourcen zu vermitteln und zu verteilen. Bei mindestens einer Ausführungsform weisen die Ausführungsressourcen eine Anordnung von Grafikausführungseinheiten zur Verarbeitung von 3D- und Medien-Threads auf. Bei mindestens einer Ausführungsform weist das 3D/Media-Subsystem 3915 einen oder mehrere interne Caches für Thread-Anweisungen und

-Daten auf. Bei mindestens einer Ausführungsform weist das Subsystem 3915 auch einen gemeinsamen Speicher auf, einschließlich Registern und adressierbarem Speicher, um Daten zwischen Threads zu teilen und Ausgabedaten zu speichern.

[0463] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform können Teile der oder die gesamte Ableitungs- und/oder Trainingslogik 1815 in den Grafikprozessor 3900 integriert sein. Zum Beispiel können bei mindestens einer Ausführungsform in dieser Schrift beschriebene Trainings- und/oder Ableitungsmethoden eine oder mehrere der in einer 3D-Pipeline 3912 realisierten ALUs verwenden. Darüber hinaus können bei mindestens einer Ausführungsform in dieser Schrift beschriebene Ableitungs- und/oder Trainingsoperationen unter Verwendung einer anderen Logik als der in den **Fig. 18A** oder **Fig. 18B** veranschaulichten Logik ausgeführt werden. Bei mindestens einer Ausführungsform können Gewichtungsparameter in chipinternen oder chipexternen Speicher und/oder Registern (gezeigt oder nicht gezeigt) gespeichert werden, die ALUs des Grafikprozessors 3900 konfigurieren, um einen oder mehrere Algorithmen maschinellen Lernens, neuronale Netzwerkarchitekturen, Anwendungsfälle oder in dieser Schrift beschriebene Trainingsmethoden auszuführen.

[0464] Bei mindestens einer Ausführungsform wird der Grafikprozessor 3910 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird der Grafikprozessor 3910 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird der Grafikprozessor 3910 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und der Grafikprozessor 3910 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts in einem Bild auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0465] Fig. 40 ist ein Blockdiagramm einer Grafikverarbeitungsmaschine 4010 eines Grafikprozessors gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform ist die Grafikverarbeitungsmaschine (GPE) 4010 eine Version der in **Fig. 39** gezeigten GPE 3910. Bei mindestens einer Ausführungsform ist die Medienpipeline 3916 optional und ist eventuell nicht ausdrücklich in der GPE 4010 vorhanden. Bei mindestens einer Ausführungsform ist ein separater Medien- und/oder Bildprozessor mit der GPE 4010 verbunden.

[0466] Bei mindestens einer Ausführungsform ist die GPE 4010 mit einem Befehlsstreamer 4003, der einen Befehlsstrom an die 3D-Pipeline 3912 und/oder die Medienpipelines 3916 liefert, gekoppelt oder beinhaltet diesen. Bei mindestens einer Ausführungsform ist der Befehlsstreamer 4003 mit einem Speicher gekoppelt, bei dem es sich um einen Systemspeicher oder um einen oder mehrere interne Cache-Speicher und gemeinsam genutzte Cache-Speicher handeln kann. Bei mindestens einer Ausführungsform empfängt der Befehlsstreamer 4003 Befehle vom Speicher und sendet Befehle an die 3D-Pipeline 3912 und/oder die Medien-Pipeline 3916. Bei mindestens einer Ausführungsform handelt es sich bei den Befehlen um Anweisungen, Primitive oder Mikrooperationen, die aus einem Ringpuffer abgerufen werden, der Befehle für die 3D-Pipeline 3912 und die Medien-Pipeline 3916 speichert. Bei mindestens einer Ausführungsform kann ein Ringpuffer zusätzlich Batch-Befehlspuffer aufweisen, die Stapel von mehreren Befehlen speichern. Bei mindestens einer Ausführungsform können die Befehle für die 3D-Pipeline 3912 auch Verweise auf im Speicher gespeicherte Daten aufweisen, wie z. B. Vertex- und Geometriedaten für die 3D-Pipeline 3912 und/oder Bilddaten und Speicherobjekte für die Medien-Pipeline 3916. Bei mindestens einer Ausführungsform verarbeiten die 3D-Pipeline 3912 und die Medien-Pipeline 3916 Befehle und Daten, indem sie Operationen ausführen oder einen oder mehrere Ausführungs-Threads an eine Grafikkernanordnung 4014 weiterleiten. Bei mindestens einer Ausführungsform beinhaltet die Grafikkernanordnung 4014 einen oder mehrere Blöcke von Grafikkernen (z. B. Grafikkern(e) 4015A, Grafikkern(e) 4015B), wobei jeder Block einen oder mehrere Grafikkerne beinhaltet. Bei mindestens einer Ausführungsform beinhaltet jeder Grafikkern einen Satz von Grafikausführungsressourcen, die Allzweck- und grafikspezifische Ausführungslogik beinhalten, um Grafik- und Rechenoperationen auszuführen, sowie feste Funktionstexturverarbeitungs- und/oder Maschinenlern- und Künstliche-Intelligenz-Beschleunigungslogik, die Ableitungs- und/oder Trainingslogik 1815 in **Fig. 18A** und **Fig. 18B** beinhaltet.

[0467] Bei mindestens einer Ausführungsform weist die 3D-Pipeline 3912 eine feste Funktion und eine programmierbare Logik auf, um ein oder mehrere Shader-Programme wie Vertex-Shader, Geometrie-Shader, Pixel-Shader, Fragment-Shader, Rechen-Shader oder andere Shader-Programme zu verarbeiten, indem Befehle verarbeitet und Ausführungs-Threads an die Grafikkernanordnung 4014 gesendet werden. Bei mindestens einer Ausführungsform stellt die Grafikkernanordnung 4014 einen einheitlichen Block von Ausführungsressourcen für die Verarbeitung von Shader-Programmen bereit. Bei mindestens einer Ausführungsform weist die Mehrzweck-Ausführungslogik (z. B. Ausführungseinheiten) in den Grafikkernen 4015A-4015B der Grafikkernanordnung 4014 Unterstützung für verschiedene 3D-API-Shader-Sprachen auf und kann mehrere gleichzeitige Ausführungs-Threads ausführen, die mehreren Shadern zugeordnet sind.

[0468] Bei mindestens einer Ausführungsform weist die Grafikkernanordnung 4014 auch eine Ausführungslogik zur Durchführung von Medienfunktionen wie Video- und/oder Bildverarbeitung auf. Bei mindestens einer Ausführungsform weisen die Ausführungseinheiten zusätzlich eine Allzwecklogik auf, die so programmierbar ist, dass sie zusätzlich zu den Grafikverarbeitungsoperationen parallele Allzweckrechenoperationen durchführt.

[0469] Bei mindestens einer Ausführungsform können Ausgabedaten, die von Threads erzeugt werden, die auf der Grafikkernanordnung 4014 ausgeführt werden, an den Speicher in einem Unified Return Buffer (URB) 4018 ausgegeben werden. Der URB 4018 kann Daten für mehrere Threads speichern. Bei mindestens einer Ausführungsform kann der URB 4018 verwendet werden, um Daten zwischen verschiedenen Threads zu senden, die auf der Grafikkernanordnung 4014 ausgeführt werden. Bei mindestens einer Ausführungsform kann der URB 4018 zusätzlich zur Synchronisation zwischen Threads auf der Grafikkernanordnung 4014 und der festen Funktionslogik innerhalb der gemeinsam genutzten Funktionslogik 4020 verwendet werden.

[0470] Bei mindestens einer Ausführungsform ist die Grafikkernanordnung 4014 skalierbar, so dass die Grafikkernanordnung 4014 eine variable Anzahl von Grafikkernen aufweist, von denen jeder eine variable Anzahl von Ausführungseinheiten hat, die auf einem angestrebten Energie- und Leistungsniveau der GPE 4010 basieren. Bei mindestens einer Ausführungsform sind die Ausführungsressourcen dynamisch skalierbar, so dass die Ausführungsressourcen je nach Bedarf aktiviert oder deaktiviert werden können.

[0471] Bei mindestens einer Ausführungsform ist die Grafikkernanordnung 4014 mit der gemeinsamen Funktionslogik 4020 gekoppelt, die mehrere Ressourcen aufweist, die von den Grafikkernen im der Grafikkernanordnung 4014 gemeinsam genutzt werden. Bei mindestens einer Ausführungsform sind die gemeinsam genutzten Funktionen, die von der gemeinsam genutzten Funktionslogik 4020 ausgeführt werden, in Hardware-Logikeinheiten verkörpert, die der Grafikkernanordnung 4014 eine spezielle Zusatzfunktionalität bieten. Bei mindestens einer Ausführungsform beinhaltet die gemeinsam genutzte Funktionslogik 4020 unter anderem einen Sampler 4021, Mathematik 4022 und eine Inter-Thread-Kommunikations- (ITC-) 4023 Logik. Bei mindestens einer Ausführungsform sind ein oder mehrere Cache(s) 4025 in der gemeinsam genutzten Funktionslogik 4020 vorhanden oder mit ihr gekoppelt.

[0472] Bei mindestens einer Ausführungsform wird eine gemeinsam genutzte Funktion verwendet, wenn die Nachfrage nach einer speziellen Funktion nicht ausreicht, um sie in die Grafikkernanordnung 4014 aufzunehmen. Bei mindestens einer Ausführungsform wird eine einzelne Instanziierung einer spezialisierten Funktion in der gemeinsam genutzten Funktionslogik 4020 verwendet und von anderen Ausführungsressourcen innerhalb der Grafikkernanordnung 4014 gemeinsam genutzt. Bei mindestens einer Ausführungsform können spezifische gemeinsam genutzte Funktionen innerhalb der gemeinsam genutzten Funktionslogik 4020, die vom der Grafikkernanordnung 4014 intensiv genutzt werden, in der gemeinsam genutzten Funktionslogik 4016 innerhalb der Grafikkernanordnung 4014 beinhaltet sein. Bei mindestens einer Ausführungsform kann die gemeinsam genutzte Funktionslogik 4016 innerhalb der Grafikkernanordnung 4014 einige oder alle Logiken der gemeinsam genutzten Funktionslogik 4020 beinhalten. Bei mindestens einer Ausführungsform können alle Logikelemente innerhalb der gemeinsam genutzten Funktionslogik 4020 innerhalb der gemeinsam genutzten Funktionslogik 4016 der Grafikkernanordnung 4014 dupliziert sein. Bei mindestens einer Ausführungsform ist die gemeinsam genutzte Funktionslogik 4020 zugunsten der gemeinsam genutzten Funktionslogik 4016 innerhalb der Grafikkernanordnung 4014 ausgeschlossen.

[0473] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform können Teile der oder die gesamte Ableitungs- und/oder Trainingslogik 1815 in den Grafikprozessor 4010 integriert sein. Zum Beispiel können bei mindestens

einer Ausführungsform in dieser Schrift beschriebene Trainings- und/oder Ableitungstechniken eine oder mehrere der in einer 3D-Pipeline 3912 realisierten ALUs, Grafikkern(e) 4015A, gemeinsam genutzte Funktionslogik 4016, Grafikkern(e) 4015B, gemeinsam genutzte Funktionslogik 4020 oder andere Logik in **Fig. 40** verwenden. Darüber hinaus können bei mindestens einer Ausführungsform in dieser Schrift beschriebene Ableitungs- und/oder Trainingsoperationen unter Verwendung einer anderen Logik als der in den **Fig. 18A** oder **Fig. 18B** veranschaulichten Logik ausgeführt werden. Bei mindestens einer Ausführungsform können Gewichtungsparameter in chipinternen oder chipexternen Speicher und/oder Registern (gezeigt oder nicht gezeigt) gespeichert werden, die ALUs des Grafikprozessors 4010 konfigurieren, um einen oder mehrere Algorithmen maschinellen Lernens, neuronale Netzwerkarchitekturen, Anwendungsfälle oder in dieser Schrift beschriebene Trainingsmethoden auszuführen.

[0474] Bei mindestens einer Ausführungsform wird der Grafikprozessor 4010 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird der Grafikprozessor 4010 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird der Grafikprozessor 4010 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und der Grafikprozessor 4010 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts in einem Bild auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0475] **Fig. 41** ist ein Blockdiagramm der Hardware-Logik eines Grafikprozessorkerns 4100, wie es hier in mindestens einer Ausführungsform beschrieben ist. Bei mindestens einer Ausführungsform ist der Grafikprozessorkern 4100 in einer Grafikkernanordnung vorhanden. Bei mindestens einer Ausführungsform kann der Grafikprozessorkern 4100, der manchmal auch als Kern-Slice bezeichnet wird, ein oder mehrere Grafikkerne innerhalb eines modularen Grafikprozessors sein. Bei mindestens einer Ausführungsform ist der Grafikprozessorkern 4100 ein Beispiel für einen Grafikkern-Slice, und ein Grafikprozessor, wie er hier beschrieben ist, kann mehrere Grafikkern-Slices aufweisen, die auf den angestrebten Energie- und Leistungshüllkurven basieren. Bei mindestens einer Ausführungsform kann jeder Grafikkern 4100 einen Festfunktionsblock 4130 beinhalten, der mit mehreren Teilkernen 4101A-4101F gekoppelt ist, die auch als Teil-Slices bezeichnet werden, die modulare Blöcke von Allzweck- und Festfunktionslogik beinhalten.

[0476] Bei mindestens einer Ausführungsform weist der Festfunktionsblock 4130 eine Geometrie-/Festfunktionspipeline 4136 auf, die von allen Unterkernen im Grafikprozessor 4100 gemeinsam genutzt werden kann, z. B. bei Grafikprozessorimplementierungen mit geringerer Leistung und/oder geringerem Energieverbrauch. Bei mindestens einer Ausführungsform weist die Geometrie/Festfunktionspipeline 4136 eine 3D-Festfunktionspipeline, eine Video-Front-End-Einheit, einen Thread-Spawner und Thread-Dispatcher sowie einen Unified-Return-Puffer-Manager auf, der Unified-Return-Puffer verwaltet.

[0477] Bei mindestens einer Ausführungsform weist der feste Funktionsblock 4130 auch eine Grafik-SoC-Schnittstelle 4137, einen Grafik-Mikrocontroller 4138 und eine Medienpipeline 4139 auf. Die Grafik-SoC-Schnittstelle 4137 stellt eine Schnittstelle zwischen dem Grafikkern 4100 und anderen Prozessorkernen innerhalb einer integrierten System-on-Chip-Schaltung bereit. Bei mindestens einer Ausführungsform ist der Grafik-Mikrocontroller 4138 ein programmierbarer Teilprozessor, der derart konfigurierbar ist, dass er verschiedene Funktionen des Grafikprozessors 4100, einschließlich Thread-Versand, Zeitplanung und Preemption, verwaltet. Bei mindestens einer Ausführungsform weist die Medienpipeline 4139 eine Logik zur Erleichterung der Decodierung, Codierung, Vorverarbeitung und/oder Nachverarbeitung von Multimediadaten, einschließlich Bild- und Videodaten, auf. Bei mindestens einer Ausführungsform implementiert die Medienpipeline 4139 Medienoperationen über Anforderungen an die Berechnungs- oder Abtastlogik innerhalb der Unterkerne 4101-4101F.

[0478] Bei mindestens einer Ausführungsform ermöglicht es die SoC-Schnittstelle 4137 dem Grafikkern 4100, mit Mehrzweck-Anwendungsprozessorkernen (z. B. CPUs) und/oder anderen Komponenten innerhalb eines SoC, einschließlich Speicherhierarchieelementen, wie einem gemeinsam genutzten Cache-Speicher der letzten Ebene, einem System-RAM und/oder einem eingebettetem On-Chip- oder On-Package-DRAM, zu kommunizieren. Bei mindestens einer Ausführungsform kann die SoC-Schnittstelle 4137 auch die Kommunikation mit Einrichtungen mit fester Funktion innerhalb eines SoCs ermöglichen, wie z. B. Kamera-Bildgebungspipelines, und sie ermöglicht die Nutzung und/oder Implementierung globaler Speicher-Atome, die von

Grafikkern 4100 und CPUs innerhalb eines SoCs gemeinsam genutzt werden können. Bei mindestens einer Ausführungsform kann die SoC-Schnittstelle 4137 auch Energieverwaltungssteuerungen für den Grafikkern 4100 implementieren und eine Schnittstelle zwischen einer Taktdomäne des Grafikkerns 4100 und anderen Taktdomänen innerhalb eines SoCs ermöglichen. Bei mindestens einer Ausführungsform ermöglicht die SoC-Schnittstelle 4137 den Empfang von Befehlspuffern von einem Befehlsstreamer und einem globalen Thread-Dispatcher, die so ausgestaltet sind, dass sie Befehle und Anweisungen für jeden von einem oder mehreren Grafikkernen innerhalb eines Grafikprozessors bereitstellen. Bei mindestens einer Ausführungsform können Befehle und Anweisungen an die Medienpipeline 4139 gesendet werden, wenn Medienoperationen ausgeführt werden sollen, oder an eine Geometrie- und Festfunktionspipeline (z. B. Geometrie- und Festfunktionspipeline 4136, Geometrie- und Festfunktionspipeline 4114) gesendet werden, wenn Grafikverarbeitungsoperationen ausgeführt werden sollen.

[0479] Bei mindestens einer Ausführungsform kann der Grafik-Mikrocontroller 4138 so ausgestaltet sein, dass er verschiedene Planungs- und Verwaltungsaufgaben für den Grafikkern 4100 ausführt. Bei mindestens einer Ausführungsform kann der Grafik-Mikrocontroller 4138 die Planung von Grafik- und/oder Rechenaufgaben auf verschiedenen parallelen Grafik-Maschinen innerhalb von Anordnungen 4102A-4102F, 4104A-4104F von Ausführungseinheiten (EU) innerhalb der Unterkerne 4101A-4101F durchführen. Bei mindestens einer Ausführungsform kann eine Host-Software, die auf einem CPU-Kern eines SoC ausgeführt wird, der den Grafikkern 4100 aufweist, Arbeitslasten an eine von mehreren Grafikprozessor-Doorbells übermitteln, die einen Planungsvorgang auf einer geeigneten Grafik-Maschine aufruft. Bei mindestens einer Ausführungsform weisen die Planungsvorgänge die Bestimmung der als Nächstes auszuführenden Arbeitslast, die Übermittlung einer Arbeitslast an einen Befehlsstreamer, das Vorziehen bestehender Arbeitslasten, die auf einer Maschine ausgeführt werden, die Überwachung des Fortschritts einer Arbeitslast und die Benachrichtigung der Host-Software nach Abschluss einer Arbeitslast auf. Bei mindestens einer Ausführungsform kann der Grafik-Mikrocontroller 4138 auch stromsparende Zustände oder Leerlaufzustände für den Grafikkern 4100 erleichtern, indem er dem Grafikkern 4100 die Möglichkeit bietet, Register innerhalb des Grafikkerns 4100 über stromsparende Zustandsübergänge unabhängig von einem Betriebssystem und/oder einer Grafiktreibersoftware auf einem System zu speichern und wiederherzustellen.

[0480] Bei mindestens einer Ausführungsform kann der Grafikkern 4100 mehr oder weniger als die dargestellten Unterkerne 4101A-4101F aufweisen, bis zu N modulare Unterkeme. Bei mindestens einer Ausführungsform kann der Grafikkern 4100 für jeden Satz von N Unterkernen auch eine gemeinsam genutzte Funktionslogik 4110, einen gemeinsam genutzten und/oder Cache-Speicher 4112, eine Geometrie-/Festfunktionspipeline 4114 sowie eine zusätzliche Festfunktionslogik 4116 aufweisen, um verschiedene Grafik- und Rechenverarbeitungsvorgänge zu beschleunigen. Bei mindestens einer Ausführungsform kann die gemeinsam genutzte Funktionslogik 4110 logische Einheiten beinhalten (z. B. Sampler, Mathematik und/oder Inter-Thread-Kommunikationslogik), die von jedem der N Teilkerne innerhalb des Grafikkerns 4100 gemeinsam genutzt werden können. Gemeinsamer und/oder Cache-Speicher 4112 kann ein Cache der letzten Ebene für N Unterkerne 4101A-4101F innerhalb des Grafikkerns 4100 sein und kann auch als gemeinsamer Speicher dienen, auf den mehrere Unterkerne zugreifen können. Bei mindestens einer Ausführungsform kann die Geometrie-/Festfunktionspipeline 4114 anstelle der Geometrie-/Festfunktionspipeline 4136 innerhalb des Festfunktionsblocks 4130 vorhanden sein und kann gleiche oder ähnliche Logikeinheiten aufweisen.

[0481] Bei mindestens einer Ausführungsform beinhaltet der Grafikkern 4100 eine zusätzliche Festfunktionslogik 4116, die verschiedene Festfunktions-Beschleunigungslogiken zur Verwendung durch den Grafikkern 4100 aufweisen kann. Bei mindestens einer Ausführungsform weist die zusätzliche Festfunktionslogik 4116 eine zusätzliche Geometrie-Pipeline zur Verwendung beim positionsgebundenen Shading auf. Bei dem positionsgebundenen Shading gibt es mindestens zwei Geometrie-Pipelines, nämlich eine vollständige Geometrie-Pipeline innerhalb der Geometrie/Festfunktions-Pipeline 4116, 4136, und eine Cull-Pipeline, die eine zusätzliche Geometrie-Pipeline ist und in der eine zusätzliche Festfunktionslogik 4116 enthalten sein kann. Bei mindestens einer Ausführungsform ist die Cull-Pipeline eine abgespeckte Version einer vollständigen Geometrie-Pipeline. Bei mindestens einer Ausführungsform können eine vollständige Pipeline und eine Cull-Pipeline verschiedene Instanzen einer Anwendung ausführen, wobei jede Instanz einen eigenen Kontext hat. Bei mindestens einer Ausführungsform kann das positionsgebundene Shading lange Cull-Läufe von verworfenen Dreiecken verbergen, so dass das Shading bei einigen Ausführungsformen früher abgeschlossen werden kann. Zum Beispiel kann bei mindestens einer Ausführungsform die Cull-Pipeline-Logik innerhalb der zusätzlichen Festfunktionslogik 4116 Positions-Shader parallel zu einer Hauptanwendung ausführen und generiert im Allgemeinen kritische Ergebnisse schneller als eine vollständige Pipeline, da die Cull-Pipeline die Positionsattribute von Vertices abruft und schattiert, ohne eine Rasterung und ein Rendering von Pixeln in einen Frame-Puffer durchzuführen. Bei mindestens einer Ausführungsform kann die Cull-Pipeline die

generierten kritischen Ergebnisse verwenden, um die Sichtbarkeitsinformationen für alle Dreiecke zu berechnen, ohne Rücksicht darauf, ob diese Dreiecke aussortiert sind. Bei mindestens einer Ausführungsform kann die vollständige Pipeline (die in diesem Fall als Wiederholungspipeline bezeichnet werden kann) Sichtbarkeitsinformationen verwenden, um aussortierte Dreiecke zu überspringen, um nur sichtbare Dreiecke zu schattieren, die schließlich an eine Rasterisierungsphase übergeben werden.

[0482] Bei mindestens einer Ausführungsform kann die zusätzliche Festfunktionslogik 4116 auch eine Logik zur Beschleunigung des maschinellen Lernens aufweisen, wie z. B. eine Logik zur Matrixmultiplikation mit fester Funktion, für Implementierungen, die Optimierungen für das Training oder Inferencing des maschinellen Lernens umfassen.

[0483] Bei mindestens einer Ausführungsform weist jeder Grafik-Unterkern 4101A-4101F einen Satz von Ausführungsressourcen auf, die verwendet werden können, um Grafik-, Medien- und Rechenoperationen als Reaktion auf Anforderungen von Grafikpipeline-, Medienpipeline- oder Shader-Programmen durchzuführen. Bei mindestens einer Ausführungsform beinhalten die Grafik-Teilkerne 4101A-4101F mehrere EU-Arrays 4102A-4102F, 4104A-4104F, eine Thread-Dispatch- und Inter-Thread-Kommunikationslogik (TD/IC) 4103A-4103F, einen 3D-Sampler (z. B. Textur) 4105A-4105F, einen Media-Sampler 4106A-4106F, einen Shader-Prozessor 4107A-4107F und einen gemeinsamen lokalen Speicher (SLM) 4108A-4108F. Die EU-Arrays 4102A-4102F, 4104A-4104F beinhalten jeweils mehrere Ausführungseinheiten, die Allzweck-Grafikverarbeitungseinheiten sind, die Gleitkomma- und Ganzzahl-/Festkomma-Logikoperationen im Dienst einer Grafik-, Medien- oder Rechenoperation ausführen können, was Grafik-, Medien- oder Rechen-Shader-Programme beinhaltet. Bei mindestens einer Ausführungsform führt die TD/IC-Logik 4103A-4103F lokale Thread-Dispatch- und Thread-Steuerungsoperationen für Ausführungseinheiten innerhalb eines Unterkerns durch und erleichtert die Kommunikation zwischen Threads, die auf Ausführungseinheiten eines Unterkerns ausgeführt werden. Bei mindestens einer Ausführungsform kann der 3D-Sampler 4105A-4105F Textur- oder andere 3D-Grafikdaten in den Speicher einlesen. Bei mindestens einer Ausführungsform kann der 3D-Sampler Texturdaten auf der Grundlage eines konfigurierten Abtaststatus und eines mit einer bestimmten Textur verbundenen Texturformats unterschiedlich lesen. Bei mindestens einer Ausführungsform kann der Mediensampler 4106A-4106F ähnliche Lesevorgänge auf der Grundlage eines Typs und Formats, die mit den Mediendaten verbunden sind, ausführen. Bei mindestens einer Ausführungsform kann jeder Grafik-Unterkern 4101A-4101F abwechselnd einen vereinheitlichten 3D- und Medien-Sampler aufweisen. Bei mindestens einer Ausführungsform können Threads, die auf Ausführungseinheiten innerhalb jedes der Unterkerne 4101A-4101F ausgeführt werden, den gemeinsamen lokalen Speicher 4108A-4108F innerhalb jedes Unterkerns nutzen, um Threads, die innerhalb einer Thread-Gruppe ausgeführt werden, die Ausführung unter Verwendung eines gemeinsamen Pools von On-Chip-Speicher zu ermöglichen.

[0484] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform können Teile der oder die gesamte Ableitungs- und/oder Trainingslogik 1815 in den Grafikprozessor 4110 integriert sein. Beispielsweise können bei mindestens einer Ausführungsform in dieser Schrift beschriebene Trainings- und/oder Ableitungstechniken eine oder mehrere der in einer 3D-Pipeline 4110, der Grafikmikrosteuerung 4138, der Geometrie- und Festfunktionspipeline 4114 und 4136 oder einer anderen Logik in **Fig. 38** realisierten ALUs verwenden. Darüber hinaus können bei mindestens einer Ausführungsform in dieser Schrift beschriebene Ableitungs- und/oder Trainingsoperationen unter Verwendung einer anderen Logik als der in den **Fig. 18A** oder **Fig. 18B** veranschaulichten Logik ausgeführt werden. Bei mindestens einer Ausführungsform können Gewichtungsparameter in chipinternen oder chipexternen Speicher und/oder Registern (gezeigt oder nicht gezeigt) gespeichert werden, die ALUs des Grafikprozessors 4100 konfigurieren, um einen oder mehrere Algorithmen maschinellen Lernens, neuronale Netzwerkarchitekturen, Anwendungsfälle oder in dieser Schrift beschriebene Trainingsmethoden auszuführen.

[0485] Bei mindestens einer Ausführungsform wird der Grafikprozessor 4110 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird der Grafikprozessor 4110 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird der Grafikprozessor 4110 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und der Grafikprozessor 4110 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trai-

nieren, um eine Ausrichtung eines Objekts in einem Bild auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0486] Die **Fig. 42A** und **Fig. 42B** veranschaulichen die Thread-Ausführungslogik 4200, die eine Anordnung von Verarbeitungselementen eines Grafikprozessorkerns gemäß mindestens einer Ausführungsform beinhaltet. **Fig. 42A** veranschaulicht mindestens eine Ausführungsform, bei der die Thread-Ausführungslogik 4200 verwendet wird. **Fig. 42B** veranschaulicht beispielhafte interne Details einer Ausführungseinheit gemäß mindestens einer Ausführungsform.

[0487] Wie in **Fig. 42A** veranschaulicht, beinhaltet die Thread-Ausführungslogik 4200 einen Shader-Prozessor 4202, einen Thread-Versender 4204, einen Anweisungs-Cache 4206, ein skalierbares Ausführungseinheiten-Array, das eine Vielzahl von Ausführungseinheiten 4208A-4208N, einen Sampler 4210, einen Datenzwischenspeicher 4212 und einen Datenanschluss 4214, beinhaltet. Bei mindestens einer Ausführungsform kann ein skalierbares Ausführungseinheiten-Array dynamisch skalieren, indem eine oder mehrere Ausführungseinheiten (z. B. eine der Ausführungseinheiten 4208A, 4208B, 4208C, 4208D bis 4208N-1 und 4208N) zum Beispiel basierend auf der Rechenanforderungen von einer Arbeitslast aktiviert oder deaktiviert werden. Bei mindestens einer Ausführungsform sind die skalierbaren Ausführungseinheiten über eine Verbindungsstruktur miteinander verbunden, die eine Verbindung zu jeder Ausführungseinheit herstellt. Bei mindestens einer Ausführungsform beinhaltet die Thread-Ausführungslogik 4200 eine oder mehrere Verbindungen mit Speicher, wie Systemspeicher oder Cache-Speicher, und zwar über eines oder mehrere der folgenden Elemente: Anweisungs-Cache 4206, Datenanschluss 4214, Sampler 4210 und Ausführungseinheiten 4208A-4208N. Bei mindestens einer Ausführungsform ist jede Ausführungseinheit (z. B. 4208A) eine eigenständige programmierbare Allzweck-Recheneinheit, die in der Lage ist, mehrere gleichzeitige Hardware-Threads auszuführen, während mehrere Datenelemente parallel für jeden Thread verarbeitet werden. Bei mindestens einer Ausführungsform ist die Anordnung der Ausführungseinheiten 4208A-4208N so skalierbar, dass sie eine beliebige Anzahl einzelner Ausführungseinheiten aufweist.

[0488] Bei mindestens einer Ausführungsform werden die Ausführungseinheiten 4208A-4208N hauptsächlich zur Ausführung von Shader-Programmen verwendet. Bei mindestens einer Ausführungsform kann der Shader-Prozessor 4202 verschiedene Shader-Programme verarbeiten und die mit den Shader-Programmen verbundenen Ausführungs-Threads über einen Thread-Dispatcher 4204 verteilen. Bei mindestens einer Ausführungsform weist der Thread-Dispatcher 4204 eine Logik auf, um Thread-Initiierungsanforderungen von Grafik- und Medienpipelines zu vermitteln und angeforderte Threads auf einer oder mehreren Ausführungseinheiten in den Ausführungseinheiten 4208A-4208N zu instanziieren. Bei mindestens einer Ausführungsform kann eine Geometrie-Pipeline beispielsweise Vertex-, Tessellierungs- oder Geometrie-Shader an die Thread-Ausführungslogik zur Verarbeitung weiterleiten. Bei mindestens einer Ausführungsform kann der Thread-Dispatcher 4204 auch Laufzeit-Thread-Erzeugungs-Anforderungen von ausführenden Shader-Programmen verarbeiten.

[0489] Bei mindestens einer Ausführungsform unterstützen die Ausführungseinheiten 4208A-4208N einen Befehlssatz, der eine native Unterstützung für viele Standard-3D-Grafik-Shader-Befehle aufweist, so dass Shader-Programme aus Grafikbibliotheken (z. B. Direct 3D und OpenGL) mit einer minimalen Übersetzung ausgeführt werden. Bei mindestens einer Ausführungsform unterstützen Ausführungseinheiten Vertex- und Geometrieverarbeitung (z. B. Vertex-Programme, Geometrieprogramme und/oder Vertex-Shader), Bildpunktverarbeitung (z. B. Bildpunkt-Shader, Fragment-Shader) und Allzweckverarbeitung (z. B. Rechen- und Medien-Shader). Bei mindestens einer Ausführungsform ist jede der Ausführungseinheiten 4208A-4208N, die eine oder mehrere arithmetische Logikeinheiten (ALUs) aufweisen, zur SIMD-Ausführung (Single Instruction Multiple Data) fähig, und der Multi-Thread-Betrieb ermöglicht trotz höherer Latenzzeiten bei Speicherzugriffen eine effiziente Ausführungsumgebung. Bei mindestens einer Ausführungsform verfügt jeder Hardware-Thread innerhalb jeder Ausführungseinheit über eine eigene Registerdatei mit hoher Bandbreite und einen zugehörigen unabhängigen Thread-Status. Bei mindestens einer Ausführungsform erfolgt die Ausführung mit mehreren Threads pro Takt auf Pipelines, die Ganzzahl-, Gleitkomma- und Doppelpräzisionsoperationen, SIMD-Verzweigungsfähigkeit, logische Operationen, transzendentale Operationen und andere verschiedene Operationen ausführen können. Bei mindestens einer Ausführungsform bewirkt die Abhängigkeitslogik in den Ausführungseinheiten 4208A-4208N, dass ein wartender Thread in den Ruhezustand versetzt wird, bis die angeforderten Daten zurückgegeben wurden, während er auf Daten aus dem Speicher oder einer der gemeinsam genutzten Funktionen wartet. Bei mindestens einer Ausführungsform können, während ein wartender Thread schläft, Hardware-Ressourcen der Verarbeitung anderer Threads gewidmet werden. Zum Beispiel kann bei mindestens einer Ausführungsform eine Ausführungseinheit während einer Verzögerung, die

mit einer Vertex-Shader-Operation verbunden ist, Operationen für einen Pixel-Shader, Fragment-Shader oder eine andere Art von Shader-Programm durchführen, das einen anderen Vertex-Shader aufweist.

[0490] Bei mindestens einer Ausführungsform arbeitet jede Ausführungseinheit in den Ausführungseinheiten 4208A-4208N mit Anordnungen von Datenelementen. Bei mindestens einer Ausführungsform ist eine Anzahl von Datenelementen die „Ausführungsgröße“ oder die Anzahl von Kanälen für eine Anweisung. Bei mindestens einer Ausführungsform ist ein Ausführungskanal eine logische Ausführungseinheit für den Zugriff auf Datenelemente, die Maskierung und die Flusssteuerung innerhalb von Anweisungen. Bei mindestens einer Ausführungsform kann die Anzahl der Kanäle unabhängig von der Anzahl der physischen Arithmetic Logic Units (ALUs) oder Floating Point Units (FPUs) für einen bestimmten Grafikprozessor sein. Bei mindestens einer Ausführungsform unterstützen die Ausführungseinheiten 4208A-4208N Ganzzahl- und Gleitkomma-Datenarten.

[0491] Bei mindestens einer Ausführungsform weist der Befehlssatz einer Ausführungseinheit SIMD-Befehle auf. Bei mindestens einer Ausführungsform können verschiedene Datenelemente als gepackter Datentyp in einem Register gespeichert werden, und die Ausführungseinheit verarbeitet verschiedene Elemente basierend auf der Datengröße der Elemente. Beispielsweise werden bei mindestens einer Ausführungsform beim Arbeiten mit einem 256 Bit breiten Vektor 256 Bits eines Vektors in einem Register gespeichert, und eine Ausführungseinheit arbeitet mit einem Vektor als vier getrennte gepackte 64-Bit-Datenelemente (Datenelemente mit der Größe Quad-Word (QW)), acht getrennte 32-Bit-gepackte Datenelemente (Datenelemente mit der Größe Double-Word (DW)), sechzehn getrennte 16-Bit-gepackte Datenelemente (Datenelemente mit der Größe Word (W)) oder zweiunddreißig getrennte 8-Bit-Datenelemente (Datenelemente mit der Größe Byte (B). Bei mindestens einer Ausführungsform sind jedoch auch andere Vektorbreiten und Registergrößen möglich.

[0492] Bei mindestens einer Ausführungsform können eine oder mehrere Ausführungseinheiten zu einer fusionierten Ausführungseinheit 4209A-4209N mit einer Thread-Steuerungslogik (4207A-4207N) kombiniert werden, die den fusionierten EUs gemeinsam ist. Bei mindestens einer Ausführungsform können mehrere EUs zu einer EU-Gruppe verschmolzen werden. Bei mindestens einer Ausführungsform kann jede EU in einer fusionierten EU-Gruppe so ausgestaltet sein, dass sie einen separaten SIMD-Hardware-Thread ausführt. Die Anzahl der EUs in einer fusionierten EU-Gruppe kann je nach Ausführungsform variieren. Bei mindestens einer Ausführungsform können verschiedene SIMD-Breiten pro EU ausgeführt werden, die unter anderem SIMD8, SIMD16 und SIMD32 beinhalten. Bei mindestens einer Ausführungsform weist jede fusionierte Grafikausführungseinheit 4209A-4209N mindestens zwei Ausführungseinheiten auf. Beispielsweise beinhaltet die fusionierte Ausführungseinheit 4209A bei mindestens einer Ausführungsform eine erste EU 4208A, eine zweite EU 4208B und eine Thread-Steuerlogik 4207A, die der ersten EU 4208A und der zweiten EU 4208B gemeinsam ist. Bei mindestens einer Ausführungsform steuert die Thread-Steuerlogik 4207A Threads, die auf der fusionierten Grafikausführungseinheit 4209A ausgeführt werden, was es jeder EU innerhalb der fusionierten Ausführungseinheiten 4209A-4209N ermöglicht, unter Verwendung eines gemeinsamen Anweisungszeigerregisters auszuführen.

[0493] Bei mindestens einer Ausführungsform weist die Thread-Ausführungslogik 4200 einen oder mehrere interne Befehls-Caches (z. B. 4206) auf, um Thread-Befehle für Ausführungseinheiten zu cachen. Bei mindestens einer Ausführungsform sind ein oder mehrere Daten-Caches (z. B 4212) vorhanden, um Thread-Daten während der Thread-Ausführung zu cachen. Bei mindestens einer Ausführungsform ist ein Sampler 4210 vorhanden, um Textur-Sampling für 3D-Operationen und Medien-Sampling für MedienOperationen bereitzustellen. Bei mindestens einer Ausführungsform weist der Sampler 4210 eine spezielle Textur- oder Mediensampling-Funktionalität auf, um Textur- oder Mediendaten während des Sampling-Prozesses zu verarbeiten, bevor er die gesampelten Daten an eine Ausführungseinheit weitergibt.

[0494] Bei mindestens einer Ausführungsform senden Grafik- und Medienpipelines während der Ausführung Thread-Initiierungsanforderungen an die Thread-Ausführungslogik 4200 über die Thread-Erzeugungs- und Versandlogik. Bei mindestens einer Ausführungsform wird, sobald eine Gruppe geometrischer Objekte verarbeitet und in Pixeldaten gerastert wurde, die Pixelprozessorlogik (z. B. Pixel-Shader-Logik, Fragment-Shader-Logik usw.) innerhalb des Shader-Prozessors 4202 aufgerufen, um weiter Ausgabeinformationen zu berechnen und zu veranlassen, dass die Ergebnisse in Ausgabeflächen (z. B. Farbpuffer, Tiefenpuffer, Schablonenpuffer usw.) geschrieben werden. Bei mindestens einer Ausführungsform berechnet ein Pixel-Shader oder Fragment-Shader die Werte verschiedener Vertex-Attribute, die über ein gerastertes Objekt interpoliert werden sollen. Bei mindestens einer Ausführungsform führt die Pixelprozessorlogik innerhalb des Shader-Prozessors 4202 dann ein über eine Anwendungsprogrammierschnittstelle (API) bereitgestelltes

Pixel- oder Fragment-Shader-Programm aus. Bei mindestens einer Ausführungsform leitet der Shader-Prozessor 4202 zur Ausführung eines Shader-Programms Threads über den Thread-Dispatcher 4204 an eine Ausführungseinheit (z. B. 4208A) weiter. Bei mindestens einer Ausführungsform verwendet der Shader-Prozessor 4202 die Texturabtastlogik im Abtaster 4210, um auf Texturdaten in den im Speicher abgelegten Texturkarten zuzugreifen. Bei mindestens einer Ausführungsform werden durch arithmetische Operationen an Texturdaten und Eingabegeometriedaten Pixelfarbdaten für jedes geometrische Fragment berechnet oder ein oder mehrere Pixel von der weiteren Verarbeitung ausgeschlossen.

[0495] Bei mindestens einer Ausführungsform stellt der Datenanschluss 4214 einen Speicherzugriffsmechanismus für die Thread-Ausführungslogik 4200 bereit, um verarbeitete Daten zur weiteren Verarbeitung auf einer Grafikprozessor-Ausgabepipeline in den Speicher auszugeben. Bei mindestens einer Ausführungsform beinhaltet der Datenanschluss 4214 einen oder mehrere Cache-Speicher (z. B den Daten-Cache 4212) oder ist mit diesen gekoppelt, um Daten für den Speicherzugriff über einen Datenanschluss zwischenzuspeichern.

[0496] Wie in **Fig. 42B** veranschaulicht ist, kann eine Grafikausführungseinheit 4208 bei mindestens einer Ausführungsform eine Anweisungsabrufeinheit 4237, eine allgemeine Registerdateianordnung (GRF) 4224, eine architektonische Registerdateianordnung (ARF) 4226, einen Thread-Arbiter 4222, eine Sendeeinheit 4230, eine Verzweigungseinheit 4232, einen Satz SIMD-Gleitkommaeinheiten (FPUs) 4234 und bei mindestens einer Ausführungsform einen Satz dedizierter ganzzahliger SIMD-ALUs 4235 beinhalten. Bei mindestens einer Ausführungsform weisen die GRF 4224 und die ARF 4226 einen Satz allgemeiner Registerdateien und Architekturregisterdateien auf, die jedem gleichzeitigen Hardware- Thread zugeordnet sind, der in der Grafikausführungseinheit 4208 aktiv sein kann. Bei mindestens einer Ausführungsform wird der architektonische Zustand pro Thread in der ARF 4226 verwaltet, während die während der Thread-Ausführung verwendeten Daten in der GRF 4224 gespeichert werden. Bei mindestens einer Ausführungsform kann der Ausführungszustand jedes Threads, der Befehlszeiger für jeden Thread aufweist, in Thread-spezifischen Registern in der ARF 4226 gehalten werden.

[0497] Bei mindestens einer Ausführungsform weist die Grafikausführungseinheit 4208 eine Architektur, die eine Kombination aus simultanem Multi-Threading (SMT) und feinkörnigem Interleaved Multi-Threading (IMT) ist, auf. Bei mindestens einer Ausführungsform weist die Architektur eine modulare Konfiguration auf, die zur Entwurfszeit auf der Grundlage einer angestrebten Anzahl gleichzeitiger Threads und der Anzahl von Registern pro Ausführungseinheit fein abgestimmt werden kann, wobei die Ressourcen der Ausführungseinheit auf die Logik aufgeteilt werden, die zur Ausführung mehrerer gleichzeitiger Threads verwendet wird.

[0498] Bei mindestens einer Ausführungsform kann die Grafikausführungseinheit 4208 mehrere Befehle gemeinsam ausgeben, die jeweils unterschiedliche Befehle sein können. Bei mindestens einer Ausführungsform kann der Thread-Arbiter 4222 des Thread der Grafikausführungseinheit 4208 Anweisungen an eine der Sendeeinheiten 4230, Verzweigungseinheiten 4242 oder SIMD-FPU(s) 4234 zur Ausführung weiterleiten. Bei mindestens einer Ausführungsform kann jeder Thread auf 128 Allzweckregister innerhalb der GRF 4224 zugreifen, wobei jedes Register 32 Byte speichern kann, die als SIMD-8-Element-Vektor von 32-Bit-Datenelementen zugänglich sind. Bei mindestens einer Ausführungsform hat jeder Thread der Ausführungseinheit Zugriff auf 4 KByte innerhalb der GRF 4224, obwohl die Ausführungsformen nicht so beschränkt sind und bei anderen Ausführungen mehr oder weniger Registerressourcen bereitgestellt werden können. Bei mindestens einer Ausführungsform können bis zu sieben Threads gleichzeitig ausgeführt werden, obwohl die Anzahl der Threads pro Ausführungseinheit je nach Ausführungsform auch variieren kann. Bei mindestens einer Ausführungsform, bei der sieben Threads auf 4 KByte zugreifen können, kann die GRF 4224 insgesamt 28 KByte speichern. Bei mindestens einer Ausführungsform können flexible Adressierungsmodi ermöglichen, dass Register gemeinsam adressiert werden, um effektiv breitere Register zu bilden oder um strided rechteckige Blockdatenstrukturen darzustellen.

[0499] Bei mindestens einer Ausführungsform werden Speicheroperationen, Abtastoperationen und andere Systemkommunikationen mit längerer Latenzzeit über „Sende“-Befehle abgewickelt, die von einer Nachrichten-Durchlass-Sendeeinheit 4230 ausgeführt werden. Bei mindestens einer Ausführungsform werden Verzweigungsbefehle an eine dedizierte Verzweigungseinheit 4232 weitergeleitet, um Divergenz und eventuelle Konvergenz bezüglich SIMD zu ermöglichen.

[0500] Bei mindestens einer Ausführungsform weist die Grafikausführungseinheit 4208 eine oder mehrere SIMD-Gleitkommaeinheiten (FPU(s)) 4234 auf, um Gleitkommaoperationen durchzuführen. Bei mindestens einer Ausführungsform unterstützen die FPU(s) 4234 auch Ganzzahlberechnungen. Bei mindestens einer Ausführungsform kann (können) die FPU(s) 4234 bis zu einer Anzahl M von 32-Bit-Gleitkomma- (oder Ganz-

zahl-) Operationen oder bis zu 2M 16-Bit-Ganzzahl- oder 16-Bit-GleitkommaOperationen bezüglich SIMD ausführen. Bei mindestens einer Ausführungsform stellt mindestens eine der FPU(s) erweiterte mathematische Fähigkeiten bereit, um transzendente mathematische Funktionen mit hohem Durchsatz und 64-Bit-Gleitkomma mit doppelter Genauigkeit zu unterstützen. Bei mindestens einer Ausführungsform ist auch ein Satz von 8-Bit-Integer-SIMD-ALUs 4235 vorhanden, die speziell für die Durchführung von Operationen im Zusammenhang mit Berechnungen zum maschinellen Lernen optimiert sein können.

[0501] Bei mindestens einer Ausführungsform können Anordnungen aus mehreren Instanzen der Grafikausführungseinheit 4208 in einer Grafik-Teilkern-Gruppierung (z. B. einem Unter-Slice) instanziiert werden. Bei mindestens einer Ausführungsform kann die Ausführungseinheit 4208 Anweisungen über eine Vielzahl von Ausführungskanälen ausführen. Bei mindestens einer Ausführungsform wird jeder Thread, der auf der Grafikausführungseinheit 4208 ausgeführt wird, auf einem anderen Kanal ausgeführt.

[0502] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform können Teile der oder die gesamte Ableitungs- und/oder Trainingslogik 1815 in die Ausführungslogik 4200 integriert sein. Darüber hinaus können bei mindestens einer Ausführungsform in dieser Schrift beschriebene Ableitungs- und/oder Trainingsoperationen unter Verwendung einer anderen Logik als der in den **Fig. 18A** oder **Fig. 18B** veranschaulichten Logik ausgeführt werden. Bei mindestens einer Ausführungsform können Gewichtungsparameter in chipinternen oder chipexternen Speicher und/oder Registern (gezeigt oder nicht gezeigt) gespeichert werden, die ALUs der Thread-Ausführungslogik 4200 konfigurieren, um einen oder mehrere Algorithmen maschinellen Lernens, neuronale Netzwerkarchitekturen, Anwendungsfälle oder in dieser Schrift beschriebene Trainingsmethoden auszuführen.

[0503] Bei mindestens einer Ausführungsform wird die Ausführungslogik 4200 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird die Ausführungslogik 4200 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird die Ausführungslogik 4200 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und die Ausführungslogik 4200 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes auf selbstüberwachte Weise zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0504] **Fig. 43** zeigt eine Parallelverarbeitungseinheit („PPU“) 4300 gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform ist die PPU 4300 mit maschinenlesbarem Code ausgestaltet, der, wenn er von der PPU 4300 ausgeführt wird, die PPU 4300 veranlasst, einige oder alle der in dieser Offenbarung beschriebenen Prozesse und Techniken durchzuführen. Bei mindestens einer Ausführungsform ist die PPU 4300 ein Multi-Thread-Prozessor, der auf einer oder mehreren integrierten Einrichtungen implementiert ist und der Multithreading als eine Technik zum Verbergen von Latenzzeiten verwendet, die dazu dient, computerlesbare Befehle (auch als maschinenlesbare Befehle oder einfach Befehle bezeichnet) auf mehreren Threads parallel zu verarbeiten. Bei mindestens einer Ausführungsform bezieht sich ein Thread auf einen Ausführungsstrang und ist eine Instanziierung eines Satzes von Anweisungen, die zur Ausführung durch die PPU 4300 konfiguriert sind. Bei mindestens einer Ausführungsform ist die PPU 4300 eine Grafikverarbeitungseinheit („GPU“), die so konfiguriert ist, dass sie eine Grafik-Rendering-Pipeline zur Verarbeitung dreidimensionaler („3D“) Grafikdaten implementiert, um zweidimensionale („2D“) Bilddaten für die Anzeige auf einer Einrichtung wie einer Flüssigkristallanzeige („LCD“) zu erzeugen. Bei mindestens einer Ausführungsform wird die PPU 4300 verwendet, um Berechnungen wie lineare Algebra-Operationen und Operationen des maschinellen Lernens durchzuführen. **Fig. 43** zeigt ein Beispiel für einen Parallelprozessor, der nur zur Veranschaulichung dient und als nicht begrenzendes Beispiel für Prozessorarchitekturen zu verstehen ist, die im Rahmen dieser Offenbarung in Betracht gezogen werden, wobei jeder geeignete Prozessor zur Ergänzung und/oder zum Ersatz desselben verwendet werden kann.

[0505] Bei mindestens einer Ausführungsform sind eine oder mehrere PPUs 4300 so ausgestaltet, dass sie Anwendungen für High Performance Computing („HPC“), Rechenzentren und maschinelles Lernen beschleunigen. Bei mindestens einer Ausführungsform ist die PPU 4300 dazu konfiguriert, Deep-Learning-Systeme

und -Anwendungen zu beschleunigen, was die folgenden nicht einschränkenden Beispiele beinhaltet: autonome Fahrzeugplattformen, Deep Learning, hochpräzise Sprache, Bilder, Texterfassungssysteme, intelligente Videoanalyse, molekulare Simulationen, Arzneimittelentdeckung, Krankheitsdiagnose, Wettervorhersage, Big-Data-Analyse, Astronomie, Molekulardynamiksimulation, Finanzmodellierung, Robotik, Fabrikautomatisierung, Echtzeit-Sprachübersetzung, Online-Suchoptimierung und personalisierte Benutzerempfehlungen und mehr.

[0506] Bei mindestens einer Ausführungsform weist die PPU 4300 ohne Einschränkung eine Input/Output (E/A-)-Einheit 4306, eine Front-End-Einheit 4310, eine Scheduler-Einheit 4312, eine Arbeitsverteilungseinheit 4314, einen Hub 4316, ein Koppelfeld („Xbar“) 4320, einen oder mehrere allgemeine Verarbeitungscluster („GPCs“) 4318 und eine oder mehrere Partitionseinheiten („Speicherpartitionseinheiten“) 4322 auf. Bei mindestens einer Ausführungsform ist die PPU 4300 mit einem Host-Prozessor oder anderen PPUs 4300 über eine oder mehrere Hochgeschwindigkeits-GPU-Verbindungen („GPU-Interconnects“) 4308 verbunden. Bei mindestens einer Ausführungsform ist die PPU 4300 mit einem Host-Prozessor oder anderen peripheren Einrichtungen über eine Zwischenverbindung 4302 verbunden. Bei mindestens einer Ausführungsform ist die PPU 4300 mit einem lokalen Speicher verbunden, der eine oder mehrere Speichereinrichtungen („Speicher“) 4304 umfasst. Bei mindestens einer Ausführungsform weisen die Speichereinrichtungen 4304 ohne Einschränkung eine oder mehrere dynamische Direktzugriffsspeicher („DRAM“)-Einrichtungen auf. Bei mindestens einer Ausführungsform sind eine oder mehrere DRAM-Vorrichtungen als Subsysteme mit Speicher mit hoher Bandbreite („HBM“) ausgestaltet und/oder konfigurierbar, wobei in jeder Einrichtung mehrere DRAM-Dies gestapelt sind.

[0507] Bei mindestens einer Ausführungsform kann das Hochgeschwindigkeits-GPU-Interconnect 4308 eine verdrahtete Mehrspur-Kommunikationsverbindung betreffen, die von Systemen verwendet wird, die skalierbar sind und eine oder mehrere PPUs 4300 beinhalten, die mit einer oder mehreren Zentraleinheiten („CPUs“) kombiniert sind, und die Cache-Kohärenz zwischen PPUs 4300 und CPUs sowie CPU-Mastering unterstützt. Bei mindestens einer Ausführungsform werden Daten und/oder Befehle durch die Hochgeschwindigkeits-GPU-Verbindung 4308 über den Hub 4316 zu/von anderen Einheiten der PPU 4300 übertragen, wie z. B. einer oder mehreren Kopiermaschinen, Video-Encodern, Video-Decodern, Energieverwaltungseinheiten und anderen Komponenten, die in **Fig. 43** möglicherweise nicht explizit dargestellt sind.

[0508] Bei mindestens einer Ausführungsform ist die E/A-Einheit 4306 so ausgestaltet, dass sie Kommunikationen (z. B. Befehle, Daten) von einem Host-Prozessor (in **Fig. 43** nicht dargestellt) über den Systembus 4302 sendet und empfängt. Bei mindestens einer Ausführungsform kommuniziert die E/A-Einheit 4306 mit dem Host-Prozessor direkt über den Systembus 4302 oder über eine oder mehrere zwischengeschaltete Einrichtungen wie z. B. eine Speicherbrücke. Bei mindestens einer Ausführungsform kann die E/A-Einheit 4306 mit einem oder mehreren anderen Prozessoren, z. B. einer oder mehreren PPUs 4300, über den Systembus 4302 kommunizieren. Bei mindestens einer Ausführungsform implementiert die E/A-Einheit 4306 eine Peripheral Component Interconnect Express („PCIe“) Schnittstelle für die Kommunikation über einen PCIe-Bus. Bei mindestens einer Ausführungsform implementiert die E/A-Einheit 4306 Schnittstellen für die Kommunikation mit externen Einrichtungen.

[0509] Bei mindestens einer Ausführungsform decodiert die E/A-Einheit 4306 über den Systembus 4302 empfangene Pakete. Bei mindestens einer Ausführungsform stellen mindestens einige Pakete Befehle dar, die so ausgestaltet sind, dass sie die PPU 4300 veranlassen, verschiedene Operationen durchzuführen. Bei mindestens einer Ausführungsform überträgt die E/A-Einheit 4306 decodierte Befehle an verschiedene andere Einheiten der PPU 4300, wie es von den Befehlen angegeben ist. Bei mindestens einer Ausführungsform werden Befehle an die Front-End-Einheit 4310 und/oder an den Hub 4316 oder andere Einheiten der PPU 4300, wie eine oder mehrere Copy-Engines, einen Video-Encoder, einen Video-Decoder, eine Energieverwaltungseinheit usw., übertragen (in **Fig. 43** nicht explizit veranschaulicht). Bei mindestens einer Ausführungsform ist die E/A-Einheit 4306 so ausgestaltet, dass sie die Kommunikation zwischen und unter verschiedenen logischen Einheiten der PPU 4300 leitet.

[0510] Bei mindestens einer Ausführungsform codiert ein vom Host-Prozessor ausgeführtes Programm einen Befehlsstrom in einem Puffer, der der PPU 4300 Arbeitslasten zur Verarbeitung bereitstellt. Bei mindestens einer Ausführungsform umfasst eine Arbeitslast Befehle und Daten, die von diesen Befehlen verarbeitet werden sollen. Bei mindestens einer Ausführungsform ist der Puffer ein Bereich in einem Speicher, auf den sowohl der Host-Prozessor als auch die PPU 4300 zugreifen können (z. B. Lese-/Schreibzugriff) - eine Host-Schnittstelleneinheit kann so ausgestaltet sein, dass sie auf den Puffer in einem mit dem Systembus 4302 verbundenen Systemspeicher über Speicheranforderungen zugreift, die von der E/A-Einheit 4306 über den

Systembus 4302 übertragen werden. Bei mindestens einer Ausführungsform schreibt ein Hostprozessor einen Befehlsstrom in einen Puffer und übermittelt dann einen Pointer zu einem Anfang eines Befehlsstroms an die PPU 4300 derart, dass die Front-End-Einheit 4310 Pointer auf einen oder mehrere Befehlsströme empfängt und einen oder mehrere Befehlsströme verwaltet, wobei Befehle aus Befehlsströmen gelesen und Befehle an verschiedene Einheiten der PPU 4300 weitergeleitet werden.

[0511] Bei mindestens einer Ausführungsform ist die Front-End-Einheit 4310 mit der Scheduler-Einheit 4312 gekoppelt, die verschiedene GPCs 4318 zur Verarbeitung von Tasks ausgestaltet, die durch einen oder mehrere Befehlsströme definiert sind. Bei mindestens einer Ausführungsform ist die Planer-Einheit 4312 derart konfiguriert, dass sie Zustandsinformationen in Bezug auf verschiedene, von der Planer-Einheit 4312 verwaltete Aufgaben verfolgt, wobei die Zustandsinformationen angeben können, welchem der GPCs 4318 eine Aufgabe zugewiesen ist, ob die Aufgabe aktiv oder inaktiv ist, welche Prioritätsstufe mit der Aufgabe assoziiert ist und so fort. Bei mindestens einer Ausführungsform verwaltet die Scheduler-Einheit 4312 die Ausführung einer Vielzahl von Tasks auf einem oder mehreren GPCs 4318.

[0512] Bei mindestens einer Ausführungsform ist die Planer-Einheit 4312 mit der Arbeitsverteilungseinheit 4314 gekoppelt, die derart konfiguriert ist, dass sie Aufgaben zur Ausführung auf den GPCs 4318 auswählt. Bei mindestens einer Ausführungsform verfolgt die Arbeitsverteilungseinheit 4314 eine Anzahl geplanter Tasks, die von der Planungseinheit 4312 empfangen wurden, und die Arbeitsverteilungseinheit 4314 verwaltet einen Pool ausstehender Tasks und einen Pool aktiver Tasks für jeden der GPCs 4318. Bei mindestens einer Ausführungsform umfasst der Pool ausstehender Aufgaben eine Anzahl von Slots (z. B 32 Slots), die Aufgaben enthalten, die zur Verarbeitung durch einen bestimmten GPC 4318 zugewiesen sind; der Pool aktiver Aufgaben kann eine Anzahl von Slots (z. B 4 Slots) für Aufgaben umfassen, die aktiv von den GPCs 4318 verarbeitet werden, so dass, während einer der GPCs 4318 die Ausführung einer Task abschließt, diese Aufgabe aus dem Pool aktiver Aufgaben für den GPC 4318 entfernt wird und eine der anderen Aufgaben aus dem Pool ausstehender Aufgaben ausgewählt und zur Ausführung auf dem GPC 4318 eingeplant wird. Bei mindestens einer Ausführungsform wird, wenn eine aktive Task auf dem GPC 4318 im Leerlauf ist, z. B während des Wartens auf die Auflösung einer Datenabhängigkeit, die aktive Task aus dem GPC 4318 entfernt und in den Pool der anstehenden Tasks zurückgeführt werden, während eine andere Task im Pool der anstehenden Tasks ausgewählt und für die Ausführung auf dem GPC 4318 eingeplant wird.

[0513] Bei mindestens einer Ausführungsform kommuniziert die Arbeitsverteilungseinheit 4314 mit einem oder mehreren GPCs 4318 über die XBar 4320. Bei mindestens einer Ausführungsform ist die XBar 4320 ein Verbindungsnetzwerk, das viele Einheiten der PPU 4300 mit anderen Einheiten der PPU 4300 verbindet und so ausgestaltet werden kann, dass es die Arbeitsverteilungseinheit 4314 mit einem bestimmten GPC 4318 verbindet. Bei mindestens einer Ausführungsform können auch eine oder mehrere andere Einheiten der PPU 4300 über den Hub 4316 mit der XBar 4320 verbunden sein.

[0514] Bei mindestens einer Ausführungsform werden die Tasks von der Scheduler-Einheit 4312 verwaltet und von der Arbeitsverteilungseinheit 4314 an einen der GPCs 4318 weitergeleitet. Der GPC 4318 ist ausgestaltet, um Tasks zu verarbeiten und Ergebnisse zu erzeugen. Bei mindestens einer Ausführungsform können die Ergebnisse von anderen Aufgaben innerhalb des GPC 4318 aufgenommen, über die XBar 4320 an einen anderen GPC 4318 geleitet oder im Speicher 4304 abgelegt werden. Bei mindestens einer Ausführungsform können die Ergebnisse in den Speicher 4304 über Partitionseinheiten 4322 geschrieben werden, die eine Speicherschnittstelle zum Lesen und Schreiben von Daten in/aus dem Speicher 4304 implementieren. Bei mindestens einer Ausführungsform können die Ergebnisse über eine Hochgeschwindigkeits-GPU-Verbindung 4308 an eine andere PPU 4304 oder CPU übertragen werden. Bei mindestens einer Ausführungsform weist die PPU 4300 ohne Einschränkung eine Anzahl U von Partitionseinheiten 4322 auf, die der Anzahl der mit der PPU 4300 verbundenen separaten und unterschiedlichen Speichereinrichtungen 4304 entspricht. Bei mindestens einer Ausführungsform wird die Partitionseinheit 4322 hier in Verbindung mit **Fig. 45** ausführlicher beschrieben.

[0515] Bei mindestens einer Ausführungsform führt ein Host-Prozessor einen Treiberkern aus, der eine Anwendungsprogrammierschnittstelle („API“) implementiert, die es einer oder mehreren auf dem Host-Prozessor ausgeführten Anwendungen ermöglicht, Operationen zur Ausführung auf der PPU 4300 zu planen. Bei mindestens einer Ausführungsform werden mehrere Rechenanwendungen gleichzeitig von der PPU 4300 ausgeführt, und die PPU 4300 stellt Isolierung, Dienstgüte („QoS“) und unabhängige Adressräume für mehrere Rechenanwendungen bereit. Bei mindestens einer Ausführungsform generiert eine Anwendung Anweisungen (z. B. in Form von API-Aufrufen), die den Treiberkern veranlassen, eine oder mehrere Tasks zur Ausführung durch die PPU 4300 zu generieren, und der Treiberkern gibt Tasks an einen oder mehrere

Streams aus, die von der PPU 4300 verarbeitet werden. Bei mindestens einer Ausführungsform umfasst jede Task eine oder mehrere Gruppen von zusammenhängenden Threads, die als Warp bezeichnet werden können. Bei mindestens einer Ausführungsform umfasst ein Warp eine Vielzahl zusammengehöriger Threads (z. B. 32 Threads), die parallel ausgeführt werden können. Bei mindestens einer Ausführungsform können sich kooperierende Threads auf eine Vielzahl von Threads beziehen, die Anweisungen zur Ausführung von Tasks aufweisen und Daten über einen gemeinsamen Speicher austauschen. Bei mindestens einer Ausführungsform werden Threads und kooperierende Threads gemäß mindestens einer Ausführungsform in Verbindung mit **Fig. 45** ausführlicher beschrieben.

[0516] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform wird der Deep-Learning-Anwendungsprozessor verwendet, um ein Modell maschinellen Lernens, wie etwa ein neuronales Netzwerk, zu trainieren, um der PPU 4300 bereitgestellte Informationen vorherzusagen oder abzuleiten. Bei mindestens einer Ausführungsform wird der Deep-Learning-Anwendungsprozessor 4300 verwendet, um Informationen basierend auf eines trainierten Modells maschinellen Lernens (z. B. neuronales Netzwerk) abzuleiten oder vorherzusagen, das von einem anderen Prozessor oder System oder von der PPU 4300 trainiert wurde. Bei mindestens einer Ausführungsform kann die PPU 4300 verwendet werden, um einen oder mehrere in dieser Schrift beschriebene Anwendungsfälle für neuronale Netzwerke auszuführen.

[0517] Bei mindestens einer Ausführungsform wird die PPU 4300 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird die PPU 4300 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird die PPU 4300 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und die PPU 4300 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes selbstüberwacht zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0518] **Fig. 44** illustriert einen allgemeinen Verarbeitungscluster („GPC“) 4400 gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform handelt es sich bei dem GPC 4400 um den GPC 4318 aus **Fig. 43**. Bei mindestens einer Ausführungsform weist jeder GPC 4400 ohne Einschränkung eine Anzahl von Hardware-Einheiten für die Verarbeitung von Tasks auf, und jeder GPC 4400 weist ohne Einschränkung einen Pipeline-Manager 4402, eine Pre-Raster-Operationseinheit („PROP“) 4404, eine Raster-Maschine 4408, ein Arbeitsverteilungs-Koppelfeld („WDX“) 4416, eine Speicherverwaltungseinheit („MMU“) 4418, einen oder mehrere Datenverarbeitungscluster („DPCs“) 4406 und jede geeignete Kombination von Teilen auf.

[0519] Bei mindestens einer Ausführungsform wird der Betrieb des GPC 4400 durch den Pipeline-Manager 4402 gesteuert. Bei mindestens einer Ausführungsform verwaltet der Pipeline-Manager 4402 die Konfiguration eines oder mehrerer DPCs 4406 für die Verarbeitung von Tasks, die dem GPC 4400 zugewiesen sind. Bei mindestens einer Ausführungsform konfiguriert der Pipeline-Manager 4402 mindestens einen von einem oder mehreren DPCs 4406, um mindestens einen Abschnitt einer Grafik-Rendering-Pipeline zu implementieren. Bei mindestens einer Ausführungsform ist der DPC 4406 derart konfiguriert, dass er ein Vertex-Shader-Programm auf einem programmierbaren Streaming-Multiprozessor („SM“) 4414 ausführt. Bei mindestens einer Ausführungsform ist der Pipeline-Manager 4402 derart konfiguriert, dass er die von einer Arbeitsverteilungseinheit empfangenen Pakete an geeignete logische Einheiten innerhalb des GPC 4400 leitet, wobei einige Pakete an Hardwareeinheiten mit fester Funktion in dem PROP 4404 und/oder in der Rastermaschine 4408 weitergeleitet werden können, während andere Pakete an DPCs 4406 zur Verarbeitung durch eine Primitivmaschine 4412 oder SM 4414 geleitet werden können. Bei mindestens einer Ausführungsform konfiguriert der Pipeline-Manager 4402 mindestens einen der DPCs 4406 zur Implementierung eines Modells eines neuronalen Netzes und/oder einer Rechenpipeline.

[0520] Bei mindestens einer Ausführungsform ist die PROP-Einheit 4404 so ausgestaltet, dass sie die von der Rastermaschine 4408 und den DPCs 4406 erzeugten Daten an eine Raster Operations („ROP“)-Einheit in der Partitionseinheit 4322 weiterleitet, die oben in Verbindung mit **Fig. 43** ausführlicher beschrieben ist.

Bei mindestens einer Ausführungsform ist die PROP-Einheit 4404 so ausgestaltet, dass sie Optimierungen für die Farbmischung durchführt, Pixeldaten organisiert, Adressübersetzungen vornimmt und vieles mehr. Bei mindestens einer Ausführungsform weist die Rastermaschine 4408 ohne Einschränkung eine Reihe von Hardware-Einheiten mit fester Funktion auf, die so konfiguriert sind, dass sie verschiedene Rasteroperationen durchführen, und die Rastermaschine 4408 weist ohne Einschränkung eine Setup-Maschine, eine Grobraster-Maschine, eine Culling-Maschine, eine Clipping-Maschine, eine Feinraster-Maschine, eine Tile-Coalescing-Maschine und eine beliebige geeignete Kombination davon auf. Bei mindestens einer Ausführungsform empfängt die Setup-Maschine transformierte Vertices und erzeugt Ebenengleichungen, die mit einer durch Vertices definierten geometrischen Primitive verbunden sind; die Ebenengleichungen werden an die Grobraster-Maschine übertragen, um Abdeckungsinformationen (z. B. eine x-, y-Abdeckungsmaske für eine Kachel) für die Primitive zu erzeugen; die Ausgabe der Grobraster-Maschine wird an die Culling-Maschine übertragen, wo Fragmente, die dem Primitive zugeordnet sind und einen z-Test nicht bestehen, aussortiert werden, und an eine Clipping-Maschine übertragen, wo Fragmente, die außerhalb eines Sichtkegelvolumens liegen, abgeschnitten werden. Bei mindestens einer Ausführungsform werden die Fragmente, die das Clipping und Culling überstehen, an eine Feinraster-Maschine weitergeleitet, um Attribute für Pixelfragmente auf der Grundlage der von der Setup-Maschine erstellten Ebenengleichungen zu erzeugen. Bei mindestens einer Ausführungsform umfasst die Ausgabe der Raster-Maschine 4408 Fragmente, die von einer beliebigen geeigneten Einheit, wie z. B. einem in DPC 4406 implementierten Fragment-Shader, verarbeitet werden.

[0521] Bei mindestens einer Ausführungsform beinhaltet jeder DPC 4406, der in dem GPC 4400 enthalten ist, ohne Einschränkung eine M-Pipe-Steuervorrichtung (M-Pipe Controller - „MPC“) 4410; eine primitive Engine 4412; ein oder mehrere SMs 4414; und eine beliebige geeignete Kombination davon. Bei mindestens einer Ausführungsform steuert die MPC 4410 den Betrieb der DPC 4406 und leitet die von dem Pipeline-Manager 4402 empfangenen Pakete an die entsprechenden Einheiten im DPC 4406. Bei mindestens einer Ausführungsform werden Pakete, die einem Vertex zugeordnet sind, an die Primitiv-Maschine 4412 weitergeleitet, die so ausgestaltet ist, dass sie Vertex-Attribute, die dem Vertex zugeordnet sind, aus dem Speicher abruft; im Gegensatz dazu können Pakete, die einem Shader-Programm zugeordnet sind, an den SM 4414 übertragen werden.

[0522] Bei mindestens einer Ausführungsform umfasst der SM 4414 ohne Einschränkung einen programmierbaren Streaming-Prozessor, der so gestaltet ist, dass er Tasks verarbeitet, die durch eine Anzahl von Threads dargestellt werden. Bei mindestens einer Ausführungsform ist der SM 4414 multi-threaded und so ausgestaltet, dass er eine Vielzahl von Threads (z. B 32 Threads) aus einer bestimmten Gruppe von Threads gleichzeitig ausführt und eine Single-Instruction, Multiple-Data („SIMD“)-Architektur implementiert, bei der jeder Thread in einer Gruppe von Threads (z. B ein Warp) so ausgestaltet ist, dass er einen anderen Datensatz auf der Grundlage desselben Satzes von Anweisungen verarbeitet. Bei mindestens einer Ausführungsform führen alle Threads in einer Gruppe von Threads dieselben Befehle aus. Bei mindestens einer Ausführungsform implementiert der SM 4414 eine Single-Instruction, Multiple Thread („SIMT“)-Architektur, bei der jeder Thread in einer Gruppe von Threads so ausgestaltet ist, dass er einen anderen Datensatz auf der Grundlage desselben Befehlssatzes verarbeitet, wobei jedoch die einzelnen Threads in der Gruppe von Threads während der Ausführung divergieren dürfen. Bei mindestens einer Ausführungsform werden ein Programmzähler, ein Aufrufstack und ein Ausführungsstatus für jeden Warp gehalten, wodurch die Gleichzeitigkeit zwischen Warps und die serielle Ausführung innerhalb von Warps ermöglicht wird, wenn Threads innerhalb eines Warps divergieren. In einer anderen Ausführungsform werden ein Programmzähler, ein Aufrufstack und ein Ausführungsstatus für jeden einzelnen Thread gehalten, was eine gleiche Nebenläufigkeit zwischen allen Threads innerhalb und zwischen Warps ermöglicht. Bei mindestens einer Ausführungsform wird der Ausführungsstatus für jeden einzelnen Thread gehalten, und Threads, die dieselben Befehle ausführen, können zur Verbesserung der Effizienz zusammengeführt und parallel ausgeführt werden. Mindestens eine Ausführungsform des SM 4414 wird hier ausführlicher beschrieben.

[0523] Bei mindestens einer Ausführungsform stellt die MMU 4418 eine Schnittstelle zwischen dem GPC 4400 und der Speicherpartitionseinheit (z. B. der Partitionseinheit 4322 in **Fig. 43**) bereit, und die MMU 4418 sorgt für die Übersetzung virtueller Adressen in physikalische Adressen, den Speicherschutz und die Konkurrenzbereinigung von Speicheranforderungen. Bei mindestens einer Ausführungsform stellt die MMU 4418 einen oder mehrere Übersetzungs-Lookaside-Puffer („TLBs“) zur Durchführung der Übersetzung virtueller Adressen in physische Adressen im Speicher bereit.

[0524] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der

Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform wird der Deep-Learning-Anwendungsprozessor verwendet, um ein Modell maschinellen Lernens, wie etwa ein neuronales Netzwerk, zu trainieren, um der GPC 4400 bereitgestellte Informationen vorherzusagen oder abzuleiten. Bei mindestens einer Ausführungsform wird die GPC 4400 verwendet, um Informationen basierend auf eines trainierten Modells maschinellen Lernens (z. B. neuronales Netzwerk), das von einem anderen Prozessor oder System oder von der GPC 4400 trainiert wurde, abzuleiten oder vorherzusagen. Bei mindestens einer Ausführungsform kann die GPC 4400 verwendet werden, um einen oder mehrere in dieser Schrift beschriebene Anwendungsfälle für neuronale Netzwerke auszuführen.

[0525] Bei mindestens einer Ausführungsform wird die GPC 4400 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird die GPC 4400 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird die GPC 4400 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und die GPC 4400 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes selbstüberwacht zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0526] Fig. 45 veranschaulicht eine Speicherpartitionseinheit 4500 einer Parallelverarbeitungseinheit („PPU“) bei mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform beinhaltet die Speicherpartitionseinheit 4500 ohne Einschränkung eine Rasteroperationseinheit („ROP“-Einheit) 4502; einen Stufe-Zwei-Cache („L2“-Cache) 4504; eine Speicherschnittstelle 4506; und eine beliebige geeignete Kombination davon; die Speicherschnittstelle 4506 ist mit dem Speicher gekoppelt; die Speicherschnittstelle 4506 kann 32-, 64-, 128-, 1024-Bit-Datenbusse oder dergleichen für eine Hochgeschwindigkeitsdatenübertragung umsetzen. Bei mindestens einer Ausführungsform umfasst die PPU U Speicherschnittstellen 4506, eine Speicherschnittstelle 4506 pro Paar von Partitionseinheiten 4500, wobei jedes Paar von Partitionseinheiten 4500 mit einer entsprechenden Speichereinrichtung verbunden ist. Bei mindestens einer Ausführungsform kann die PPU beispielsweise mit bis zu Y Speichereinrichtungen verbunden sein, wie mit Speicherstapeln mit hoher Bandbreite oder mit einem synchronen dynamischen wahlfreien Grafikspeicher mit doppelter Datenrate, Version 5 („GDDR5 SDRAM“).

[0527] Bei mindestens einer Ausführungsform implementiert die Speicherschnittstelle 4506 eine Speicherschnittstelle der zweiten Generation mit hoher Bandbreite („HBM2“), und Y ist gleich der Hälfte von U. Bei mindestens einer Ausführungsform befinden sich die HBM2-Speicherstacks auf demselben physischen Gehäuse wie die PPU, was im Vergleich zu herkömmlichen GDDR5-SDRAM-Systemen erhebliche Energie- und Flächeneinsparungen ermöglicht. Bei mindestens einer Ausführungsform weist jeder HBM2-Stack, ohne Einschränkung, vier Speicherchips auf und Y ist gleich 4, wobei jeder HBM2-Stack zwei 128-Bit-Kanäle pro Chip für insgesamt 8 Kanäle und eine Datenbusbreite von 1024 Bit aufweist. Bei mindestens einer Ausführungsform unterstützt der Speicher den Single-Error Correcting Double-Error Detecting („SECDED“) Error Correction Code („ECC“) zum Schutz der Daten. ECC bietet eine höhere Zuverlässigkeit für Datenverarbeitungsanwendungen, die empfindlich auf Datenverfälschung reagieren.

[0528] Bei mindestens einer Ausführungsform implementiert die PPU eine mehrstufige Speicherhierarchie. Bei mindestens einer Ausführungsform unterstützt die Speicherpartitionierungseinheit 4500 einen einheitlichen Speicher, um einen einzigen einheitlichen virtuellen Adressraum für die Zentraleinheit („CPU“) und den PPU-Speicher bereitzustellen, was die gemeinsame Nutzung von Daten zwischen virtuellen Speichersystemen ermöglicht. Bei mindestens einer Ausführungsform wird die Häufigkeit der Zugriffe einer PPU auf Speicher auf anderen Prozessoren verfolgt, um sicherzustellen, dass Speicherseiten in den physischen Speicher der PPU, die häufiger Zugriffe auf Seiten vornimmt, verschoben werden. Bei mindestens einer Ausführungsform unterstützt das Hochgeschwindigkeits-GPU-Interconnect 4308 Adressübersetzungsdienste, die es der PPU ermöglichen, direkt auf die Seitentabellen der CPU zuzugreifen und der PPU vollen Zugriff auf den CPU-Speicher bereitzustellen.

[0529] Bei mindestens einer Ausführungsform übertragen Kopiermodule Daten zwischen mehreren PPUs oder zwischen PPUs und CPUs. Bei mindestens einer Ausführungsform können Kopiermodule Seitenfehler für Adressen erzeugen, die nicht in Seitentabellen abgebildet sind, und die Speicherpartitionierungseinheit

4500 bearbeitet dann die Seitenfehler, indem sie die Adressen in die Seitentabelle abbildet, woraufhin das Kopiermodul die Übertragung durchführt. Bei mindestens einer Ausführungsform ist der Speicher für mehrere Copy-Engine-Operationen zwischen mehreren Prozessoren gepinnt (d. h. nicht auslagerbar), wodurch der verfügbare Speicher erheblich verringert wird. Bei mindestens einer Ausführungsform können mit Hardware für Seitenfehler Adressen an Kopiermaschinen weitergegeben werden, ohne Rücksicht darauf, ob Speicherseiten resident sind, und der Kopiervorgang ist transparent.

[0530] Daten aus dem Speicher 4304 der **Fig. 43** oder einem anderen Systemspeicher werden von der Speicherpartitionseinheit 4500 abgerufen und in dem L2-Cache 4504, der sich auf dem Chip befindet und bei mindestens einer Ausführungsform von verschiedenen GPCs gemeinsam genutzt wird, gespeichert. Jede Speicherpartitionseinheit 4500 weist bei mindestens einer Ausführungsform ohne Einschränkung mindestens einen Abschnitt des L2-Cache auf, der einer entsprechenden Einrichtung zugeordnet ist. Bei mindestens einer Ausführungsform sind Caches der unteren Ebene in verschiedenen Einheiten innerhalb von GPCs implementiert. Bei mindestens einer Ausführungsform kann jeder der SMs 4414 einen Cache der Ebene eins („L1“) implementieren, wobei der L1-Cache ein privater Speicher ist, der einem bestimmten SM 4414 zugeordnet ist, und Daten aus dem L2-Cache 4504 abgerufen und in jedem der L1-Caches zur Verarbeitung in Funktionseinheiten der SMs 4414 gespeichert werden. Bei mindestens einer Ausführungsform ist der L2-Cache 4504 mit der Speicherschnittstelle 4506 und der XBar 4320 verbunden.

[0531] Die ROP-Einheit 4502 führt bei mindestens einer Ausführungsform Grafikrasteroperationen durch, die sich auf die Pixelfarbe beziehen, wie z. B. Farbkomprimierung, Pixelüberblendung und mehr. Bei mindestens einer Ausführungsform setzt OP-Einheit 4502 eine Tiefenprüfung in Verbindung mit der Raster-Engine 4408 um, die eine Tiefe für eine Abtastposition, die mit einem Pixelfragment assoziiert ist, von der Culling-Maschine der Rastermaschine 4408 empfängt. Bei mindestens einer Ausführungsform wird die Tiefe gegen eine entsprechende Tiefe in einem Tiefenpuffer für einen mit dem Fragment verbundenen Probenort getestet. Bei mindestens einer Ausführungsform aktualisiert die ROP-Einheit 4502 den Tiefenpuffer und überträgt das Ergebnis des Tiefentests an die Rastermaschine 4408, wenn das Fragment den Tiefentest für den Probenort besteht. Es wird deutlich, dass die Anzahl der Partitionseinheiten 4500 von der Anzahl der GPCs abweichen kann, und daher kann jede ROP-Einheit 4502 bei mindestens einer Ausführungsform mit jedem der GPCs gekoppelt sein. Bei mindestens einer Ausführungsform verfolgt die ROP-Einheit 4502 die von verschiedenen GPCs empfangenen Pakete und bestimmt, an welche ein von der ROP-Einheit 4502 erzeugtes Ergebnis über das XBar 4320 weitergeleitet wird.

[0532] Fig. 46 zeigt einen Streaming-Multiprozessor („SM“) 4600 gemäß mindestens einer Ausführungsform. Bei mindestens einer Ausführungsform ist der SM 4600 der SM von **Fig. 44**. Bei mindestens einer Ausführungsform beinhaltet der SM 4600 ohne Einschränkung einen Anweisungs-Cache 4602; eine oder mehrere Planereinheiten 4604; eine Registerdatei 4608; einen oder mehrere Verarbeitungskerne („Kerne“) 4610; eine oder mehrere Spezialfunktionseinheiten („SFUs“) 4612; eine oder mehrere Lade-/Speichereinheiten („LSUs“) 4614; ein Interconnect-Netzwerk 4616; einen gemeinsam genutzten Speicher/Stufe-Eins-Cache („L1“-Cache) 4618; und eine beliebige geeignete Kombination davon. Bei mindestens einer Ausführungsform verteilt eine Arbeitsverteilungseinheit Tasks zur Ausführung auf allgemeinen Verarbeitungsclustern („GPCs“) von Parallelverarbeitungseinheiten („PPUs“), und jede Task wird einem bestimmten Datenverarbeitungscluster („DPC“) innerhalb eines GPCs zugewiesen, und wenn die Task mit einem Shader-Programm verbunden ist, wird die Task einem der SMs 4600 zugewiesen. Bei mindestens einer Ausführungsform empfängt die Scheduler-Einheit 4604 Tasks von der Arbeitsverteilungseinheit und verwaltet die Befehlsplanung für einen oder mehrere Thread-Blöcke, die dem SM 4600 zugewiesen sind. Bei mindestens einer Ausführungsform plant die Scheduler-Einheit 4604 Thread-Blöcke für die Ausführung als Warps von parallelen Threads, wobei jedem Thread-Block mindestens ein Warp zugewiesen wird. Bei mindestens einer Ausführungsform führt jeder Warp Threads aus. Bei mindestens einer Ausführungsform verwaltet die Planer-Einheit 4604 eine Vielzahl unterschiedlicher Thread-Blöcke, indem sie den unterschiedlichen Thread-Blöcken Warps zuordnet und dann während jedes Taktzyklus Anweisungen aus einer Vielzahl unterschiedlicher kooperativer Gruppen an verschiedene Funktionseinheiten (z. B. Verarbeitungskerne 4610, SFUs 4612 und LSUs 4614) verteilt.

[0533] Bei mindestens einer Ausführungsform können sich kooperative Gruppen auf ein Programmiermodell zum Organisieren von Gruppen kommunizierender Threads beziehen, das es Entwicklern ermöglicht, die Granularität auszudrücken, mit der Threads kommunizieren, und um so reichhaltigere, effizientere parallele Dekompositionen zu ermöglichen. Bei mindestens einer Ausführungsform unterstützen kooperative Start-APIs die Synchronisierung zwischen Thread-Blöcken zur Ausführung paralleler Algorithmen. Bei mindestens einer Ausführungsform stellen Anwendungen herkömmlicher Programmiermodelle ein einziges, einfaches Konstrukt für die Synchronisierung kooperierender Threads bereit: eine Barriere über alle Threads eines

Thread-Blocks (z. B. die Funktion syncthreads()). Bei mindestens einer Ausführungsform können Programmierer jedoch Gruppen von Threads mit kleineren als Thread-Block-Granularitäten definieren und innerhalb definierter Gruppen synchronisieren, um größere Leistung, Ausgestaltungsflexibilität und Softwarewiederverwendung in Form von gemeinschaftlichen gruppenweiten Funktionsschnittstellen zu ermöglichen. Bei mindestens einer Ausführungsform ermöglichen es kooperative Gruppen Programmierern, Gruppen von Threads explizit in Teilblock- (d. h. so klein wie ein einzelner Thread) und Multiblock-Granularitäten zu definieren und kollektive Operationen, wie etwa Synchronisation an Threads, in einer kooperativen Gruppe auszuführen, und das Programmiermodell unterstützt eine saubere Komposition über Softwaregrenzen derart hinweg, dass Bibliotheken und Hilfsfunktionen innerhalb ihres lokalen Kontexts sicher synchronisiert werden können, ohne Annahmen über Konvergenz treffen zu müssen. Bei mindestens einer Ausführungsform ermöglichen die Primitives für kooperative Gruppen neue Muster kooperativer Parallelität, die ohne Einschränkung Erzeuger-Verbraucher-Parallelität, opportunistische Parallelität und globale Synchronisierung über ein ganzes Raster von Thread-Blöcken einschließen.

[0534] Bei mindestens einer Ausführungsform ist eine Dispatcher-Einheit 4606 ausgestaltet, um Anweisungen an eine oder mehrere Funktionseinheiten zu übertragen, und die Scheduler-Einheit 4604 weist ohne Einschränkung zwei Dispatcher-Einheiten 4606 auf, die es ermöglichen, dass zwei verschiedene Anweisungen aus demselben Warp während jedes Taktzyklus versandt werden. Bei mindestens einer Ausführungsform beinhaltet jede Planer-Einheit 4604 eine einzelne Dispatcher-Einheit 4606 oder mehrere Dispatcher-Einheiten 460.

[0535] Bei mindestens einer Ausführungsform beinhaltet jeder SM 4600 ohne Einschränkung eine Registerdatei 4608, die einen Satz von Registern für Funktionseinheiten des SM 4600 bereitstellt. Bei mindestens einer Ausführungsform ist die Registerdatei 4608 zwischen den einzelnen Funktionseinheiten aufgeteilt, so dass jeder Funktionseinheit ein eigener Abschnitt der Registerdatei 4608 zugewiesen ist. Bei mindestens einer Ausführungsform ist die Registerdatei 4608 zwischen verschiedenen Warps aufgeteilt, die von dem SM 4600 ausgeführt werden, und die Registerdatei 4608 stellt einen temporären Speicher für Operanden bereit, die mit Datenpfaden von Funktionseinheiten verbunden sind. Bei mindestens einer Ausführungsform umfasst jeder SM 4600 ohne Einschränkung eine Vielzahl von L-Verarbeitungskernen 4610. Bei mindestens einer Ausführungsform weist der SM 4600 ohne Einschränkung eine große Anzahl (z. B. 128 oder mehr) unterschiedlicher Verarbeitungskerne 4610 auf. In mindestens einer Ausführungsform weist jeder Verarbeitungskern 4610 bei mindestens einer Ausführungsform ohne Einschränkung eine Vollpipeline-, Einzelpräzisions-, Doppelpräzisions- und/oder gemischte Präzisionsverarbeitungseinheit auf, die ohne Einschränkung eine arithmetische Gleitkomma-Logikeinheit und eine arithmetische Ganzzahl-Logikeinheit umfasst. Bei mindestens einer Ausführungsform implementieren die arithmetischen Gleitkomma-Logikeinheiten den Standard IEEE 754-2008 für Gleitkomma-Arithmetik. Bei mindestens einer Ausführungsform beinhalten die Verarbeitungskerne 4610 ohne Einschränkung 64 Gleitkomma-Kerne mit einfacher Genauigkeit (32 Bit), 64 Ganzzahl-Kerne, 32 Gleitkomma-Kerne mit doppelter Genauigkeit (64 Bit) und 8 Tensor-Kerne.

[0536] Tensorkerne sind gemäß mindestens einer Ausführungsform für die Durchführung von Matrixoperationen ausgestaltet. Bei mindestens einer Ausführungsform sind ein oder mehrere Tensorkerne in den Verarbeitungskernen 4610 vorhanden. Bei mindestens einer Ausführungsform sind Tensorkerne so ausgestaltet, dass sie Deep-Learning-Matrixarithmetik durchführen, wie z. B. Faltungsoperationen für das Training und Inferencing von neuronalen Netzen. Bei mindestens einer Ausführungsform arbeitet jeder Tensorkern mit einer 4x4-Matrix und führt eine Matrixmultiplikations- und Akkumulationsoperation D = A X B + C durch, wobei A, B, C und D 4x4-Matrizen sind.

[0537] Bei mindestens einer Ausführungsform sind die Matrixmultiplikationseingänge A und B 16-Bit-Gleitkommamatrizen und die Akkumulationsmatrizen C und D sind 16-Bit-Gleitkomma- oder 32-Bit-Gleitkommamatrizen. Bei mindestens einer Ausführungsform arbeiten die Tensorkerne mit 16-Bit-Gleitkomma-Eingangsdaten und 32-Bit-Gleitkomma-Akkumulation. Bei mindestens einer Ausführungsform werden für die 16-Bit-Gleitkommamultiplikation Operationen verwendet 64, die ein Produkt mit voller Genauigkeit ergeben, das dann unter Verwendung einer 32-Bit-Gleitkomma-Addition mit anderen Zwischenprodukten zu einer 4x4x4-Matrixmultiplikation akkumuliert wird. Bei mindestens einer Ausführungsform werden Tensor-Kerne verwendet, um viel größere zweidimensionale oder höherdimensionale Matrixoperationen durchzuführen, die aus diesen kleineren Elementen aufgebaut sind. Bei mindestens einer Ausführungsform stellt eine API wie die CUDA 9 C++ API spezialisierte Operationen zum Laden, Multiplizieren und Akkumulieren von Matrizen sowie zum Speichern von Matrizen bereit, um Tensorkerne von einem CUDA-C++-Programm aus effizient zu nutzen. Bei mindestens einer Ausführungsform auf CUDA-Ebene geht die Schnittstelle auf Warp-Ebene von Matrizen der Größe 16x16 aus, die sich über alle 32 Threads des Warp erstrecken.

[0538] Bei mindestens einer Ausführungsform umfasst jeder SM 4600 ohne Einschränkung M SFUs 4612, die spezielle Funktionen ausführen (z. B. Attributauswertung, reziproke Quadratwurzel und dergleichen). Bei mindestens einer Ausführungsform weisen die SFUs 4612 ohne Einschränkung eine Baum-Traversierungs-Einheit auf, die so ausgestaltet ist, dass sie eine hierarchische Baumdatenstruktur durchläuft. Bei mindestens einer Ausführungsform weisen die SFUs 4612 ohne Einschränkung eine Textureinheit auf, die so konfiguriert ist, dass sie Filteroperationen für die Texturabbildung durchführt. Bei mindestens einer Ausführungsform sind die Textureinheiten so ausgestaltet, dass sie Texturkarten (z. B. eine 2D-Anordnung von Texeln) aus dem Speicher laden und Texturkarten abtasten, um abgetastete Texturwerte zur Verwendung in von dem SM 4600 ausgeführten Shader-Programmen zu erzeugen. Bei mindestens einer Ausführungsform werden die Texturkarten im gemeinsamen Speicher/L1-Cache 4618 gespeichert. Bei mindestens einer Ausführungsform implementieren die Textureinheiten Texturoperationen wie Filteroperationen unter Verwendung von Mip-Maps (z. B. Texturkarten mit unterschiedlichen Detailstufen). Bei mindestens einer Ausführungsform weist jeder SM 4600, ohne Einschränkung, zwei Textureinheiten auf.

[0539] Jeder SM 4600 umfasst, ohne Einschränkung, N LSUs 4614, die bei mindestens einer Ausführungsform Lade- und Speicheroperationen zwischen dem gemeinsamen Speicher/L1-Cache 4618 und der Registerdatei 4608 implementieren. Jeder SM 4600 weist ohne Einschränkung ein Verbindungsnetzwerk 4616 auf, das bei mindestens einer Ausführungsform jede der Funktionseinheiten mit der Registerdatei 4608 und die LSU 4614 mit der Registerdatei 4608 und dem gemeinsamen Speicher/L1-Cache 4618 verbindet. Bei mindestens einer Ausführungsform ist das Verbindungsnetzwerk 4616 ein Koppelfeld, das so ausgestaltet sein kann, dass es jede der Funktionseinheiten mit jedem der Register in der Registerdatei 4608 verbindet und die LSUs 4614 mit der Registerdatei 4608 und den Speicherplätzen im gemeinsamen Speicher/L1-Cache 4618 verbindet.

[0540] In mindestens einer Ausführungsform ist der gemeinsam genutzte Speicher/L1 -Cache 4618 eine Anordnung von On-Chip-Speicher, der bei mindestens einer Ausführungsform die Datenspeicherung und die Kommunikation zwischen dem SM 4600 und der Primitiv-Maschine und zwischen Threads im SM 4600 ermöglicht. Bei mindestens einer Ausführungsform umfasst der gemeinsam genutzte Speicher/L1-Cache 4618 ohne Einschränkung 128 KB Speicherkapazität und liegt auf dem Pfad von dem SM 4600 zu der Partitionseinheit. Bei mindestens einer Ausführungsform wird der gemeinsame Speicher/L1-Cache 4618 zum Zwischenspeichern von Lese- und Schreibvorgängen verwendet. Bei mindestens einer Ausführungsform sind einer oder mehrere von gemeinsamem Speicher/L1-Cache 4618, L2-Cache und Arbeitsspeicher Zusatzspeicher (Backing-Stores).

[0541] Die Kombination von Daten-Cache und gemeinsam genutzter Speicherfunktionalität in einem einzigen Speicherblock bietet bei mindestens einer Ausführungsform eine verbesserte Leistung für beide Arten von Speicherzugriffen. Bei mindestens einer Ausführungsform wird die Kapazität von Programmen, die den gemeinsam genutzten Speicher nicht verwenden, als Cache genutzt oder kann von diesen genutzt werden, z. B. wenn der gemeinsam genutzte Speicher so ausgestaltet ist, dass er die Hälfte der Kapazität nutzt, können Textur- und Lade-/Speicheroperationen die verbleibende Kapazität nutzen. Durch Integration in den gemeinsam genutzten Speicher/L1-Cache 4618 kann der gemeinsam genutzte Speicher/L1-Cache 4618 gemäß mindestens einer Ausführungsform als durchsatzstarke Leitung für Streaming-Daten fungieren und gleichzeitig Zugriff auf häufig wiederverwendete Daten mit hoher Bandbreite und geringer Latenz bereitstellen. Bei mindestens einer Ausführungsform kann, wenn sie für allgemeine parallele Berechnungen ausgestaltet ist, eine einfachere Konfiguration im Vergleich zur Grafikverarbeitung verwendet werden. Bei mindestens einer Ausführungsform werden Grafikverarbeitungseinheiten mit festen Funktionen umgangen, wodurch ein wesentlich einfacheres Programmiermodell entsteht. Bei mindestens einer Ausführungsform weist die Arbeitsverteilungseinheit in der Konfiguration für allgemeine parallele Berechnungen Blöcke von Threads direkt den DPCs zu und verteilt sie. Bei mindestens einer Ausführungsform führen Threads in einem Block dasselbe Programm aus, wobei eine eindeutige Thread-ID in der Berechnung verwendet wird, um sicherzustellen, dass jeder Thread eindeutige Ergebnisse erzeugt, wobei der SM 4600 zur Ausführung des Programms und zur Durchführung von Berechnungen, der gemeinsame Speicher/L1-Cache 4618 zur Kommunikation zwischen Threads und die LSU 4614 zum Lesen und Schreiben des globalen Speichers über den gemeinsamen Speicher/L1-Cache 4618 und die Speicherpartitionseinheit verwendet werden. Bei mindestens einer Ausführungsform schreibt der SM 4600, wenn er für allgemeine parallele Berechnungen ausgestaltet ist, Befehle, die die Scheduler-Einheit 4604 verwenden kann, um neue Arbeiten auf DPCs zu starten.

[0542] Bei mindestens einer Ausführungsform ist die PPU in einem Desktop-Computer, einem Laptop-Computer, einem Tablet-Computer, Servern, Supercomputern, einem Smartphone (z. B. einer drahtlosen Handheld-Einrichtung), einem persönlichen digitalen Assistenten („PDA“), einer Digitalkamera, einem Fahrzeug,

einer am Kopf montierten Anzeige, einer elektronischen in der Hand gehaltenen Einrichtung usw. vorhanden oder damit verbunden. Bei mindestens einer Ausführungsform ist die PPU auf einem einzigen Halbleitersubstrat untergebracht. Bei mindestens einer Ausführungsform ist die PPU in einem System-ona-Chip („SoC“) zusammen mit einer oder mehreren anderen Einrichtungen wie zusätzlichen PPUs, Speicher, einer CPU mit reduziertem Befehlssatz („RISC“), einer Speicherverwaltungseinheit („MMU“), einem Digital-Analog-Wandler („DAC“) und dergleichen vorhanden.

[0543] Bei mindestens einer Ausführungsform kann die PPU auf einer Grafikkarte enthalten sein, die ein oder mehrere Speichergeräte beinhaltet, und die Grafikkarte kann dazu konfiguriert sein, eine Schnittstelle mit einem PCIe-Steckplatz auf einer Hauptplatine eines Desktop-Computers zu bilden. Bei mindestens einer Ausführungsform kann die PPU eine integrierte Grafikverarbeitungseinheit („iGPU“) sein, die im Chipsatz der Hauptplatine vorhanden ist.

[0544] Ableitungs- und/oder Trainingslogik 1815 wird verwendet, um Ableitungs- und/oder Trainingsoperationen, die mit einer oder mehreren Ausführungsformen assoziiert sind, auszuführen. Details bezüglich der Ableitungs- und/oder Trainingslogik 1815 werden hierin in Verbindung mit den **Fig. 18A** und/oder **Fig. 18B** bereitgestellt. Bei mindestens einer Ausführungsform wird der Deep-Learning-Anwendungsprozessor verwendet, um ein Modell maschinellen Lernens, wie etwa ein neuronales Netzwerk, zu trainieren, um der SM 4600 bereitgestellte Informationen vorherzusagen oder abzuleiten. Bei mindestens einer Ausführungsform wird die SM 4600 verwendet, um Informationen basierend auf eines trainierten Modells maschinellen Lernens (z. B. neuronales Netzwerk), das von einem anderen Prozessor oder System oder von der SM 4600 trainiert wurde, abzuleiten oder vorherzusagen. Bei mindestens einer Ausführungsform kann die SM 4600 verwendet werden, um einen oder mehrere in dieser Schrift beschriebene Anwendungsfälle für neuronale Netzwerke auszuführen.

[0545] Bei mindestens einer Ausführungsform wird die SM 4600 eingesetzt, um einen Diskriminator umzusetzen, der dazu trainiert ist, einen Sichtpunkt und einen Satz von Erscheinungsbildattributen aus einem Eingabebild abzuleiten. Bei mindestens einer Ausführungsform wird die SM 4600 eingesetzt, um einen Generator umzusetzen, der dazu trainiert ist, ein Bild basierend auf einem eingegebenen Sichtpunkt und einem eingegebenen Satz von Erscheinungsbildparametern zu erzeugen. Bei mindestens einer Ausführungsform wird die SM 4600 eingesetzt, um ein oder mehrere neuronale Netzwerke umzusetzen, die einen Diskriminator und einen Generator umfassen, und die SM 4600 wird in Verbindung mit einem oder mehreren Prozessen eingesetzt, die ein oder mehrere neuronale Netzwerke trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes selbstüberwacht zu identifizieren, indem mindestens eine oder mehrere Verlustfunktionen, die eine oder mehrere Charakteristiken von Bildern eines Trainingssatzes bewerten, als Teil des Trainings berechnet werden.

[0546] Mindestens eine Ausführungsform der Offenbarung kann angesichts der folgenden Klauseln beschrieben werden:

Klausel 1. Prozessor, umfassend: eine oder mehrere Schaltungen, um dabei zu helfen, ein oder mehrere neuronale Netzwerke zu trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes mindestens zum Teil basierend auf einer oder mehreren anderen Charakteristiken des Objekts als der Ausrichtung des Objekts zu identifizieren.

Klausel 2. Prozessor nach Klausel 1, wobei die eine oder mehreren Schaltungen dabei helfen sollen, das eine oder die mehreren neuronalen Netzwerke auf eine Sammlung von Bildern derselben Kategorie wie das Bild zu trainieren.

Klausel 3. Prozessor nach Klausel 1 oder 2, wobei Ground-Truth-Anmerkungen für mindestens einen Teil der Sammlung von Bildern nicht verfügbar sind.

Klausel 4. Prozessor nach einer der Klauseln 1-3, wobei das eine oder die mehreren Merkmale des Objekts symmetrische Konsistenz zwischen dem Bild des Objekts und einem gespiegelten Bild des Objekts beinhalten.

Klausel 5. Prozessor nach einem der Abschnitte 1-4, wobei eine oder mehrere Schaltungen dabei helfen sollen, das eine oder die mehreren neuronalen Netzwerke zu trainieren, um ein zweites Bild des Objekts mit einer zweiten Ausrichtung zu erzeugen.

Klausel 6. Prozessor nach einem der Abschnitte 1-5, wobei die Ausrichtung des Objekts in einem Satz von Parametern codiert ist, der einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfasst.

Klausel 7. System umfassend:

einen oder mehrere Prozessoren zum Berechnen von Parametern, um dabei zu helfen, ein oder mehrere neuronale Netzwerke zu trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes mindestens zum Teil basierend auf einer oder mehreren anderen Charakteristiken des Objekts als der Ausrichtung des Objekts zu identifizieren; und

einen oder mehrere Speicher zum Speichern der Parameter.

Klausel 8. System nach Klausel 7, wobei der eine oder die mehreren Prozessoren zum Berechnen der Parameter zum Helfen beim Trainieren des einen oder der mehreren neuronalen Netzwerke dabei helfen sollen, das eine oder die mehreren neuronalen Netzwerke mit einer Sammlung von Bildern unterschiedlicher Objekte derselben Kategorie wie das Objekt zu trainieren.

Klausel 9. System nach Klausel 7 oder 8, wobei der eine oder die mehreren Prozessoren das eine oder die mehreren neuronalen Netzwerke trainieren sollen durch mindestens:

Erhalten eines Eingabebildes;

Verwenden eines Diskriminators, um mindestens einen vorhergesagten Sichtpunkt und einen vorhergesagten Satz von Erscheinungsbildparametern zu bestimmen;

Verwenden eines Generators zum Anlegen eines synthetischen Bildes mindestens zum Teil basierend auf dem vorhergesagten Sichtpunkt und dem vorhergesagten Satz von Erscheinungsbildparametern; und

Berechnen eines Sichtpunktkonsistenzverlusts mindestens zum Teil basierend auf dem Eingabebild und dem synthetischen Bild.

Klausel 10. System nach einer der Klauseln 7-9, wobei das Eingabebild ein echtes Bild ist.

Klausel 11. Das System nach einer der Klauseln 7-10, wobei der eine oder die mehreren Prozessoren das eine oder die mehreren neuronalen Netzwerke trainieren sollen durch mindestens:

Erhalten eines ersten Sichtpunkts und eines ersten Satzes von Erscheinungsbildparametern;

Verwenden eines Generators zum Anlegen eines synthetischen Bildes mindestens zum Teil basierend auf dem ersten Sichtpunkt und dem ersten Satz von Erscheinungsbildparametern;

Verwenden eines Diskriminators, um basierend auf dem synthetischen Bild einen zweiten Sichtpunkt und einen zweiten Satz von Erscheinungsbildparametern vorherzusagen;

Berechnen eines Sichtpunktkonsistenzverlusts mindestens zum Teil basierend auf dem ersten Sichtpunkt und dem zweiten Sichtpunkt; und

Berechnen eines Rekonstruktionsverlusts mindestens zum Teil basierend auf dem ersten Bild und dem erzeugten synthetischen Bild.

Klausel 12. System nach einer der Klauseln 7-11, wobei der eine oder die mehreren Prozessoren das eine oder die mehreren neuronalen Netzwerke trainieren sollen durch mindestens:

Verwenden eines Generators zum Anlegen eines ersten synthetischen Bildes mindestens zum Teil basierend auf einem ersten Sichtpunkt und einem Satz von Erscheinungsbildparametern;

Ausführen einer Transformation an dem ersten Sichtpunkt, um einen zweiten Sichtpunkt zu erhalten;

Verwenden des Generators zum Anlegen eines zweiten synthetischen Bildes mindestens zum Teil basierend auf dem zweiten Sichtpunkt und dem Satz von Erscheinungsbildparametern; und

Berechnen eines Symmetrieverlusts mindestens zum Teil basierend auf dem ersten synthetischen Bild und dem zweiten synthetischen Bild.

Klausel 13. System nach einer der Klauseln 7-12, wobei die Transformation den ersten Sichtpunkt horizontal spiegelt, um den zweiten Sichtpunkt zu erhalten.

Klausel 14. Verfahren, das das Trainieren eines oder mehrerer neuronaler Netzwerke umfasst, um eine Ausrichtung eines Objekts innerhalb eines Bildes mindestens zum Teil basierend auf einer oder mehrerer Charakteristiken des Objekts außer der Ausrichtung des Objekts zu identifizieren.

Klausel 15. Verfahren nach Klausel 14, wobei das Trainieren des einen oder der mehreren neuronalen Netzwerke das Trainieren des einen oder der mehreren neuronalen Netzwerke auf selbstüberwachte

Weise an einer Sammlung von Bildern unterschiedlicher Objekte derselben Kategorie wie das Objekt innerhalb des Bildes umfasst.

Klausel 16. Verfahren nach Klause 14 oder 15, wobei das Trainieren des einen oder der mehreren neuronalen Netzwerke auf die selbstüberwachte Weise das Verwenden eines Satzes von Verlustfunktionen umfasst, um die eine oder die mehreren Charakteristiken des Objekts außer der Ausrichtung des Objekts zu bewerten.

Klausel 17. Verfahren nach einer der Klauseln 14-16, wobei das Objekt zu einer ersten Kategorie gehört, und das Verfahren weiter das Trainieren des einen oder der mehreren neuronalen Netzwerke zum Identifizieren einer zweiten Ausrichtung eines zweiten Objekts unter Verwendung einer zweiten Sammlung von Bildern umfasst, wobei:

das zweite Objekt zu einer zweiten Kategorie, die sich von der ersten Kategorie unterscheidet, gehört; und

die zweite Sammlung von Bildern aus Objekten der zweiten Kategorie, die sich von dem zweiten Objekt unterscheidet, besteht.

Klausel 18. Verfahren nach einer der Klauseln 14-17, wobei das Trainieren des einen oder der mehreren neuronalen Netzwerke auf die selbstüberwachte Weise das Trainieren des einen oder der mehreren neuronalen Netzwerke mit mindestens Folgendes umfasst:

Erhalten eines Eingabebildes;

Verwenden eines Diskriminators, um aus dem Eingabebild einen Sichtpunkt und einen Satz von Parametern vorherzusagen;

Verwenden eines Generators, um ein synthetisches Bild mindestens zum Teil basierend auf dem Sichtpunkt und dem Satz von Parametern anzulegen; und

Berechnen eines oder mehrerer Gradienten und Aktualisieren von Parametern des Diskriminators mindestens zum Teil basierend auf dem synthetischen Bild.

Klausel 19. Verfahren nach einer der Klauseln 14-18, wobei der Generator ein tiefes generatives Modell ist.

Klausel 20. Verfahren nach einer der Klauseln 14-19, wobei das tiefe generative Modell ein Renderer, ein Variational Autoencoder oder ein Generative Adversarial Network (GAN) ist.

Klausel 21. Verfahren nach einer der Klauseln 14-20, wobei das Objekt ein Fahrzeug ist.

Klausel 22. Prozessor, umfassend: eine oder mehrere Schaltungen zum Identifizieren einer oder mehrerer Ausrichtungen eines Objekts innerhalb eines Bildes, mindestens zum Teil basierend auf einer oder mehreren anderen Charakteristiken des Objekts als der Ausrichtung des Objekts.

Klausel 23. Prozessor nach Klausel 22, wobei die eine oder mehreren Schaltungen ein oder mehrere neuronale Netzwerke trainieren sollen, um die eine oder mehreren Ausrichtungen des Objekts innerhalb des Bildes zu identifizieren.

Klausel 24. Prozessor nach Klausel 22 oder 23, wobei das eine oder die mehreren neuronalen Netzwerke mit einer Sammlung von Bildern unterschiedlicher Objekte derselben Kategorie wie das Objekt trainiert werden.

Klausel 25. Prozessor einer der Klauseln 22-24, wobei Ground-Truth-Anmerkungen für die Sammlung von Bildern nicht verfügbar sind.

Klausel 26. Prozessor nach einer der Klauseln 22-25, wobei das eine oder die mehreren Merkmale des Objekts symmetrische Konsistenz zwischen dem Bild des Objekts und einem gespiegelten Bild des Objekts beinhalten.

Klausel 27. Prozessor nach einem der Abschnitte 22-26, wobei die Ausrichtung des Objekts in einem Satz von Parametern codiert ist, der einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfasst.

Klausel 28. System umfassend:

einen oder mehrere Speicher; und

einen oder mehrere Prozessoren zum Identifizieren einer oder mehrerer Ausrichtungen eines Objekts innerhalb eines Bildes mindestens zum Teil basierend auf einer oder mehreren anderen Charakteristiken des Objekts als der Ausrichtung des Objekts.

Klausel 29. System nach Klausel 28, wobei der eine oder die mehreren Prozessoren ein oder mehrere neuronale Netzwerke trainieren sollen, um die eine oder mehreren Ausrichtungen des Objekts innerhalb des Bildes mindestens zum Teil basierend auf der einen oder mehreren Charakteristiken des anderen Objekts zu identifizieren als die Ausrichtung des Objekts.

Klausel 30. System nach Klausel 28 oder 29, wobei der eine oder die mehreren Prozessoren zum Trainieren des einen oder der mehreren neuronalen Netzwerke helfen sollen, das eine oder die mehreren neuronalen Netzwerke mit einer Sammlung von Bildern mit unterschiedlichen Objekten zu trainieren, wobei die unterschiedlichen Objekte zu derselben Kategorie wie das Objekt gehören.

Klausel 31. System nach einer der Klauseln 28-30, wobei der eine oder die mehreren Prozessoren zum Trainieren des einen oder der mehreren neuronalen Netzwerke das eine oder die mehreren neuronalen Netzwerke trainieren sollen durch mindestens:

Berechnen eines ersten Satzes von Gradienten zum Aktualisieren eines ersten Satzes von Parametern eines Generators; und

Berechnen eines zweiten Satzes von Gradienten, um einen zweiten Satz von Parametern für einen Diskriminator zu aktualisieren.

Klausel 32. System nach einer der Klauseln 28-31, wobei der eine oder die mehreren Prozessoren zum Trainieren des einen oder der mehreren neuronalen Netzwerke das eine oder die mehreren neuronalen Netzwerke trainieren sollen, indem sie mindestens einen Entflechtungsverlust berechnen durch mindestens:

Verwenden eines ersten Sichtpunkts und eines ersten Satzes von Erscheinungsbildparametern, um ein erstes synthetisches Bild zu erzeugen;

Verwenden des ersten Sichtpunkts und eines zweiten Satzes von Erscheinungsbildparametern, um ein zweites synthetisches Bild zu erzeugen; und

Verwenden eines zweiten Sichtpunkts und des ersten Satzes von Erscheinungsbildparametern, um ein drittes synthetisches Bild zu erzeugen.

Klausel 33. System nach einer der Klauseln 28-32, wobei die eine oder mehreren Ausrichtungen relativ zu einer kanonischen Ausrichtung sind.

Klausel 34. System nach einer der Klauseln 28-33, wobei die eine oder mehreren Ausrichtungen jeweils einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfassen.

Klausel 35. Verfahren, umfassend: Identifizieren einer oder mehrerer Ausrichtungen eines Objekts innerhalb eines Bildes mindestens zum Teil basierend auf einer oder mehreren anderen Charakteristiken des Objekts als der Ausrichtung des Objekts.

Klausel 36. Verfahren nach Klausel 35, wobei ein oder mehrere neuronale Netzwerke dazu trainiert sind, das Identifizieren der einen oder mehreren Ausrichtungen des Objekts innerhalb des Bildes mindestens zum Teil basierend auf der einen oder mehreren anderen Charakteristiken des Objekts als der Ausrichtung des Objekts auszuführen.

Klausel 37. Verfahren nach Klausel 35 oder 36, wobei das eine oder die mehreren neuronalen Netzwerke auf selbstüberwachte Weise mit einer Sammlung von Bildern trainiert werden, dasselbe Label wie das Bild haben, wobei das Label eine andere Charakteristik als die Ausrichtung des Objekts angibt.

Klausel 38. Verfahren nach einer der Klauseln 35-37, wobei das eine oder die mehreren neuronalen Netzwerke auf die selbstüberwachte Weise dazu trainiert werden, Ausrichtungen der Sammlung von Bildern basierend auf anderen Labels als Ausrichtungen der Sammlung von Bildern zu identifizieren.

Klausel 39. Verfahren nach einer der Klauseln 35-38, wobei das eine oder die mehreren neuronalen Netzwerke umfassen:

einen Generator, um synthetische Bilder mindestens zum Teil basierend auf einem spezifizierten Sichtpunkt und einem spezifizierten Satz von Erscheinungsbildparametern anzulegen; und

einen Diskriminator, um aus einem oder mehreren Bildern einen vorhergesagten Sichtpunkt und einen vorhergesagten Satz von Erscheinungsbildparametern zu bestimmen.

Klausel 40. Verfahren nach einer der Klauseln 35-39, wobei der Generator ein tiefes generatives Modell ist.

Klausel 41. Verfahren nach einer der Klauseln 35-40, wobei das Objekt ein Mensch ist.

Klausel 42. Verfahren nach einer der Klauseln 35-41, wobei die Ausrichtung des Objekts in einem Satz von Parametern codiert ist, der einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfasst.

Klausel 43. Auto, umfassend: eine oder mehrere Kameras zum Erfassen von Bildern von einem oder mehreren Objekten und ein oder mehrere neuronale Netzwerke zum Identifizieren einer oder mehrerer Ausrichtungen des einen oder der mehreren Objekte mindestens zum Teil basierend auf einer oder mehreren Charakteristiken des anderen Objekts als die Ausrichtung des Objekts.

Klausel 44. Auto nach Klausel 43, wobei das eine oder die mehreren neuronalen Netzwerke auf selbstüberwachte Weise mit einer Sammlung von Bildern trainiert werden, die dasselbe Label wie das Bild aufweisen, wobei das Label eine andere Charakteristik als die Ausrichtung des Objekts angibt.

Klausel 45. Auto nach Klausel 43 oder 44, wobei die eine oder die mehreren Charakteristiken des Objekts symmetrische Konsistenz zwischen dem Bild des Objekts und einem gespiegelten Bild des Objekts beinhalten.

Klausel 46. Auto nach einer der Klauseln 43-45, wobei ein oder mehrere neuronale Netzwerke dazu trainiert sind, ein zweites Bild mit der Ausrichtung des Objekts zu erzeugen.

Klausel 47. Auto nach einer der Klauseln 43-46, wobei der eine oder die mehreren Prozessoren das eine oder die mehreren neuronalen Netzwerke trainieren sollen durch mindestens:

Erhalten eines Eingabebildes;

Verwenden eines Diskriminators, um mindestens einen vorhergesagten Sichtpunkt und einen vorhergesagten Satz von Erscheinungsbildparametern zu bestimmen;

Verwenden eines Generators zum Anlegen eines synthetischen Bildes mindestens zum Teil basierend auf dem vorhergesagten Sichtpunkt und dem vorhergesagten Satz von Erscheinungsbildparametern; und

Berechnen eines Sichtpunktkonsistenzverlusts mindestens zum Teil basierend auf dem Eingabebild und dem synthetischen Bild.

Klausel 48. Auto nach einer der Klauseln 43-47, wobei die Ausrichtung des Objekts eine dreidimensionale Ausrichtung ist.

Klausel 49. Auto nach einer der Klauseln 43-48, wobei das Objekt ein Mensch ist.

Klausel 50. Auto nach einer der Klauseln 43-49, wobei das Objekt ein anderes Fahrzeug als das Auto ist.

[0547] Bei mindestens einer Ausführungsform kann sich eine einzelne Halbleiterplattform auf eine einzige einheitliche halbleiterbasierte integrierte Schaltung oder einen Chip beziehen. Bei mindestens einer Ausführungsform können Multi-Chip-Module mit erhöhter Konnektivität verwendet werden, die einen On-Chip-Betrieb simulieren und wesentliche Verbesserungen gegenüber der Verwendung einer herkömmlichen Zentraleinheit („CPU“) und einer Bus-Implementierung bieten. Bei mindestens einer Ausführungsform können verschiedene Module auch separat oder in verschiedenen Kombinationen von Halbleiterplattformen je nach Wunsch des Benutzers angeordnet sein.

[0548] Bei mindestens einer Ausführungsform sind Computerprogramme in Form von maschinenlesbarem, ausführbarem Code oder Computersteuerungslogik-Algorithmen im Hauptspeicher 2404 und/oder im Sekundärspeicher gespeichert. Computerprogramme, die, falls sie von einem oder mehreren Prozessoren ausgeführt werden, es dem System 2400 ermöglichen, verschiedene Funktionen gemäß mindestens einer Ausführungsform auszuführen. Bei mindestens einer Ausführungsform sind Speicher 2404, Speicher und/oder jeder andere Speicher mögliche Beispiele für computerlesbare Medien. Bei mindestens einer Ausführungsform kann der Sekundärspeicher eine beliebige geeignete Speichervorrichtung oder -system betreffen, wie eine Festplatte und/oder ein entfernbares Speicherlaufwerk, das ein Diskettenlaufwerk, ein Magnetbandlaufwerk,

ein Compact-Disk-Laufwerk, ein DVD-Laufwerk (Digital Versatile Disk), eine Aufnahmevorrichtung, einen universellen seriellen Bus-Flash-Speicher („USB“) usw. darstellt. Bei mindestens einer Ausführungsform werden die Architektur und/oder die Funktionalität verschiedener vorheriger Figuren in dem Kontext der CPU 2402 umgesetzt; des Parallelverarbeitungssystems 2412; einer integrierte Schaltung mit der Fähigkeit von mindestens einem Teil der Fähigkeiten der CPU 2402; des Parallelverarbeitungssystems 2412; eines Chipsatzes (z. B. eine Gruppe von integrierten Schaltungen, die ausgelegt sind, um als eine Einheit zum Ausführen von diesbezüglichen Funktionen zu arbeiten und gehandelt zu werden usw.); und einer beliebigen geeigneten Kombination von (einer) integrierten Schaltung(en) umgesetzt.

[0549] Bei mindestens einer Ausführungsform ist die Architektur und/oder Funktionalität verschiedener vorhergehender Figuren im Zusammenhang mit einem allgemeinen Computersystem, einem Leiterplattensystem, einem Spielkonsolensystem für Unterhaltungszwecke, einem anwendungsspezifischen System und mehr implementiert. Bei mindestens einer Ausführungsform kann das Computersystem 2400 die Form eines Desktop-Computers, eines Laptops, eines Tablet-Computers, eines Servers, eines Supercomputers, eines Smartphones (z. B einer drahtlosen, in der Hand gehaltenen Einrichtung), eines persönlichen digitalen Assistenten („PDA“), einer Digitalkamera, eines Fahrzeugs, einer auf dem Kopf montierten Anzeige, einer in der Hand gehaltenen elektronischen Einrichtung, einer Mobiltelefoneinrichtung, eines Fernsehers, einer Workstation, von Spielkonsolen, eines eingebetteten Systems und/oder jeder anderen Art von Logik annehmen.

[0550] Bei mindestens einer Ausführungsform weist das Parallelverarbeitungssystem 2412 ohne Einschränkung eine Vielzahl von Parallelverarbeitungseinheiten („PPUs“) 2414 und zugehörige Speicher 2416 auf. Bei mindestens einer Ausführungsform sind die PPUs 2414 mit einem Host-Prozessor oder anderen peripheren Einrichtungen über eine Zwischenverbindung 2418 und einen Switch 2420 oder Multiplexer verbunden. Bei mindestens einer Ausführungsform verteilt das Parallelverarbeitungssystem 2412 Rechenaufgaben auf PPUs 2414, die parallelisierbar sein können - beispielsweise als Teil der Verteilung von Rechenaufgaben auf mehrere Thread-Blöcke der Grafikverarbeitungseinheit („GPU“). Bei mindestens einer Ausführungsform wird der Speicher gemeinsam genutzt und ist für einige oder alle PPUs 2414 zugänglich (z. B. für Lese- und/oder Schreibzugriffe), obwohl ein solcher gemeinsam genutzter Speicher zu Leistungseinbußen im Vergleich zur Nutzung von lokalem Speicher und Registern führen kann, die in einer PPU 2414 resident sind. Bei mindestens einer Ausführungsform wird der Betrieb der PPUs 2414 durch Verwendung eines Befehls wie _syncthreads() synchronisiert, wobei alle Threads in einem Block (z. B. über mehrere PPUs 2414 ausgeführt) einen bestimmten Punkt der Codeausführung erreichen müssen, bevor sie fortfahren.

[0551] Andere Variationen sind im Sinne der vorliegenden Offenbarung. Während die offenbarten Verfahren für verschiedene Modifikationen und alternative Konstruktionen anfällig sind, sind bestimmte dargestellte Ausführungsformen davon in den Zeichnungen gezeigt und oben im Detail beschrieben worden. Es ist jedoch klar, dass es nicht beabsichtigt ist, die Offenbarung auf eine bestimmte Form oder bestimmte Formen zu beschränken, sondern dass im Gegenteil beabsichtigt ist, alle Modifikationen, alternativen Konstruktionen und Äquivalente abzudecken, die in den Geist und den Umfang der Offenbarung fallen, wie es in den beigefügten Ansprüchen definiert ist.

[0552] Die Verwendung der Begriffe „ein“ und „eine“ sowie „der“, „die“ und „das“ und ähnlicher Bezeichnungen im Zusammenhang mit der Beschreibung offengelegter Ausführungsformen (insbesondere im Zusammenhang mit den folgenden Ansprüchen) ist so auszulegen, dass sie sowohl die Einzahl als auch die Mehrzahl umfasst, sofern es hier nicht anders angegeben oder durch den Kontext eindeutig widerlegt ist, und nicht als Definition eines Begriffs. Die Begriffe „umfassend“, „mit“, „einschließlich“, „aufweisend“ und „enthaltend“ sind, sofern es nicht anders angegeben ist, als offene Begriffe zu verstehen (d.h. „einschließlich, aber nicht beschränkt auf”). Der Begriff „verbunden“ ist, wenn er unverändert bleibt und sich auf physikalische Verbindungen bezieht, als teilweise oder ganz in einem Teil enthalten, an ihm angebracht oder mit ihm verbunden zu verstehen, auch wenn etwas dazwischen liegt. Die Aufzählung von Wertebereichen soll lediglich als Kurzform dienen, um sich individuell auf jeden einzelnen Wert zu beziehen, der in den Bereich fällt, sofern hier nichts anderes angegeben ist, und jeder einzelne Wert ist in die Beschreibung aufgenommen, als wäre er hier einzeln aufgeführt. Bei mindestens einer Ausführungsform ist die Verwendung des Begriffs „Menge“ (z. B. „eine Menge von Gegenständen“) oder „Teilmenge“, sofern es nicht anders angegeben oder durch den Kontext widerlegt ist, als eine nicht leere Sammlung zu verstehen, die ein oder mehrere Elemente umfasst. Darüber hinaus bezeichnet der Begriff „Teilmenge“ einer entsprechenden Menge, sofern es nicht anders vermerkt oder durch den Kontext widerlegt ist, nicht notwendigerweise eine echte Teilmenge der entsprechenden Menge, sondern die Teilmenge und die entsprechende Menge können gleich sein.

[0553] Konjunktive Ausdrücke, wie z. B. Sätze der Form „mindestens eines von A, B und C“ oder „mindestens eines von A, B und C“, werden, sofern es nicht ausdrücklich anders angegeben oder durch den Kontext eindeutig widerlegt ist, mit dem Kontext so verstanden, wie sie im Allgemeinen verwendet werden, um darzustellen, dass ein Element, ein Begriff usw, entweder A oder B oder C oder eine beliebige, nicht leere Teilmenge der Menge von A und B und C sein kann. In einem Beispiel für eine Menge mit drei Mitgliedern beziehen sich die konjunktiven Ausdrücke „mindestens eines von A, B, und C“ und „mindestens eines von A, B und C“ auf eine der folgenden Mengen: {A}, {B}, {C}, {A, B}, {A, C}, {B, C}, {A, B, C}. Derartige konjunktive Formulierungen sind also nicht generell so zu verstehen, dass bei bestimmten Ausführungsformen jeweils mindestens eines von A, mindestens eines von B und mindestens eines von C vorhanden sein muss. Sofern es nicht anders vermerkt oder durch den Kontext widerlegt ist, bezeichnet der Begriff „Mehrzahl“ einen Zustand der Pluralität (z. B. „eine Mehrzahl von Elementen“ bezeichnet mehrere Elemente). Bei mindestens einer Ausführungsform beträgt die Anzahl der Gegenstände in einer Mehrzahl mindestens zwei, kann aber auch mehr sein, wenn dies entweder ausdrücklich oder durch den Kontext angegeben ist. Darüber hinaus bedeutet „basierend auf“, sofern es nicht anders angegeben oder anderweitig aus dem Kontext klar ist, „zumindest teilweise basierend auf” und nicht „ausschließlich basierend auf”.

[0554] Die Vorgänge der hier beschriebenen Verfahren können in jeder geeigneten Reihenfolge durchgeführt werden, sofern dies hier nicht anders angegeben ist oder sich aus dem Kontext eindeutig ergibt. Bei mindestens einer Ausführungsform wird ein Verfahren wie die hier beschriebenen Verfahren (oder Variationen und/oder Kombinationen davon) unter der Steuerung eines oder mehrerer Computersysteme durchgeführt, die mit ausführbaren Anweisungen konfiguriert sind und als Code (z. B. ausführbare Anweisungen, ein oder mehrere Computerprogramme oder eine oder mehrere Anwendungen) implementiert sind, die gemeinsam auf einem oder mehreren Prozessoren, durch Hardware oder Kombinationen davon ausgeführt werden. Bei mindestens einer Ausführungsform ist der Code auf einem computerlesbaren Speichermedium gespeichert, z. B. in Form eines Computerprogramms, das eine Vielzahl von Anweisungen umfasst, die von einem oder mehreren Prozessoren ausgeführt werden können. Bei mindestens einer Ausführungsform handelt es sich bei einem computerlesbaren Speichermedium um ein nicht flüchtiges computerlesbares Speichermedium, was transitorische Signale (z. B. eine sich ausbreitende transiente elektrische oder elektromagnetische Übertragung) ausschließt, jedoch nicht flüchtige Datenspeicherschaltungen (z. B. Puffer, Cache und Warteschlangen) innerhalb von Transceivern für transitorische Signale einschließt. Bei mindestens einer Ausführungsform ist ein Code (z. B. ein ausführbarer Code oder Quellcode) auf einem Satz von einem oder mehreren nicht flüchtigen, computerlesbaren Speichermedien gespeichert, auf denen ausführbare Befehle gespeichert sind (oder ein anderer Speicher zum Speichern ausführbarer Befehle), die, wenn sie von einem oder mehreren Prozessoren eines Computersystems ausgeführt werden (d. h. als Ergebnis der Ausführung), das Computersystem veranlassen, hier beschriebene Operationen durchzuführen. Bei mindestens einer Ausführungsform umfasst der Satz von nichtflüchtigen computerlesbaren Speichermedien mehrere nichtflüchtige computerlesbare Speichermedien und einem oder mehreren einzelnen nichtflüchtigen Speichermedien von mehreren nichtflüchtigen computerlesbaren Speichermedien fehlen jeglicher Code, während mehrere nichtflüchtige computerlesbare Speichermedien gemeinsam den gesamten Code speichern. Bei mindestens einer Ausführungsform werden ausführbare Befehle so ausgeführt, dass verschiedene Befehle von verschiedenen Prozessoren ausgeführt werden - zum Beispiel speichert ein nicht flüchtiges computerlesbares Speichermedium Befehle und eine Hauptzentraleinheit („CPU“) führt einige der Befehle aus, während eine Grafikverarbeitungseinheit („GPU“) andere Befehle ausführt. Bei mindestens einer Ausführungsform haben verschiedene Komponenten eines Computersystems getrennte Prozessoren, und verschiedene Prozessoren führen verschiedene Teilmengen von Befehlen aus.

[0555] Dementsprechend sind bei mindestens einer Ausführungsform Computersysteme ausgestaltet, um einen oder mehrere Dienste zu implementieren, die einzeln oder gemeinsam Operationen der hier beschriebenen Verfahren durchführen, und solche Computersysteme sind mit anwendbarer Hardware und/oder Software ausgestaltet, die die Durchführung der Operationen ermöglichen. Darüber hinaus ist ein Computersystem, das mindestens eine Ausführungsform der vorliegenden Offenbarung implementiert, eine einzelne Einrichtung und in einer anderen Ausführungsform ein verteiltes Computersystem, das mehrere Einrichtungen umfasst, die unterschiedlich arbeiten, so dass das verteilte Computersystem die hier beschriebenen Operationen durchführt und dass eine einzelne Einrichtung nicht alle Operationen durchführt.

[0556] Die Verwendung von Beispielen oder beispielhaften Formulierungen (z. B „wie z. B“) dient lediglich der besseren Veranschaulichung von Ausführungsformen der Offenbarung und stellt keine Einschränkung des Umfangs der Offenbarung dar, sofern nicht etwas anderes beansprucht wird. Keine Formulierung in der Beschreibung sollte so ausgelegt werden, dass ein nicht beanspruchtes Element als wesentlich für die Durchführung der Offenbarung angesehen wird.

[0557] Alle Referenzen, einschließlich Veröffentlichungen, Patentanmeldungen und Patente, die hierin zitiert werden, werden hiermit durch Bezugnahme in demselben Ausmaß einbezogen, als ob jede Referenz einzeln und ausdrücklich als durch Bezugnahme einbezogen angegeben wäre und hier in ihrer Gesamtheit wiedergegeben würde.

[0558] In der Beschreibung und den Ansprüchen können die Begriffe „gekoppelt“ und „verbunden“ sowie deren Derivate verwendet werden. Es sollte verstanden werden, dass diese Begriffe nicht als Synonyme füreinander zu verstehen sind. Vielmehr kann in bestimmten Beispielen „verbunden“ oder „gekoppelt“ verwendet werden, um anzuzeigen, dass zwei oder mehr Elemente in direktem oder indirektem physischem oder elektrischem Kontakt zueinander stehen. „Gekoppelt“ kann auch bedeuten, dass zwei oder mehr Elemente nicht in direktem Kontakt zueinander stehen, aber dennoch miteinander kooperieren oder interagieren.

[0559] Sofern nicht ausdrücklich etwas anderes angegeben ist, beziehen sich Begriffe wie „Verarbeiten“, „Rechnen“, „Berechnen“, „Bestimmen“ oder dergleichen in der gesamten Beschreibung auf Aktionen und/oder Verfahren bzw. Prozesse eines Computers oder eines Rechensystems oder einer ähnlichen elektronischen Recheneinrichtung, die Daten, die als physikalische, z. B. elektronische, Größen in den Registern und/oder Speichern des Rechensystems dargestellt sind, manipulieren und/oder in andere Daten umwandeln, die in ähnlicher Weise als physikalische Größen in den Speichern, Registern oder anderen derartigen Einrichtungen zur Speicherung, Übertragung oder Anzeige von Informationen des Rechensystems dargestellt werden.

[0560] In ähnlicher Weise kann sich der Begriff „Prozessor“ auf eine Einrichtung oder einen Abschnitt einer Einrichtung beziehen, die elektronische Daten aus Registern und/oder Speichern verarbeitet und diese elektronischen Daten in andere elektronische Daten umwandelt, die in Registern und/oder Speichern gespeichert werden können. Als nicht einschränkende Beispiele kann der „Prozessor“ eine CPU oder eine GPU sein. Eine „Datenverarbeitungsplattform“ kann einen oder mehrere Prozessoren umfassen. Der hier verwendete Begriff „Software“-Prozesse kann z. B. Software- und/oder Hardware-Einheiten aufweisen, die im Laufe der Zeit Arbeit verrichten, wie z. B. Tasks, Threads und intelligente Agenten. Jeder Prozess bzw. jedes Verfahren kann sich auch auf mehrere Prozesse bzw. Verfahren beziehen, um Anweisungen nacheinander oder parallel, kontinuierlich oder intermittierend auszuführen. Die Begriffe „System“ und „Verfahren“ werden in dieser Schrift insoweit austauschbar verwendet, als das System ein oder mehrere Verfahren verkörpern kann und Verfahren als ein System betrachtet werden können.

[0561] Im vorliegenden Dokument kann auf das Gewinnen, Erfassen, Empfangen oder Eingeben analoger oder digitaler Daten in ein Teilsystem, ein Computersystem oder eine computerimplementierte Maschine Bezug genommen werden. Bei mindestens einer Ausführungsform kann ein Prozess des Erhaltens, Ermittelns, Empfangens oder Eingebens analoger und digitaler Daten auf Vielfalt von Weisen ausgeführt werden, wie durch Empfangen von Daten als ein Parameter eines Funktionsaufrufs oder eines Aufrufs an eine Anwendungsprogrammierschnittstelle. In einigen Implementierungen kann das Verfahren des Erhaltens, Erfassens, Empfangens oder Eingebens von analogen oder digitalen Daten durch die Übertragung von Daten über eine serielle oder parallele Schnittstelle durchgeführt werden. In einer anderen Implementierung kann das Verfahren des Erhaltens, Erfassens, Empfangens oder Eingebens analoger oder digitaler Daten durch die Übertragung von Daten über ein Computernetzwerk von der bereitstellenden Einheit zur erfassenden Einheit durchgeführt werden. Es kann auch auf ein Bereitstellen, Ausgeben, Übertragen, Senden oder Präsentieren analoger oder digitaler Daten Bezug genommen werden. In verschiedenen Beispielen kann das Verfahren eines Bereitstellens, Ausgebens, Übertragens, Sendens oder Darstellens analoger oder digitaler Daten durch die Übertragung von Daten als Eingabe- oder Ausgabeparameter eines Funktionsaufrufs, eines Parameters einer Anwendungsprogrammierschnittstelle oder eines Interprozess-Kommunikationsmechanismus durchgeführt werden.

[0562] Obwohl die obige Diskussion beispielhafte Implementierungen der beschriebenen Verfahren bzw. Techniken darlegt, können andere Architekturen verwendet werden, um die beschriebene Funktionalität zu implementieren, und diese sollen in den Anwendungsbereich dieser Offenbarung fallen. Darüber hinaus, obwohl spezifische Verteilungen von Verantwortlichkeiten oben zu Diskussionszwecken definiert sind, können verschiedene Funktionen und Verantwortlichkeiten auf unterschiedliche Weise verteilt und aufgeteilt werden, abhängig von den Umständen.

[0563] Obwohl der Gegenstand in einer Sprache beschrieben wurde, die sich auf strukturelle Merkmale und/oder methodische Handlungen bezieht, ist zu verstehen, dass der in den beigefügten Ansprüchen beanspruchte Gegenstand nicht notwendigerweise auf die beschriebenen spezifischen Merkmale oder Handlun-

gen beschränkt ist. Vielmehr sind bestimmte Merkmale und Aktionen als beispielhafte Formen der Umsetzung der Ansprüche offengelegt.

ZITATE ENTHALTEN IN DER BESCHREIBUNG

Zitierte Patentliteratur

- US 16690015 B [0001]
- JP 3016201806 [0170]
- JP 3016201609 [0170]

Patentansprüche

1. Prozessor, umfassend: eine oder mehrere Schaltungen, um dabei zu helfen, ein oder mehrere neuronale Netzwerke zu trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes mindestens zum Teil basierend auf einer oder mehreren anderen Charakteristiken des Objekts als der Ausrichtung des Objekts zu identifizieren.

2. Prozessor nach Anspruch 1, wobei die eine oder mehreren Schaltungen dabei helfen sollen, das eine oder die mehreren neuronalen Netzwerke mit einer Sammlung von Bildern derselben Kategorie wie das Bild zu trainieren.

3. Prozessor nach Anspruch 2, wobei Ground-Truth-Anmerkungen für mindestens einen Teil der Sammlung von Bildern nicht verfügbar sind.

4. Prozessor nach Anspruch 1, wobei die eine oder mehreren Charakteristiken des Objekts symmetrische Konsistenz zwischen dem Bild des Objekts und einem gespielten Bild des Objekts beinhalten.

5. Prozessor nach Anspruch 1, wobei eine oder mehrere Schaltungen helfen sollen, das eine oder die mehreren neuronalen Netzwerke zu trainieren, um ein zweites Bild des Objekts mit einer zweiten Ausrichtung zu erzeugen.

6. Prozessor nach Anspruch 1, wobei die Ausrichtung des Objekts in einem Satz von Parametern codiert ist, der einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfasst.

7. System umfassend:
einen oder mehrere Prozessoren zum Berechnen von Parametern, um dabei zu helfen, ein oder mehrere neuronale Netzwerke zu trainieren, um eine Ausrichtung eines Objekts innerhalb eines Bildes mindestens zum Teil basierend auf einer oder mehreren anderen Charakteristiken des Objekts als der Ausrichtung des Objekts zu identifizieren, und
einen oder mehrere Speicher zum Speichern der Parameter.

8. System nach Anspruch 7, wobei der eine oder die mehreren Prozessoren zum Berechnen der Parameter zum Unterstützen des Trainierens des einen oder der mehreren neuronalen Netzwerke dazu dienen sollen, das eine oder die mehreren neuronalen Netzwerke mit einer Sammlung von Bildern unterschiedlicher Objekte derselben Kategorie wie das Objekt zu trainieren.

9. System nach Anspruch 8, wobei der eine oder die mehreren Prozessoren das eine oder die mehreren neuronalen Netzwerke trainieren sollen durch mindestens:
Erhalten eines Eingabebildes;
Verwenden eines Diskriminators, um mindestens einen vorhergesagten Sichtpunkt und einen vorhergesagten Satz von Erscheinungsbildparametern zu bestimmen;
Verwenden eines Generators zum Anlegen eines synthetischen Bildes mindestens zum Teil basierend auf dem vorhergesagten Sichtpunkt und dem vorhergesagten Satz von Erscheinungsbildparametern; und
Berechnen eines Sichtpunktkonsistenzverlusts mindestens zum Teil basierend auf dem Eingabebild und dem synthetischen Bild.

10. System nach Anspruch 9, wobei das Eingabebild ein echtes Bild ist.

11. System nach Anspruch 8, wobei der eine oder die mehreren Prozessoren das eine oder die mehreren neuronalen Netzwerke trainieren sollen durch mindestens:
Erhalten eines ersten Sichtpunkts und eines ersten Satzes von Erscheinungsbildparametern;
Verwenden eines Generators zum Anlegen eines synthetischen Bildes mindestens zum Teil basierend auf dem ersten Sichtpunkt und dem ersten Satz von Erscheinungsbildparametern;
Verwenden eines Diskriminators, um basierend auf dem synthetischen Bild einen zweiten Sichtpunkt und einen zweiten Satz von Erscheinungsbildparametern vorherzusagen;
Berechnen eines Sichtpunktkonsistenzverlusts mindestens zum Teil basierend auf dem ersten Sichtpunkt und dem zweiten Sichtpunkt; und
Berechnen eines Rekonstruktionsverlusts mindestens zum Teil basierend auf dem ersten Bild und dem erzeugten synthetischen Bild.

12. System nach Anspruch 8, wobei der eine oder die mehreren Prozessoren das eine oder die mehreren neuronalen Netzwerke trainieren sollen durch mindestens:
Verwenden eines Generators zum Anlegen eines ersten synthetischen Bildes mindestens zum Teil basierend auf einem ersten Sichtpunkt und einem Satz von Erscheinungsbildparametern;
Ausführen einer Transformation an dem ersten Sichtpunkt, um einen zweiten Sichtpunkt zu erhalten;
Verwenden des Generators zum Anlegen eines zweiten synthetischen Bildes mindestens zum Teil basierend auf dem zweiten Sichtpunkt und dem Satz von Erscheinungsbildparametern; und
Berechnen eines Symmetrieverlusts mindestens zum Teil basierend auf dem ersten synthetischen Bild und dem zweiten synthetischen Bild.

13. System nach Anspruch 12, wobei die Transformation den ersten Sichtpunkt horizontal spiegelt, um den zweiten Sichtpunkt zu erhalten.

14. Verfahren, das das Trainieren eines oder mehrerer neuronaler Netzwerke umfasst, um eine Ausrichtung eines Objekts innerhalb eines Bildes mindestens zum Teil basierend auf einer oder mehrerer Charakteristiken des Objekts außer der Ausrichtung des Objekts zu identifizieren.

15. Verfahren nach Anspruch 14, wobei das Trainieren des einen oder der mehreren neuronalen Netzwerke das Trainieren des einen oder der mehreren neuronalen Netzwerke auf selbstüberwachte Weise an einer Sammlung von Bildern unterschiedlicher Objekte derselben Kategorie wie das Objekt innerhalb des Bildes umfasst.

16. Verfahren nach Anspruch 15, wobei das Trainieren des einen oder der mehreren neuronalen Netzwerke auf die selbstüberwachte Weise das Verwenden eines Satzes von Verlustfunktionen umfasst, um die eine oder mehreren Charakteristiken des Objekts außer der Ausrichtung des Objekts zu bewerten.

17. Verfahren nach Anspruch 15, wobei das Objekt zu einer ersten Kategorie gehört, und das Verfahren weiter das Trainieren des einen oder der mehreren neuronalen Netzwerke zum Identifizieren einer zweiten Ausrichtung eines zweiten Objekts unter Verwendung einer zweiten Sammlung von Bildern umfasst, wobei:
das zweite Objekt zu einer zweiten Kategorie, die sich von der ersten Kategorie unterscheidet, gehört; und
die zweite Sammlung von Bildern aus Objekten der zweiten Kategorie, die sich von dem zweiten Objekt unterscheidet, besteht.

18. Verfahren nach Anspruch 15, wobei das Trainieren des einen oder der mehreren neuronalen Netzwerke auf die selbstüberwachte Weise das Trainieren des einen oder der mehreren neuronalen Netzwerke auf mindestens Folgendes umfasst:
Erhalten eines Eingabebildes;
Verwenden eines Diskriminators, um aus dem Eingabebild einen Sichtpunkt und einen Satz von Parametern vorherzusagen;
Verwenden eines Generators, um ein synthetisches Bild mindestens zum Teil basierend auf dem Sichtpunkt und dem Satz von Parametern anzulegen; und
Berechnen eines oder mehrerer Gradienten und Aktualisieren von Parametern des Diskriminators mindestens zum Teil basierend auf dem synthetischen Bild.

19. Verfahren nach Anspruch 18, wobei der Generator ein tiefes generatives Modell ist.

20. Verfahren nach Anspruch 19, wobei das tiefe generative Modell ein Renderer, ein Variational Autoencoder oder ein Generative Adversarial Network (GAN) ist.

21. Verfahren nach Anspruch 14, wobei das Objekt ein Fahrzeug ist.

22. Prozessor, umfassend: eine oder mehrere Schaltungen zum Identifizieren einer oder mehrerer Ausrichtungen eines Objekts innerhalb eines Bildes mindestens zum Teil basierend auf einer oder mehreren anderen Charakteristiken des Objekts als der Ausrichtung des Objekts.

23. Prozessor nach Anspruch 22, wobei die eine oder mehreren Schaltungen ein oder mehrere neuronale Netzwerke trainieren sollen, um die eine oder mehreren Ausrichtungen des Objekts innerhalb des Bildes zu identifizieren.

24. Prozessor nach Anspruch 23, wobei das eine oder die mehreren neuronalen Netzwerke mit einer Sammlung von Bildern verschiedener Objekte derselben Kategorie wie das Objekt trainiert werden.

25. Prozessor nach Anspruch 23, wobei Ground-Truth-Anmerkungen für die Sammlung von Bildern nicht verfügbar sind.

26. Prozessor nach Anspruch 22, wobei die eine oder mehreren Charakteristiken des Objekts symmetrische Konsistenz zwischen dem Bild des Objekts und einem gespielten Bild des Objekts beinhalten.

27. Prozessor nach Anspruch 22, wobei die Ausrichtung des Objekts in einem Satz von Parametern codiert ist, der einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfasst.

28. System umfassend:
einen oder mehrere Speicher; und
einen oder mehrere Prozessoren zum Identifizieren einer oder mehrerer Ausrichtungen eines Objekts innerhalb eines Bildes mindestens zum Teil basierend auf einer oder mehreren anderen Charakteristiken des Objekts als der Ausrichtung des Objekts.

29. System nach Anspruch 28, wobei der eine oder die mehreren Prozessoren ein oder mehrere neuronale Netzwerke trainieren sollen, um die eine oder mehreren Ausrichtungen des Objekts innerhalb des Bildes mindestens zum Teil basierend auf der einen oder mehreren Charakteristiken des anderen Objekts als die Ausrichtung des Objekts zu identifizieren.

30. System nach Anspruch 29, wobei der eine oder die mehreren Prozessoren zum Trainieren des einen oder der mehreren neuronalen Netzwerke dabei helfen sollen, das eine oder die mehreren neuronalen Netzwerke auf eine Sammlung von Bildern mit unterschiedlichen Objekten zu trainieren, wobei die unterschiedlichen Objekte zu derselben Kategorie wie das Objekt gehören.

31. System nach Anspruch 29, wobei der eine oder die mehreren Prozessoren zum Trainieren des einen oder der mehreren neuronalen Netzwerke das eine oder die mehreren neuronalen Netzwerke trainieren sollen durch mindestens:
Berechnen eines ersten Satzes von Gradienten zum Aktualisieren eines ersten Satzes von Parametern eines Generators; und
Berechnen eines zweiten Satzes von Gradienten, um einen zweiten Satz von Parametern für einen Diskriminator zu aktualisieren.

32. System nach Anspruch 28, wobei der eine oder die mehreren Prozessoren zum Trainieren des einen oder der mehreren neuronalen Netzwerke das eine oder die mehreren neuronalen Netzwerke trainieren sollen, indem sie mindestens einen Entflechtungsverlust berechnen durch mindestens:
Verwenden eines ersten Sichtpunkts und eines ersten Satzes von Erscheinungsbildparametern, um ein erstes synthetisches Bild zu erzeugen;
Verwenden des ersten Sichtpunkts und eines zweiten Satzes von Erscheinungsbildparametern, um ein zweites synthetisches Bild zu erzeugen; und
Verwenden eines zweiten Sichtpunkts und des ersten Satzes von Erscheinungsbildparametern, um ein drittes synthetisches Bild zu erzeugen.

33. System nach Anspruch 28, wobei die eine oder mehreren Ausrichtungen relativ zu einer kanonischen Ausrichtung sind.

34. System nach Anspruch 28, wobei die eine oder mehreren Ausrichtungen jeweils einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfassen.

35. Verfahren, umfassend: Identifizieren einer oder mehrerer Ausrichtungen eines Objekts innerhalb eines Bildes mindestens zum Teil basierend auf einer oder mehreren anderen Charakteristiken des Objekts als der Ausrichtung des Objekts.

36. Verfahren nach Anspruch 35, wobei ein oder mehrere neuronale Netzwerke dazu trainiert sind, das Identifizieren der einen oder mehreren Ausrichtungen des Objekts innerhalb des Bildes mindestens zum Teil basierend auf der einen oder mehreren anderen Charakteristiken des Objekts als der Ausrichtung des Objekts auszuführen.

37. Verfahren nach Anspruch 36, wobei das eine oder die mehreren neuronalen Netzwerke auf selbstüberwachte Weise mit einer Sammlung von Bildern trainiert werden, die dasselbe Label wie das Bild aufweisen, wobei das Label eine andere Charakteristik als die Ausrichtung des Objekts angibt.

38. Verfahren nach Anspruch 37, wobei das eine oder die mehreren neuronalen Netzwerke auf die selbstüberwachte Weise trainiert werden, um Ausrichtungen der Sammlung von Bildern basierend auf anderen Labels als Ausrichtungen der Sammlung von Bildern zu identifizieren.

39. Verfahren nach Anspruch 37, wobei das eine oder die mehreren neuronalen Netzwerke umfassen:
einen Generator, um synthetische Bilder mindestens zum Teil basierend auf einem spezifizierten Sichtpunkt und einem spezifizierten Satz von Erscheinungsbildparametern anzulegen; und
einen Diskriminator, um aus einem oder mehreren Bildern einen vorhergesagten Sichtpunkt und einen vorhergesagten Satz von Erscheinungsbildparametern zu bestimmen.

40. Verfahren nach Anspruch 39, wobei der Generator ein tiefes generatives Modell ist.

41. Verfahren nach Anspruch 37, wobei das Objekt ein Mensch ist.

42. Verfahren nach Anspruch 37, wobei die Ausrichtung des Objekts auf einem Satz von Parametern codiert ist, der einen Azimutparameter, einen Elevationsparameter und einen Kippparameter umfasst.

43. Auto, umfassend: eine oder mehrere Kameras zum Erfassen von Bildern von einem oder mehreren Objekten und ein oder mehrere neuronale Netzwerke zum Identifizieren einer oder mehrerer Ausrichtungen des einen oder der mehreren Objekte mindestens zum Teil basierend auf einer oder mehrerer anderen Charakteristiken des Objekts als die Ausrichtung des Objekts.

44. Auto nach Anspruch 43, wobei das eine oder die mehreren neuronalen Netzwerke auf selbstüberwachte Weise mit einer Sammlung von Bildern trainiert werden, die dasselbe Label wie das Bild nutzen, wobei das Label eine andere Charakteristik als die Ausrichtung des Objekts angibt.

45. Auto nach Anspruch 43, wobei die eine oder mehreren Charakteristiken des Objekts symmetrische Konsistenz zwischen dem Bild des Objekts und einem gespiegelten Bild des Objekts beinhalten.

46. Auto nach Anspruch 43, wobei ein oder mehrere neuronale Netzwerke dazu trainiert sind, ein zweites Bild mit der Ausrichtung des Objekts zu erzeugen.

47. Auto nach Anspruch 43, wobei der eine oder die mehreren Prozessoren das eine oder die mehreren neuronalen Netzwerke trainieren sollen durch mindestens:
Erhalten eines Eingabebildes;
Verwenden eines Diskriminators, um mindestens einen vorhergesagten Sichtpunkt und einen vorhergesagten Satz von Erscheinungsbildparametern zu bestimmen;
Verwenden eines Generators zum Anlegen eines synthetischen Bildes mindestens zum Teil basierend auf dem vorhergesagten Sichtpunkt und dem vorhergesagten Satz von Erscheinungsbildparametern; und
Berechnen eines Sichtpunktkonsistenzverlusts mindestens zum Teil basierend auf dem Eingabebild und dem synthetischen Bild.

48. Auto nach Anspruch 43, wobei die Ausrichtung des Objekts eine dreidimensionale Ausrichtung ist.

49. Auto nach Anspruch 43, wobei das Objekt ein Mensch ist.

50. Auto nach Anspruch 43, wobei das Objekt ein anderes Fahrzeug als das Auto ist.

Es folgen 62 Seiten Zeichnungen

Anhängende Zeichnungen

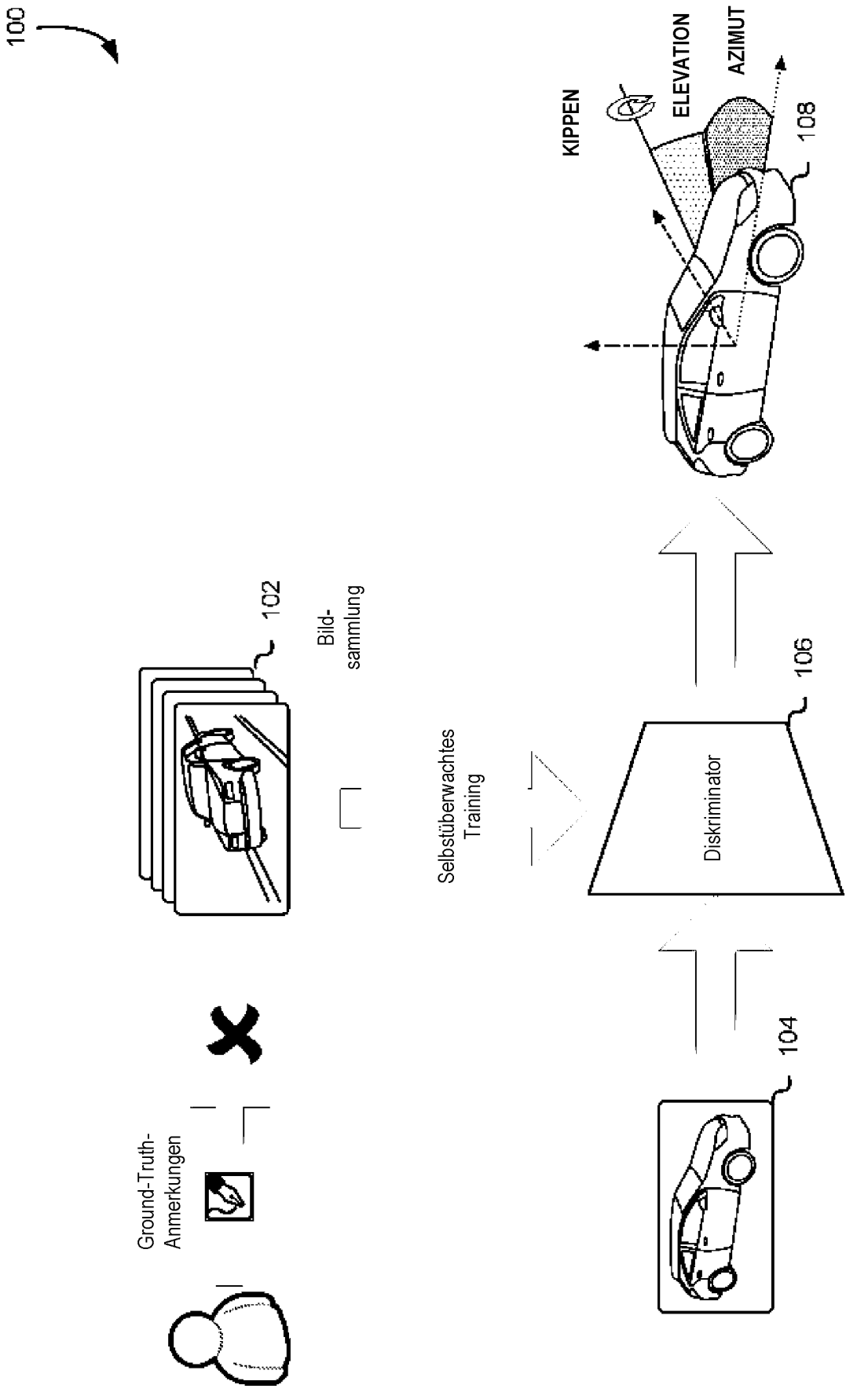


FIG. 1

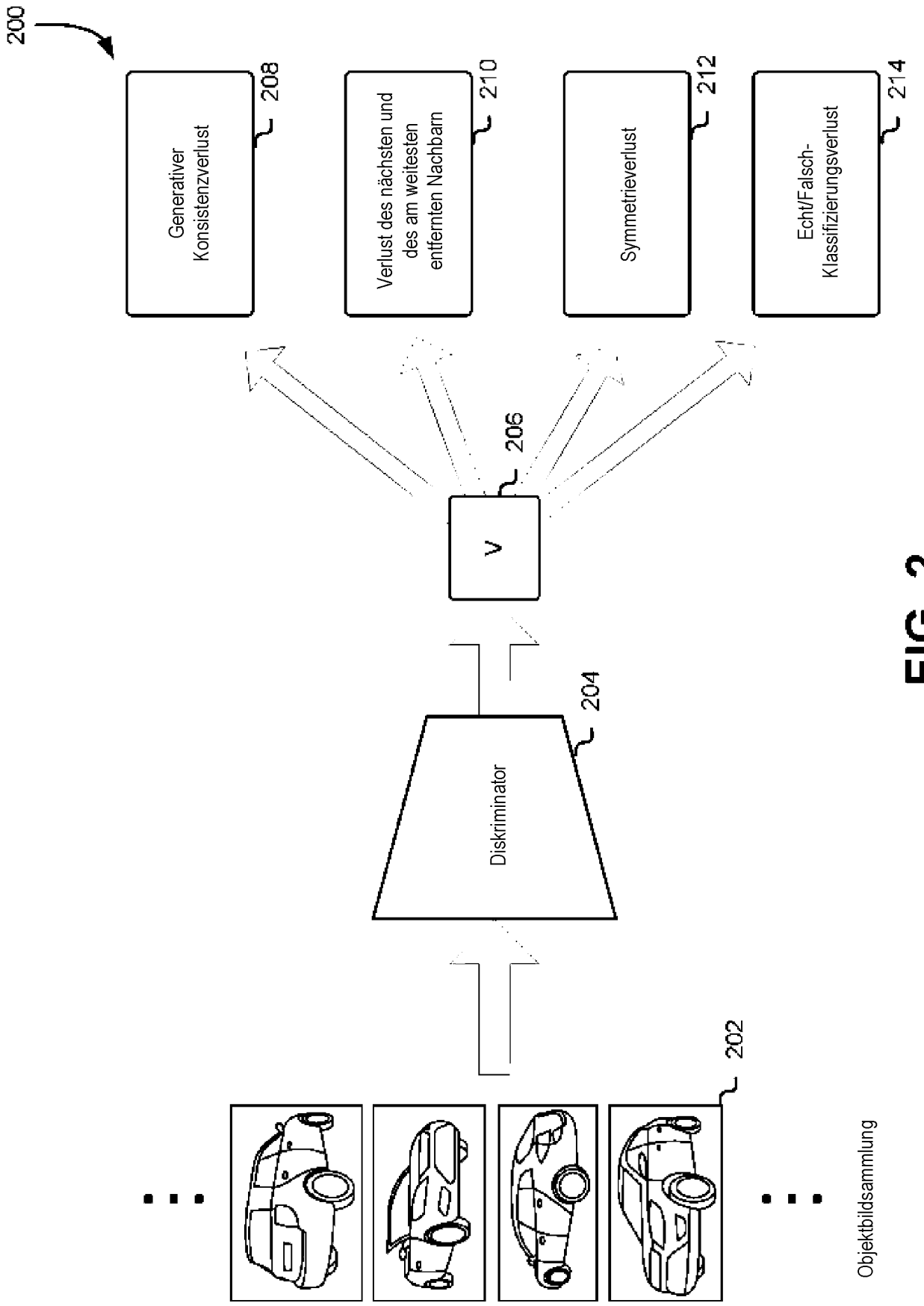

200
Generativer Konsistenzverlust
208
Verlust des nächsten und des am weitesten entfernten Nachbarn
210
Symmetrieverlust
212
Echt/Falsch-Klassifizierungsverlust
214
206
V
204
Diskriminator
202
Objektbildsammlung

FIG. 2

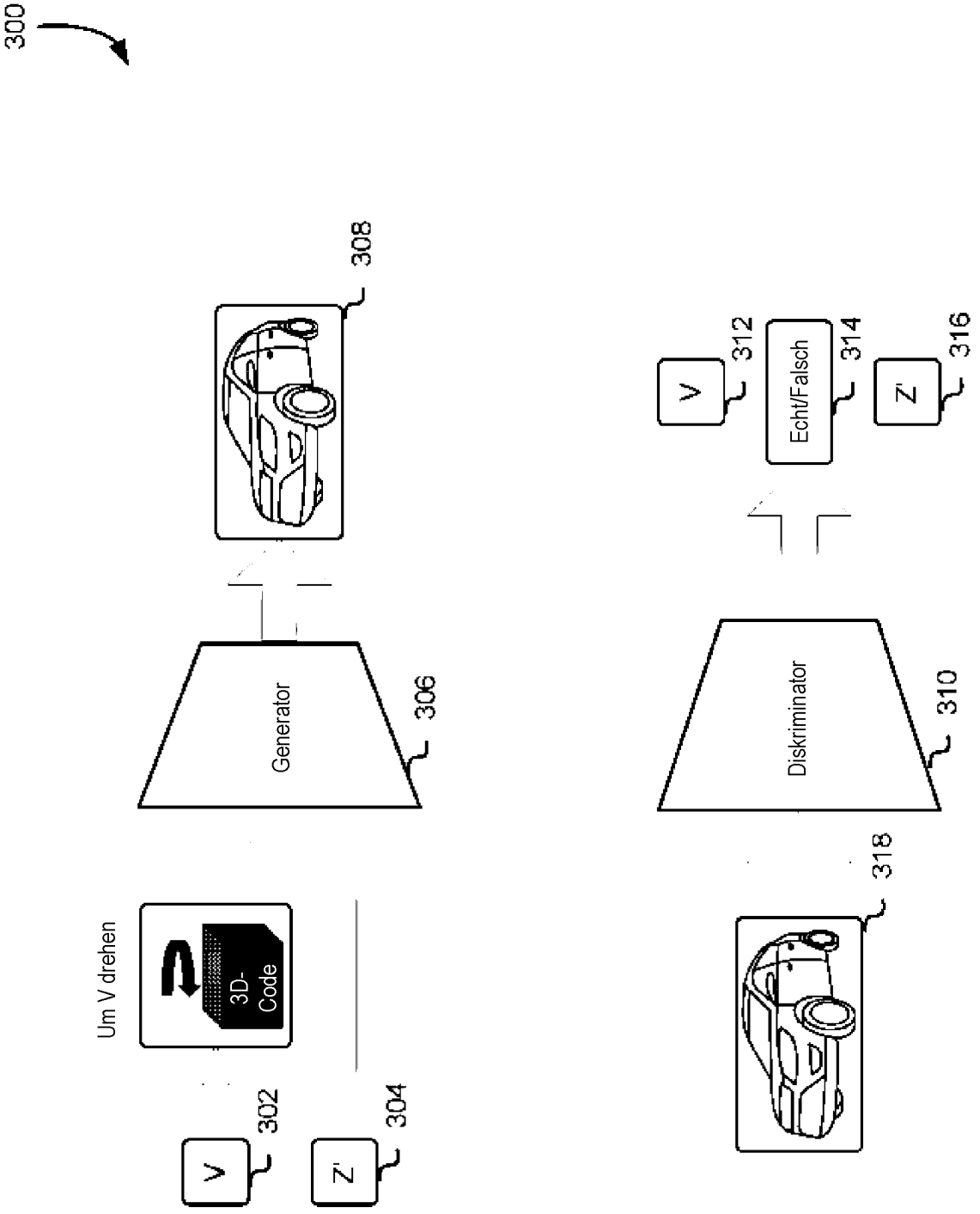

300
308
Generator
306
Um V drehen
3D-Code
V
302
Z'
304
V
312
Echt/Falsch
314
Z'
316
Diskriminator
310
318

FIG. 3

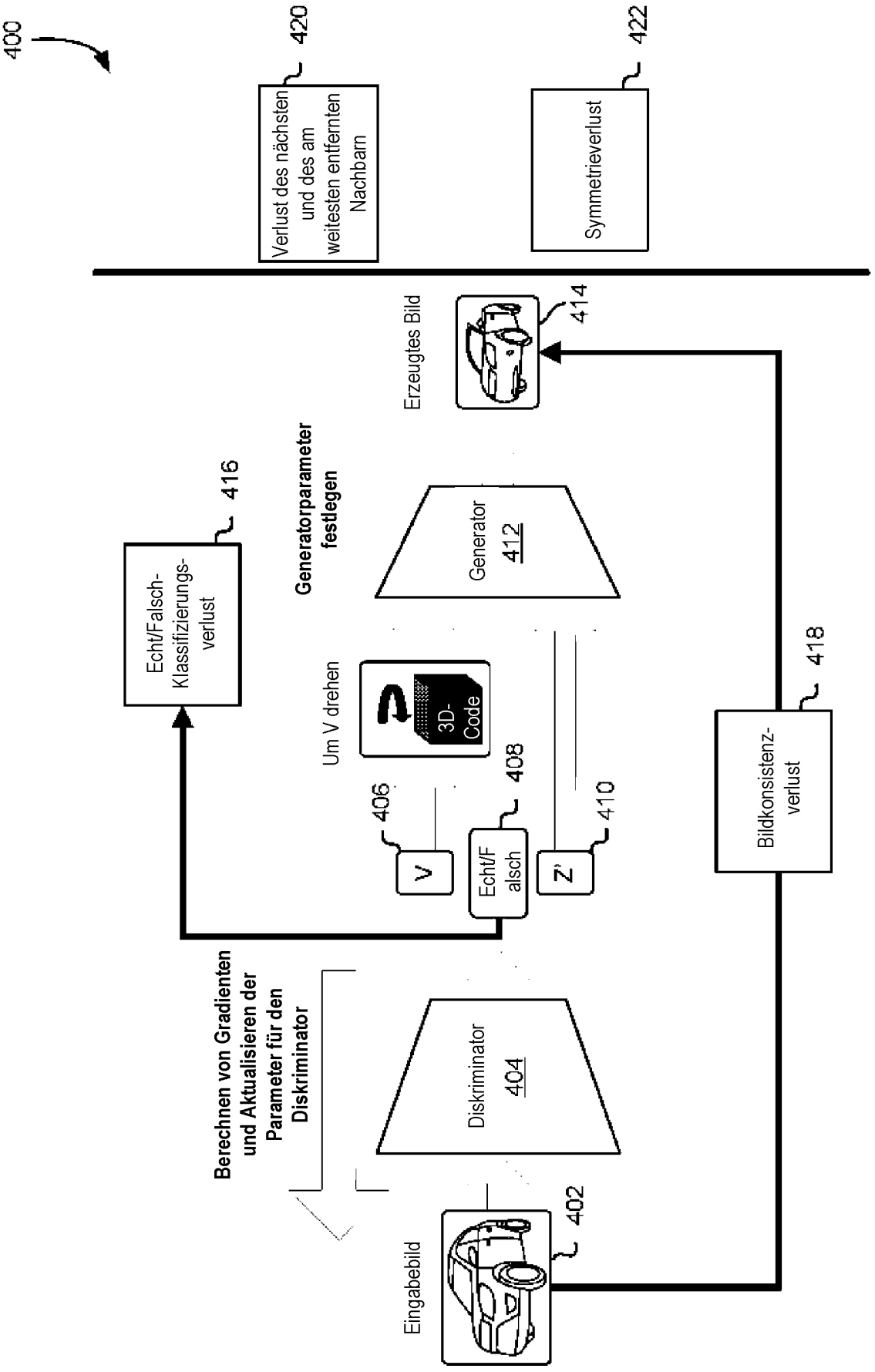

400
420
Verlust des nächsten und des am weitesten entfernten Nachbarn
422
Symmetrieverlust
Erzeugtes Bild
414
Generatorparameter festlegen
Generator 412
416
Echt/Falsch-Klassifizierungs-verlust
Um V drehen
3D-Code
406
V
408
Echt/Falsch
410
Z'
418
Bildkonsistenz-verlust
Berechnen von Gradienten und Aktualisieren der Parameter für den Diskriminator
Diskriminator 404
Eingabebild
402

FIG. 4

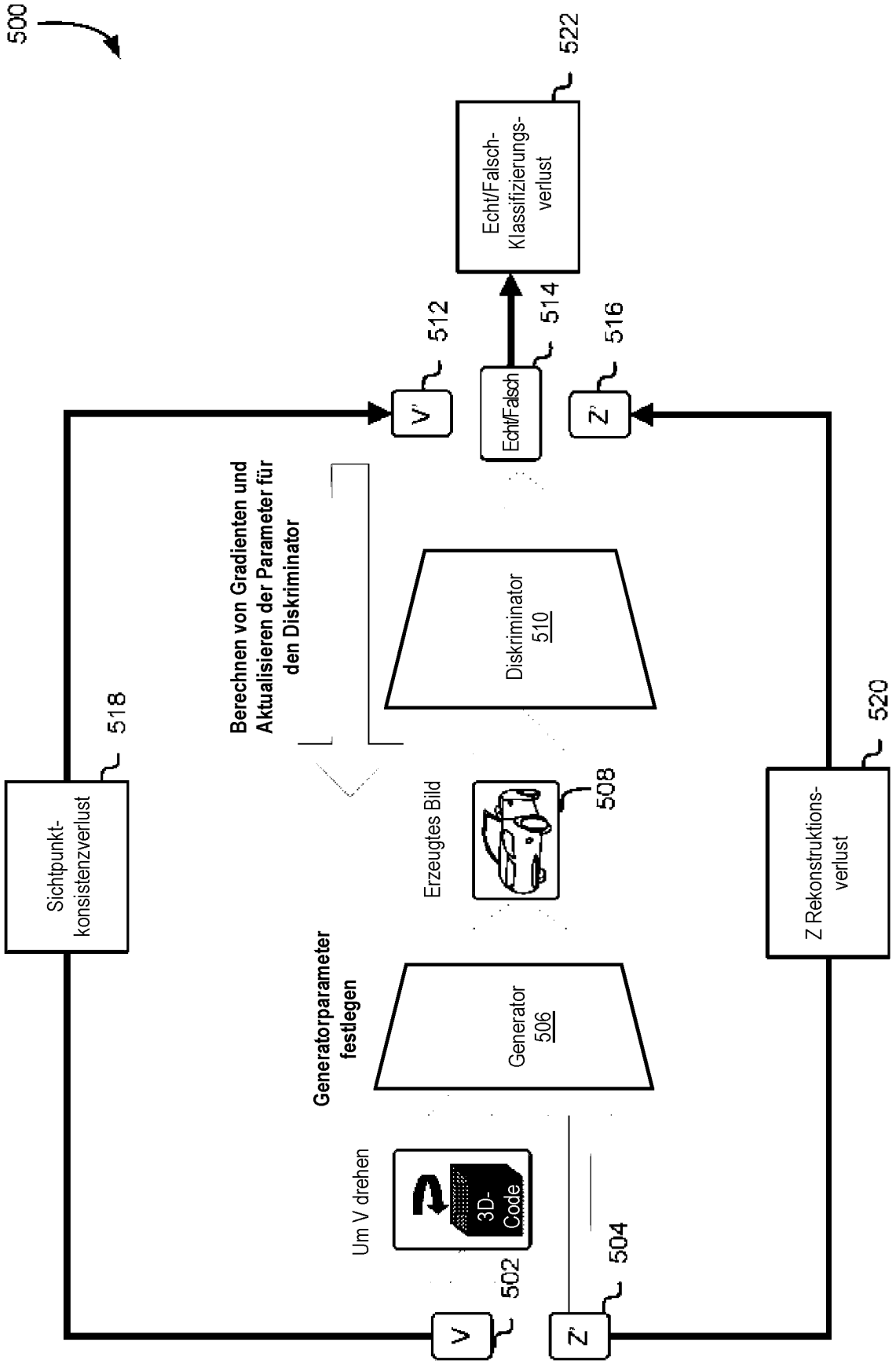

500
522
Echt/Falsch-Klassifizierungs-verlust
512
514
516
V'
Echt/Falsch
Z'
Berechnen von Gradienten und Aktualisieren der Parameter für den Diskriminator
Diskriminator 510
518
520
Sichtpunkt-konsistenzverlust
Erzeugtes Bild
508
Z Rekonstruktions-verlust
Generatorparameter festlegen
Generator 506
Um V drehen
3D-Code
502
504
V
Z'

FIG. 5

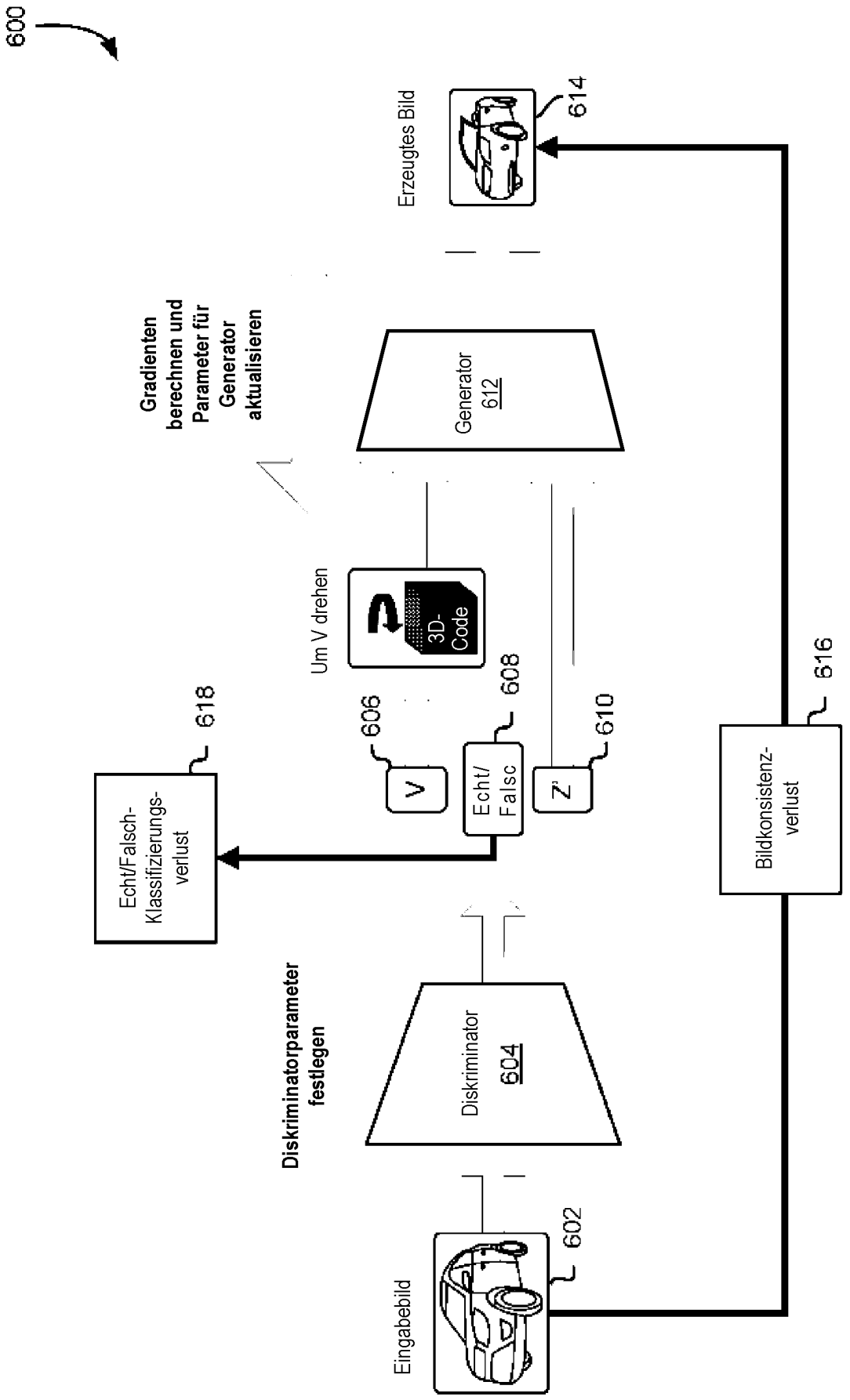

600
Erzeugtes Bild
614
Gradienten berechnen und Parameter für Generator aktualisieren
Generator 612
Um V drehen
3D-Code
606
V
608
Echt/ Falsc
610
Z'
618
Echt/Falsch-Klassifizierungs-verlust
616
Bildkonsistenz-verlust
Diskriminatorparameter festlegen
Diskriminator 604
602
Eingabebild

FIG. 6

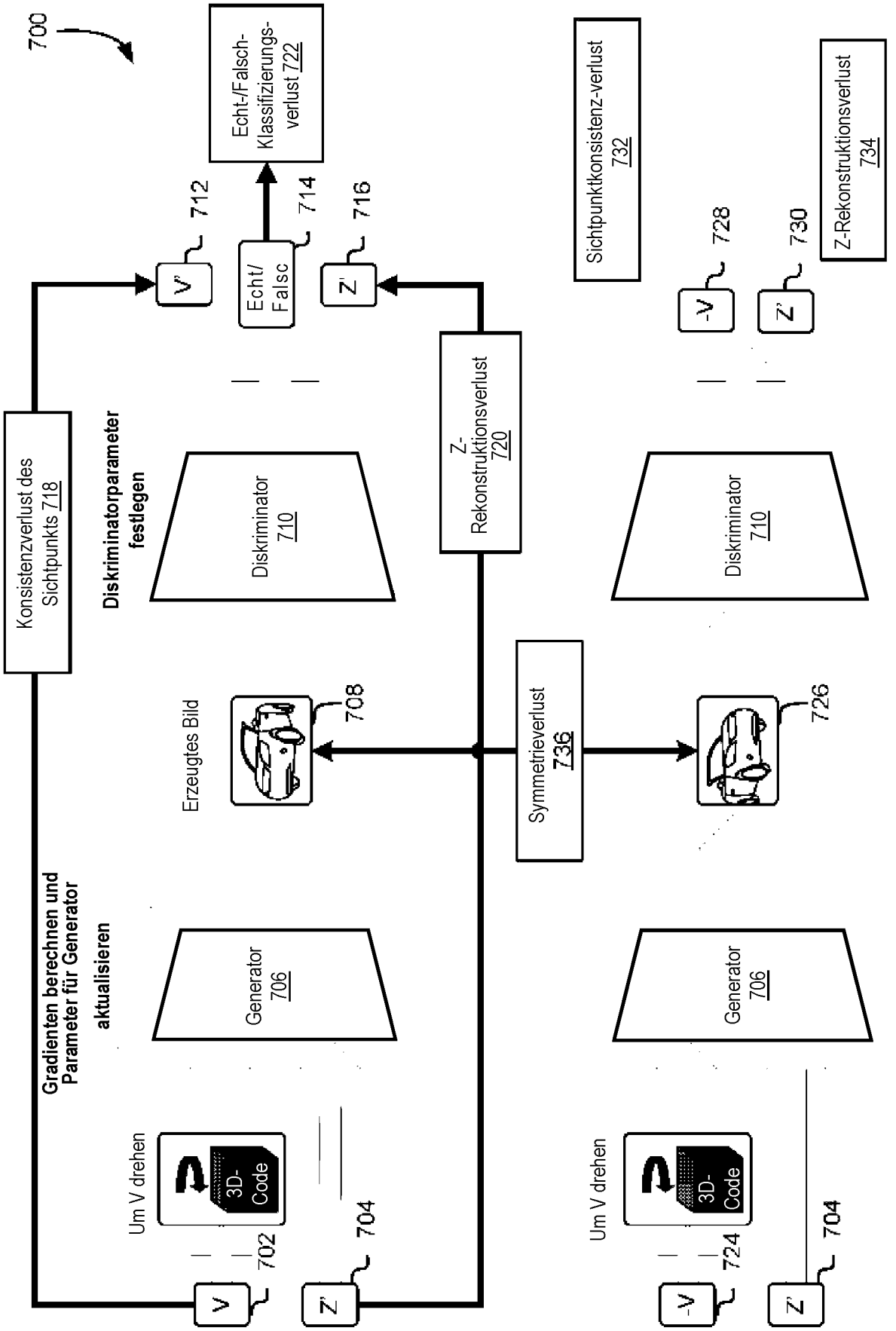

700
Echt-/Falsch-Klassifizierungs-verlust 722
712
714
716
V'
Echt/ Falsc
Z'
Sichtpunktkonsistenz-verlust 732
728
730
-V
Z'
Z-Rekonstruktionsverlust 734
Konsistenzverlust des Sichtpunkts 718
Diskriminatorparameter festlegen
Diskriminator 710
Z-Rekonstruktionsverlust 720
Diskriminator 710
Erzeugtes Bild
708
Symmetrieverlust 736
726
Gradienten berechnen und Parameter für Generator aktualisieren
Generator 706
Generator 706
Um V drehen
3D-Code
702
704
Um V drehen
3D-Code
724
704
V
Z'
-V
Z'

FIG. 7

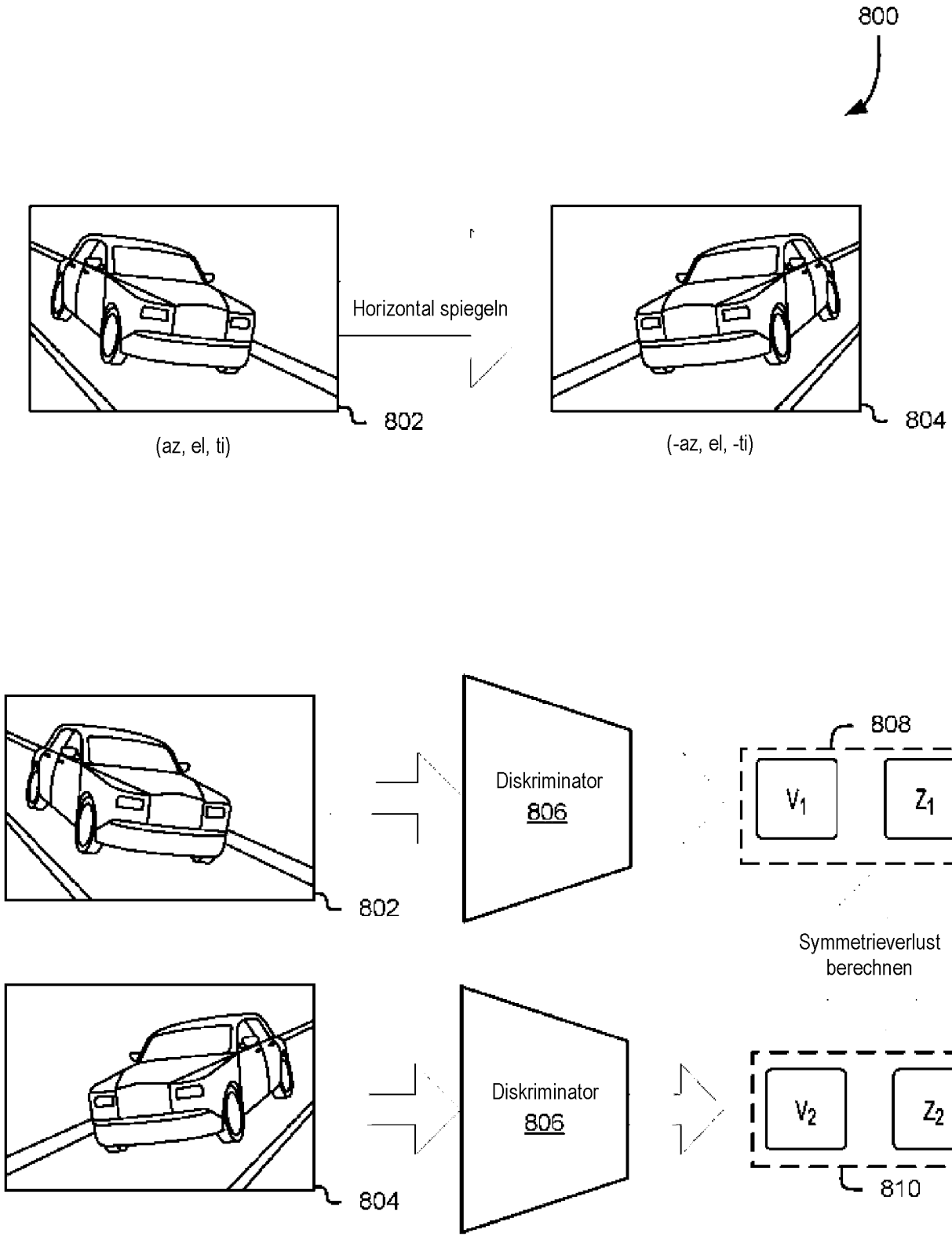

800
Horizontal spiegeln
802
804
(az, el, ti)
(-az, el, -ti)
Diskriminator
806
808
V_1
Z_1
Symmetrieverlust
berechnen
Diskriminator
806
V_2
Z_2
810

FIG. 8

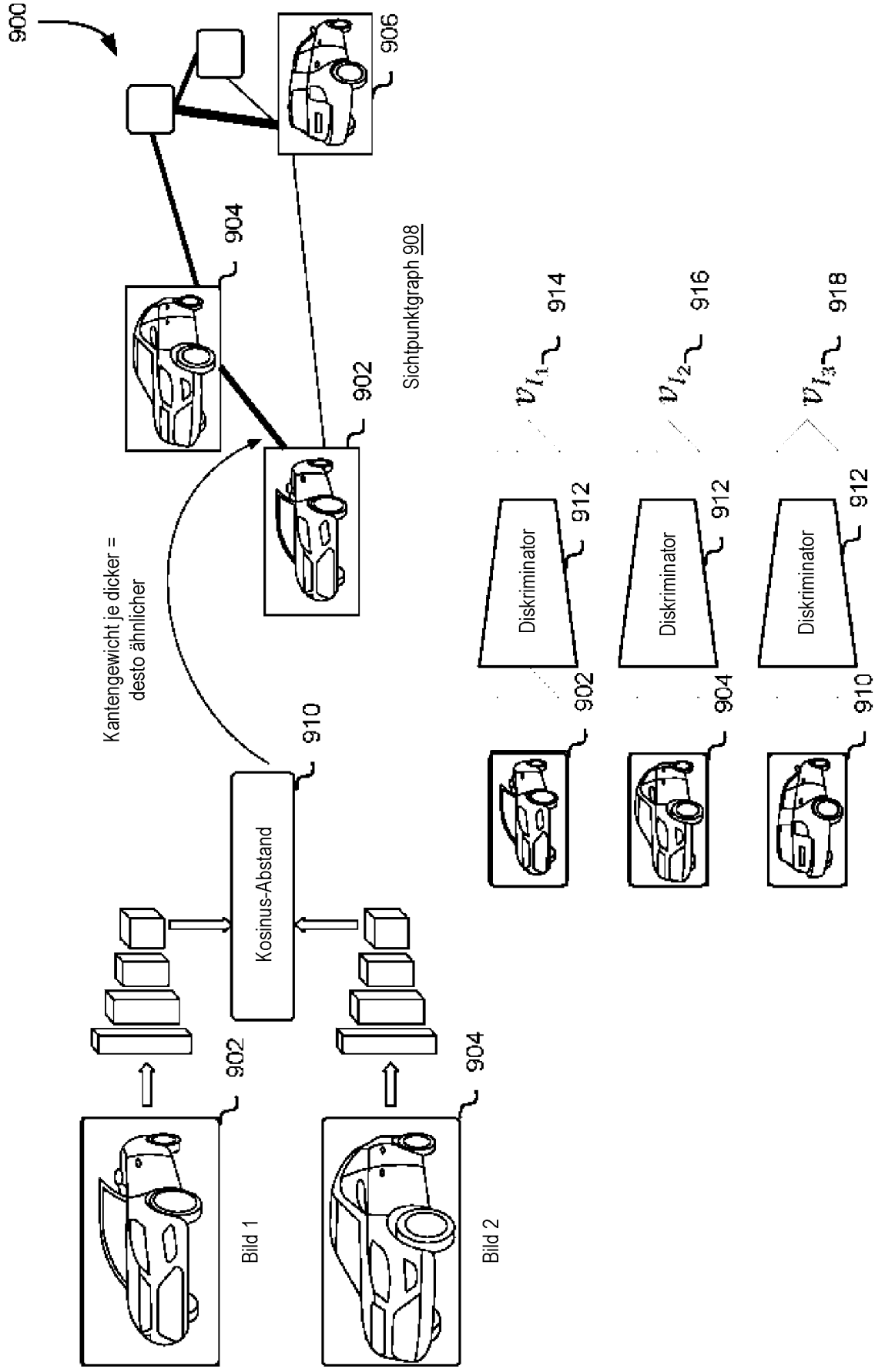

900
906
904
902
Sichtpunktgraph 908
Kantengewicht je dicker = desto ähnlicher
910
Kosinus-Abstand
902
904
Bild 1
Bild 2
v_{I_1} 914
v_{I_2} 916
v_{I_3} 918
Diskriminator 912
Diskriminator 912
Diskriminator 912
902
904
910

FIG. 9

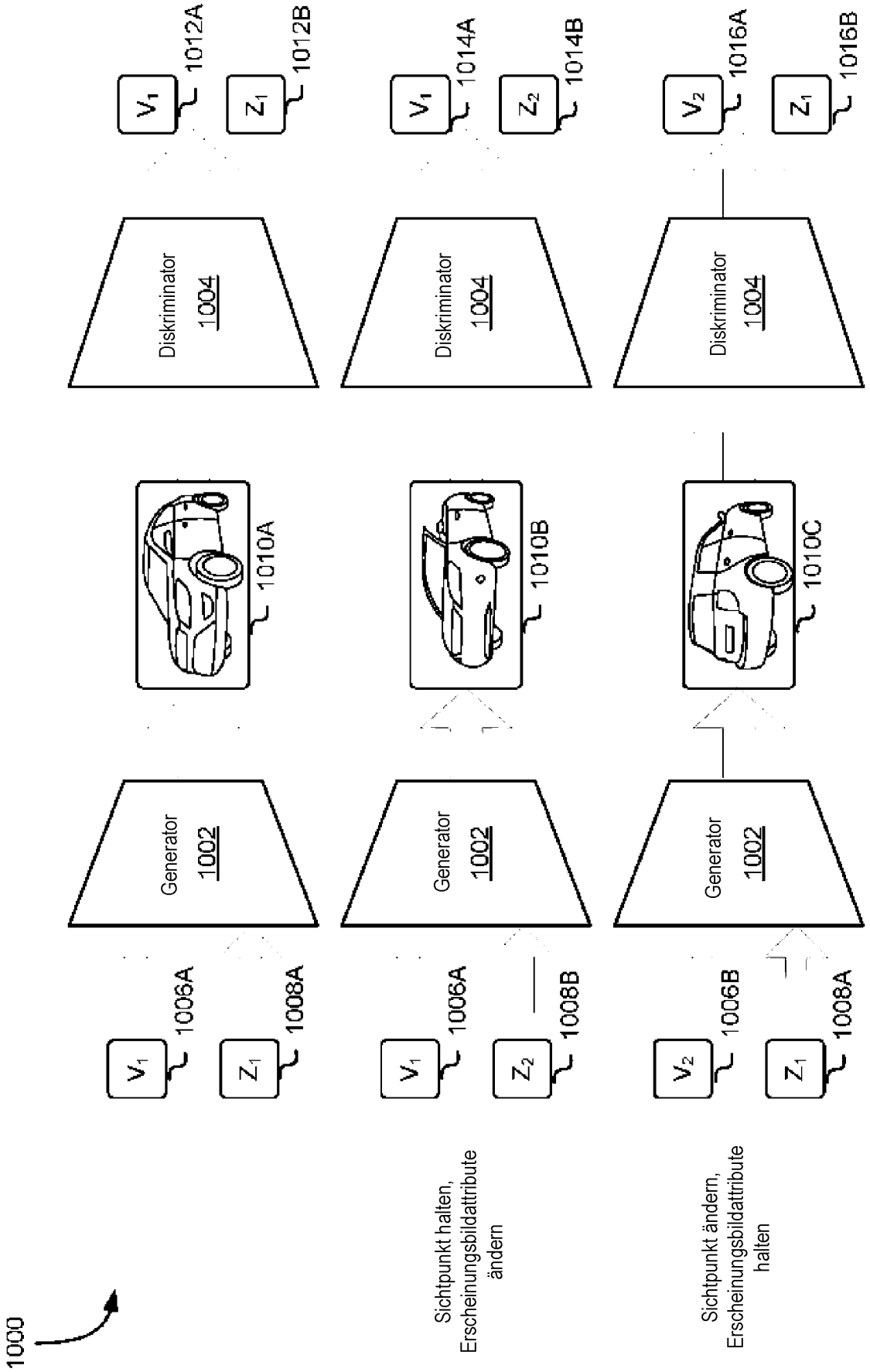

1000
Sichtpunkt halten, Erscheinungsbildattribute ändern
Sichtpunkt ändern, Erscheinungsbildattribute halten
Generator 1002
Diskriminator 1004
v_1 1006A
z_1 1008A
v_1 1006A
z_2 1008B
v_2 1006B
z_1 1008A
1010A
1010B
1010C
v_1 1012A
z_1 1012B
v_1 1014A
z_2 1014B
v_2 1016A
z_1 1016B

FIG. 10

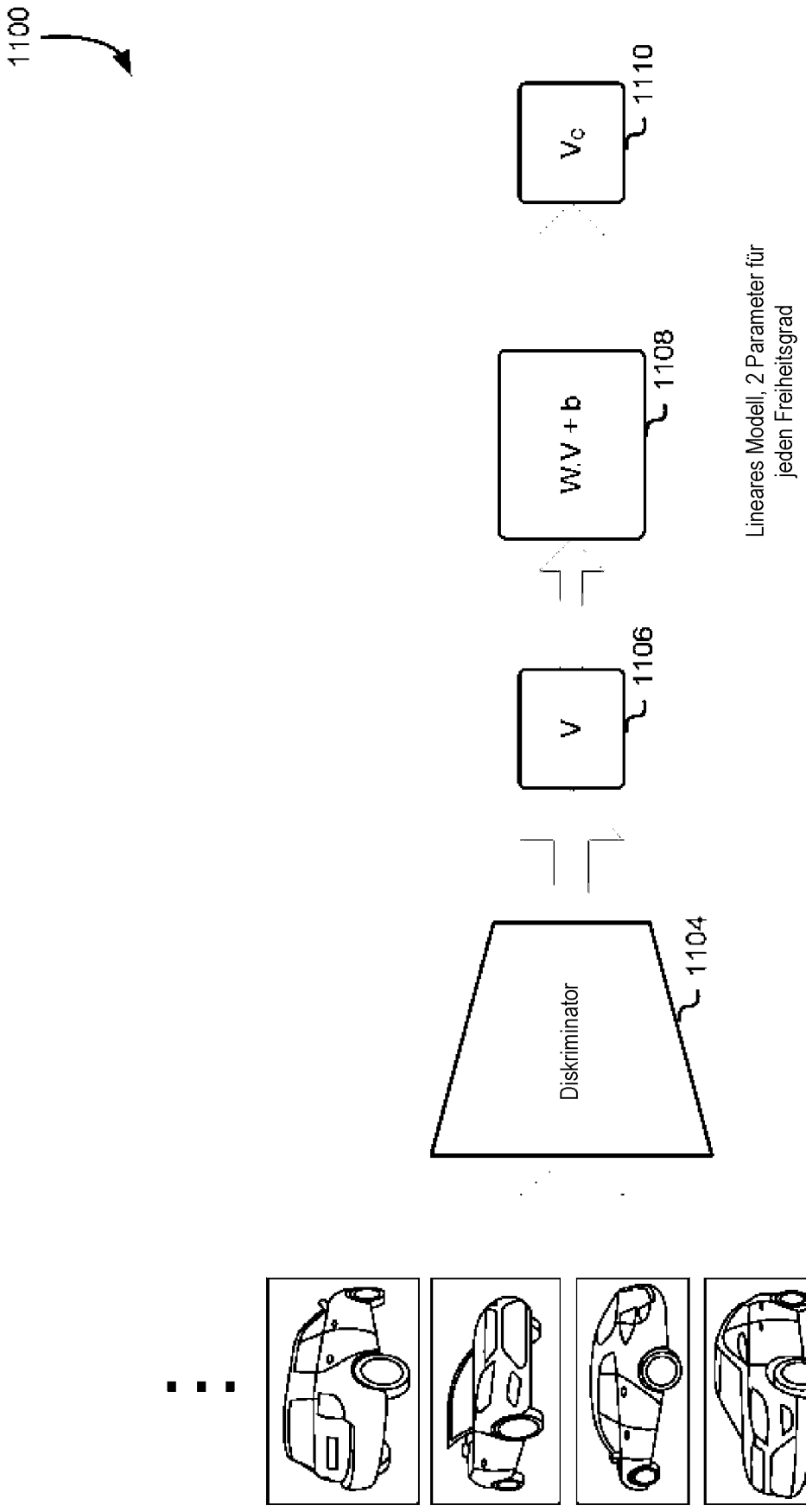

1100
1110
V_c
1108
W.V + b
Lineares Modell, 2 Parameter für jeden Freiheitsgrad
1106
V
1104
Diskriminator
1102
Objektbildsammlung

FIG. 11

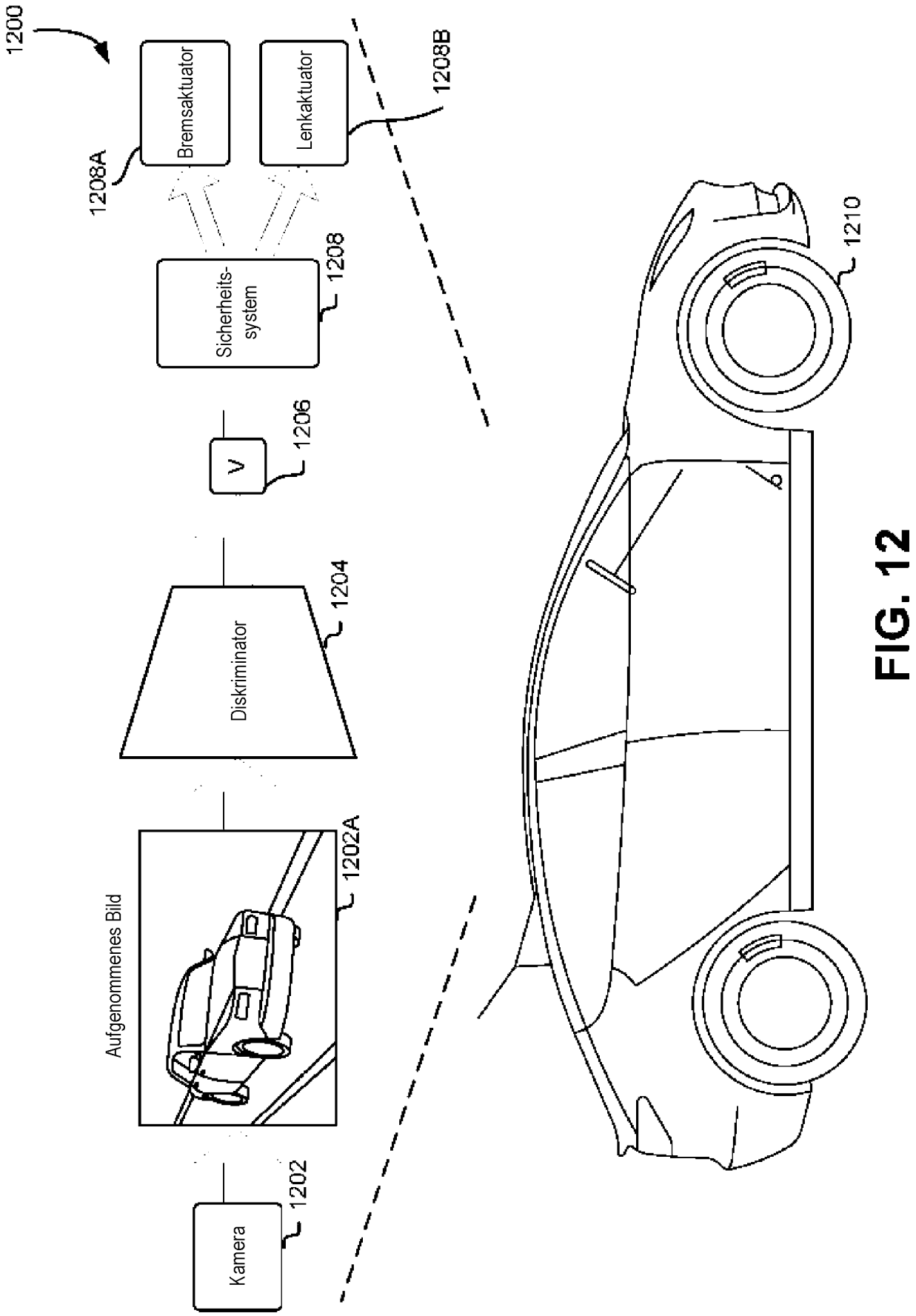

1200
Bremsaktuator
Lenkaktuator
1208A
1208B
Sicherheits-system
1208
V
1206
Diskriminator
1204
Aufgenommenes Bild
1202A
Kamera
1202
1210

FIG. 12

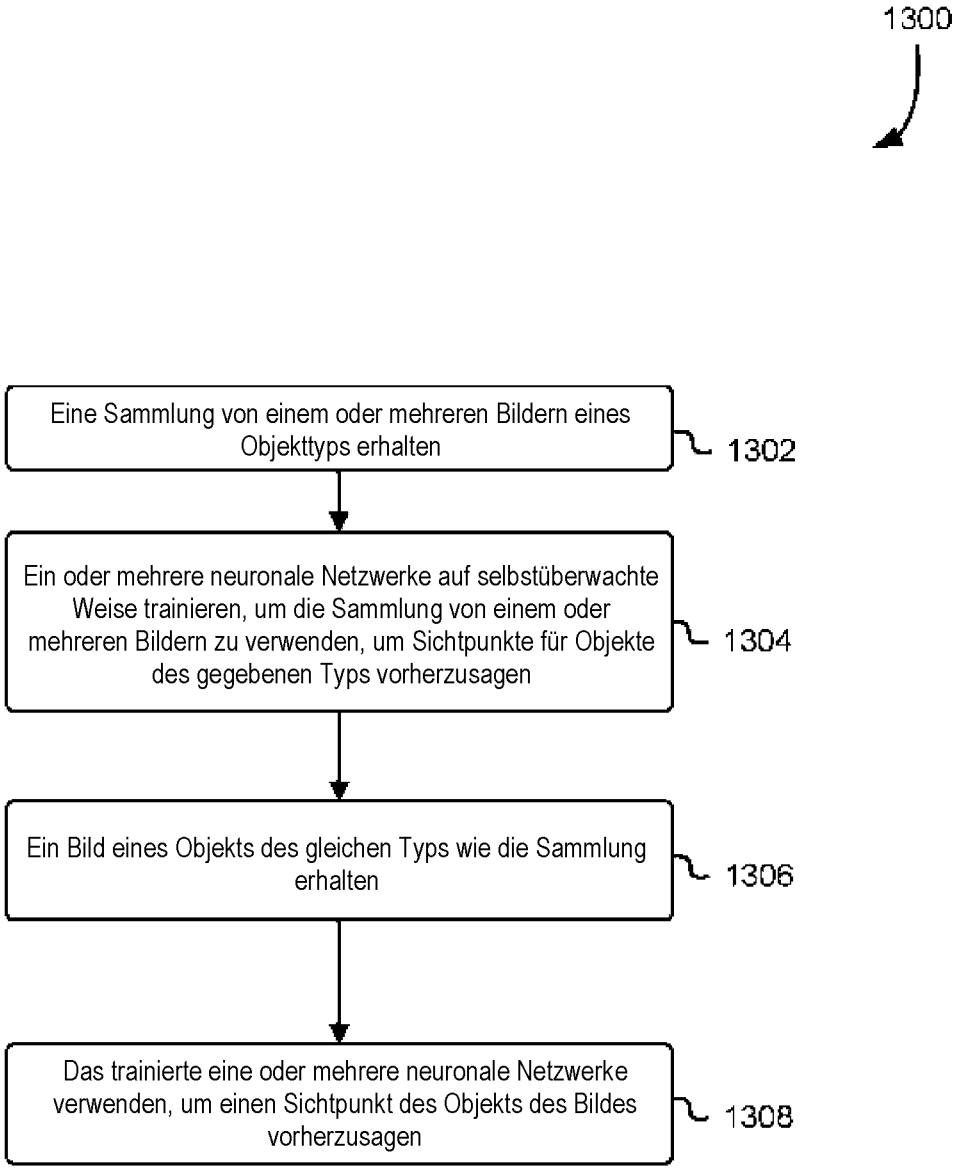

1300
Eine Sammlung von einem oder mehreren Bildern eines Objekttyps erhalten
1302
Ein oder mehrere neuronale Netzwerke auf selbstüberwachte Weise trainieren, um die Sammlung von einem oder mehreren Bildern zu verwenden, um Sichtpunkte für Objekte des gegebenen Typs vorherzusagen
1304
Ein Bild eines Objekts des gleichen Typs wie die Sammlung erhalten
1306
Das trainierte eine oder mehrere neuronale Netzwerke verwenden, um einen Sichtpunkt des Objekts des Bildes vorherzusagen
1308

FIG. 13

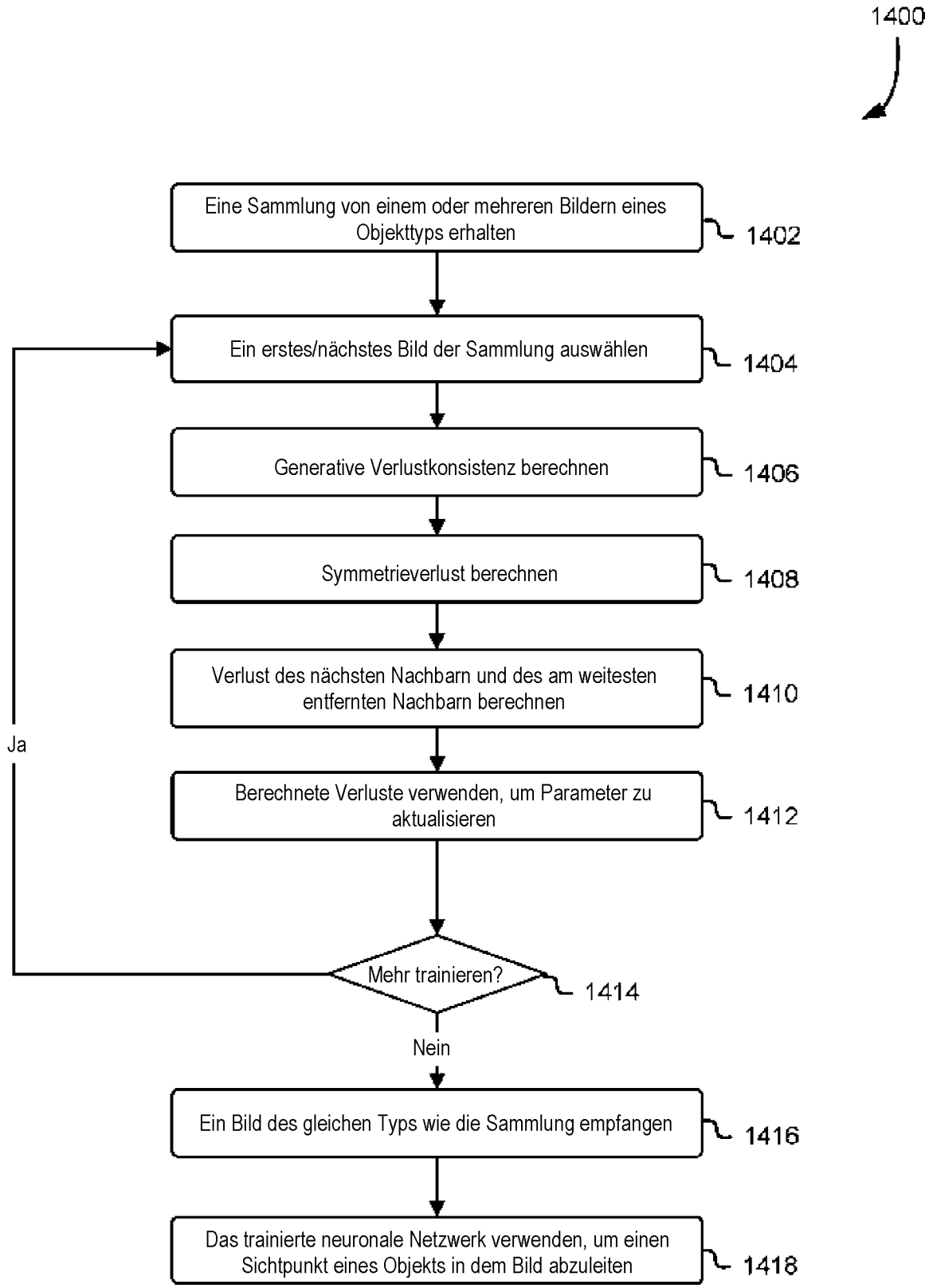

1400
Eine Sammlung von einem oder mehreren Bildern eines Objekttyps erhalten
1402
Ein erstes/nächstes Bild der Sammlung auswählen
1404
Generative Verlustkonsistenz berechnen
1406
Symmetrieverlust berechnen
1408
Verlust des nächsten Nachbarn und des am weitesten entfernten Nachbarn berechnen
1410
Berechnete Verluste verwenden, um Parameter zu aktualisieren
1412
Mehr trainieren?
1414
Ja
Nein
Ein Bild des gleichen Typs wie die Sammlung empfangen
1416
Das trainierte neuronale Netzwerk verwenden, um einen Sichtpunkt eines Objekts in dem Bild abzuleiten
1418

FIG. 14

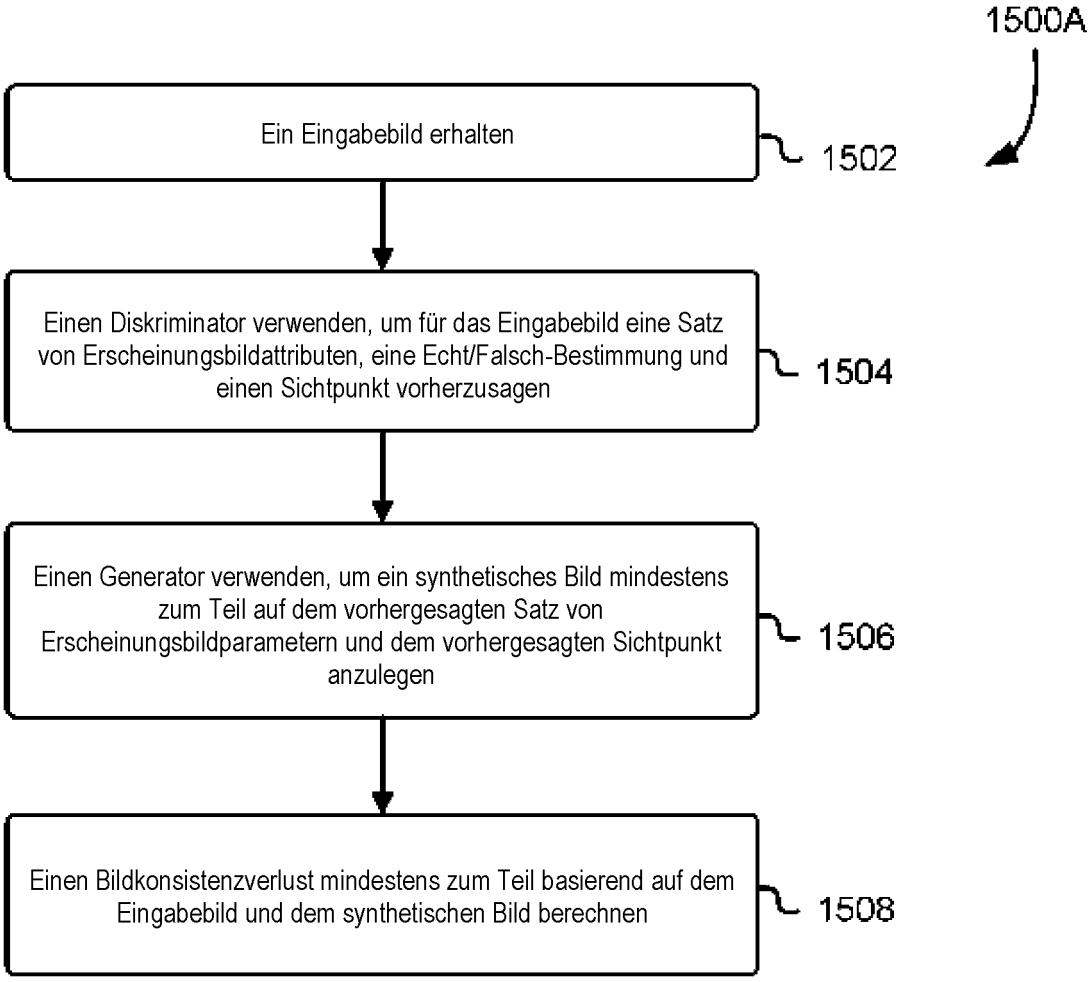

1500A
Ein Eingabebild erhalten
1502
Einen Diskriminator verwenden, um für das Eingabebild eine Satz von Erscheinungsbildattributen, eine Echt/Falsch-Bestimmung und einen Sichtpunkt vorherzusagen
1504
Einen Generator verwenden, um ein synthetisches Bild mindestens zum Teil auf dem vorhergesagten Satz von Erscheinungsbildparametern und dem vorhergesagten Sichtpunkt anzulegen
1506
Einen Bildkonsistenzverlust mindestens zum Teil basierend auf dem Eingabebild und dem synthetischen Bild berechnen
1508

FIG. 15A

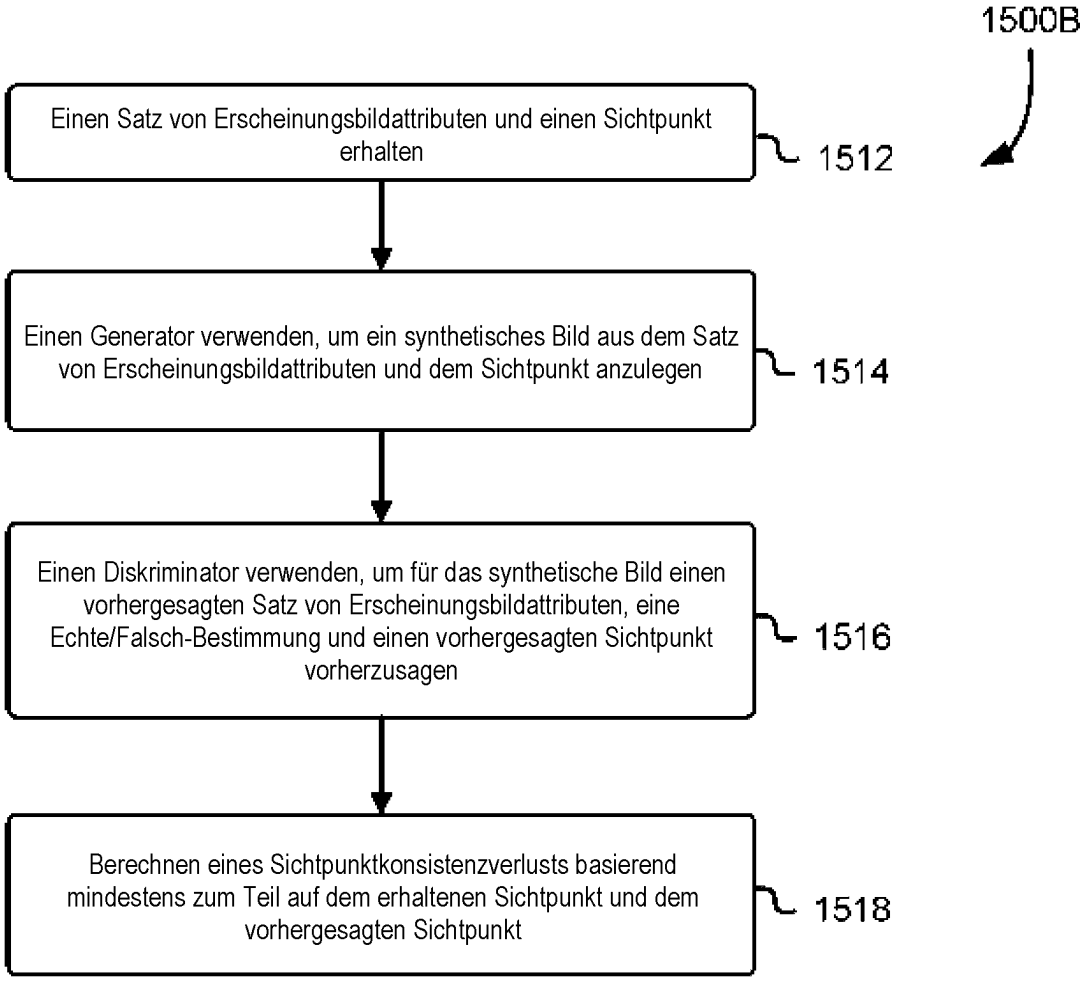

1500B
Einen Satz von Erscheinungsbildattributen und einen Sichtpunkt erhalten
1512
Einen Generator verwenden, um ein synthetisches Bild aus dem Satz von Erscheinungsbildattributen und dem Sichtpunkt anzulegen
1514
Einen Diskriminator verwenden, um für das synthetische Bild einen vorhergesagten Satz von Erscheinungsbildattributen, eine Echte/Falsch-Bestimmung und einen vorhergesagten Sichtpunkt vorherzusagen
1516
Berechnen eines Sichtpunktkonsistenzverlusts basierend mindestens zum Teil auf dem erhaltenen Sichtpunkt und dem vorhergesagten Sichtpunkt
1518

FIG. 15B

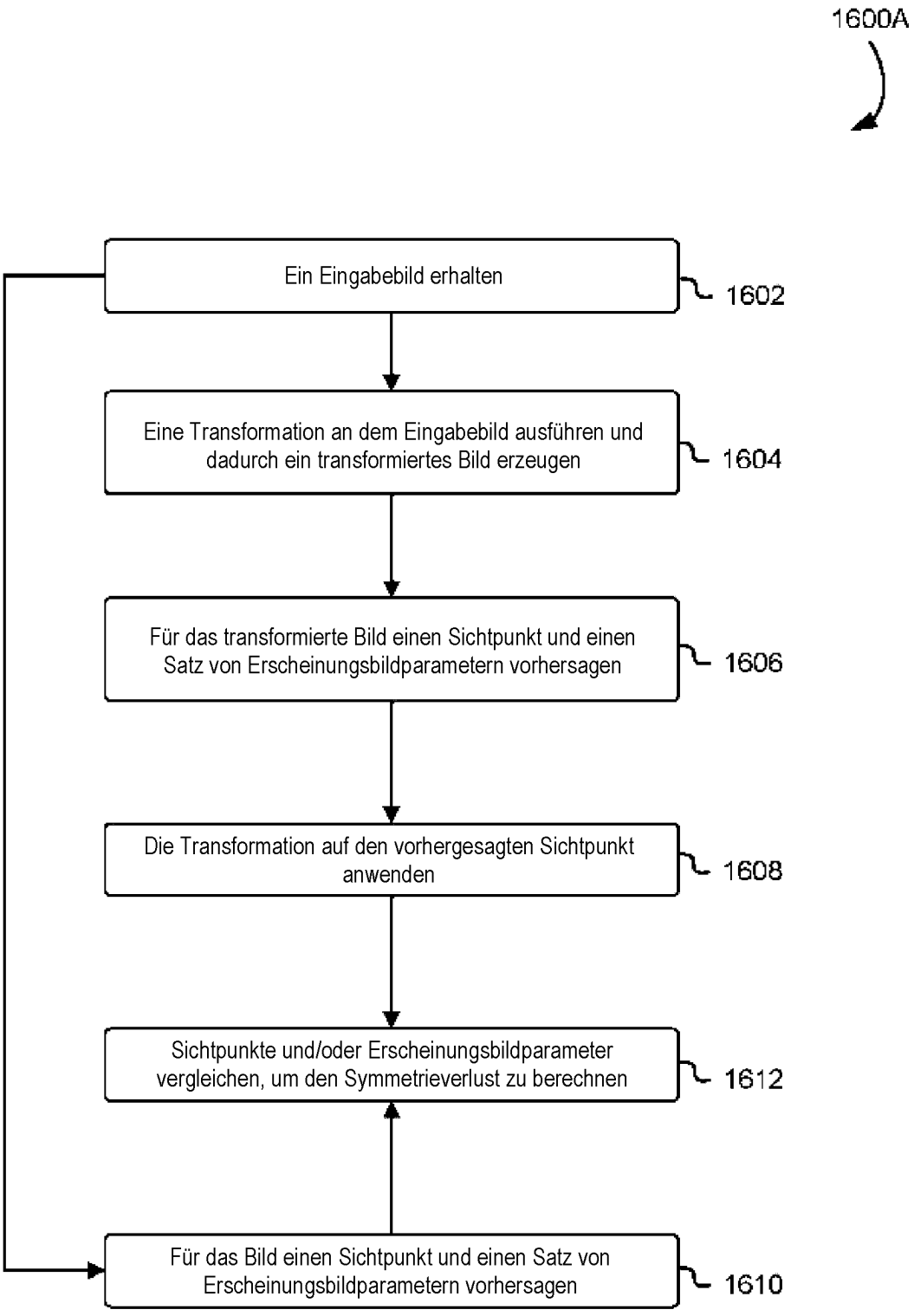

1600A
Ein Eingabebild erhalten
1602
Eine Transformation an dem Eingabebild ausführen und dadurch ein transformiertes Bild erzeugen
1604
Für das transformierte Bild einen Sichtpunkt und einen Satz von Erscheinungsbildparametern vorhersagen
1606
Die Transformation auf den vorhergesagten Sichtpunkt anwenden
1608
Sichtpunkte und/oder Erscheinungsbildparameter vergleichen, um den Symmetrieverlust zu berechnen
1612
Für das Bild einen Sichtpunkt und einen Satz von Erscheinungsbildparametern vorhersagen
1610

FIG. 16A

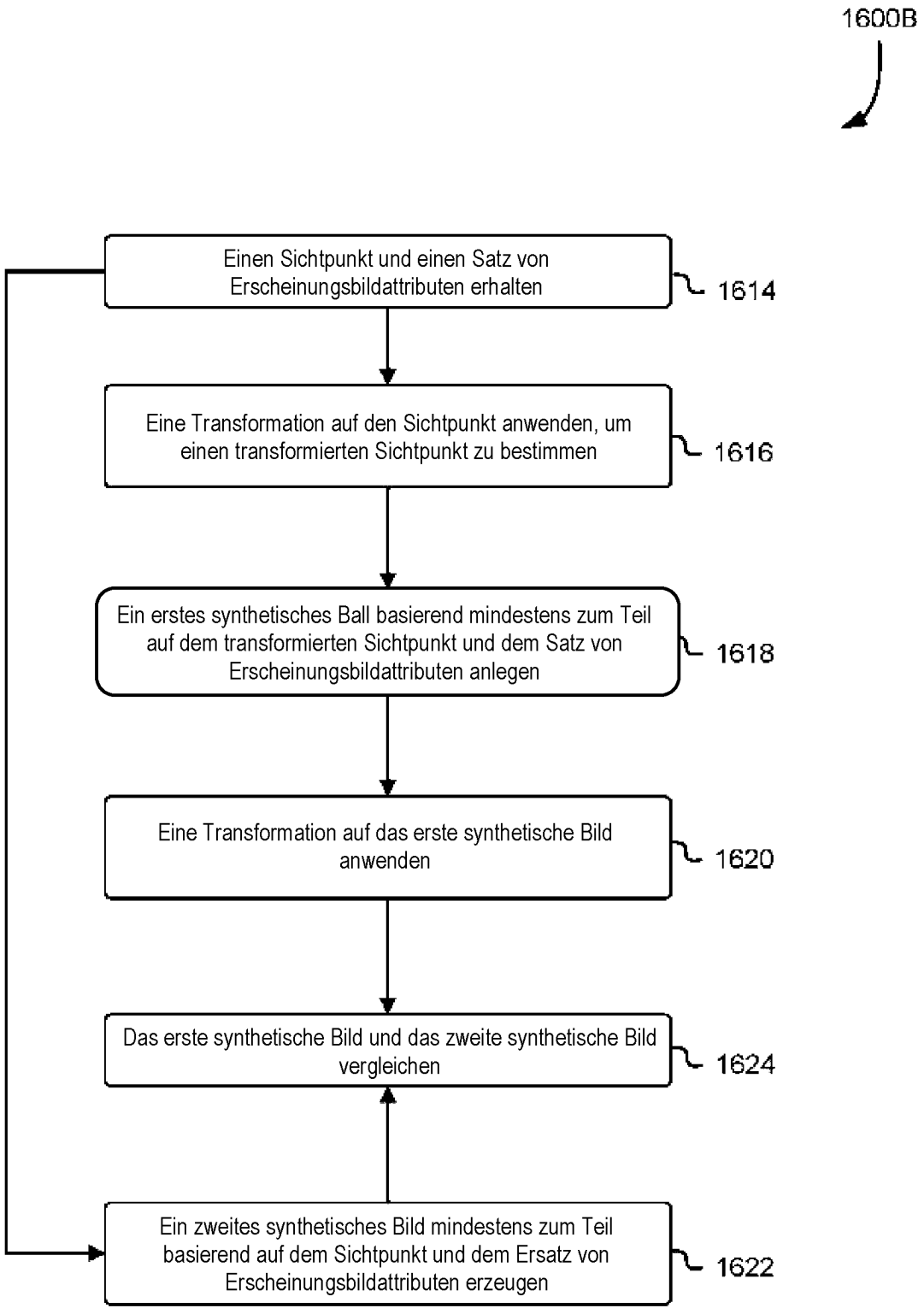

1600B
Einen Sichtpunkt und einen Satz von Erscheinungsbildattributen erhalten
1614
Eine Transformation auf den Sichtpunkt anwenden, um einen transformierten Sichtpunkt zu bestimmen
1616
Ein erstes synthetisches Ball basierend mindestens zum Teil auf dem transformierten Sichtpunkt und dem Satz von Erscheinungsbildattributen anlegen
1618
Eine Transformation auf das erste synthetische Bild anwenden
1620
Das erste synthetische Bild und das zweite synthetische Bild vergleichen
1624
Ein zweites synthetisches Bild mindestens zum Teil basierend auf dem Sichtpunkt und dem Ersatz von Erscheinungsbildattributen erzeugen
1622

FIG. 16B

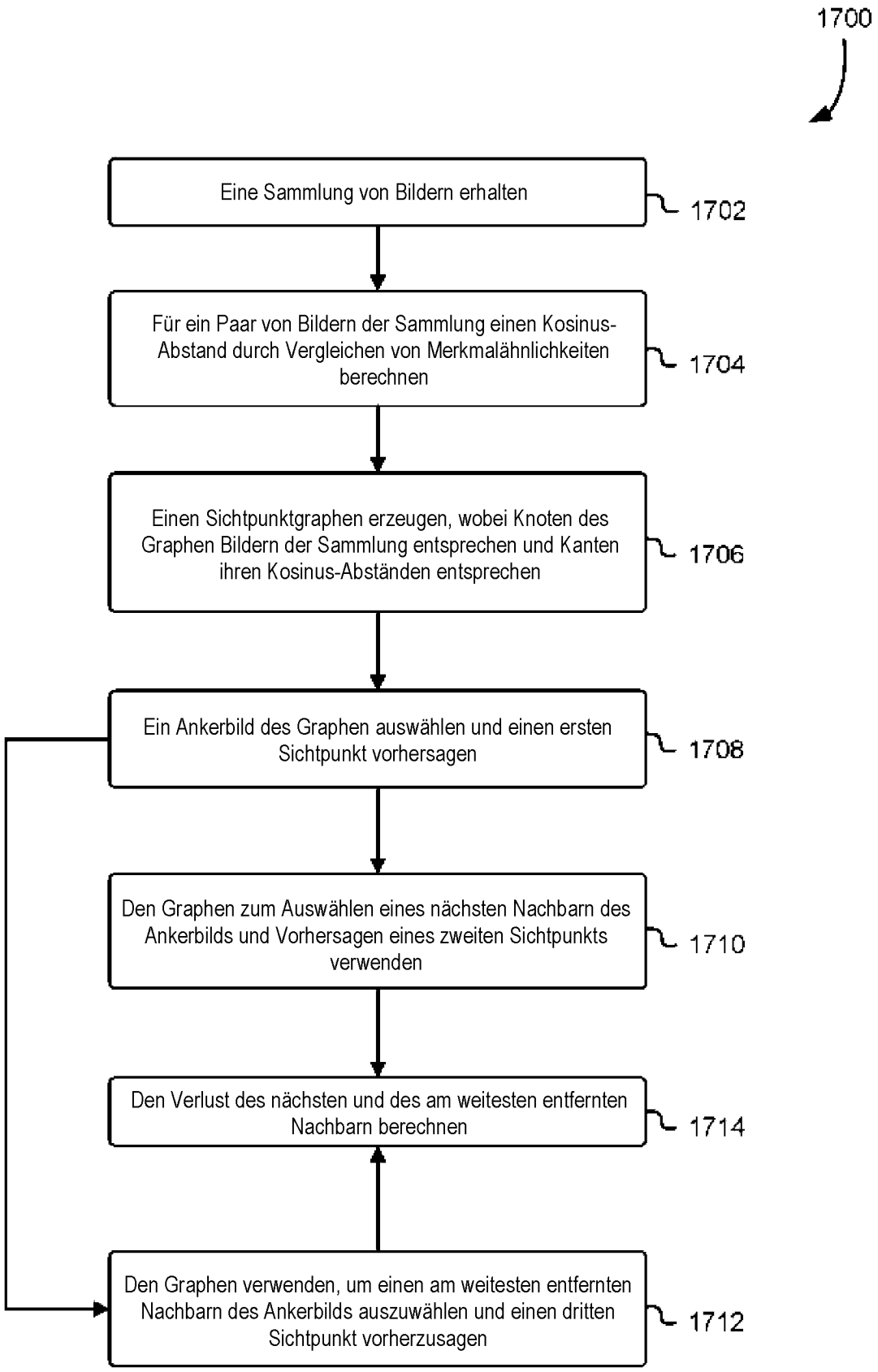

1700
Eine Sammlung von Bildern erhalten
1702
Für ein Paar von Bildern der Sammlung einen Kosinus-Abstand durch Vergleichen von Merkmalähnlichkeiten berechnen
1704
Einen Sichtpunktgraphen erzeugen, wobei Knoten des Graphen Bildern der Sammlung entsprechen und Kanten ihren Kosinus-Abständen entsprechen
1706
Ein Ankerbild des Graphen auswählen und einen ersten Sichtpunkt vorhersagen
1708
Den Graphen zum Auswählen eines nächsten Nachbarn des Ankerbilds und Vorhersagen eines zweiten Sichtpunkts verwenden
1710
Den Verlust des nächsten und des am weitesten entfernten Nachbarn berechnen
1714
Den Graphen verwenden, um einen am weitesten entfernten Nachbarn des Ankerbilds auszuwählen und einen dritten Sichtpunkt vorherzusagen
1712

FIG. 17

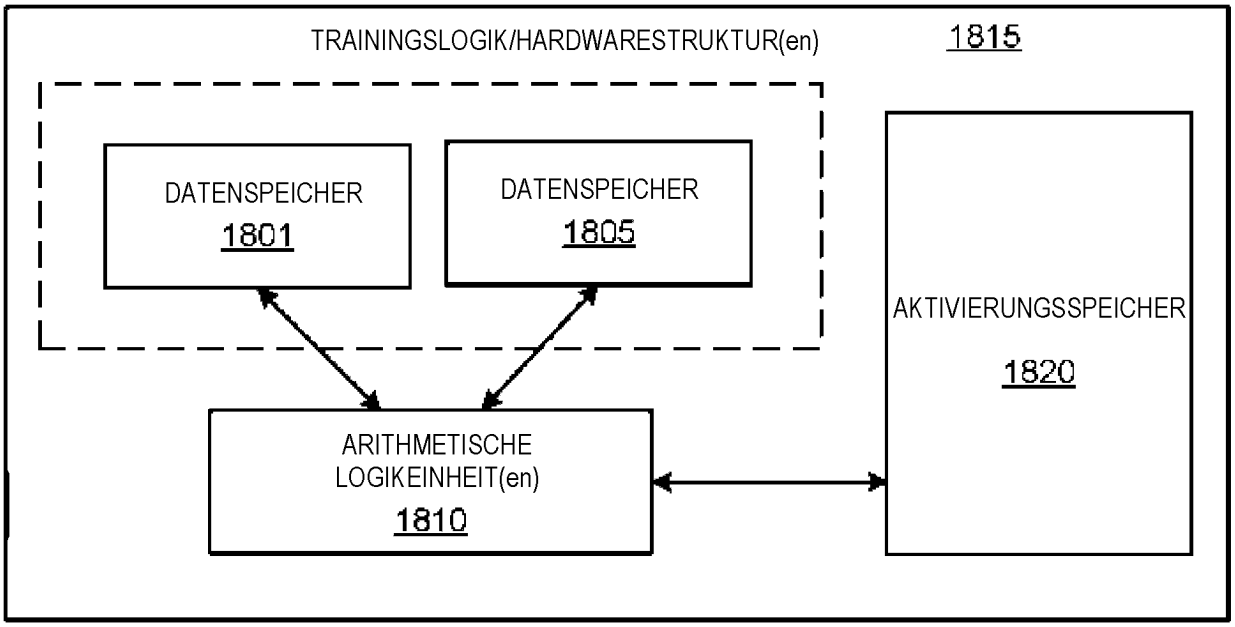


FIG. 18A

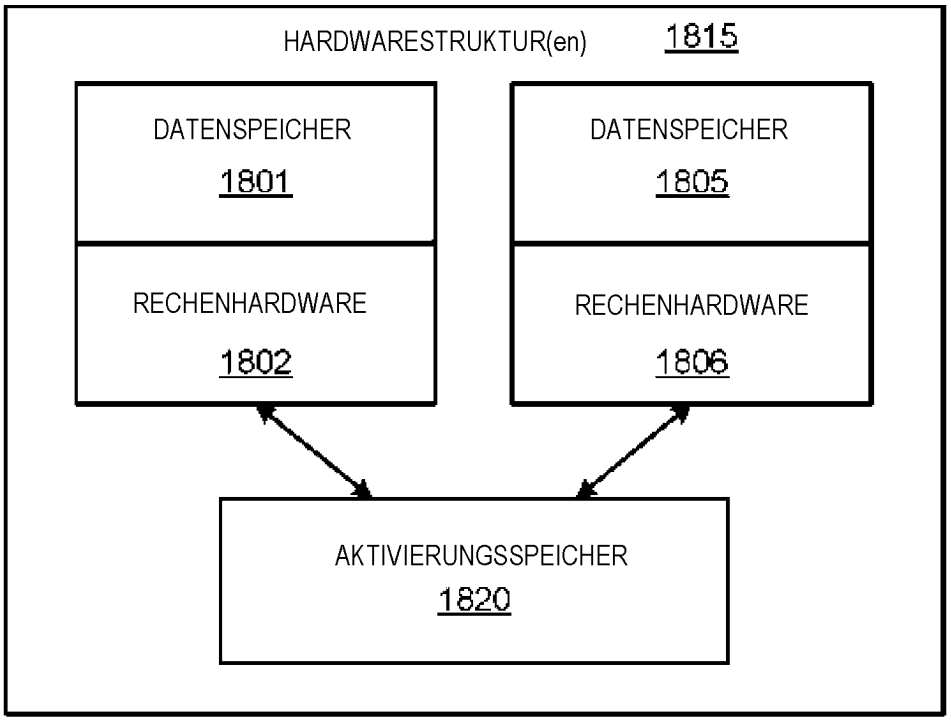


FIG. 18B

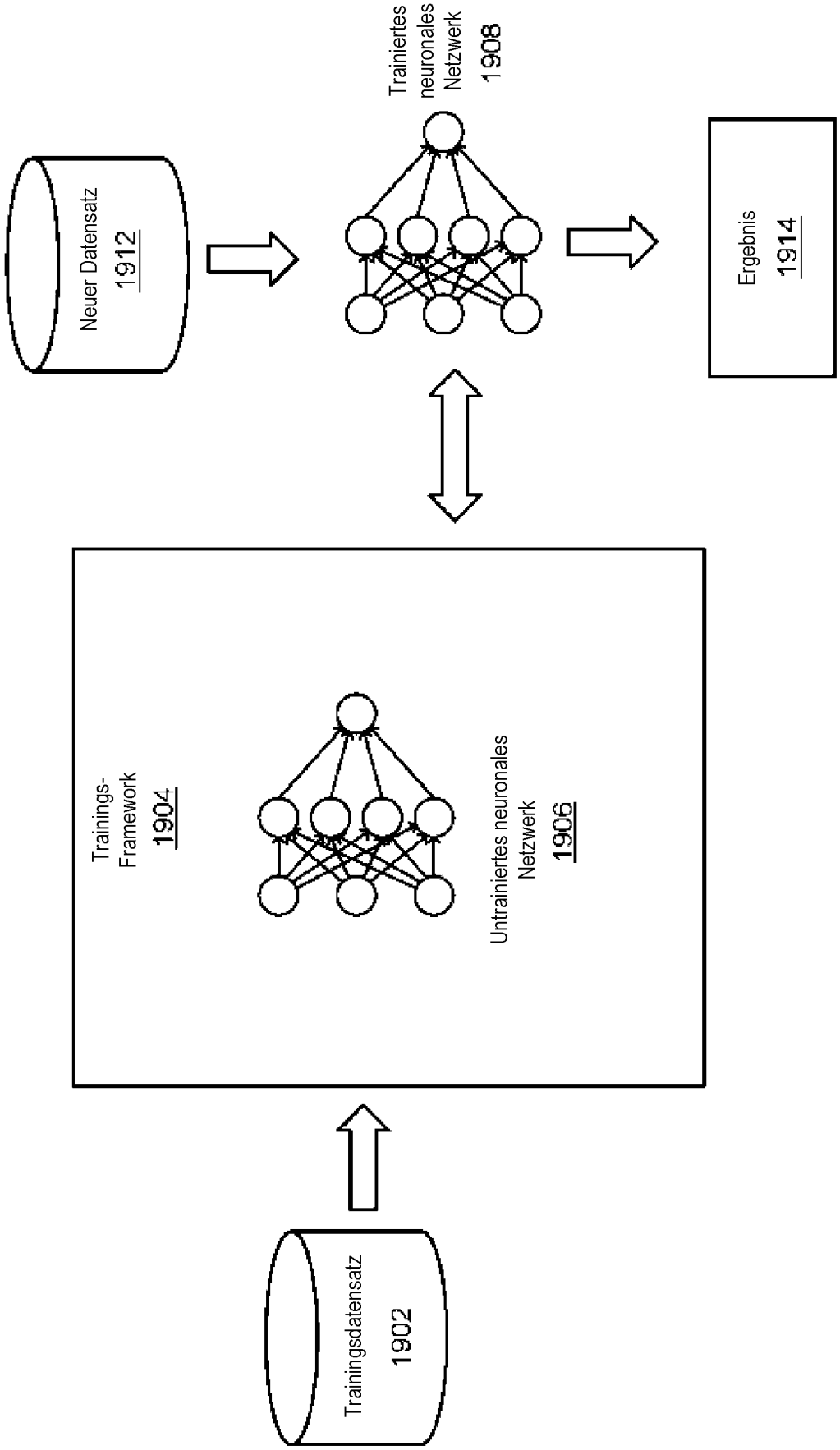

Trainingsdatensatz
1902
Trainings-
Framework
1904
Untrainiertes neuronales
Netzwerk
1906
Trainiertes
neuronales
Netzwerk
1908
Neuer Datensatz
1912
Ergebnis
1914

FIG. 19

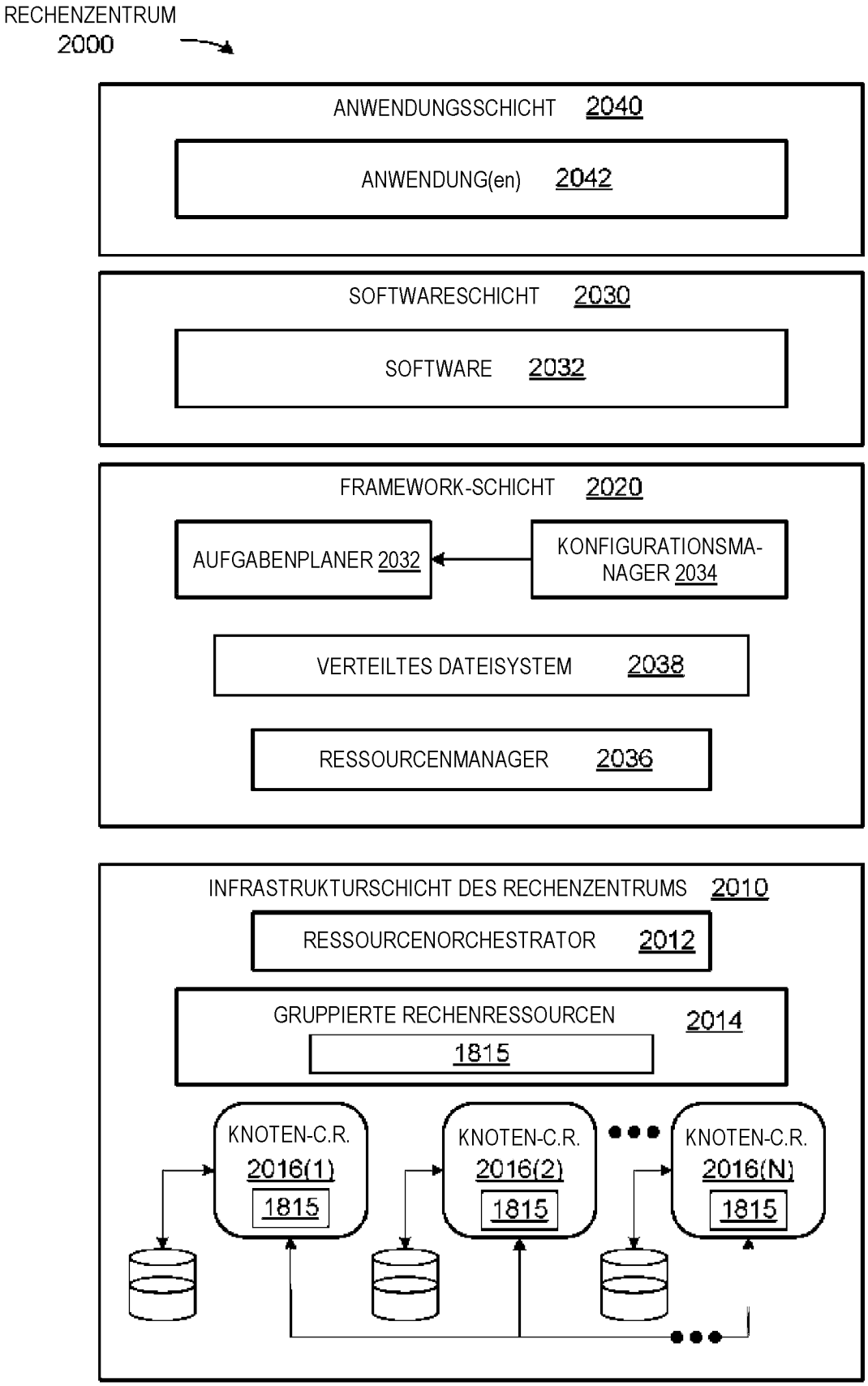

RECHENZENTRUM
2000
ANWENDUNGSSCHICHT 2040
ANWENDUNG(en) 2042
SOFTWARESCHICHT 2030
SOFTWARE 2032
FRAMEWORK-SCHICHT 2020
AUFGABENPLANER 2032
KONFIGURATIONSMA-NAGER 2034
VERTEILTES DATEISYSTEM 2038
RESSOURCENMANAGER 2036
INFRASTRUKTURSCHICHT DES RECHENZENTRUMS 2010
RESSOURCENORCHESTRATOR 2012
GRUPPIERTE RECHENRESSOURCEN 2014
1815
KNOTEN-C.R. 2016(1) 1815
KNOTEN-C.R. 2016(2) 1815
KNOTEN-C.R. 2016(N) 1815

FIG. 20

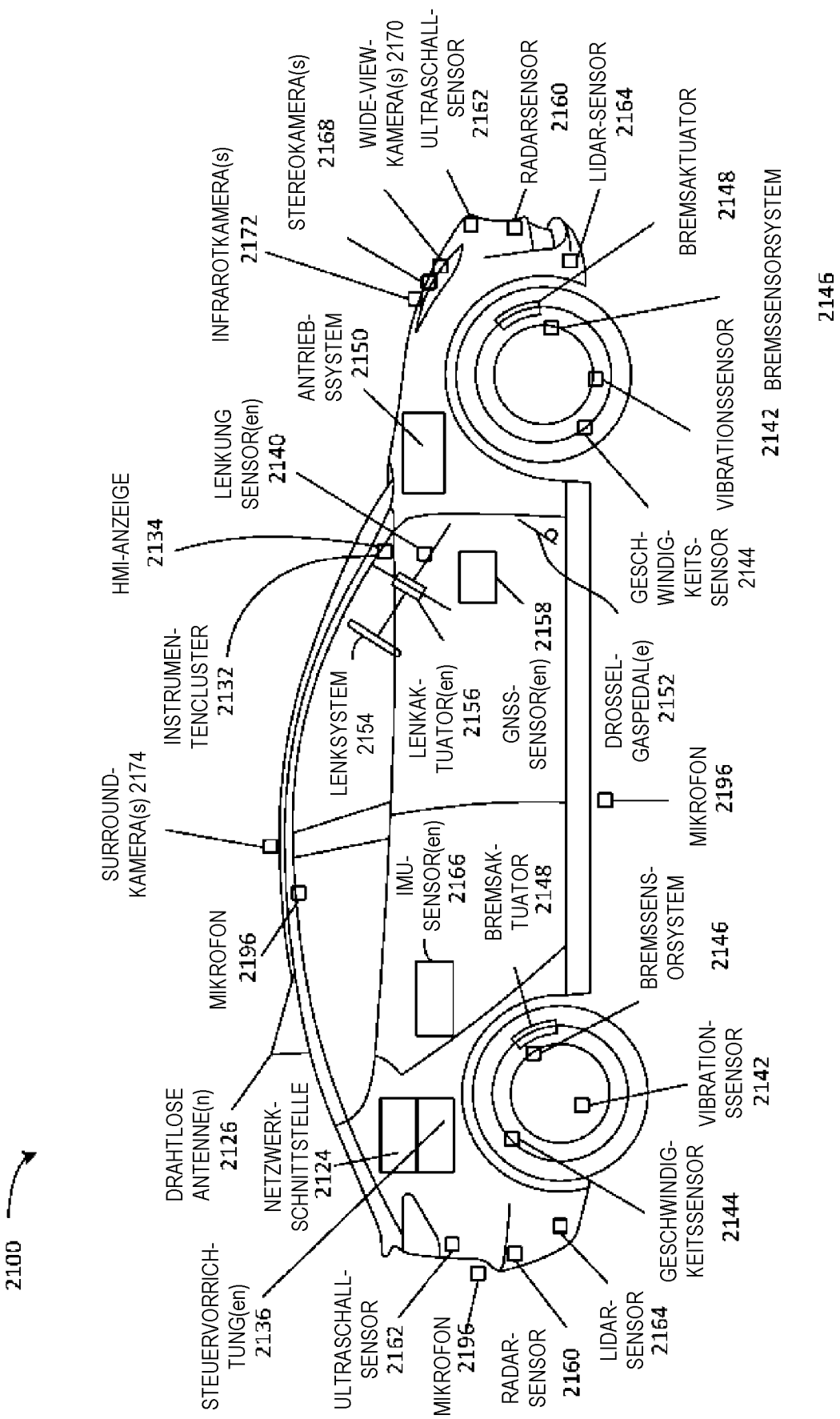

2100
STEUERVORRICHTUNG(en) 2136
DRAHTLOSE ANTENNE(n) 2126
NETZWERKSCHNITTSTELLE 2124
SURROUNDKAMERA(s) 2174
MIKROFON 2196
HMI-ANZEIGE 2134
INSTRUMENTENCLUSTER 2132
LENKUNG SENSOR(en) 2140
INFRAROTKAMERA(s) 2172
STEREOKAMERA(s) 2168
WIDE-VIEWKAMERA(s) 2170
ANTRIEBSSYSTEM 2150
ULTRASCHALLSENSOR 2162
RADARSENSOR 2160
LIDAR-SENSOR 2164
BREMSAKTUATOR 2148
BREMSSENSORSYSTEM 2146
VIBRATIONSSENSOR 2142
GESCHWINDIGKEITSSENSOR 2144
LENKSYSTEM 2154
LENKAKTUATOR(en) 2156
GNSSSENSOR(en) 2158
DROSSELGASPEDAL(e) 2152
MIKROFON 2196
IMUSENSOR(en) 2166
BREMSAKTUATOR 2148
BREMSSENSORSYSTEM 2146
VIBRATIONSSENSOR 2142
GESCHWINDIGKEITSSENSOR 2144
ULTRASCHALLSENSOR 2162
MIKROFON 2196
RADARSENSOR 2160
LIDARSENSOR 2164

FIG. 21A

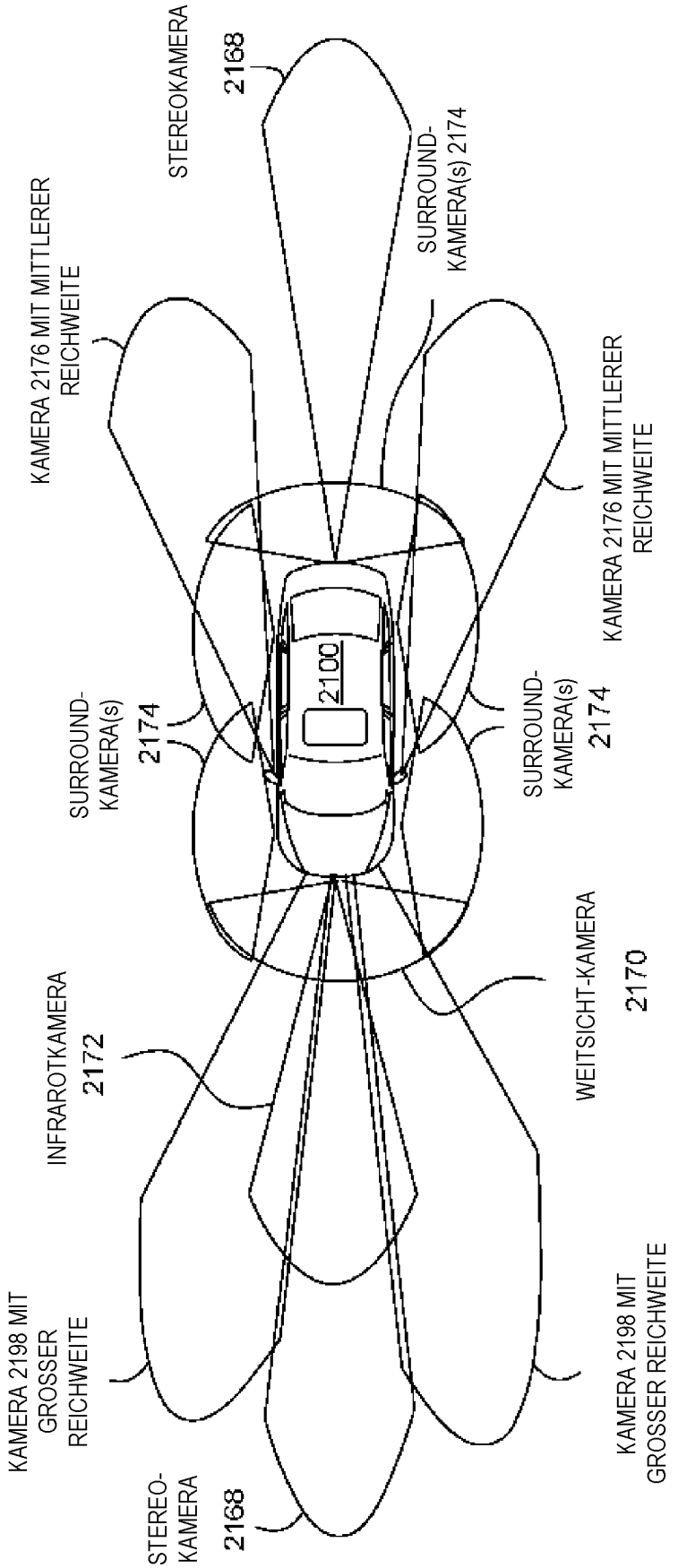

STEREOKAMERA
2168
KAMERA 2176 MIT MITTLERER REICHWEITE
SURROUND-KAMERA(s) 2174
2100
SURROUND-KAMERA(s) 2174
KAMERA 2176 MIT MITTLERER REICHWEITE
INFRAROTKAMERA 2172
WEITSICHT-KAMERA 2170
KAMERA 2198 MIT GROSSER REICHWEITE
STEREO-KAMERA 2168
KAMERA 2198 MIT GROSSER REICHWEITE

FIG. 21B

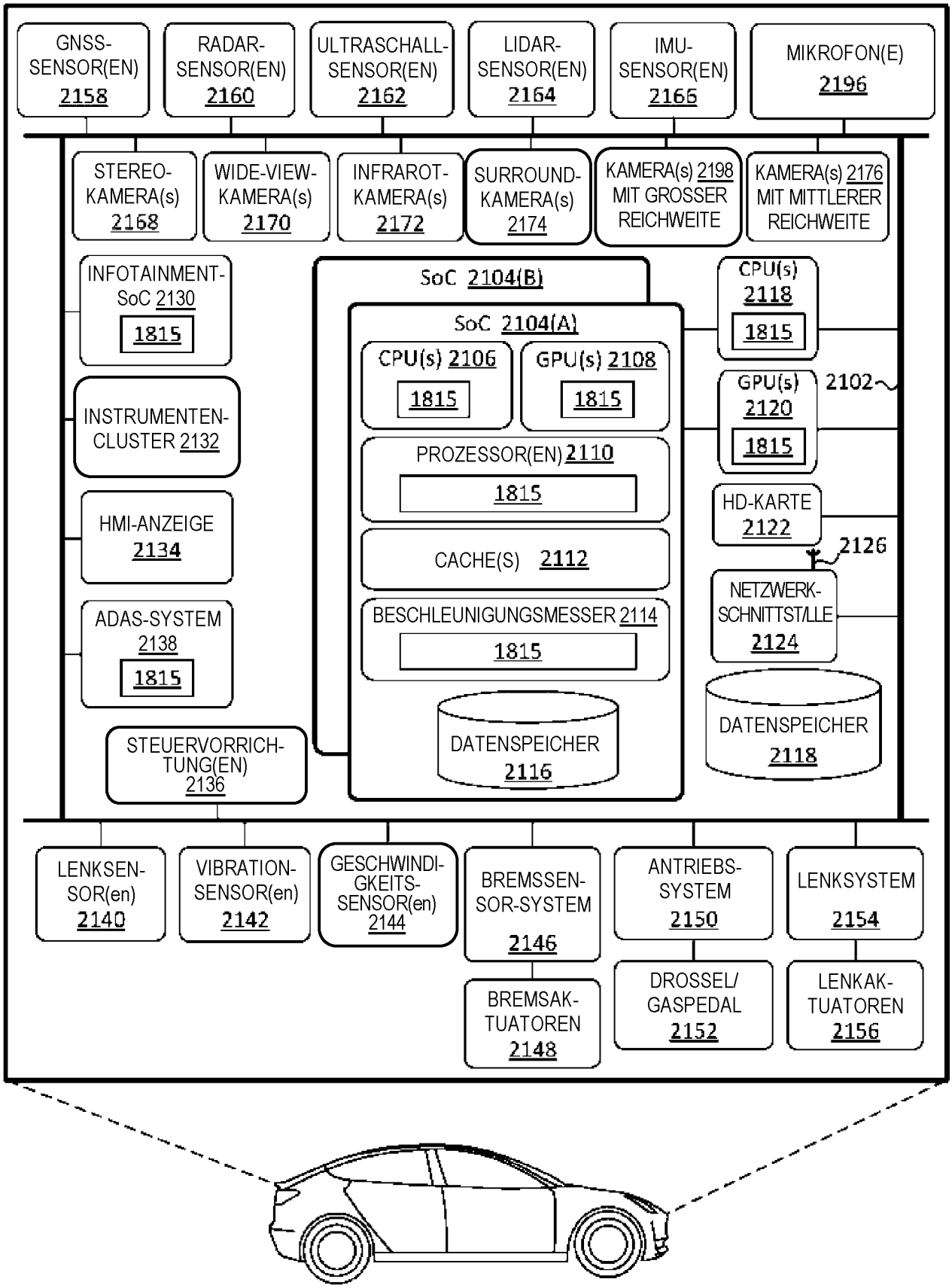

GNSS-SENSOR(EN) 2158
RADAR-SENSOR(EN) 2160
ULTRASCHALL-SENSOR(EN) 2162
LIDAR-SENSOR(EN) 2164
IMU-SENSOR(EN) 2166
MIKROFON(E) 2196
STEREO-KAMERA(s) 2168
WIDE-VIEW-KAMERA(s) 2170
INFRAROT-KAMERA(s) 2172
SURROUND-KAMERA(s) 2174
KAMERA(s) 2198 MIT GROSSER REICHWEITE
KAMERA(s) 2176 MIT MITTLERER REICHWEITE
INFOTAINMENT-SoC 2130
1815
INSTRUMENTEN-CLUSTER 2132
HMI-ANZEIGE 2134
ADAS-SYSTEM 2138
1815
STEUERVORRICH-TUNG(EN) 2136
SoC 2104(B)
SoC 2104(A)
CPU(s) 2106
1815
GPU(s) 2108
1815
PROZESSOR(EN) 2110
1815
CACHE(S) 2112
BESCHLEUNIGUNGSMESSER 2114
1815
DATENSPEICHER 2116
CPU(s) 2118
1815
GPU(s) 2120
1815
2102
HD-KARTE 2122
2126
NETZWERK-SCHNITTST/LLE 2124
DATENSPEICHER 2118
LENKSEN-SOR(en) 2140
VIBRATION-SENSOR(en) 2142
GESCHWINDI-GKEITS-SENSOR(en) 2144
BREMSSEN-SOR-SYSTEM 2146
BREMSAK-TUATOREN 2148
ANTRIEBS-SYSTEM 2150
DROSSEL/ GASPEDAL 2152
LENKSYSTEM 2154
LENKAK-TUATOREN 2156

FIG. 21C

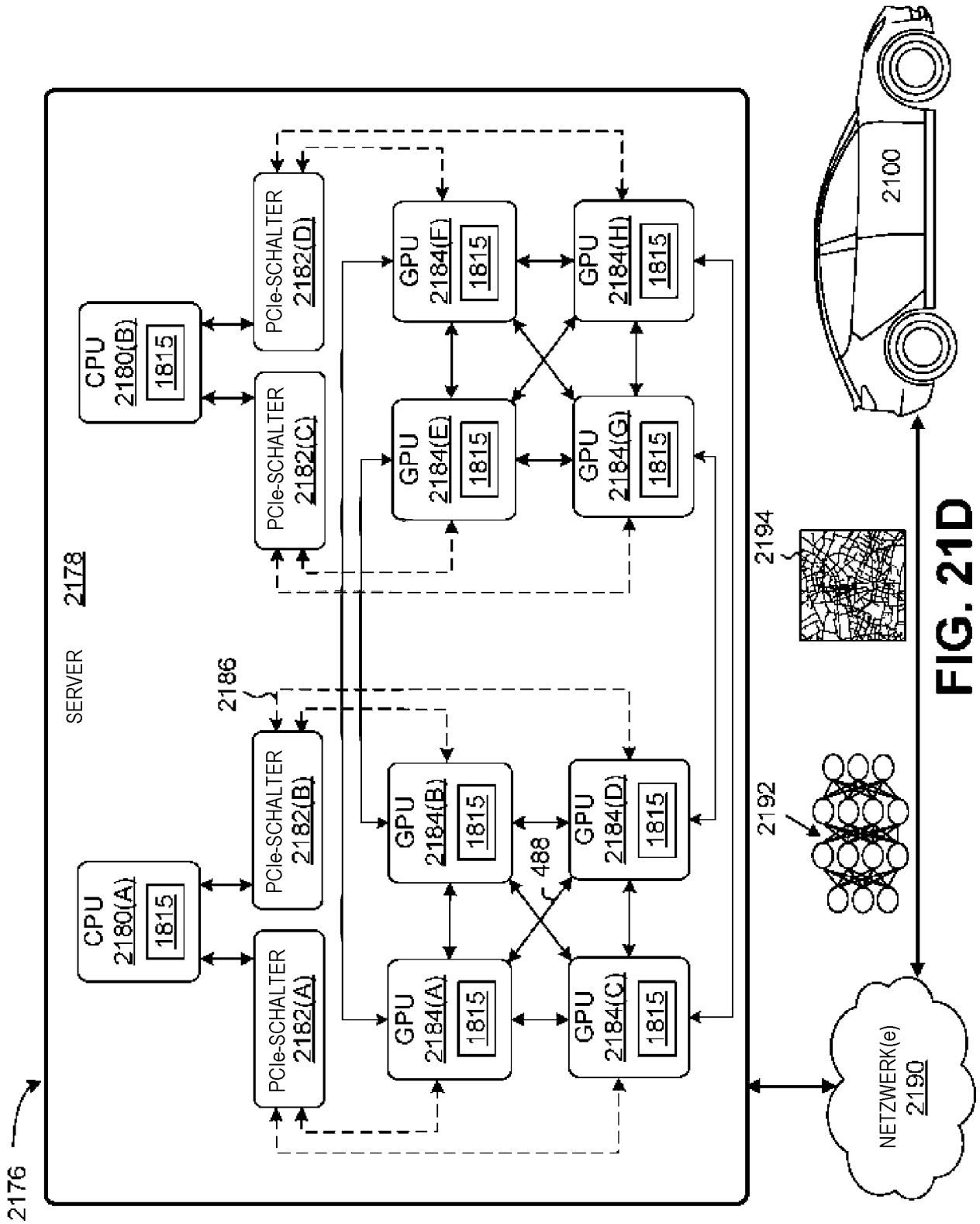

2176
SERVER 2178
CPU 2180(A) 1815
CPU 2180(B) 1815
PCIe-SCHALTER 2182(A)
PCIe-SCHALTER 2182(B)
PCIe-SCHALTER 2182(C)
PCIe-SCHALTER 2182(D)
2186
GPU 2184(A) 1815
GPU 2184(B) 1815
GPU 2184(C) 1815
GPU 2184(D) 1815
GPU 2184(E) 1815
GPU 2184(F) 1815
GPU 2184(G) 1815
GPU 2184(H) 1815
488
2192
2194
2100
NETZWERK(e) 2190

FIG. 21D

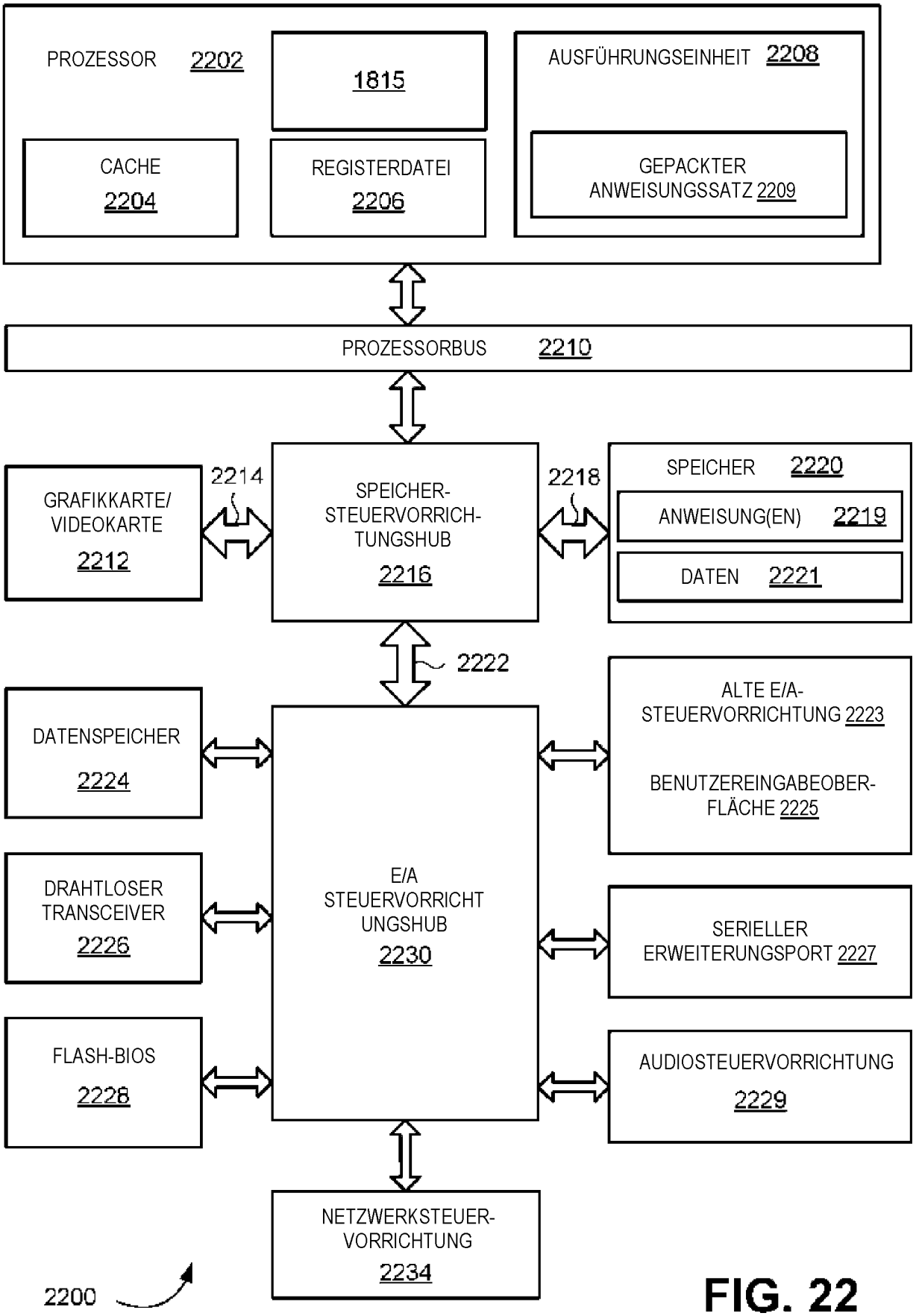

PROZESSOR 2202
1815
AUSFÜHRUNGSEINHEIT 2208
CACHE 2204
REGISTERDATEI 2206
GEPACKTER ANWEISUNGSSATZ 2209
PROZESSORBUS 2210
GRAFIKKARTE/ VIDEOKARTE 2212
2214
SPEICHER-STEUERVORRICH-TUNGSHUB 2216
2218
SPEICHER 2220
ANWEISUNG(EN) 2219
DATEN 2221
2222
DATENSPEICHER 2224
ALTE E/A-STEUERVORRICHTUNG 2223
BENUTZEREINGABEOBER-FLÄCHE 2225
DRAHTLOSER TRANSCEIVER 2226
E/A STEUERVORRICHTUNGSHUB 2230
SERIELLER ERWEITERUNGSPORT 2227
FLASH-BIOS 2228
AUDIOSTEUERVORRICHTUNG 2229
NETZWERKSTEUER-VORRICHTUNG 2234
2200

FIG. 22

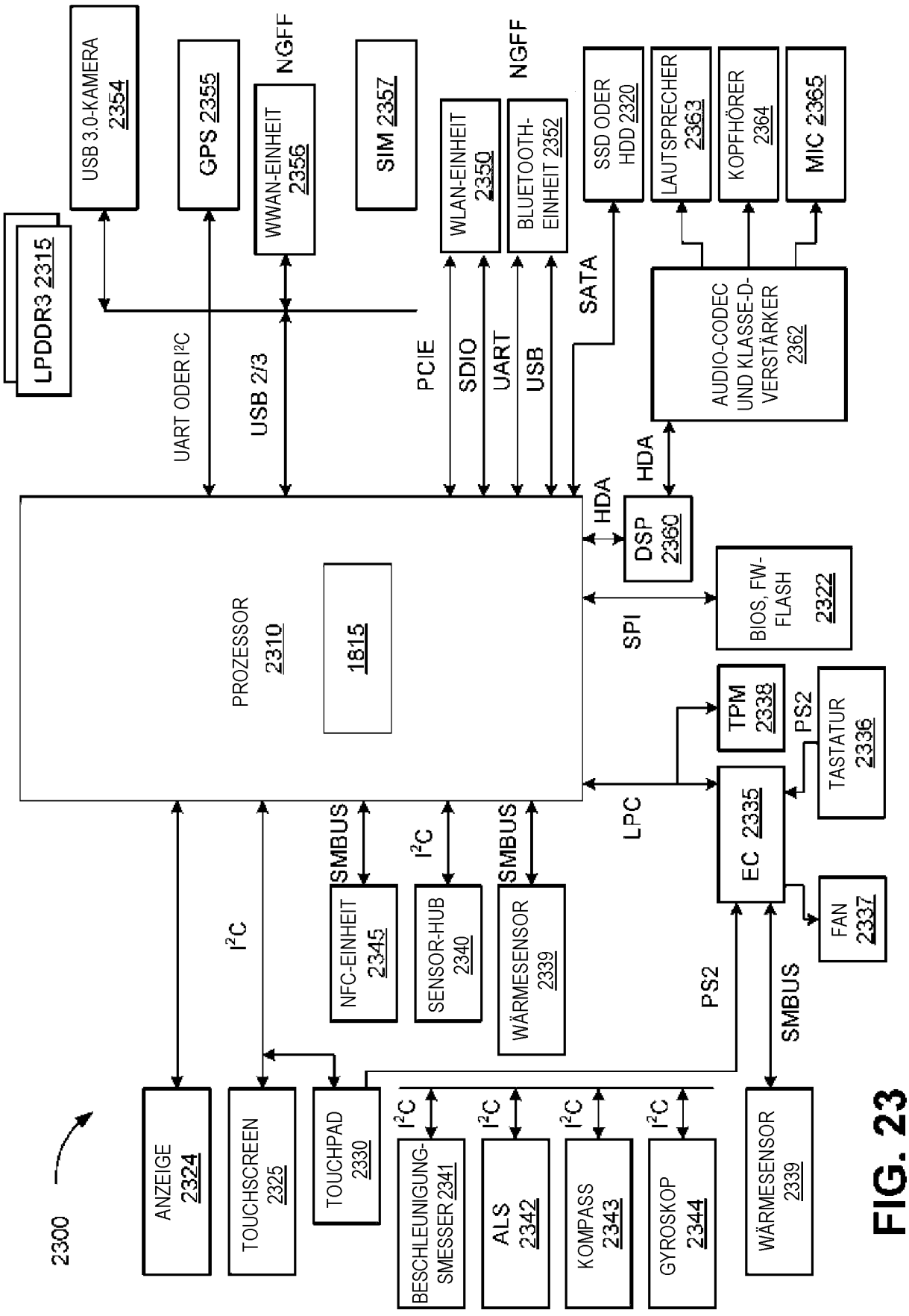

2300
PROZESSOR 2310
1815
LPDDR3 2315
USB 3.0-KAMERA 2354
GPS 2355
NGFF
WWAN-EINHEIT 2356
SIM 2357
NGFF
WLAN-EINHEIT 2350
BLUETOOTH-EINHEIT 2352
SSD ODER HDD 2320
LAUTSPRECHER 2363
KOPFHÖRER 2364
MIC 2365
AUDIO-CODEC UND KLASSE-D-VERSTÄRKER 2362
UART ODER I²C
USB 2/3
PCIE
SDIO
UART
USB
SATA
HDA
HDA
DSP 2360
SPI
BIOS, FW-FLASH 2322
TPM 2338
PS2
TASTATUR 2336
LPC
EC 2335
FAN 2337
SMBUS
I²C
SMBUS
I²C
SMBUS
NFC-EINHEIT 2345
SENSOR-HUB 2340
WÄRMESENSOR 2339
PS2
SMBUS
ANZEIGE 2324
TOUCHSCREEN 2325
TOUCHPAD 2330
I²C
I²C
I²C
I²C
BESCHLEUNIGUNGSMESSER 2341
ALS 2342
KOMPASS 2343
GYROSKOP 2344
WÄRMESENSOR 2339

FIG. 23

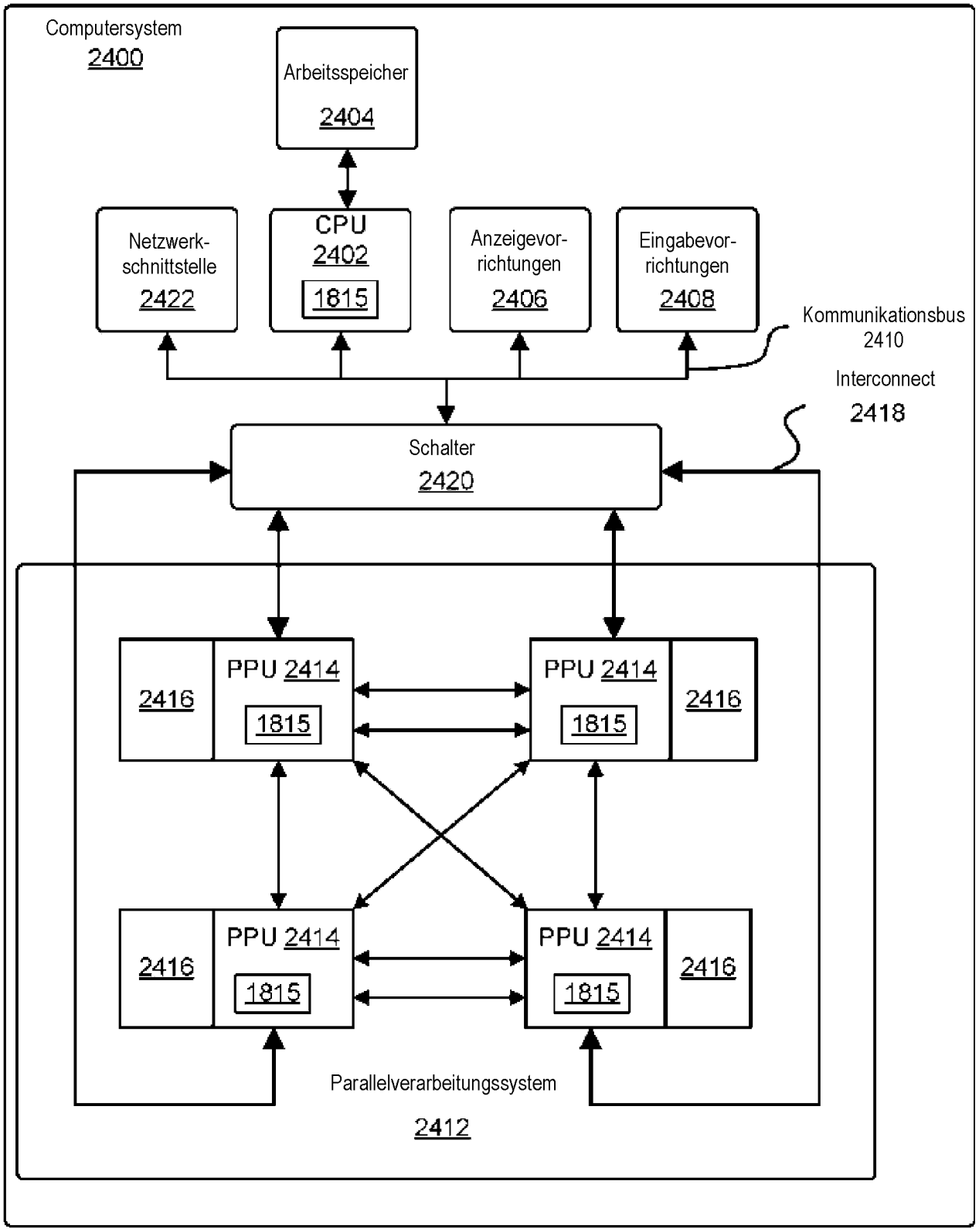

Computersystem
2400
Arbeitsspeicher
2404
Netzwerk-
schnittstelle
2422
CPU
2402
1815
Anzeigevor-
richtungen
2406
Eingabevor-
richtungen
2408
Kommunikationsbus
2410
Interconnect
2418
Schalter
2420
2416
PPU 2414
1815
PPU 2414
1815
2416
2416
PPU 2414
1815
PPU 2414
1815
2416
Parallelverarbeitungssystem
2412

FIG. 24

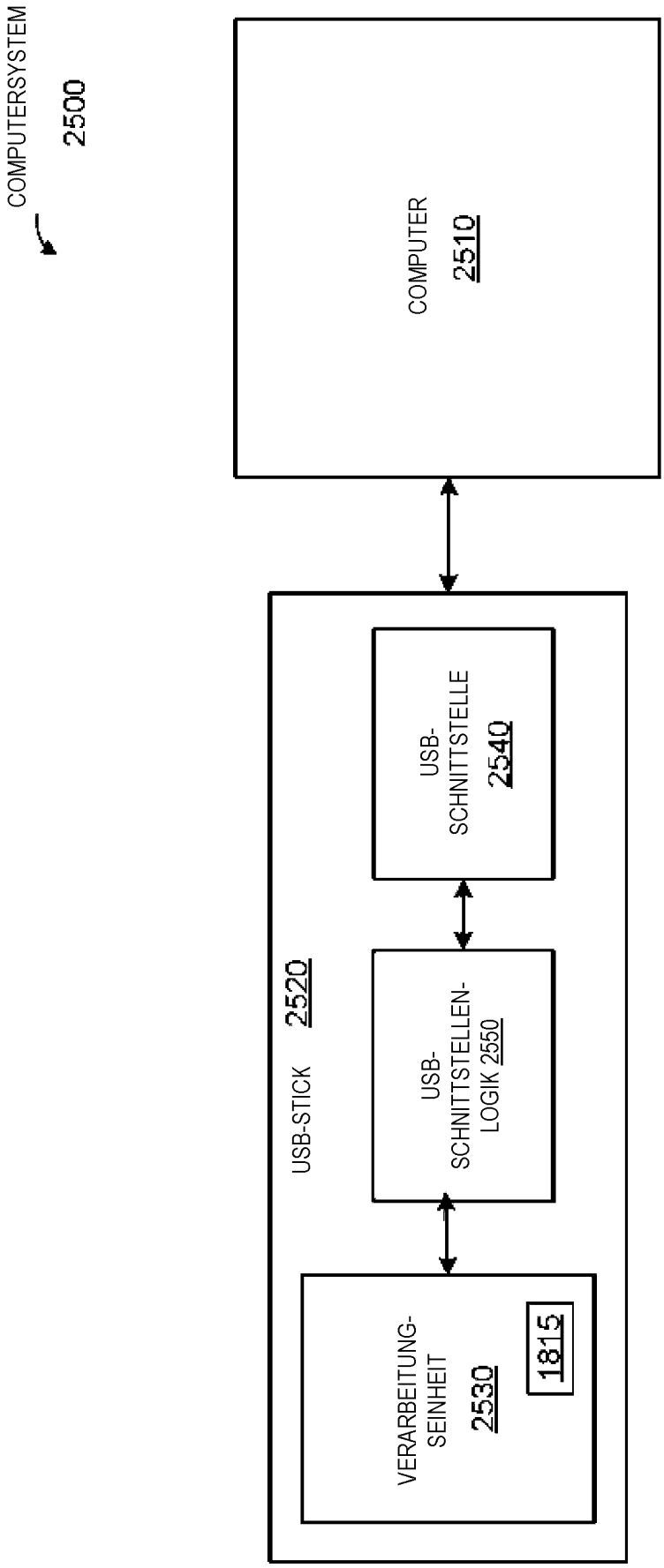

COMPUTERSYSTEM
2500
COMPUTER
2510
USB-STICK
2520
USB-SCHNITTSTELLE
2540
USB-SCHNITTSTELLEN-LOGIK
2550
VERARBEITUNG-SEINHEIT
2530
1815

FIG. 25

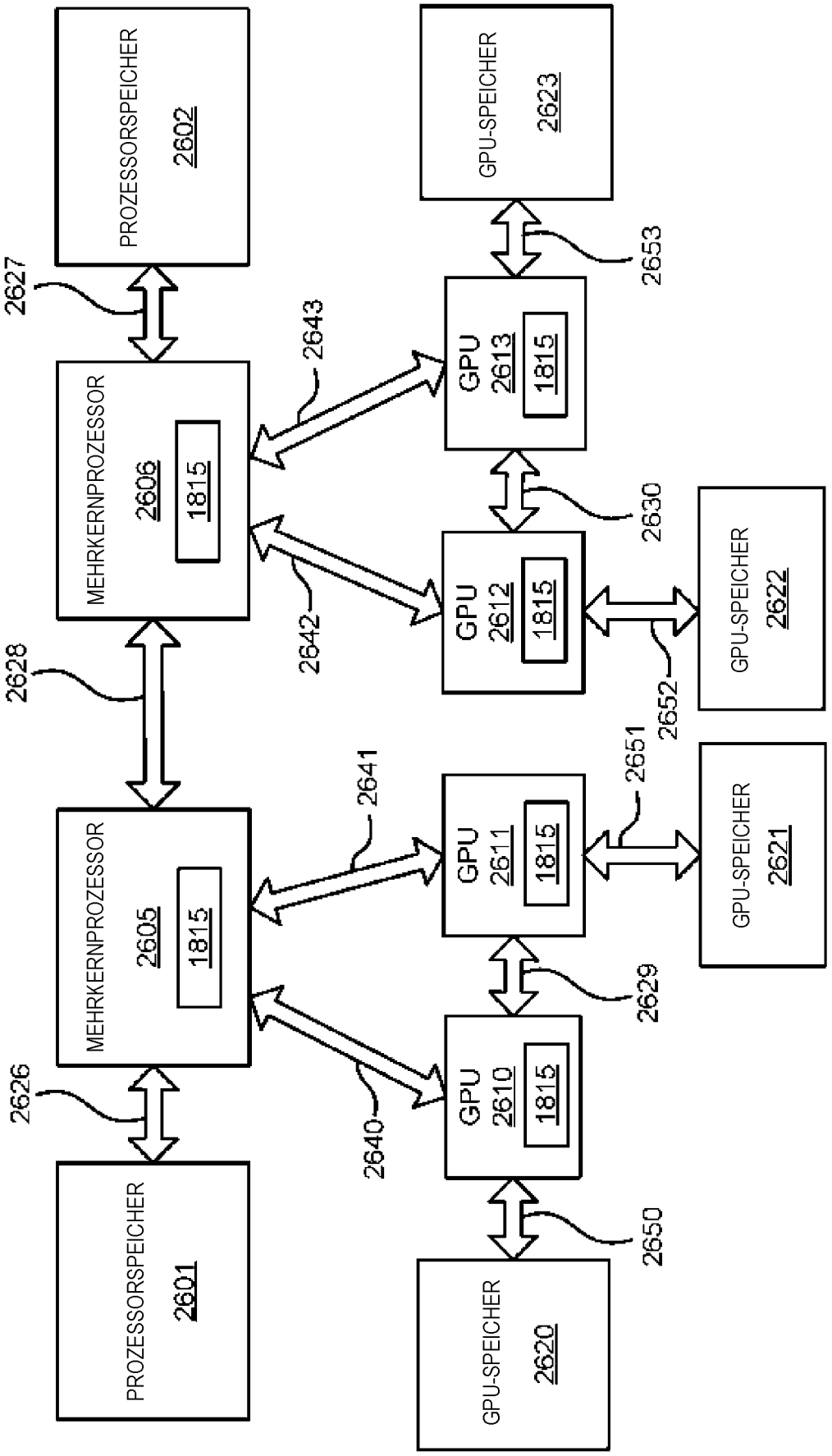

PROZESSORSPEICHER
2602
GPU-SPEICHER
2623
2627
2653
2643
MEHRKERNPROZESSOR
2606
1815
GPU
2613
1815
2630
2628
2642
GPU
2612
1815
GPU-SPEICHER
2622
2652
2651
2641
MEHRKERNPROZESSOR
2605
1815
GPU
2611
1815
GPU-SPEICHER
2621
2629
2626
2640
GPU
2610
1815
PROZESSORSPEICHER
2601
2650
GPU-SPEICHER
2620

FIG. 26A

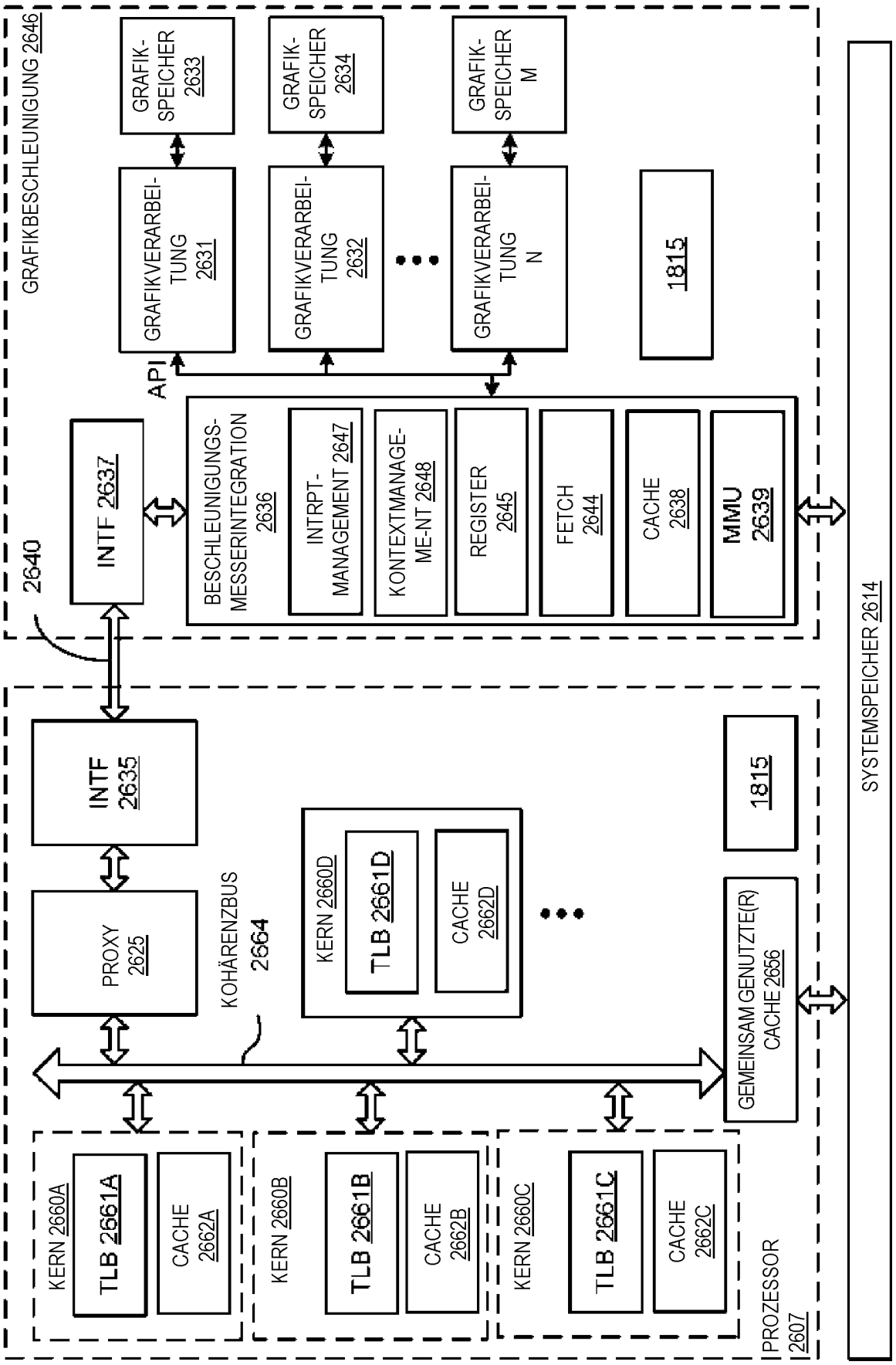

GRAFIKBESCHLEUNIGUNG 2646
GRAFIK-SPEICHER 2633
GRAFIK-SPEICHER 2634
GRAFIK-SPEICHER M
GRAFIKVERARBEI-TUNG 2631
GRAFIKVERARBEI-TUNG 2632
GRAFIKVERARBEI-TUNG N
1815
API
2640
INTF 2637
BESCHLEUNIGUNGS-MESSERINTEGRATION 2636
INTRPT-MANAGEMENT 2647
KONTEXTMANAGE-ME-NT 2648
REGISTER 2645
FETCH 2644
CACHE 2638
MMU 2639
INTF 2635
PROXY 2625
KOHÄRENZBUS 2664
KERN 2660D
TLB 2661D
CACHE 2662D
1815
GEMEINSAM GENUTZTE(R) CACHE 2656
SYSTEMSPEICHER 2614
KERN 2660A
TLB 2661A
CACHE 2662A
KERN 2660B
TLB 2661B
CACHE 2662B
KERN 2660C
TLB 2661C
CACHE 2662C
PROZESSOR 2607

FIG. 26B

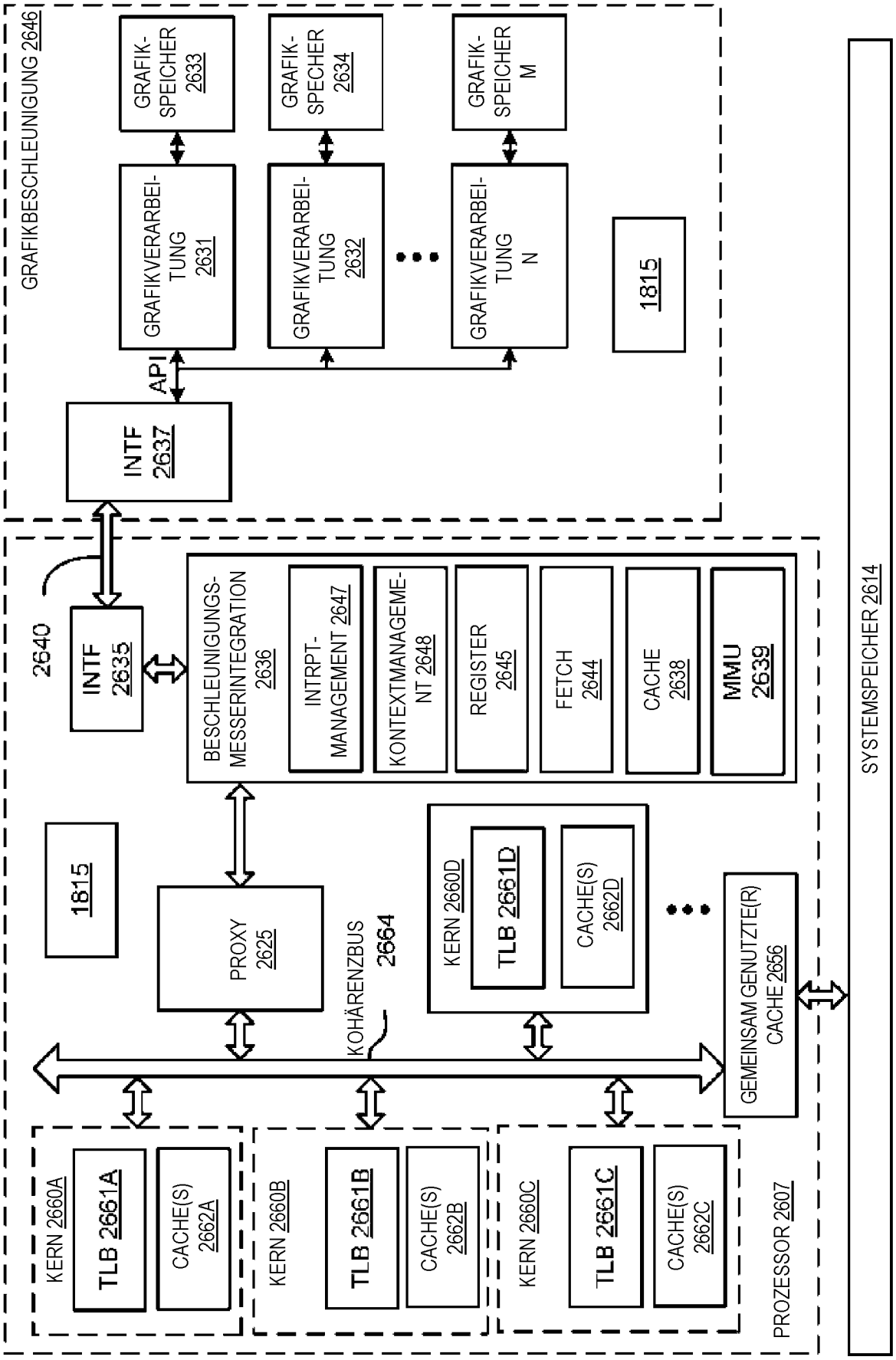

GRAFIKBESCHLEUNIGUNG 2646
GRAFIK-SPEICHER 2633
GRAFIK-SPECHER 2634
GRAFIK-SPEICHER M
GRAFIKVERARBEI-TUNG 2631
GRAFIKVERARBEI-TUNG 2632
GRAFIKVERARBEI-TUNG N
1815
API
INTF 2637
2640
INTF 2635
BESCHLEUNIGUNGS-MESSERINTEGRATION 2636
INTRPT-MANAGEMENT 2647
KONTEXTMANAGEME-NT 2648
REGISTER 2645
FETCH 2644
CACHE 2638
MMU 2639
1815
PROXY 2625
KOHÄRENZBUS 2664
KERN 2660D
TLB 2661D
CACHE(S) 2662D
GEMEINSAM GENUTZTE(R) CACHE 2656
SYSTEMSPEICHER 2614
KERN 2660A
TLB 2661A
CACHE(S) 2662A
KERN 2660B
TLB 2661B
CACHE(S) 2662B
KERN 2660C
TLB 2661C
CACHE(S) 2662C
PROZESSOR 2607

FIG. 26C

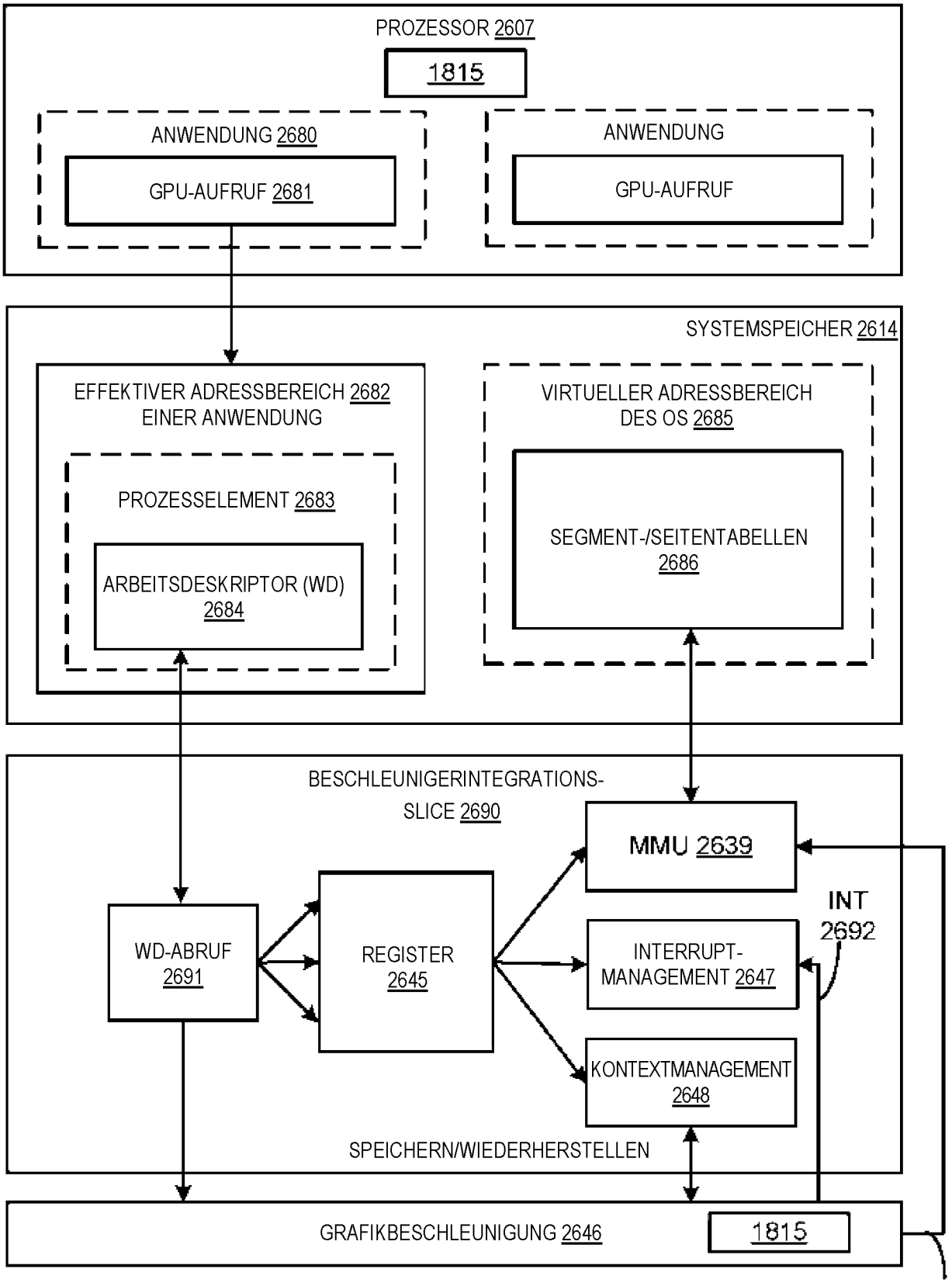

PROZESSOR 2607
1815
ANWENDUNG 2680
GPU-AUFRUF 2681
ANWENDUNG
GPU-AUFRUF
SYSTEMSPEICHER 2614
EFFEKTIVER ADRESSBEREICH 2682 EINER ANWENDUNG
PROZESSELEMENT 2683
ARBEITSDESKRIPTOR (WD) 2684
VIRTUELLER ADRESSBEREICH DES OS 2685
SEGMENT-/SEITENTABELLEN 2686
BESCHLEUNIGERINTEGRATIONS-SLICE 2690
MMU 2639
INT 2692
WD-ABRUF 2691
REGISTER 2645
INTERRUPT-MANAGEMENT 2647
KONTEXTMANAGEMENT 2648
SPEICHERN/WIEDERHERSTELLEN
GRAFIKBESCHLEUNIGUNG 2646
1815
EFFEKTIVE ADRESSE 2693

FIG. 26D

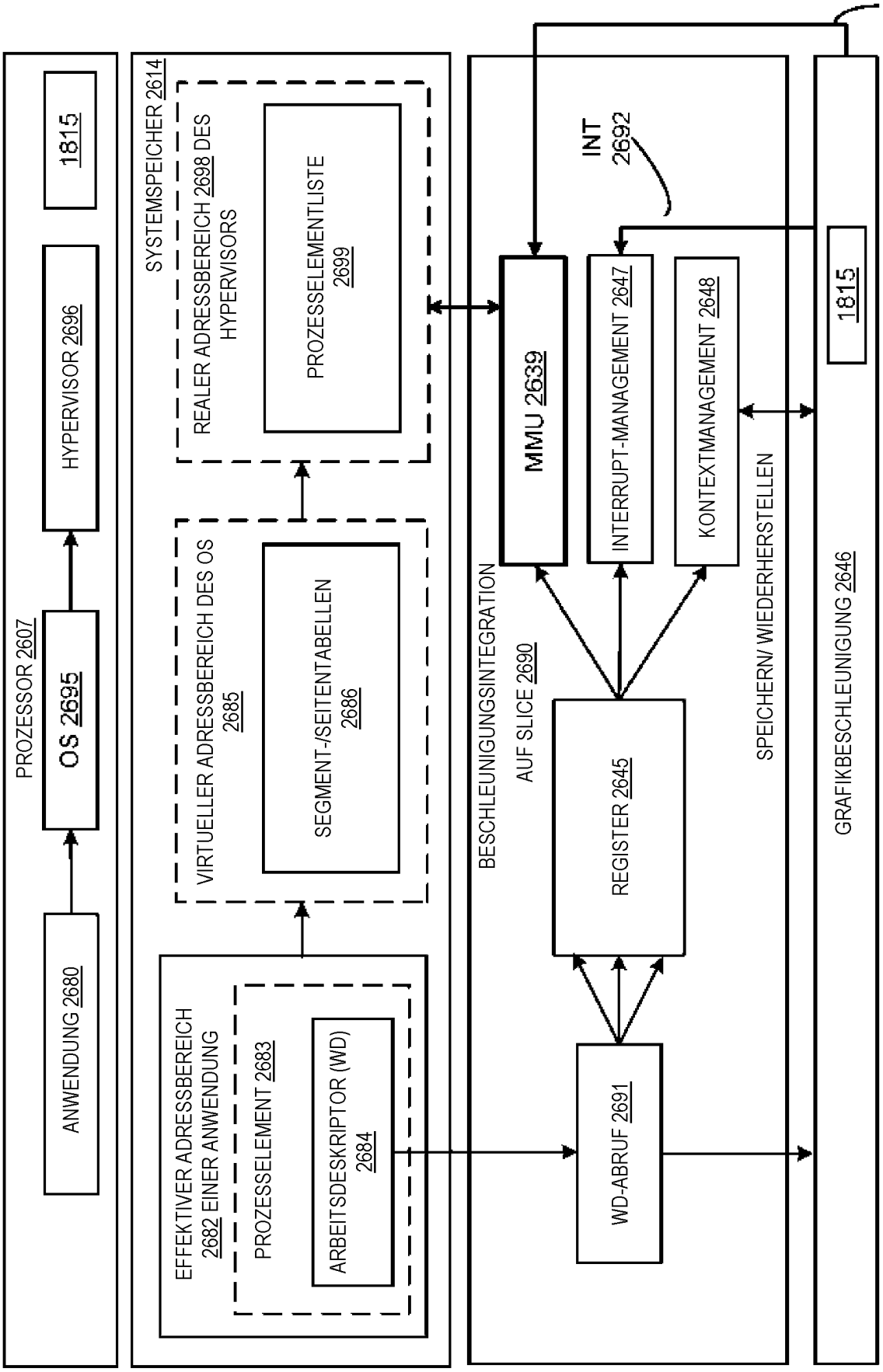

PROZESSOR 2607
ANWENDUNG 2680
OS 2695
HYPERVISOR 2696
1815
SYSTEMSPEICHER 2614
REALER ADRESSBEREICH 2698 DES HYPERVISORS
PROZESSELEMENTLISTE 2699
VIRTUELLER ADRESSBEREICH DES OS 2685
SEGMENT-/SEITENTABELLEN 2686
EFFEKTIVER ADRESSBEREICH 2682 EINER ANWENDUNG
PROZESSELEMENT 2683
ARBEITSDESKRIPTOR (WD) 2684
BESCHLEUNIGUNGSINTEGRATION AUF SLICE 2690
MMU 2639
INT 2692
INTERRUPT-MANAGEMENT 2647
KONTEXTMANAGEMENT 2648
REGISTER 2645
WD-ABRUF 2691
SPEICHERN/ WIEDERHERSTELLEN
EFFEKTIVE ADRESSE 2693
1815
GRAFIKBESCHLEUNIGUNG 2646

FIG. 26E

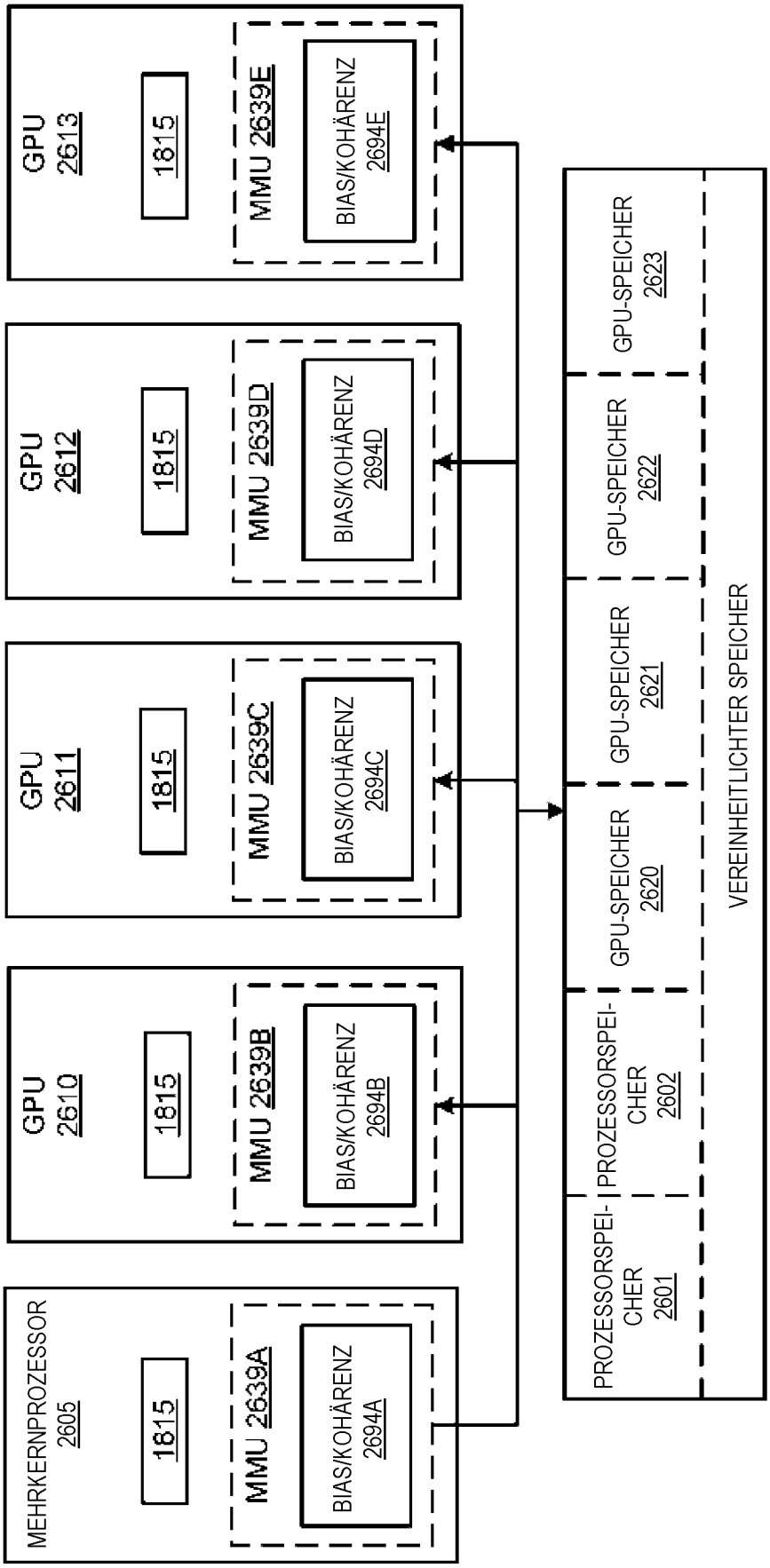

MEHRKERNPROZESSOR 2605
1815
MMU 2639A
BIAS/KOHÄRENZ 2694A
GPU 2610
1815
MMU 2639B
BIAS/KOHÄRENZ 2694B
GPU 2611
1815
MMU 2639C
BIAS/KOHÄRENZ 2694C
GPU 2612
1815
MMU 2639D
BIAS/KOHÄRENZ 2694D
GPU 2613
1815
MMU 2639E
BIAS/KOHÄRENZ 2694E
PROZESSORSPEI-CHER 2601
PROZESSORSPEI-CHER 2602
GPU-SPEICHER 2620
GPU-SPEICHER 2621
GPU-SPEICHER 2622
GPU-SPEICHER 2623
VEREINHEITLICHTER SPEICHER

FIG. 26F

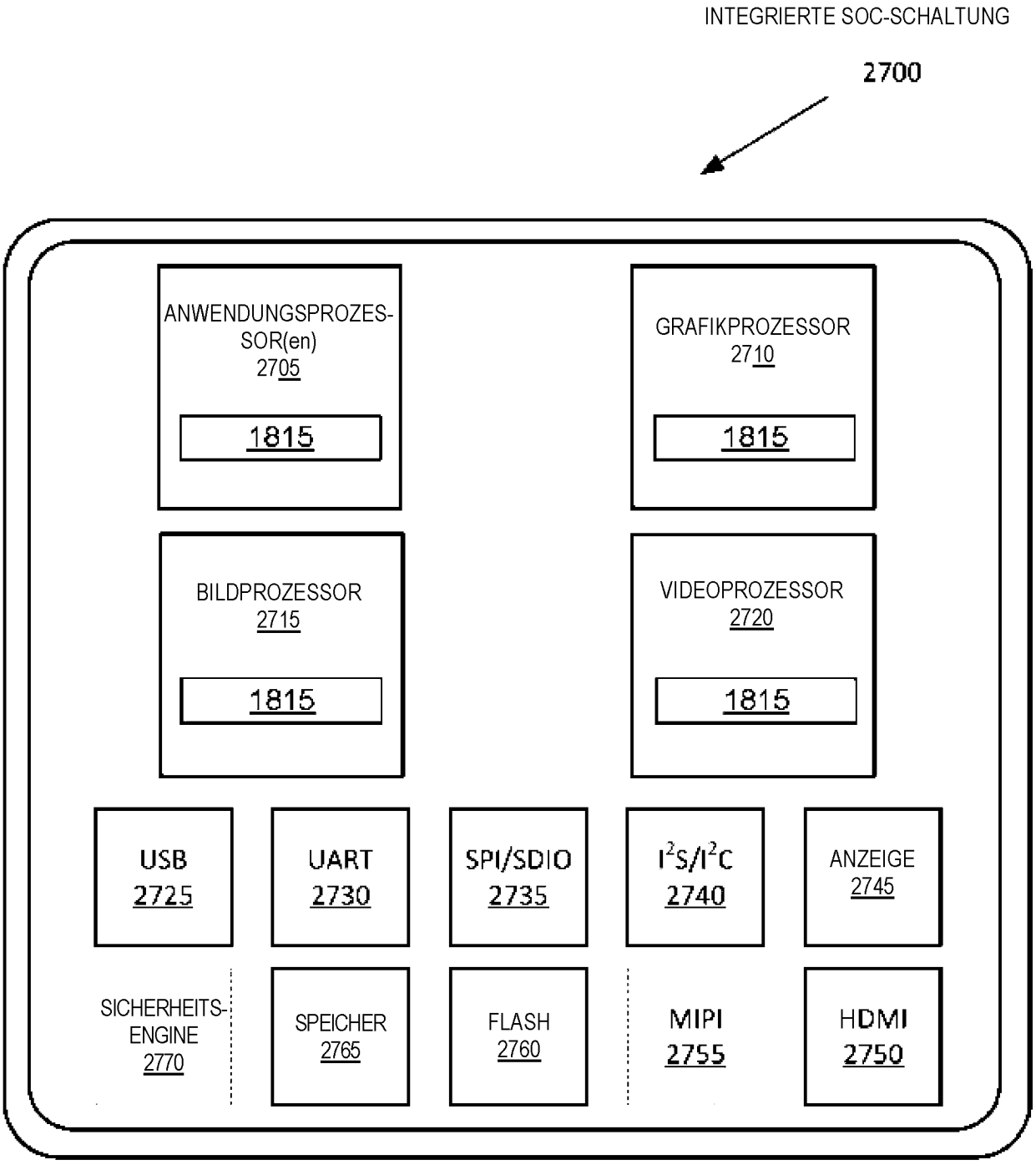

INTEGRIERTE SOC-SCHALTUNG
2700
ANWENDUNGSPROZES-SOR(en) 2705
1815
GRAFIKPROZESSOR 2710
1815
BILDPROZESSOR 2715
1815
VIDEOPROZESSOR 2720
1815
USB 2725
UART 2730
SPI/SDIO 2735
I2S/I2C 2740
ANZEIGE 2745
SICHERHEITS-ENGINE 2770
SPEICHER 2765
FLASH 2760
MIPI 2755
HDMI 2750

FIG. 27

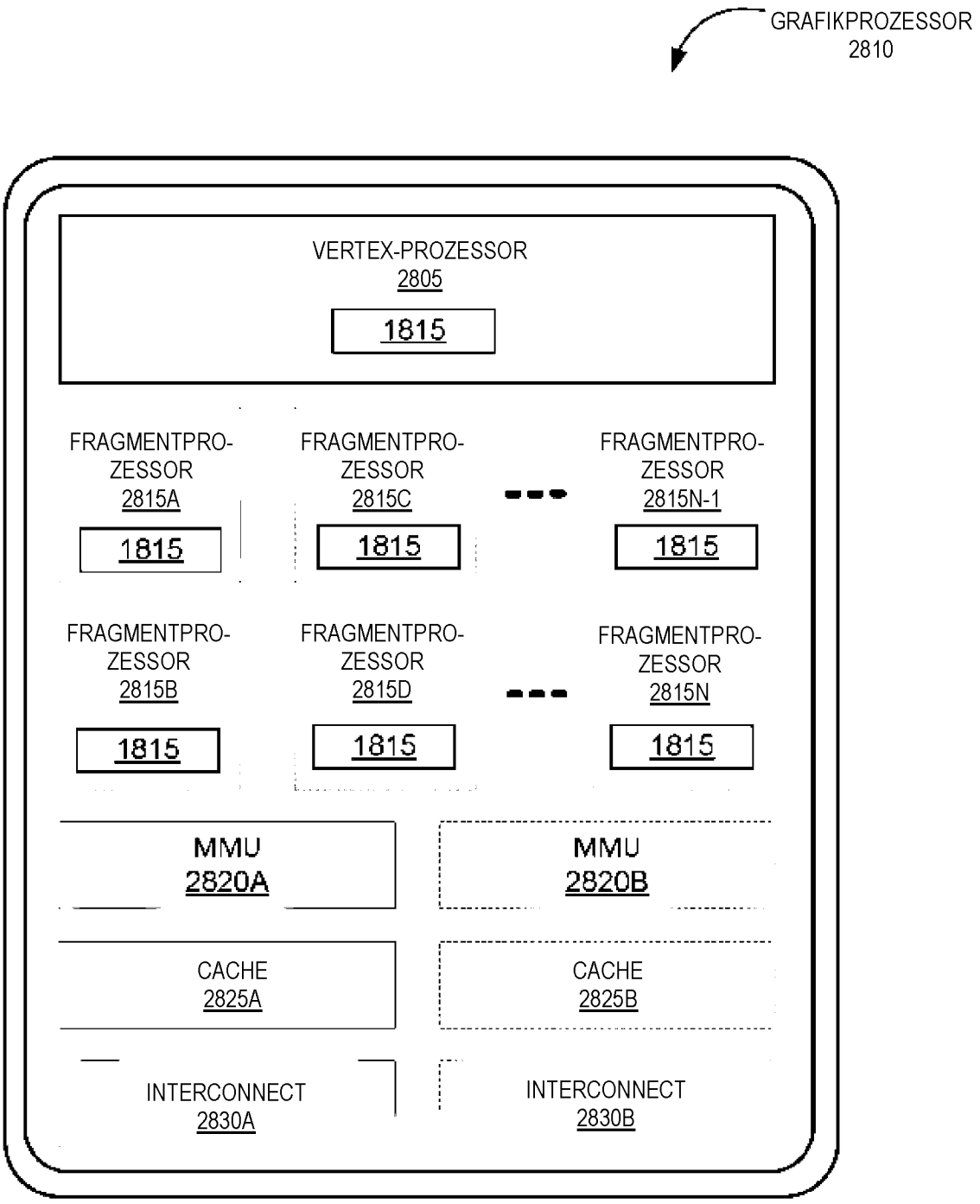

GRAFIKPROZESSOR
2810
VERTEX-PROZESSOR
2805
1815
FRAGMENTPRO-
ZESSOR
2815A
1815
FRAGMENTPRO-
ZESSOR
2815C
1815
FRAGMENTPRO-
ZESSOR
2815N-1
1815
FRAGMENTPRO-
ZESSOR
2815B
1815
FRAGMENTPRO-
ZESSOR
2815D
1815
FRAGMENTPRO-
ZESSOR
2815N
1815
MMU
2820A
MMU
2820B
CACHE
2825A
CACHE
2825B
INTERCONNECT
2830A
INTERCONNECT
2830B

FIG. 28A

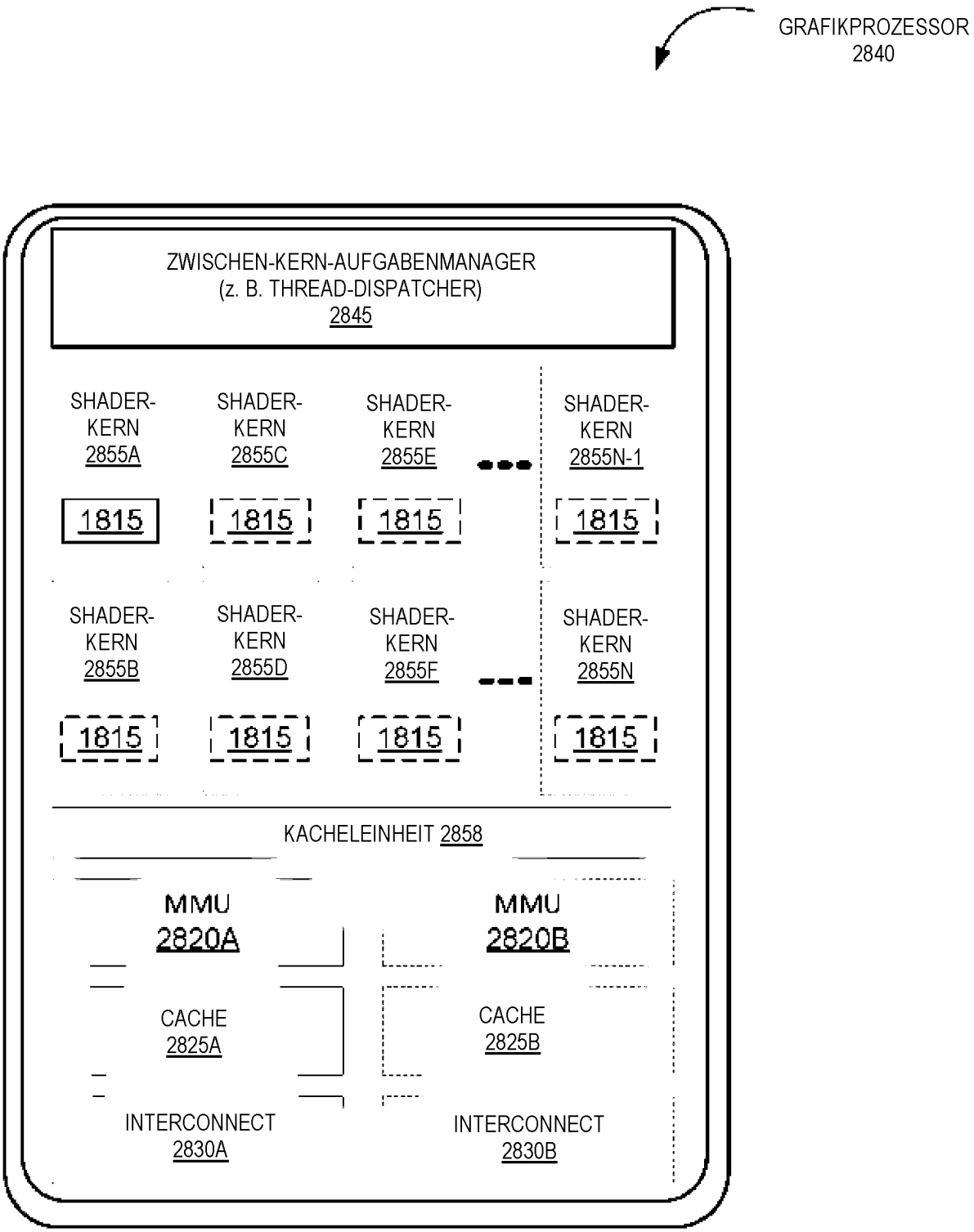

GRAFIKPROZESSOR
2840
ZWISCHEN-KERN-AUFGABENMANAGER
(z. B. THREAD-DISPATCHER)
2845
SHADER-KERN 2855A
SHADER-KERN 2855C
SHADER-KERN 2855E
SHADER-KERN 2855N-1
1815
SHADER-KERN 2855B
SHADER-KERN 2855D
SHADER-KERN 2855F
SHADER-KERN 2855N
KACHELEINHEIT 2858
MMU 2820A
MMU 2820B
CACHE 2825A
CACHE 2825B
INTERCONNECT 2830A
INTERCONNECT 2830B

FIG. 28B

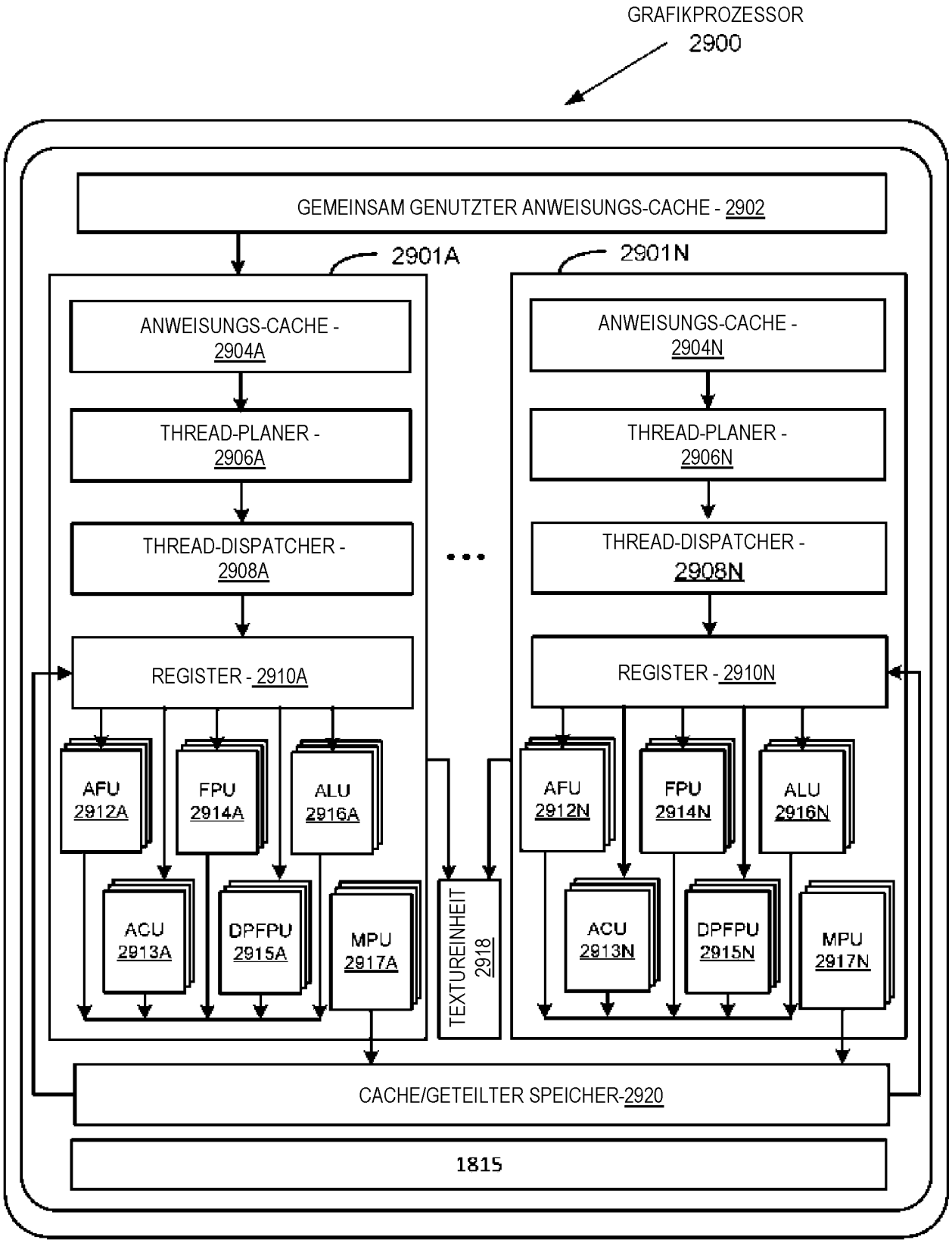

GRAFIKPROZESSOR
2900
GEMEINSAM GENUTZTER ANWEISUNGS-CACHE - 2902
2901A
2901N
ANWEISUNGS-CACHE - 2904A
ANWEISUNGS-CACHE - 2904N
THREAD-PLANER - 2906A
THREAD-PLANER - 2906N
THREAD-DISPATCHER - 2908A
THREAD-DISPATCHER - 2908N
REGISTER - 2910A
REGISTER - 2910N
AFU 2912A
FPU 2914A
ALU 2916A
ACU 2913A
DPFPU 2915A
MPU 2917A
TEXTUREINHEIT 2918
AFU 2912N
FPU 2914N
ALU 2916N
ACU 2913N
DPFPU 2915N
MPU 2917N
CACHE/GETEILTER SPEICHER-2920
1815

FIG. 29A

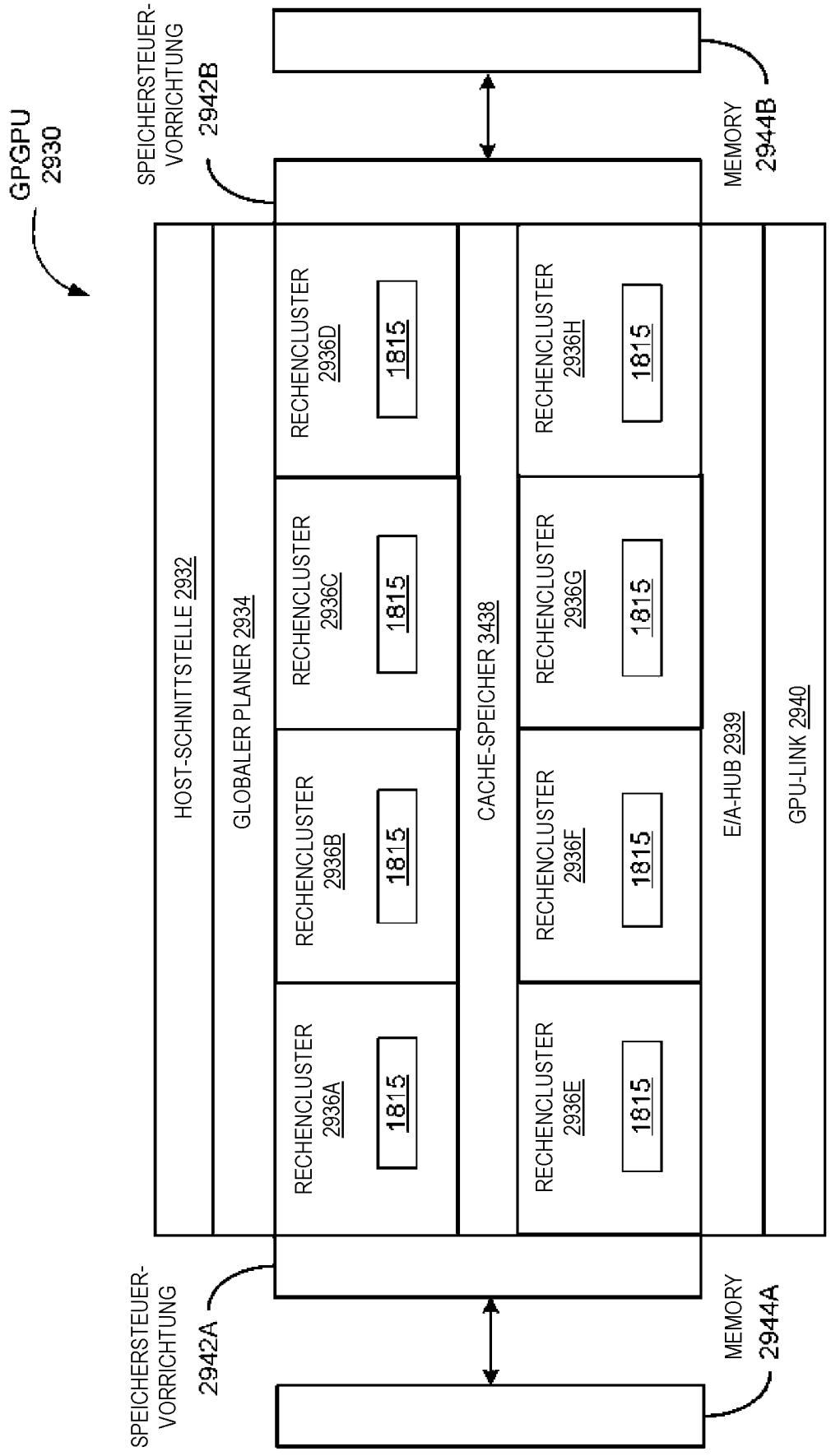

GPGPU
2930
SPEICHERSTEUER-VORRICHTUNG
2942B
MEMORY
2944B
HOST-SCHNITTSTELLE 2932
GLOBALER PLANER 2934
RECHENCLUSTER
2936D
1815
RECHENCLUSTER
2936C
1815
RECHENCLUSTER
2936B
1815
RECHENCLUSTER
2936A
1815
CACHE-SPEICHER 3438
RECHENCLUSTER
2936H
1815
RECHENCLUSTER
2936G
1815
RECHENCLUSTER
2936F
1815
RECHENCLUSTER
2936E
1815
E/A-HUB 2939
GPU-LINK 2940
SPEICHERSTEUER-VORRICHTUNG
2942A
MEMORY
2944A

FIG. 29B

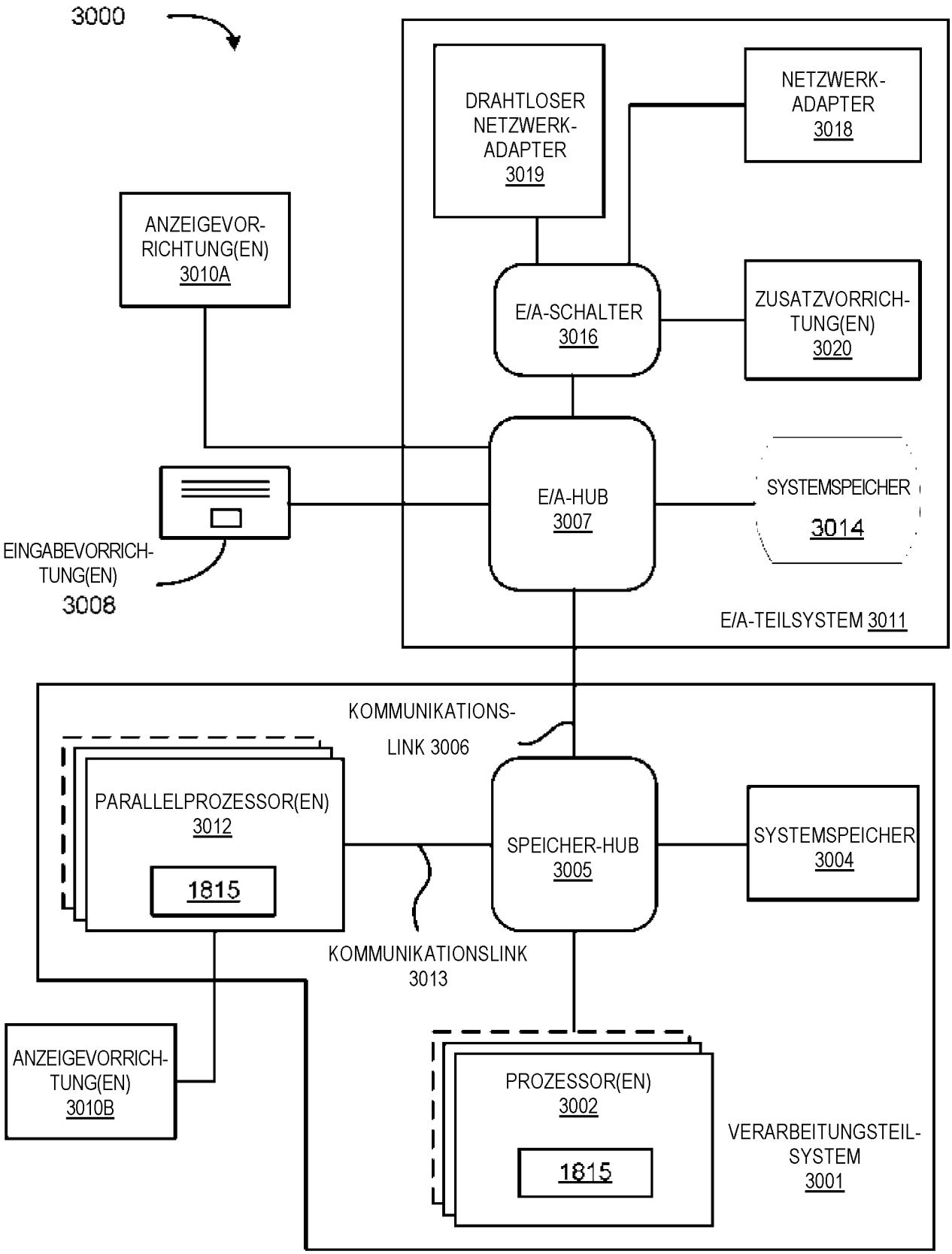

3000
DRAHTLOSER NETZWERK-ADAPTER 3019
NETZWERK-ADAPTER 3018
ANZEIGEVOR-RICHTUNG(EN) 3010A
E/A-SCHALTER 3016
ZUSATZVORRICH-TUNG(EN) 3020
E/A-HUB 3007
SYSTEMSPEICHER 3014
EINGABEVORRICH-TUNG(EN) 3008
E/A-TEILSYSTEM 3011
KOMMUNIKATIONS-LINK 3006
PARALLELPROZESSOR(EN) 3012
1815
SPEICHER-HUB 3005
SYSTEMSPEICHER 3004
KOMMUNIKATIONSLINK 3013
ANZEIGEVORRICH-TUNG(EN) 3010B
PROZESSOR(EN) 3002
1815
VERARBEITUNGSTEIL-SYSTEM 3001

FIG. 30

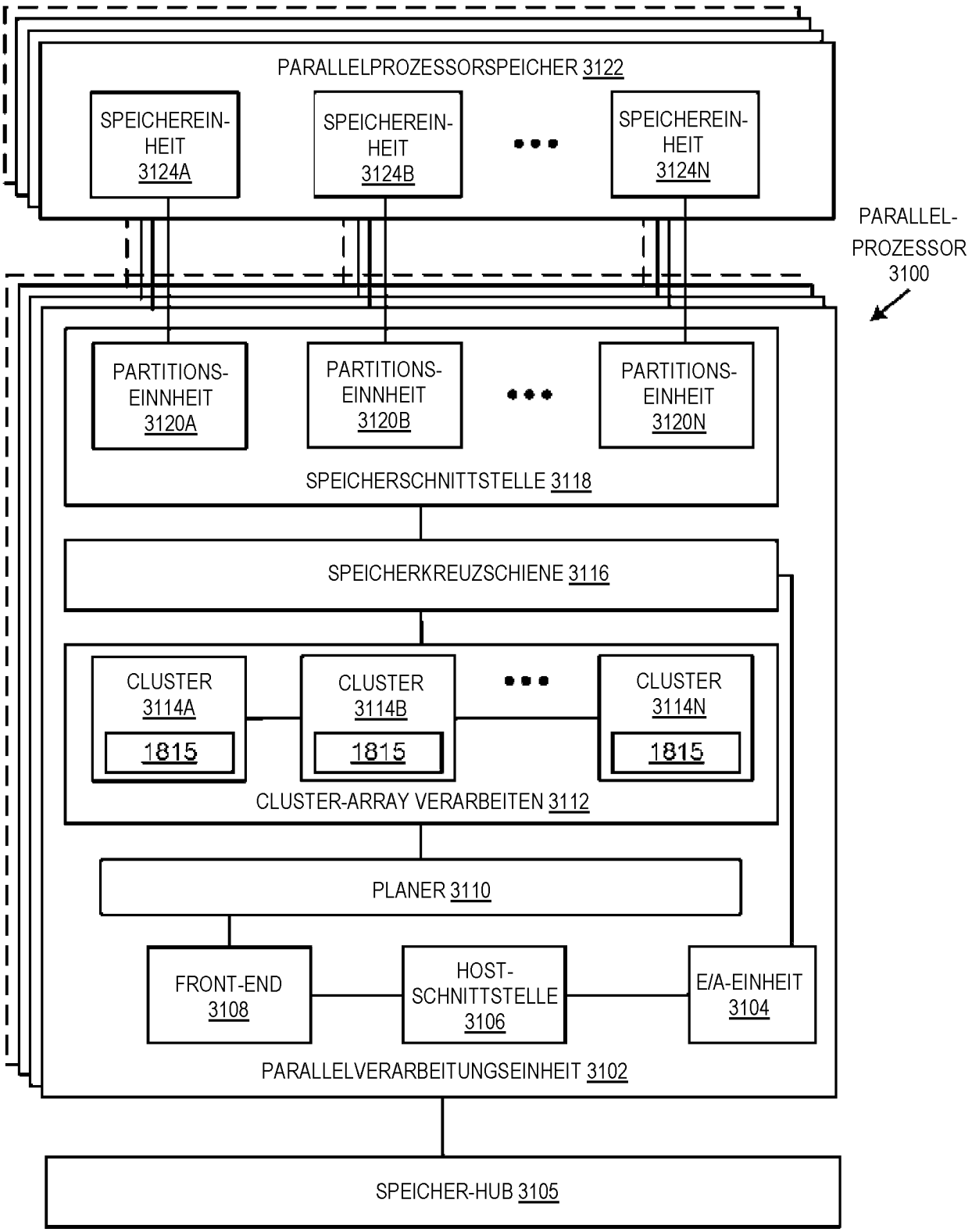

PARALLELPROZESSORSPEICHER 3122
SPEICHEREIN-HEIT 3124A
SPEICHEREIN-HEIT 3124B
SPEICHEREIN-HEIT 3124N
PARALLEL-PROZESSOR 3100
PARTITIONS-EINNHEIT 3120A
PARTITIONS-EINNHEIT 3120B
PARTITIONS-EINHEIT 3120N
SPEICHERSCHNITTSTELLE 3118
SPEICHERKREUZSCHIENE 3116
CLUSTER 3114A
1815
CLUSTER 3114B
1815
CLUSTER 3114N
1815
CLUSTER-ARRAY VERARBEITEN 3112
PLANER 3110
FRONT-END 3108
HOST-SCHNITTSTELLE 3106
E/A-EINHEIT 3104
PARALLELVERARBEITUNGSEINHEIT 3102
SPEICHER-HUB 3105

FIG. 31A

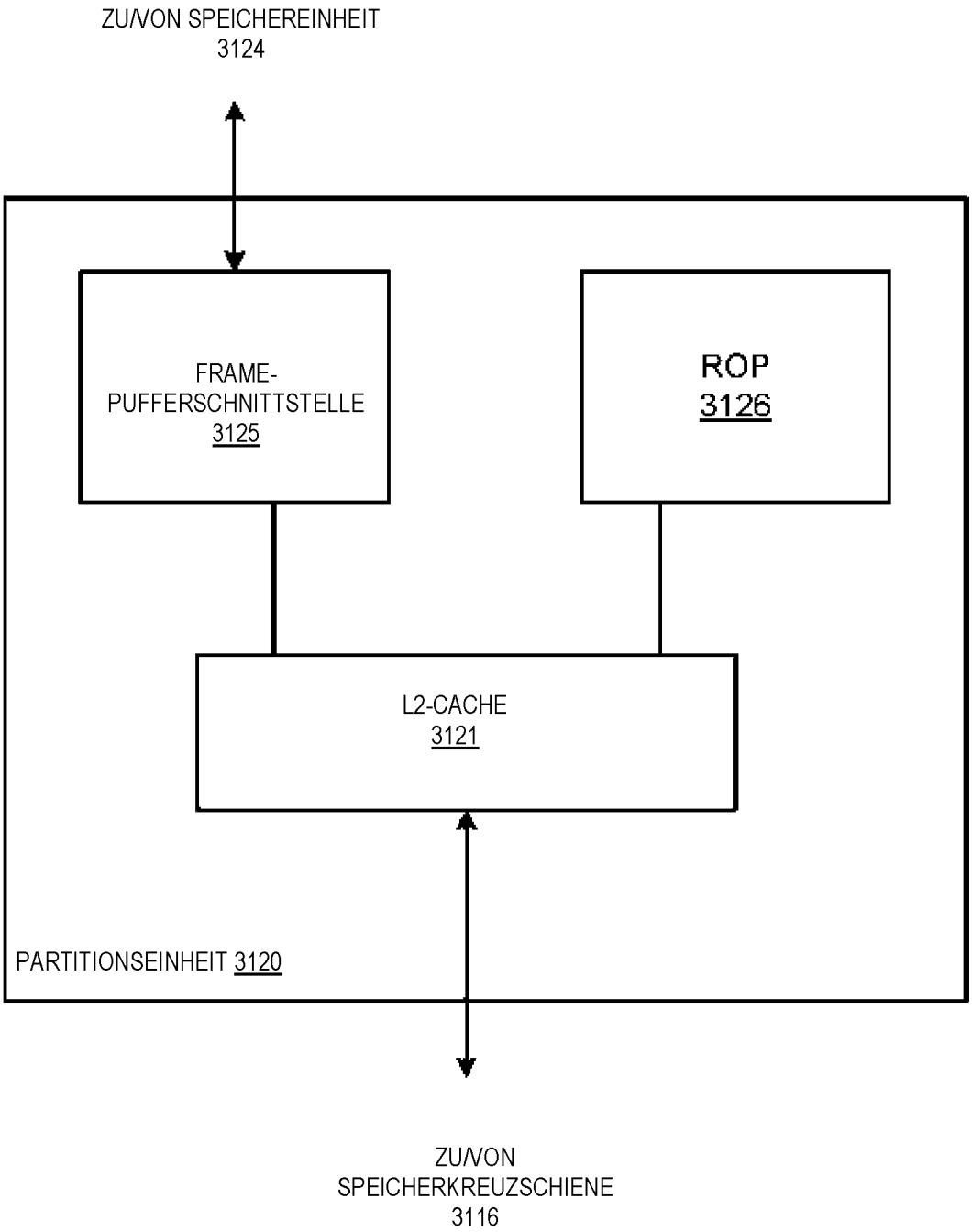

ZU/VON SPEICHEREINHEIT
3124
FRAME-
PUFFERSCHNITTSTELLE
3125
ROP
3126
L2-CACHE
3121
PARTITIONSEINHEIT 3120
ZU/VON
SPEICHERKREUZSCHIENE
3116

FIG. 31B

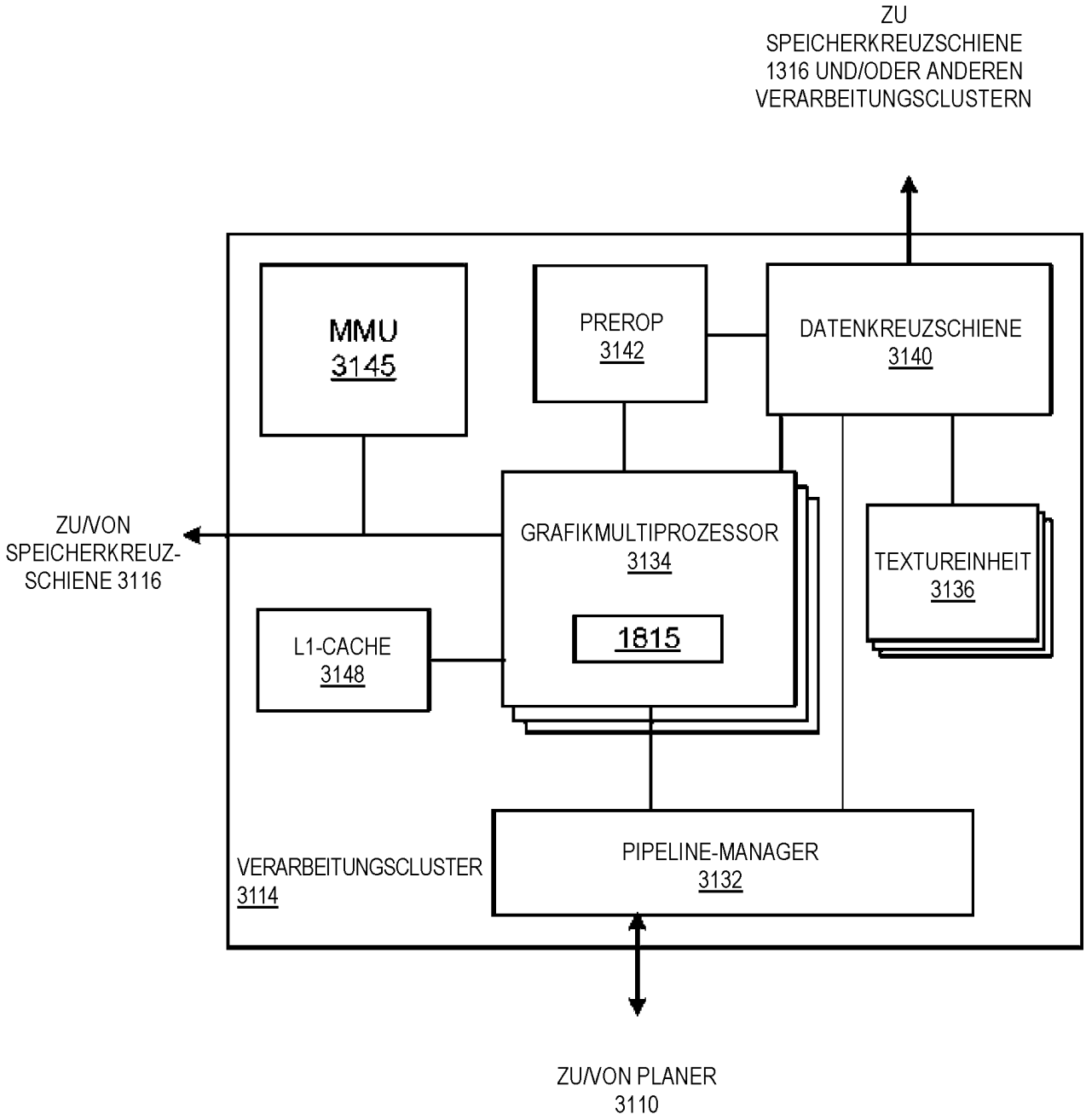

ZU
SPEICHERKREUZSCHIENE
1316 UND/ODER ANDEREN
VERARBEITUNGSCLUSTERN
MMU
3145
PREROP
3142
DATENKREUZSCHIENE
3140
ZU/VON
SPEICHERKREUZ-
SCHIENE 3116
GRAFIKMULTIPROZESSOR
3134
1815
TEXTUREINHEIT
3136
L1-CACHE
3148
VERARBEITUNGSCLUSTER
3114
PIPELINE-MANAGER
3132
ZU/VON PLANER
3110

FIG. 31C

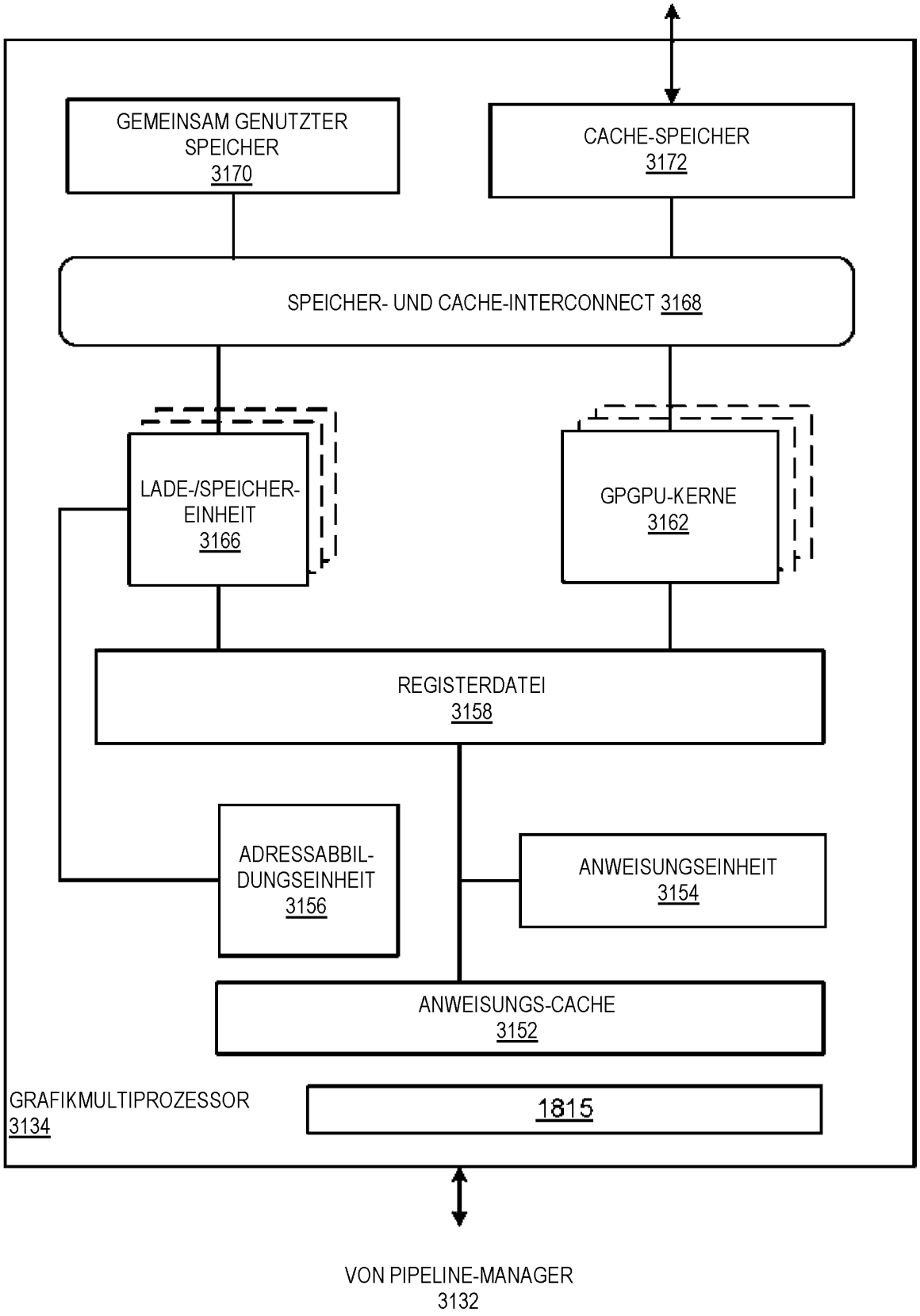

GEMEINSAM GENUTZTER SPEICHER 3170
CACHE-SPEICHER 3172
SPEICHER- UND CACHE-INTERCONNECT 3168
LADE-/SPEICHER-EINHEIT 3166
GPGPU-KERNE 3162
REGISTERDATEI 3158
ADRESSABBIL-DUNGSEINHEIT 3156
ANWEISUNGSEINHEIT 3154
ANWEISUNGS-CACHE 3152
GRAFIKMULTIPROZESSOR 3134
1815
VON PIPELINE-MANAGER 3132

FIG. 31D

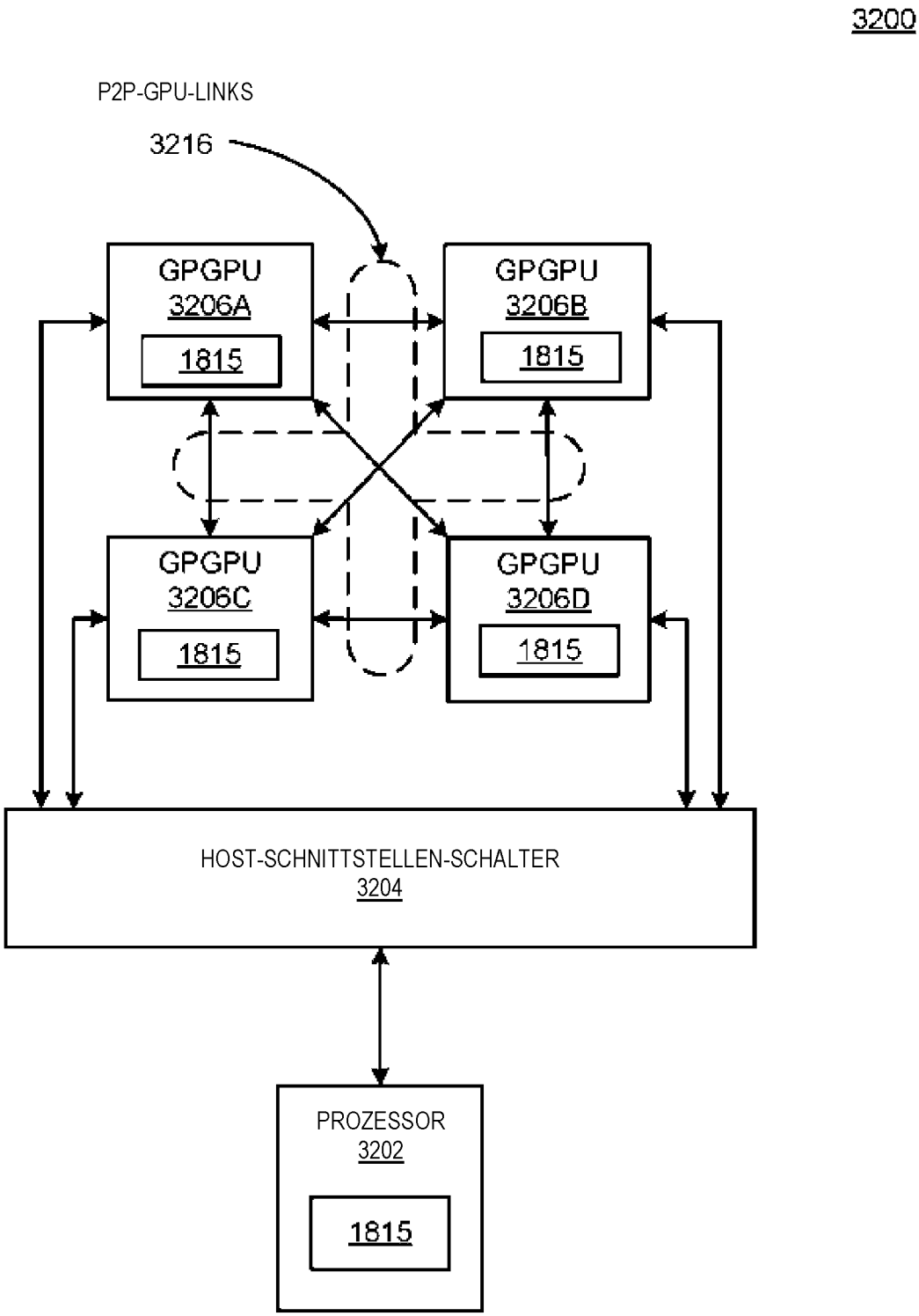

3200
P2P-GPU-LINKS
3216
GPGPU
3206A
1815
GPGPU
3206B
1815
GPGPU
3206C
1815
GPGPU
3206D
1815
HOST-SCHNITTSTELLEN-SCHALTER
3204
PROZESSOR
3202
1815

FIG. 32

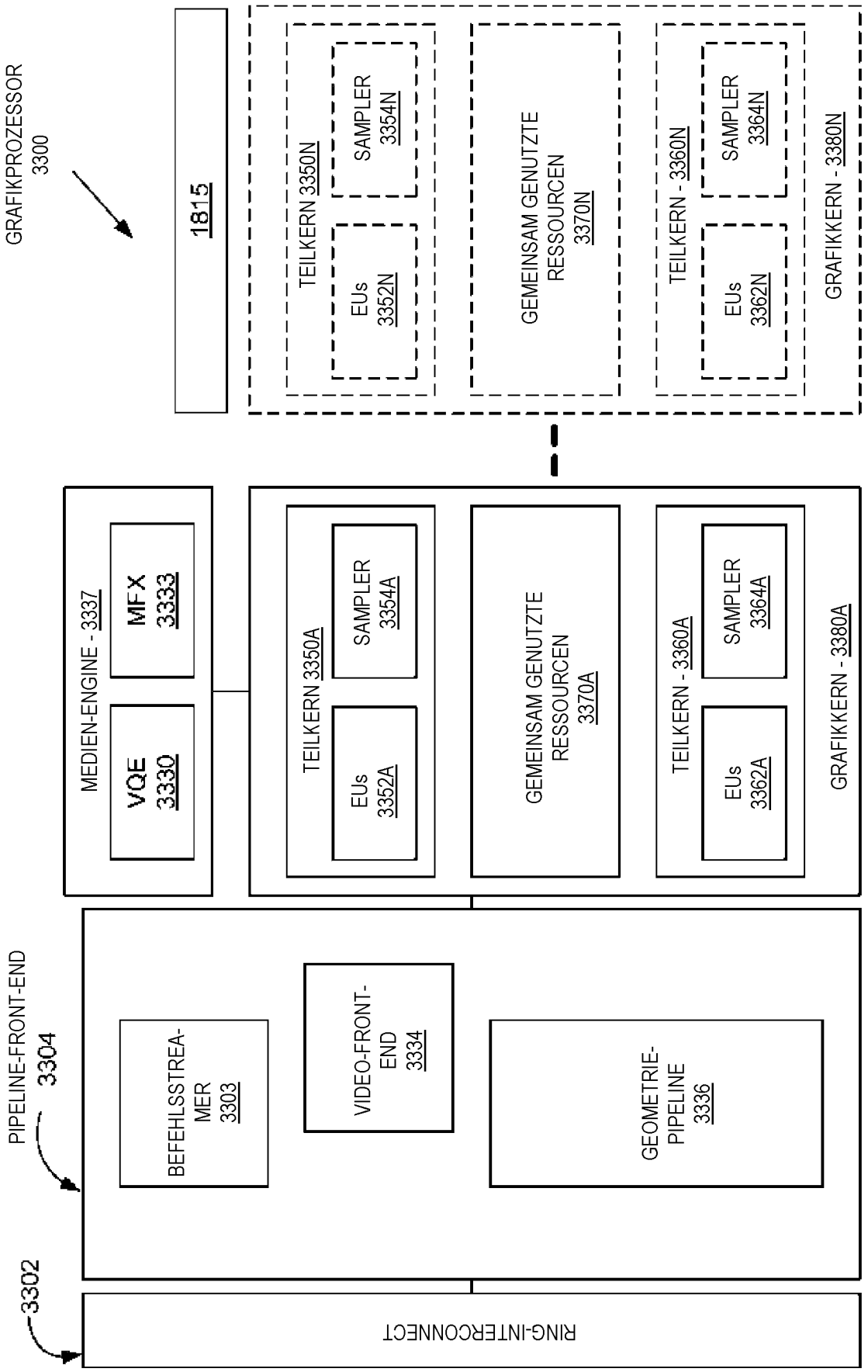

GRAFIKPROZESSOR 3300
1815
TEILKERN 3350N
EUs 3352N
SAMPLER 3354N
GEMEINSAM GENUTZTE RESSOURCEN 3370N
TEILKERN - 3360N
EUs 3362N
SAMPLER 3364N
GRAFIKKERN - 3380N
MEDIEN-ENGINE - 3337
VQE 3330
MFX 3333
TEILKERN 3350A
EUs 3352A
SAMPLER 3354A
GEMEINSAM GENUTZTE RESSOURCEN 3370A
TEILKERN - 3360A
EUs 3362A
SAMPLER 3364A
GRAFIKKERN - 3380A
PIPELINE-FRONT-END 3304
BEFEHLSSTREA-MER 3303
VIDEO-FRONT-END 3334
GEOMETRIE-PIPELINE 3336
3302
RING-INTERCONNECT

FIG. 33

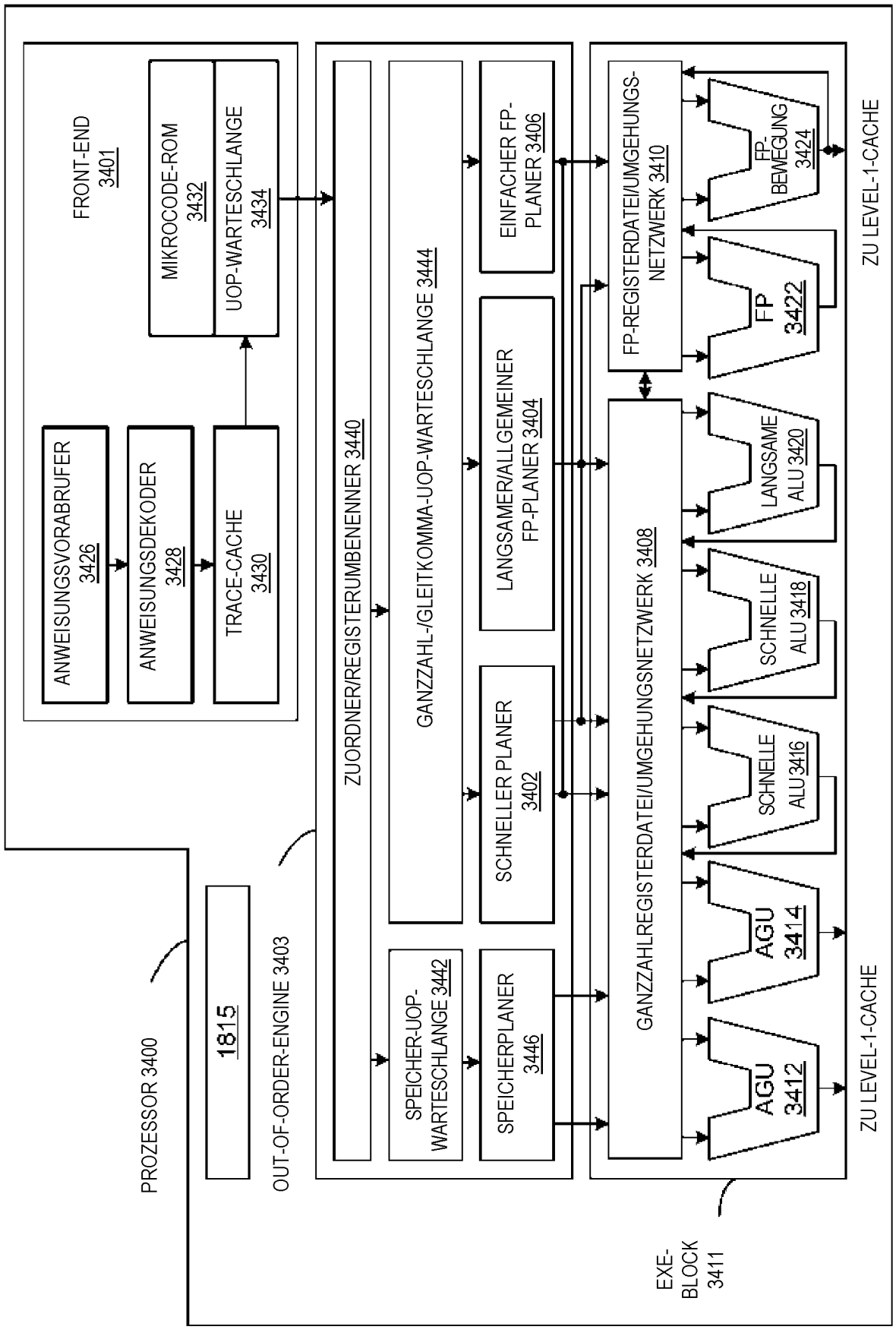

PROZESSOR 3400
1815
OUT-OF-ORDER-ENGINE 3403
FRONT-END 3401
ANWEISUNGSVORABRUFER 3426
ANWEISUNGSDEKODER 3428
TRACE-CACHE 3430
MIKROCODE-ROM 3432
UOP-WARTESCHLANGE 3434
ZUORDNER/REGISTERUMBENENNER 3440
SPEICHER-UOP-WARTESCHLANGE 3442
SPEICHERPLANER 3446
GANZZAHL-/GLEITKOMMA-UOP-WARTESCHLANGE 3444
SCHNELLER PLANER 3402
LANGSAMER/ALLGEMEINER FP-PLANER 3404
EINFACHER FP-PLANER 3406
GANZZAHLREGISTERDATEI/UMGEHUNGSNETZWERK 3408
FP-REGISTERDATEI/UMGEHUNGS-NETZWERK 3410
AGU 3412
AGU 3414
SCHNELLE ALU 3416
SCHNELLE ALU 3418
LANGSAME ALU 3420
FP 3422
FP-BEWEGUNG 3424
ZU LEVEL-1-CACHE
ZU LEVEL-1-CACHE
EXE-BLOCK 3411

FIG. 34

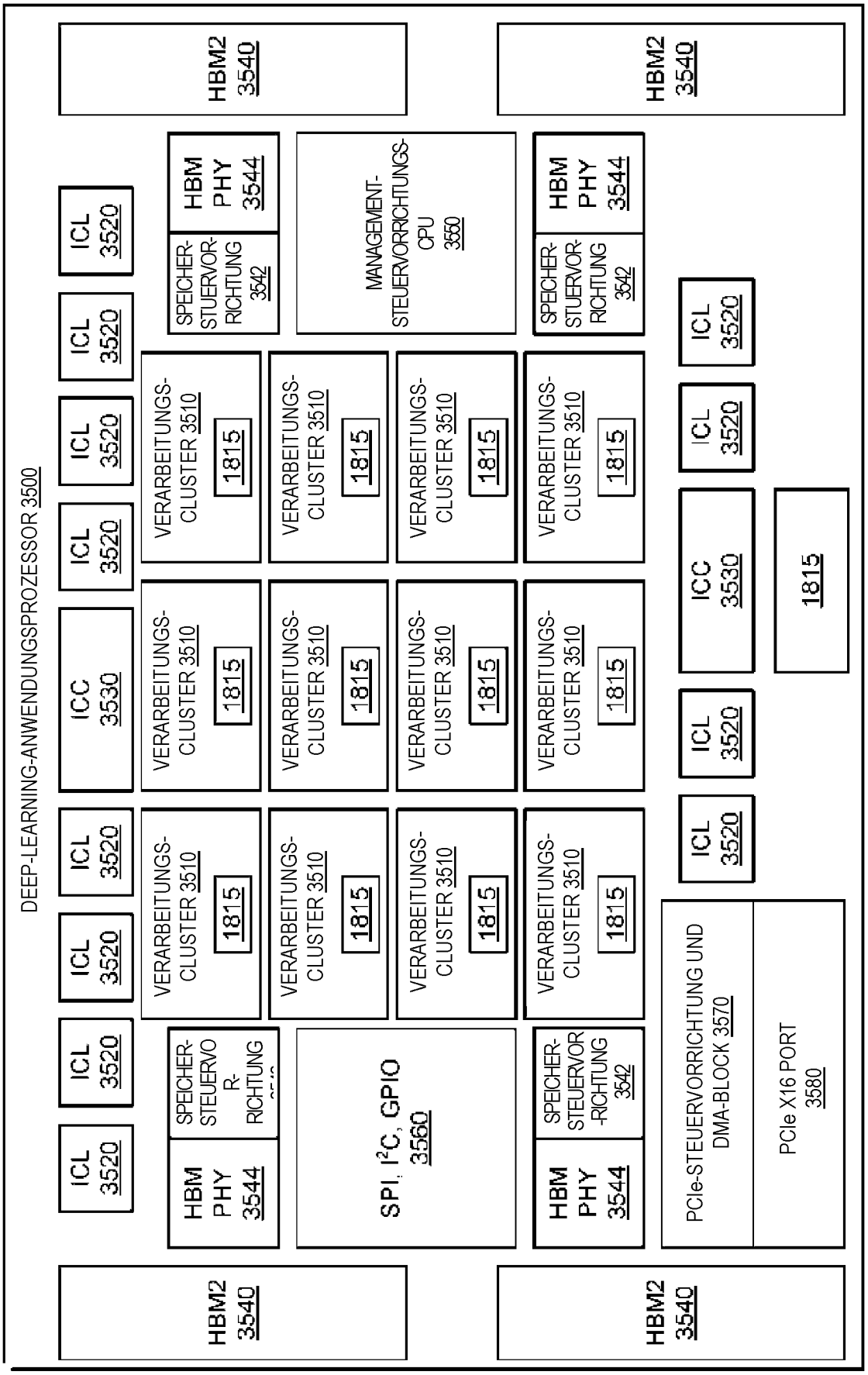

DEEP-LEARNING-ANWENDUNGSPROZESSOR 3500
HBM2 3540
HBM2 3540
HBM2 3540
HBM2 3540
HBM PHY 3544
SPEICHER-STEUERVO R-RICHTUNG
SPEICHER-STEUERVOR-RICHTUNG 3542
SPI, I2C, GPIO 3560
PCIe-STEUERVORRICHTUNG UND DMA-BLOCK 3570
PCIe X16 PORT 3580
ICL 3520
ICC 3530
VERARBEITUNGS-CLUSTER 3510
1815
MANAGEMENT-STEUERVORRICHTUNGS-CPU 3550

FIG. 35

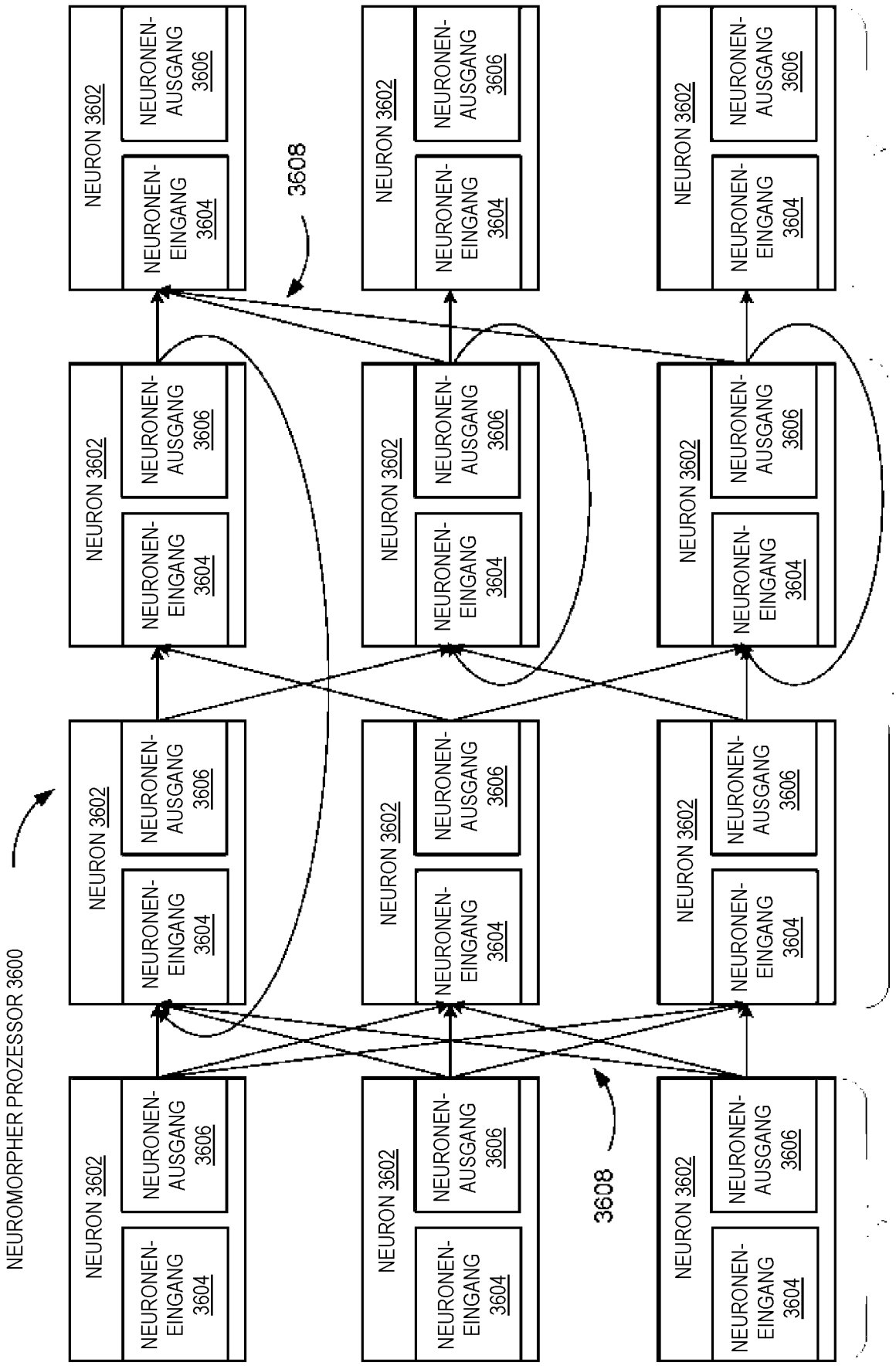

NEUROMORPHER PROZESSOR 3600
NEURON 3602
NEURONEN-EINGANG 3604
NEURONEN-AUSGANG 3606
3608
3610
3612
3614

FIG. 36

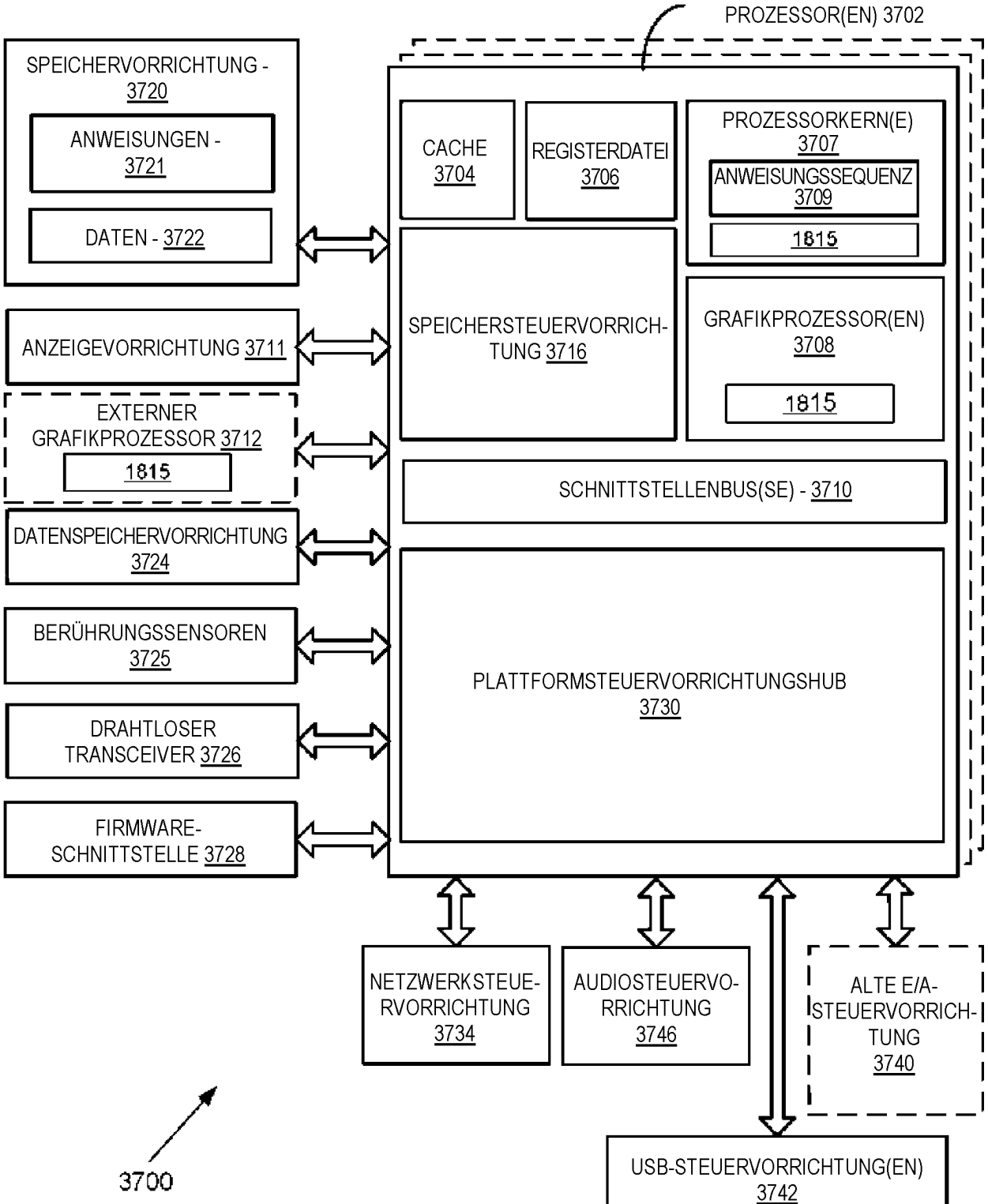

PROZESSOR(EN) 3702
SPEICHERVORRICHTUNG - 3720
ANWEISUNGEN - 3721
DATEN - 3722
CACHE 3704
REGISTERDATEI 3706
PROZESSORKERN(E) 3707
ANWEISUNGSSEQUENZ 3709
1815
SPEICHERSTEUERVORRICHTUNG 3716
GRAFIKPROZESSOR(EN) 3708
1815
ANZEIGEVORRICHTUNG 3711
EXTERNER GRAFIKPROZESSOR 3712
1815
SCHNITTSTELLENBUS(SE) - 3710
DATENSPEICHERVORRICHTUNG 3724
BERÜHRUNGSSENSOREN 3725
PLATTFORMSTEUERVORRICHTUNGSHUB 3730
DRAHTLOSER TRANSCEIVER 3726
FIRMWARE-SCHNITTSTELLE 3728
NETZWERKSTEUERVORRICHTUNG 3734
AUDIOSTEUERVORRICHTUNG 3746
ALTE E/A-STEUERVORRICHTUNG 3740
3700
USB-STEUERVORRICHTUNG(EN) 3742
TASTATUR/MAUS 3743
KCAMERA 3744

FIG. 37

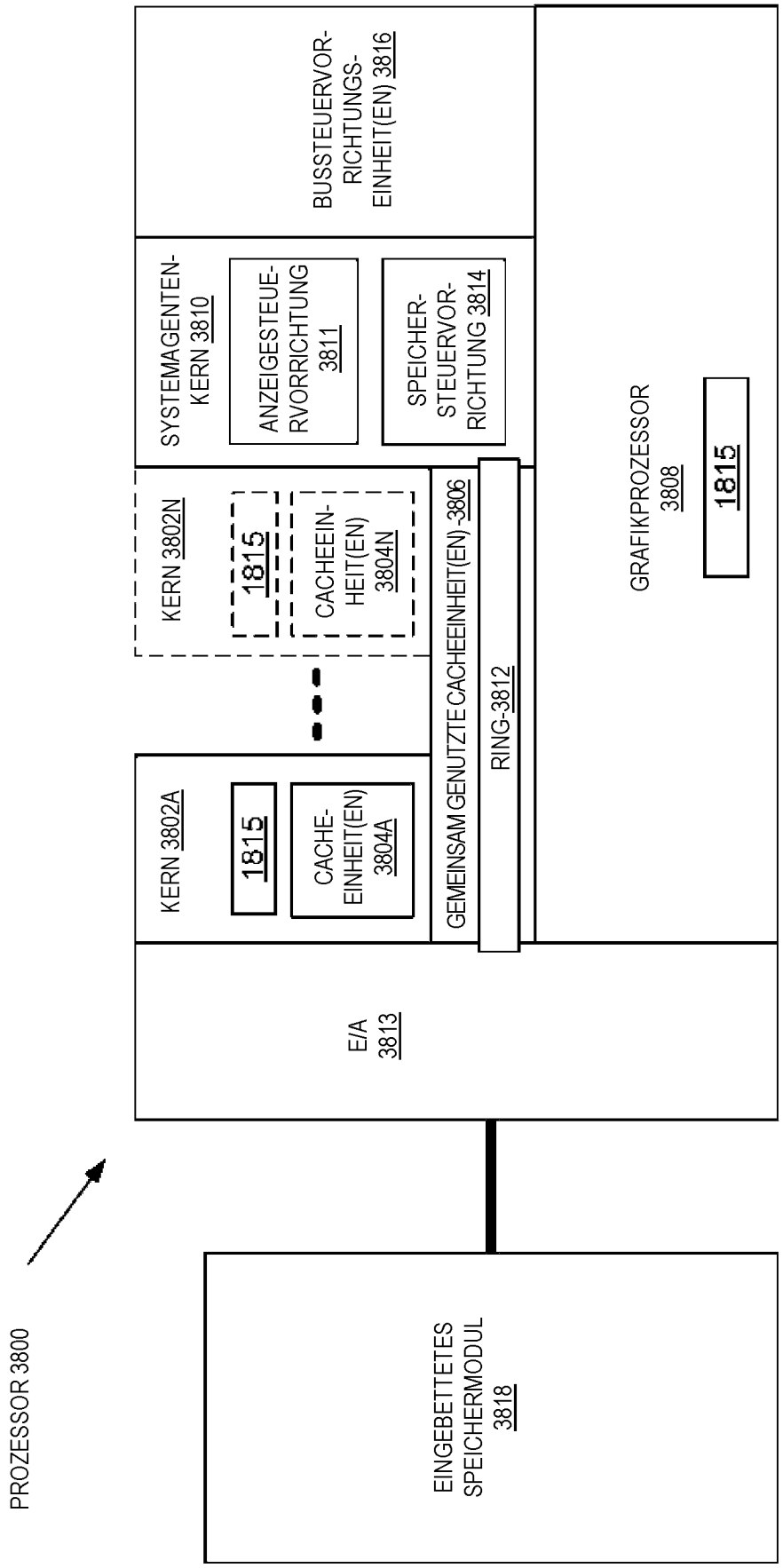

PROZESSOR 3800
EINGEBETTETES SPEICHERMODUL 3818
E/A 3813
KERN 3802A
1815
CACHE-EINHEIT(EN) 3804A
KERN 3802N
1815
CACHEEIN-HEIT(EN) 3804N
GEMEINSAM GENUTZTE CACHEEINHEIT(EN) -3806
RING-3812
SYSTEMAGENTEN-KERN 3810
ANZEIGESTEUE-RVORRICHTUNG 3811
SPEICHER-STEUERVOR-RICHTUNG 3814
BUSSTEUERVOR-RICHTUNGS-EINHEIT(EN) 3816
GRAFIKPROZESSOR 3808
1815

FIG. 38

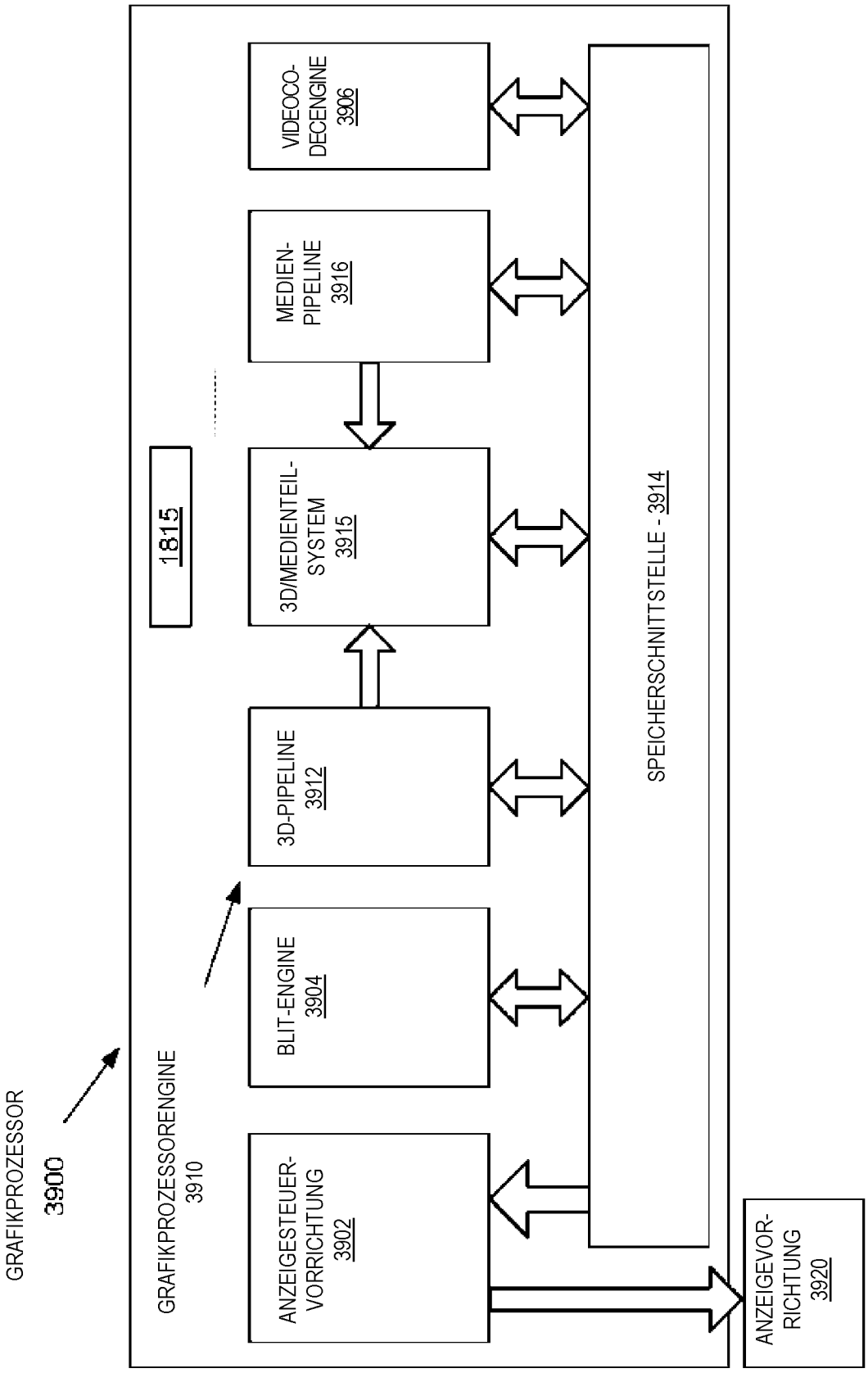

GRAFIKPROZESSOR
3900
GRAFIKPROZESSORENGINE
3910
1815
ANZEIGESTEUER-VORRICHTUNG
3902
BLIT-ENGINE
3904
3D-PIPELINE
3912
3D/MEDIENTEIL-SYSTEM
3915
MEDIEN-PIPELINE
3916
VIDEOCO-DECENGINE
3906
SPEICHERSCHNITTSTELLE - 3914
ANZEIGEVOR-RICHTUNG
3920

FIG. 39

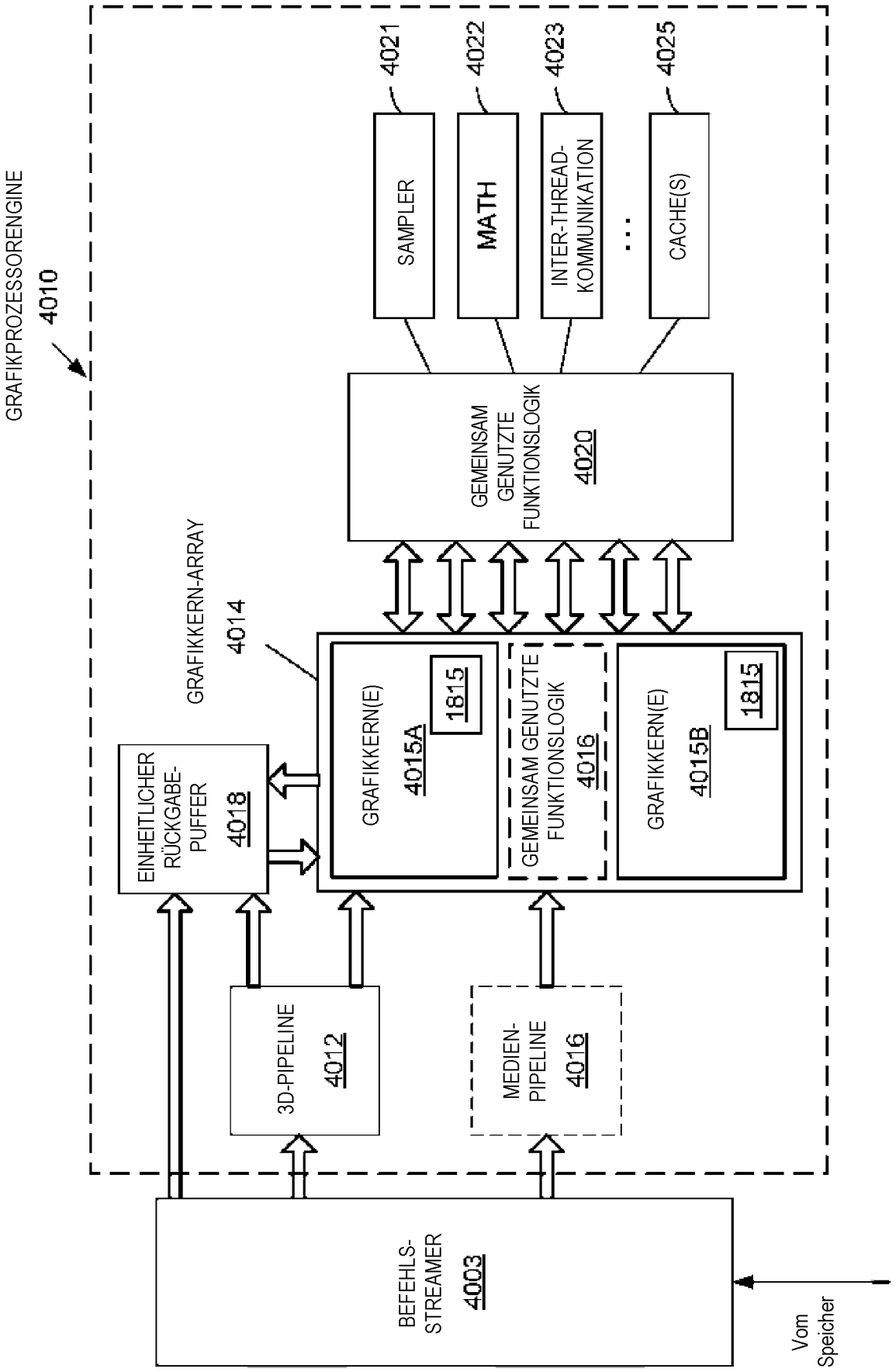

GRAFIKPROZESSORENGINE
4010
4021
4022
4023
4025
SAMPLER
MATH
INTER-THREAD-KOMMUNIKATION
...
CACHE(S)
GEMEINSAM GENUTZTE FUNKTIONSLOGIK
4020
GRAFIKKERN-ARRAY
4014
GRAFIKKERN(E)
4015A
1815
GEMEINSAM GENUTZTE FUNKTIONSLOGIK
4016
GRAFIKKERN(E)
4015B
1815
EINHEITLICHER RÜCKGABE-PUFFER
4018
3D-PIPELINE
4012
MEDIEN-PIPELINE
4016
BEFEHLS-STREAMER
4003
Vom Speicher

FIG. 40

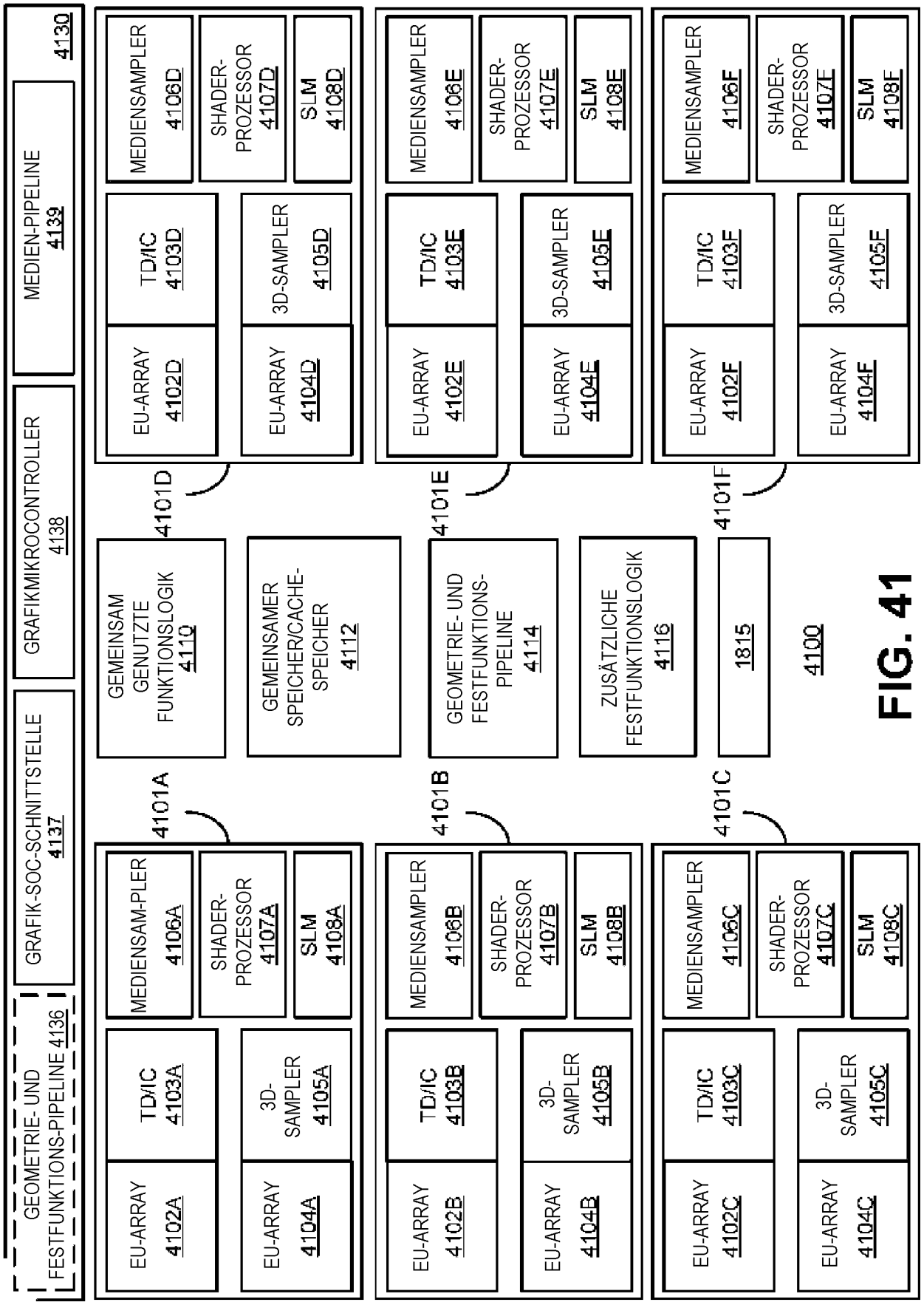

GEOMETRIE- UND FESTFUNKTIONS-PIPELINE 4136
GRAFIK-SOC-SCHNITTSTELLE 4137
GRAFIKMIKROCONTROLLER 4138
MEDIEN-PIPELINE 4139
4130
4101A
4101B
4101C
4101D
4101E
4101F
EU-ARRAY 4102A
TD/IC 4103A
3D-SAMPLER 4105A
EU-ARRAY 4104A
MEDIENSAM-PLER 4106A
SHADER-PROZESSOR 4107A
SLM 4108A
EU-ARRAY 4102B
TD/IC 4103B
3D-SAMPLER 4105B
EU-ARRAY 4104B
MEDIENSAMPLER 4106B
SHADER-PROZESSOR 4107B
SLM 4108B
EU-ARRAY 4102C
TD/IC 4103C
3D-SAMPLER 4105C
EU-ARRAY 4104C
MEDIENSAMPLER 4106C
SHADER-PROZESSOR 4107C
SLM 4108C
EU-ARRAY 4102D
TD/IC 4103D
3D-SAMPLER 4105D
EU-ARRAY 4104D
MEDIENSAMPLER 4106D
SHADER-PROZESSOR 4107D
SLM 4108D
EU-ARRAY 4102E
TD/IC 4103E
3D-SAMPLER 4105E
EU-ARRAY 4104E
MEDIENSAMPLER 4106E
SHADER-PROZESSOR 4107E
SLM 4108E
EU-ARRAY 4102F
TD/IC 4103F
3D-SAMPLER 4105F
EU-ARRAY 4104F
MEDIENSAMPLER 4106F
SHADER-PROZESSOR 4107F
SLM 4108F
GEMEINSAM GENUTZTE FUNKTIONSLOGIK 4110
GEMEINSAMER SPEICHER/CACHE-SPEICHER 4112
GEOMETRIE- UND FESTFUNKTIONS-PIPELINE 4114
ZUSÄTZLICHE FESTFUNKTIONSLOGIK 4116
1815
4100

FIG. 41

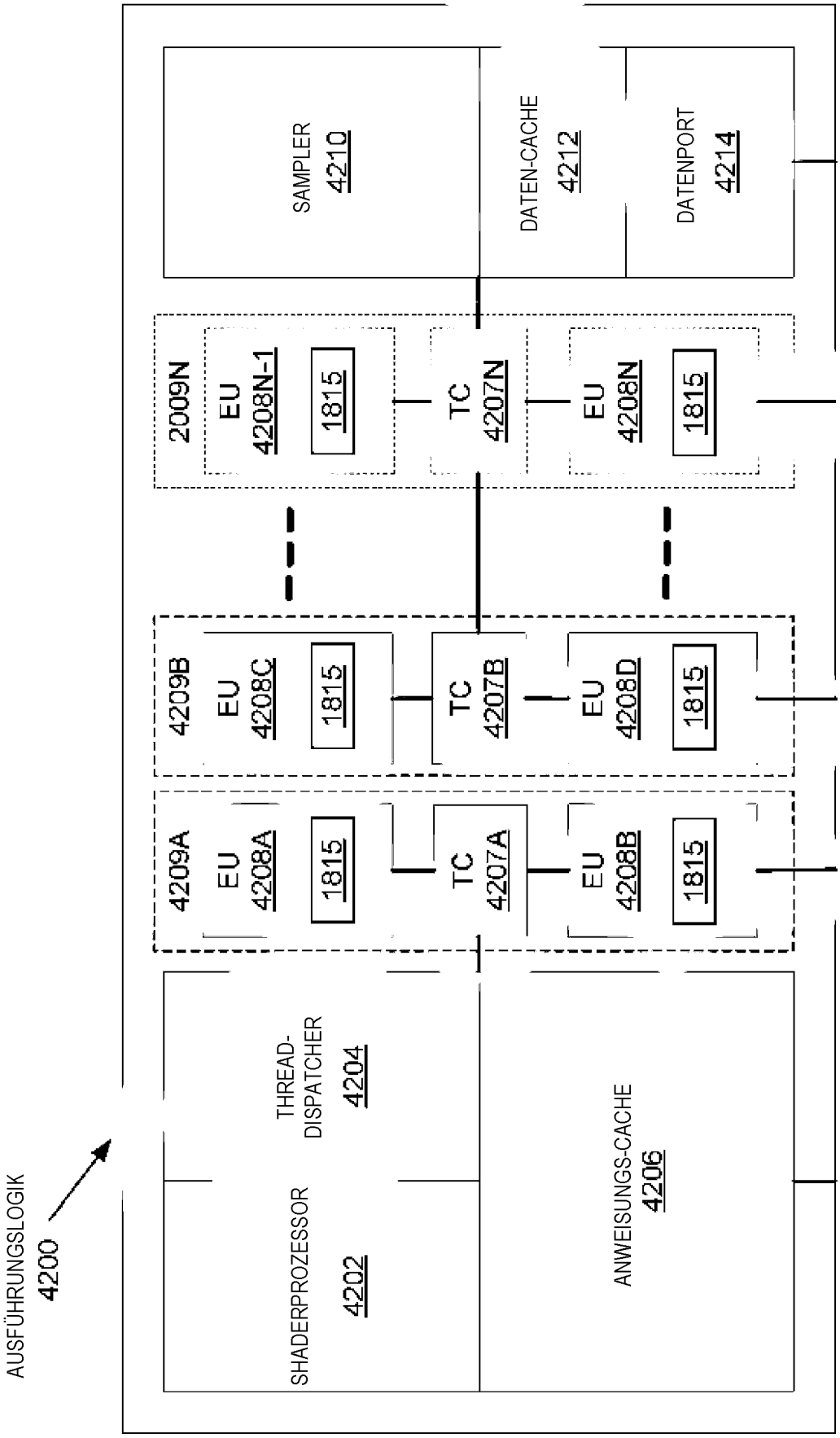

AUSFÜHRUNGSLOGIK
4200
SHADERPROZESSOR
4202
THREAD-
DISPATCHER
4204
ANWEISUNGS-CACHE
4206
4209A
EU
4208A
1815
TC
4207A
EU
4208B
1815
4209B
EU
4208C
1815
TC
4207B
EU
4208D
1815
2009N
EU
4208N-1
1815
TC
4207N
EU
4208N
1815
SAMPLER
4210
DATEN-CACHE
4212
DATENPORT
4214

FIG. 42A

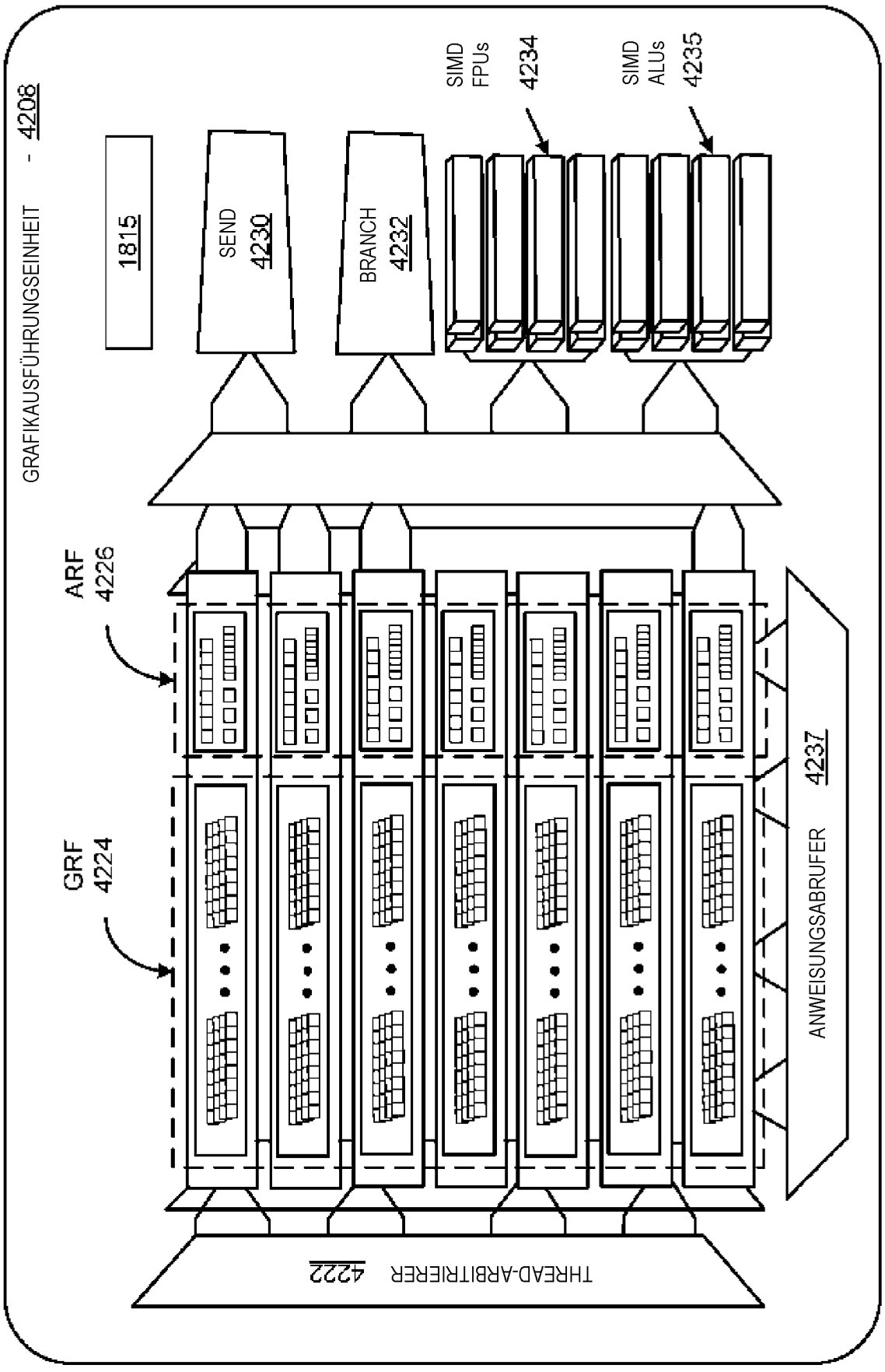

GRAFIKAUSFÜHRUNGSEINHEIT - 4208
1815
SEND 4230
BRANCH 4232
SIMD FPUs 4234
SIMD ALUs 4235
ARF 4226
GRF 4224
ANWEISUNGSABRUFER 4237
THREAD-ARBITRIERER 4222

FIG. 42B

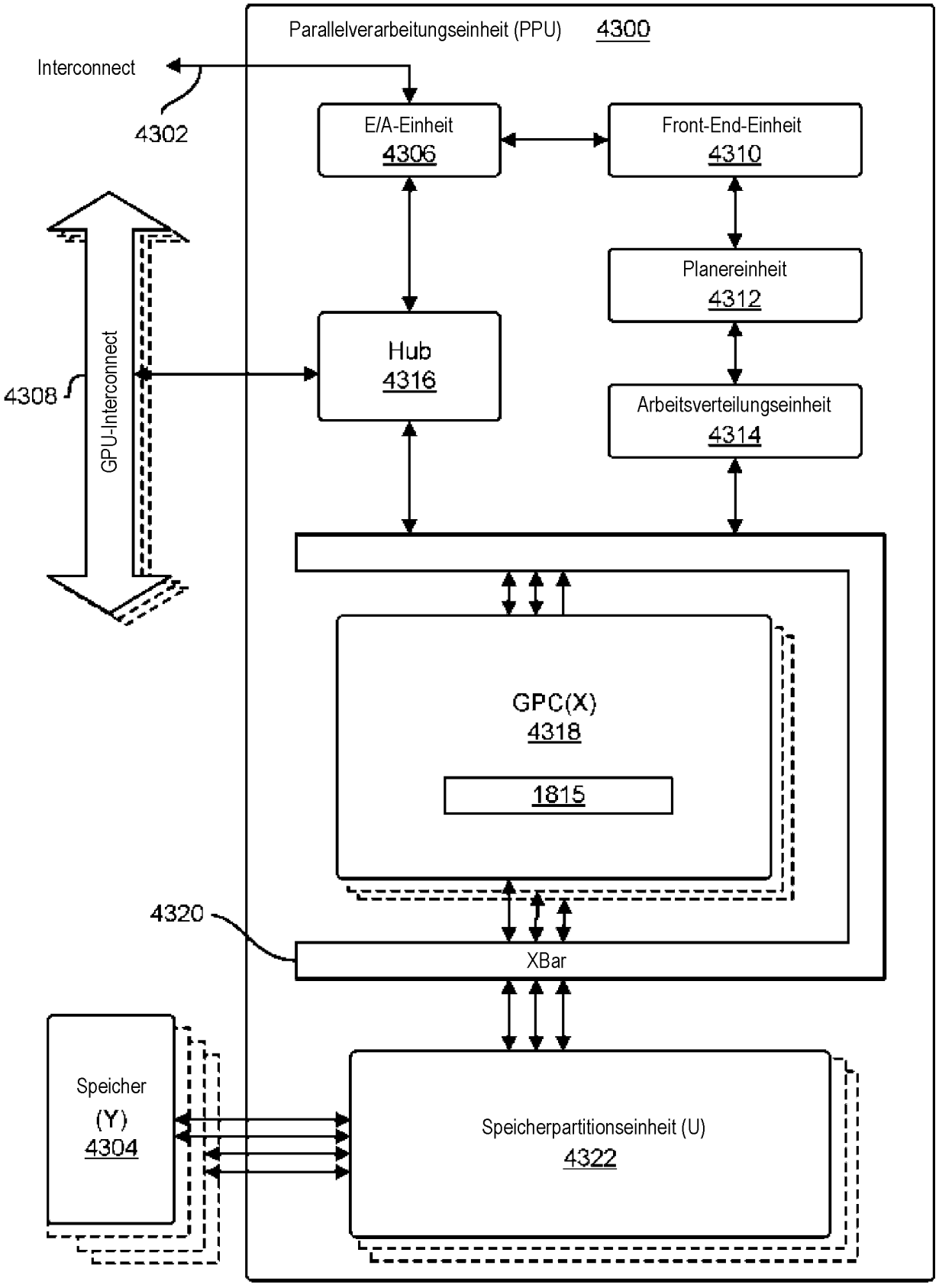

Parallelverarbeitungseinheit (PPU) 4300
Interconnect
4302
E/A-Einheit
4306
Front-End-Einheit
4310
Planereinheit
4312
Hub
4316
Arbeitsverteilungseinheit
4314
4308
GPU-Interconnect
GPC(X)
4318
1815
4320
XBar
Speicher
(Y)
4304
Speicherpartitionseinheit (U)
4322

FIG. 43

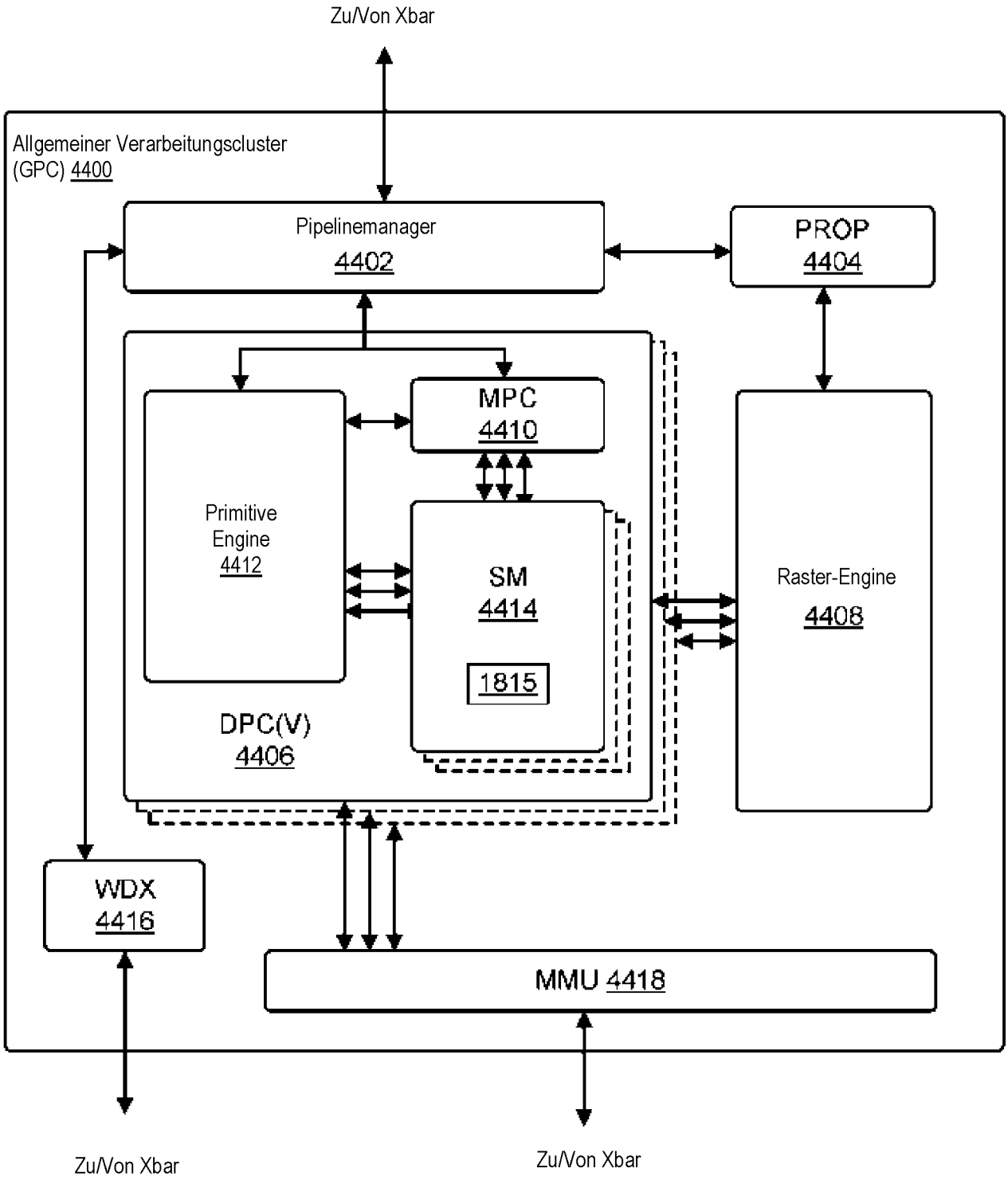

Zu/Von Xbar
Allgemeiner Verarbeitungscluster (GPC) 4400
Pipelinemanager 4402
PROP 4404
MPC 4410
Primitive Engine 4412
SM 4414
1815
DPC(V) 4406
Raster-Engine 4408
WDX 4416
MMU 4418
Zu/Von Xbar
Zu/Von Xbar

FIG. 44

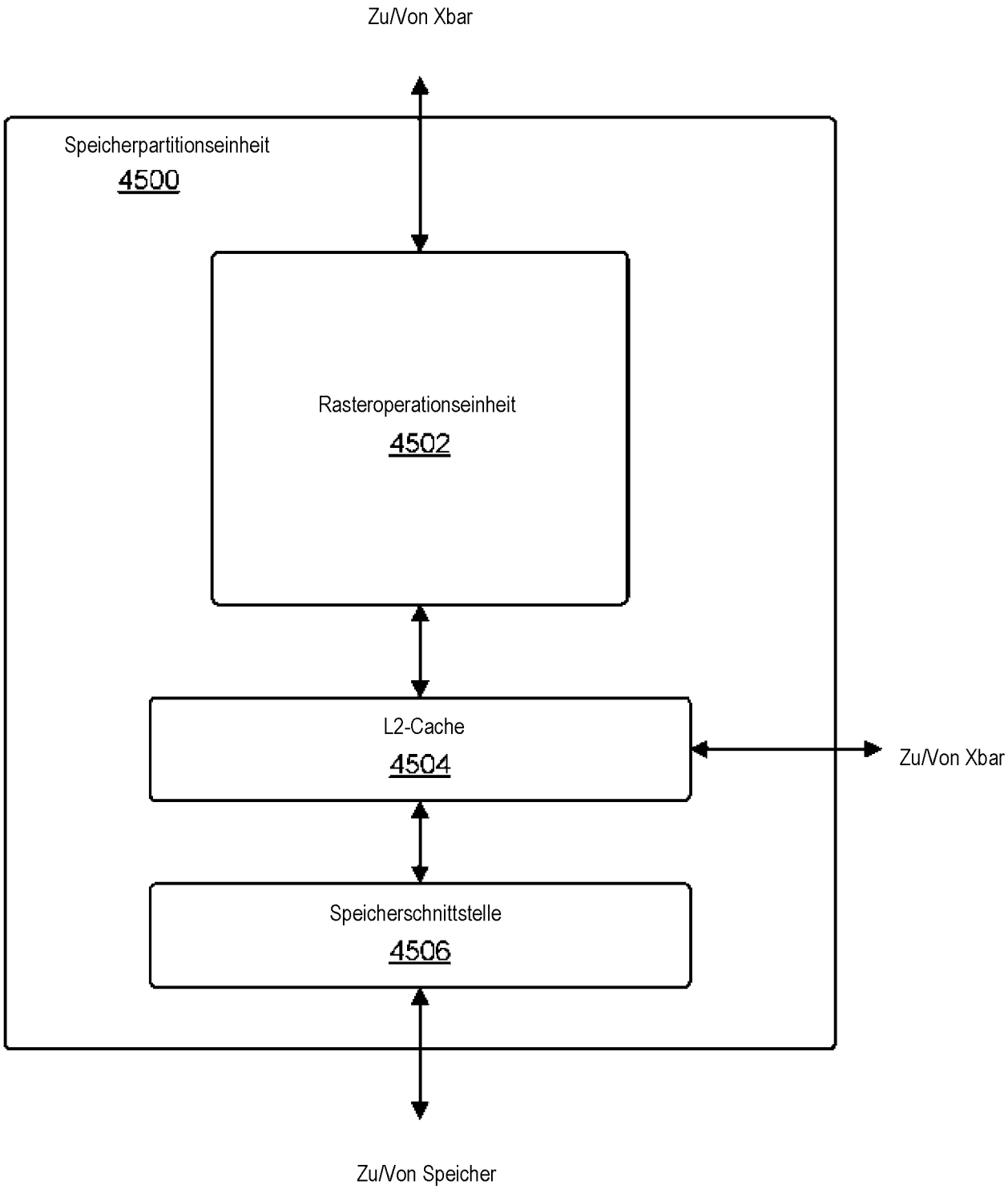

Zu/Von Xbar
Speicherpartitionseinheit
4500
Rasteroperationseinheit
4502
L2-Cache
4504
Zu/Von Xbar
Speicherschnittstelle
4506
Zu/Von Speicher

FIG. 45

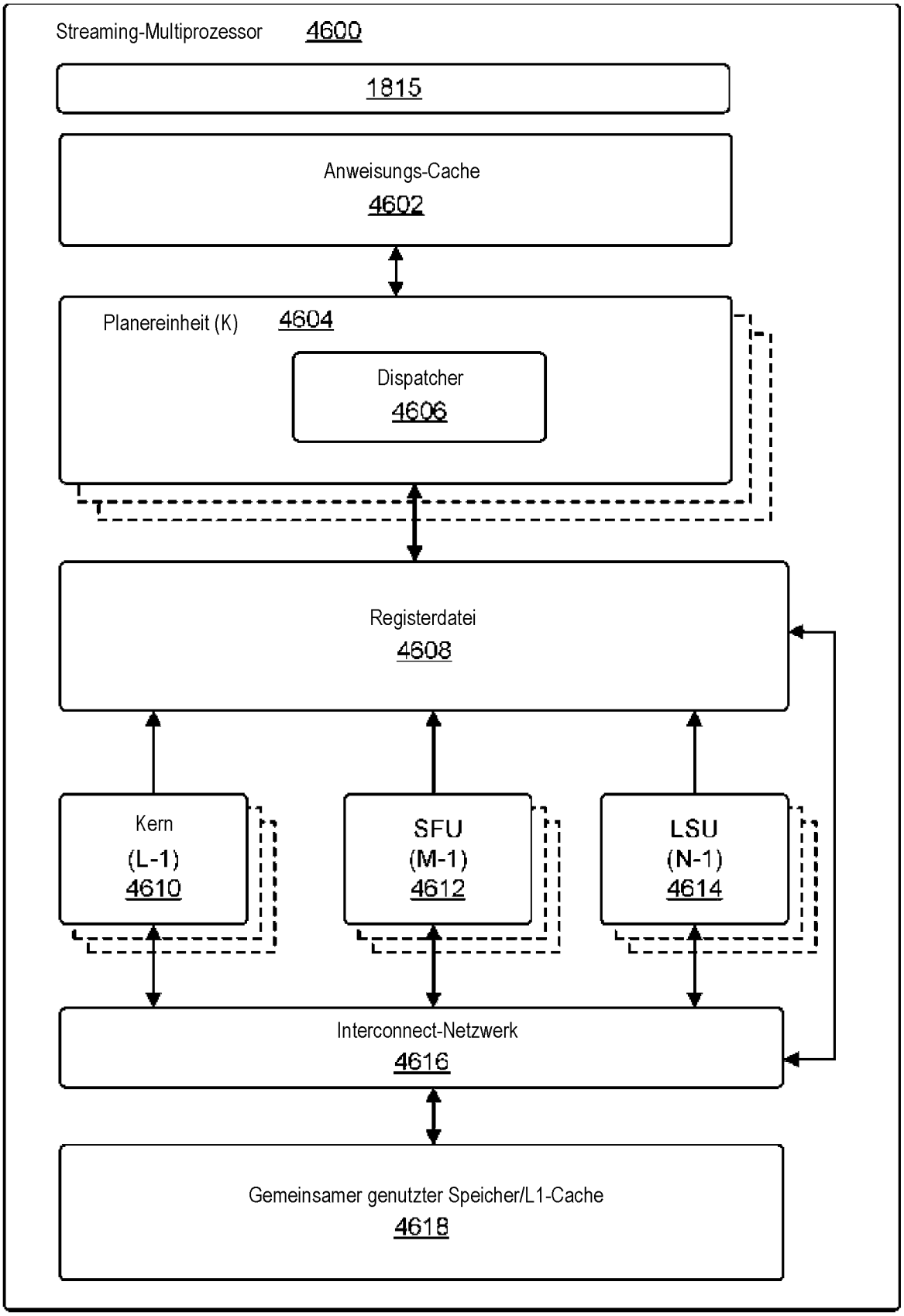

Streaming-Multiprozessor 4600
1815
Anweisungs-Cache
4602
Planereinheit (K) 4604
Dispatcher
4606
Registerdatei
4608
Kern
(L-1)
4610
SFU
(M-1)
4612
LSU
(N-1)
4614
Interconnect-Netzwerk
4616
Gemeinsamer genutzter Speicher/L1-Cache
4618

FIG. 46