



(10) **DE 11 2021 005 482 T5** 2023.09.14

(12)

Veröffentlichung

der internationalen Anmeldung mit der
(87) Veröffentlichungs-Nr.: **WO 2022/111282**
in der deutschen Übersetzung (Art. III § 8 Abs. 2
IntPatÜbkG)
(21) Deutsches Aktenzeichen: **11 2021 005 482.1**
(86) PCT-Aktenzeichen: **PCT/CN2021/129740**
(86) PCT-Anmeldetag: **10.11.2021**
(87) PCT-Veröffentlichungstag: **02.06.2022**
(43) Veröffentlichungstag der PCT Anmeldung
in deutscher Übersetzung: **14.09.2023**

(51) Int Cl.: **G06F 3/16** (2006.01)
G06F 3/01 (2006.01)
G10L 17/00 (2013.01)
G10L 15/00 (2013.01)

(30) Unionspriorität:
17/102,687 **24.11.2020** **US**

(71) Anmelder:
International Business Machines Corporation,
Armonk, NY, US

(74) Vertreter:
Richardt Patentanwälte
Partnerschaftsgesellschaft mit beschränkter
Berufshaftung, 65185 Wiesbaden, DE

(72) Erfinder:
Decrop, Clement, Washington, DC, US; Agrawal,
Tushar, Fargo, ND, US; Fox, Jeremy R., Austin, TX,
US; Rakshit, Sarbajit K., Kolkata, IN

Prüfungsantrag gemäß § 44 PatG ist gestellt.

Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen.

(54) Bezeichnung: **AR- (AUGMENTED REALITY) GESTÜTZTE SELEKTIVE GERÄUSCHEINBINDUNG AUS DER
UMGEBUNG WÄHREND DER AUSFÜHRUNG EINES SPRACHBEFEHLS**

(57) Zusammenfassung: Angegeben sind ein Verfahren und eine Vorrichtung zum Erzeugen eines erweiterten Sprachbefehls, umfassend Identifizieren einer Mehrzahl von Geräuschen von einer jeweiligen Mehrzahl von Wandlern für eine intelligente Lautsprecher- (smart speaker) Einheit, Erzeugen einer Visualisierung der Geräusche unter Verwendung einer AR-Einheit, wobei eines oder mehrere der Geräusche unter Verwendung der Visualisierung ausgewählt werden können, und Erzeugen des erweiterten Sprachbefehls für die intelligente Lautsprechereinheit, wobei der erweiterte Sprachbefehl das eine oder die mehreren Geräusche aufweist, die unter Verwendung der Visualisierung der AR-Einheit ausgewählt wurden.

Übersetzen des Sprachbefehls an AR-Brille 302

Auswählen externer Befehle im AR-Raum 304

Ausführung des Sprachbefehls im AR-Raum 306

Historisches Korpus zum Auswählen von Befehlen 308

Beschreibung**HINTERGRUND**

[0001] Die vorliegende Erfindung betrifft eine Ausführungsform eines Verfahrens, einer Vorrichtung und eines Systems zur selektiven Geräuscheinbindung und insbesondere, aber nicht einschränkend, ein Verfahren, eine Vorrichtung und ein System für AR- (augmented reality, erweiterte Realität) gestützte selektive Geräuscheinbindung aus der Umgebung während eines Ausführens eines Sprachbefehls.

[0002] KI- (künstliche Intelligenz) Sprachassistenzsysteme wie AMAZON ALEXA, GOOGLE HOME usw. werden immer beliebter. Beispielsweise wird erwartet, dass der weltweite Markt für Sprachassistenz ein noch nie dagewesenes Ausmaß erreichen wird. Eine der treibenden Kräfte ist eine steigende Nachfrage nach einem verbesserten Kundenerlebnis, eine andere die zunehmenden Anwendungsfälle im Gesundheitswesen (Patientenbeteiligung).

[0003] In solchen Systemen kann ein Benutzer einen Sprachbefehl eingeben, der dann entsprechend ausgeführt wird.

[0004] Mehrere Personen können Sprachbefehle an ein einziges KI-Sprachassistenzsystem übermitteln, und das KI-Sprachassistenzsystem kann die Stimme erkennen und benutzerspezifische Sprachbefehle ausführen.

[0005] Es gibt jedoch Fälle, in denen ein Benutzer einen beliebigen Sprachbefehl abgibt und andere Benutzer in der Umgebung weitere Vorschläge oder zusätzliche Rückmeldungen zum Sprachbefehl des ersten Benutzers geben. In diesem Fall kann es vorkommen, dass der ursprüngliche Benutzer einige der Vorschläge anderer Benutzer in seinem Sprachbefehl nicht akzeptieren möchte. Derzeit ist das KI-Sprachassistenzsystem nicht in der Lage zu verstehen, welche Sprachbefehle es ausführen soll und welche es ignorieren kann.

[0006] Daher besteht ein Bedarf an einem KI-Sprachassistenzsystem, das komplexere Eingaben verarbeiten kann. So besteht beispielsweise Bedarf an einem Verfahren und System, mit dem ein Benutzer selektiv identifizieren kann, welche gesprochenen oder nicht gesprochenen Inhalte aus der Umgebung berücksichtigt werden sollen, während der Benutzer Sprachbefehle abgibt.

KURZDARSTELLUNG

[0007] Vor dem Hintergrund der vorgenannten und anderer Probleme, Nachteile und Mängel des vorgenannten Standes der Technik stellt ein beispielhafter

Aspekt der offenbarten Erfindung ein Verfahren, eine Vorrichtung und ein System für AR-gestützte selektive Geräuscheinbindung aus der Umgebung während des Ausführens eines Sprachbefehls bereit.

[0008] Gemäß einer Ausführungsform der vorliegenden Erfindung umfasst ein Verfahren zum Erzeugen eines erweiterten Sprachbefehls Identifizieren einer Mehrzahl von Geräuschen von einer jeweiligen Mehrzahl von Wählern für eine intelligente Lautsprecher-(smart speaker) Einheit, Erzeugen einer Visualisierung der Geräusche unter Verwendung einer AR-Einheit, wobei eines oder mehrere der Geräusche unter Verwendung der Visualisierung ausgewählt werden können, und Erzeugen des erweiterten Sprachbefehls für die intelligente Lautsprechereinheit, wobei der erweiterte Sprachbefehl das eine oder die mehreren Geräusche aufweist, die unter Verwendung der Visualisierung der AR-Einheit ausgewählt wurden.

[0009] Gemäß einer anderen Ausführungsform der vorliegenden Erfindung enthält ein System zum Erzeugen eines erweiterten Sprachbefehls einen Speicher, der Computeranweisungen speichert, und einen Prozessor, der so konfiguriert ist, dass er die Computeranweisungen ausführt, um eine Mehrzahl von Geräuschen aus einer jeweiligen Mehrzahl von Wählern für eine intelligente Lautsprechereinheit zu identifizieren, eine Visualisierung der Geräusche unter Verwendung einer AR-Einheit zu erzeugen, wobei eines oder mehrere der Geräusche unter Verwendung der Visualisierung ausgewählt werden können, und den erweiterten Sprachbefehl für die intelligente Lautsprechereinheit zu erzeugen, wobei der erweiterte Sprachbefehl das eine oder die mehreren Geräusche aufweist, die unter Verwendung der Visualisierung der AR-Einheit ausgewählt wurden.

[0010] Gemäß einer weiteren Ausführungsform der vorliegenden Erfindung enthält ein Computerprogrammprodukt, das ein durch einen Computer lesbares Speichermedium mit darauf befindlichen Programmanweisungen aufweist, ein Computerprogrammprodukt, das ein durch einen Computer lesbares Speichermedium mit darauf befindlichen Programmanweisungen aufweist, wobei die Programmanweisungen durch einen Computer lesbar und ausführbar sind, um zu bewirken, dass der Computer ein Verfahren durchführt, das Identifizieren einer Mehrzahl von Geräuschen aus einer jeweiligen Mehrzahl von Wählern für eine intelligente Lautsprechereinheit und Erzeugen einer Visualisierung der Geräusche unter Verwendung einer AR-Einheit umfasst. Eines oder mehrere der Geräusche können unter Verwendung der Visualisierung ausgewählt werden, und es erfolgt ein Erzeugen des erweiterten Sprachbefehls für die intelligente Lautsprechereinheit, wobei der erweiterte Sprachbefehl das eine oder die mehreren Geräusche

sche aufweist, die unter Verwendung der Visualisierung der AR-Einheit ausgewählt wurden.

[0011] Zum besseren Verständnis der ausführlichen Beschreibung der Erfindung und des vorliegenden Beitrags zu diesem Fachgebiet wurden einige Ausführungsformen der Erfindung in groben Zügen umrissen. Es gibt natürlich noch weitere Ausführungsformen der Erfindung, die nachstehend beschrieben werden und den Gegenstand der beiliegenden Ansprüche bilden.

[0012] Es versteht sich, dass die Erfindung in ihrer Anwendung nicht auf die Konstruktionsdetails und die Anordnungen der in der folgenden Beschreibung aufgeführten oder in den Zeichnungen dargestellten Komponenten beschränkt ist. Die Erfindung kann zusätzlich zu den beschriebenen Ausführungsformen weitere Ausführungsformen enthalten und auf verschiedene Art und Weise umgesetzt und ausgeführt werden. Zudem versteht es sich, dass die vorliegend verwendete Phraseologie und Terminologie sowie die Zusammenfassung Beschreibungszwecken dienen und nicht als einschränkend angesehen werden sollten.

[0013] Ein Fachmann versteht somit, dass das Konzept, auf dem diese Offenbarung beruht, ohne Weiteres als Grundlage für die Gestaltung anderer Strukturen, Verfahren und Systeme zur Durchführung der verschiedenen Zwecke der vorliegenden Erfindung verwendet werden kann. Es ist daher wichtig, die Ansprüche als solche äquivalenten Konstruktionen umfassend zu betrachten, soweit sie nicht vom Umfang der vorliegenden Erfindung abweichen.

KURZBESCHREIBUNG DER ZEICHNUNGEN

[0014] Die beispielhaften Aspekte der Erfindung werden aus der folgenden ausführlichen Beschreibung der beispielhaften Ausführungsformen der Erfindung unter Bezugnahme auf die Zeichnungen besser verständlich.

Fig. 1 veranschaulicht ein Kontextschaubild einer Ausführungsform der vorliegenden Erfindung.

Fig. 2 veranschaulicht das System einer Ausführungsform der vorliegenden Erfindung.

Fig. 3 zeigt ein Prozessablauf-/Arbeitsablauf-Schaubild einer Ausführungsform der vorliegenden Erfindung.

Fig. 4 veranschaulicht die Bestätigung von Befehlen durch das System aus **Fig. 3**.

Fig. 5 veranschaulicht das Erkennen eines Standortes eines Geräusches durch das System aus **Fig. 3**.

Fig. 6 veranschaulicht das Anzeigen von Befehlen auf der AR-Brille aus **Fig. 3**.

Fig. 7 veranschaulicht das Auswählen von Befehlen auf der AR-Brille in einer Ausführungsform der vorliegenden Erfindung aus **Fig. 3**.

Fig. 8 veranschaulicht, wie ein KI-Sprachassistent den Befehl aus der AR-Auswahl aus **Fig. 3** ausführt.

Fig. 9 zeigt eine Rückmeldungsschleife zum historischen Korpus aus **Fig. 3**.

Fig. 10 veranschaulicht eine zusammenfassende Übersicht einer Ausführungsform der vorliegenden Erfindung.

Fig. 11 veranschaulicht ein beispielhaftes Hardware-/ Informationsverarbeitungssystem, in das die beispielhafte Ausführungsform der vorliegenden Erfindung integriert werden kann.

Fig. 12 veranschaulicht ein signaltragendes Speichermedium zum Speichern von durch eine Maschine lesbaren Anweisungen eines Programms, das das Verfahren gemäß der beispielhaften Ausführungsform der vorliegenden Erfindung implementiert.

Fig. 13 zeigt einen Cloud-Computing-Knoten gemäß einer beispielhaften Ausführungsform der vorliegenden Erfindung.

Fig. 14 zeigt eine Cloud-Computing-Umgebung gemäß einer beispielhaften Ausführungsform der vorliegenden Erfindung.

Fig. 15 zeigt Abstraktionsmodellschichten gemäß einer beispielhaften Ausführungsform der vorliegenden Erfindung.

AUSFÜHRLICHE BESCHREIBUNG

[0015] Die Erfindung wird nun unter Bezugnahme auf die Figuren beschrieben, in denen sich gleiche Bezugszeichen durchgehend auf gleiche Teile beziehen. Es wird darauf hingewiesen, dass gemäß üblicher Praxis die verschiedenen Merkmale der Zeichnungen nicht unbedingt maßstabsgetreu sind, wie es in der Praxis üblich ist. Vielmehr können die Abmessungen der verschiedenen Merkmale aus Gründen der Übersichtlichkeit beliebig vergrößert oder verkleinert sein. Nachfolgend werden beispielhafte Ausführungsformen zur Veranschaulichung bereitgestellt, die die Ansprüche nicht einschränken. Darüber hinaus ist zu beachten, dass jeder der Schritte in unterschiedlicher Reihenfolge, kombiniert oder gleichzeitig durchgeführt werden kann. Darüber hinaus kann jede der gezeigten Strukturen und Ausführungsformen modifiziert oder kombiniert werden.

[0016] Wie vorstehend erwähnt, gibt es Fälle, in denen ein Benutzer einen beliebigen Sprachbefehl abgibt und andere Benutzer in der Umgebung wei-

tere Vorschläge oder zusätzliche Rückmeldungen zum Sprachbefehl des ersten Benutzers geben. In diesem Fall kann es vorkommen, dass der ursprüngliche Benutzer einige der Vorschläge anderer Benutzer in seinem Sprachbefehl nicht akzeptieren möchte. Derzeit ist das KI-Sprachassistenzsystem nicht in der Lage zu verstehen, welche Sprachbefehle es ausführen soll und welche es ignorieren kann. Die vorliegende Erfindung bietet eine Lösung wie folgt.

[0017] Fig. 1 veranschaulicht ein Kontextschaubild einer Ausführungsform der vorliegenden Erfindung.

[0018] Die vorliegende Erfindung liefert ein Verfahren und ein System, mit denen ein Benutzer 10 selektiv identifizieren kann, welche gesprochenen oder nicht gesprochenen Inhalte aus der Umgebung berücksichtigt werden sollen, während der Benutzer seinen Sprachbefehl abgibt.

[0019] Während ein Benutzer 10 einen Sprachbefehl abgibt, machen Personen (Benutzer 12, 14, 16) in der Umgebung gleichzeitig zusätzliche Vorschläge zu den Sprachbefehlen. Dementsprechend kann der Benutzer 10 selektiv ein oder mehrere Geräusche aus der Umgebung einbeziehen.

[0020] Beispielsweise gibt der Benutzer B 12 eine Sprachansage „Milch“, der Benutzer C 14 eine Sprachansage „Apfel“ und der Benutzer D 16 eine Sprachansage „Spielzeug“. Andererseits gibt der Hauptnutzer A 10 einen Sprachbefehl „Bitte bestelle Reis A, Kekse B und ein Kaltgetränk“ an das System 100.

[0021] Das KI- (künstliche Intelligenz) Sprachassistenzsystem 30 empfängt den Sprachbefehl „Bitte bestelle Reis, Kekse, ein Kaltgetränk, Milch und einen Apfel“. Der Benutzer verwendet eine AR-Brille und sieht die Sprachbefehle und andere Geräusche in der AR-Umgebung, um „Milch“ 22 und „Apfel“ 24 auszuwählen, während er „Spielzeug“ 26 in der AR-Auswahl 18 nicht auswählt, und wählt zudem den Befehl „Bitte bestelle Reis, Kekse und ein Kaltgetränk“ 20 aus, um die endgültige Ausgabe „Bitte bestelle Reis, Kekse, ein Kaltgetränk aus Milch und einen Apfel“ bereitzustellen. Die Auswahl des Spielzeugs 26 wird entfernt.

[0022] Fig. 2 veranschaulicht ein beispielhaftes System einer Ausführungsform der vorliegenden Erfindung. Der Benutzer A 10, der Benutzer B 12, der Benutzer C 14 und der Benutzer D 16 können mit dem System 100 über verbundene Einheiten des AR-Systems 40, des KI-Assistenten 30 und der IoT- (Internet of Things, Internet der Dinge) Einheit 50, die nachstehend näher beschrieben wird, Daten austauschen. Das System 100 kann einen Systemcomputer 102 enthalten, der einen Prozessor 64 und

einen Speicher 66 enthält. Ein historisches Korpus 60 kann im Speicher 66 enthalten sein oder das historische Korpus kann sich separat in einer Speichereinheit 62 befinden. Das AR-System 40 enthält ebenfalls einen Prozessor 42 mit Speicher 44, während der KI-Assistent 30 einen Prozessor 32 mit Speicher 34 enthält. Eine Mehrzahl von IoT-Einheiten 50 kann mit dem System 100 verbunden werden. Das AR-System 40 kann eine AR-Brille 46 enthalten (oder es kann sich dabei um eine separate IoT-Einheit 50 handeln). Die IoT-Einheiten 50 können einen Prozessor 52 mit einem Speicher 54 und Wandler 56 wie Mikrofone und Lautsprecher enthalten. Bei den IoT-Einheiten 50 kann es sich um eine Mehrzahl von Einheiten handeln, die jeweils auch eine Anzeige 58 und eine Mehrzahl von Sensoren 57 enthalten. Die Sensoren 57 können z.B. biometrische Daten oder andere Informationen erfassen, wie z.B. in einer Smartwatch oder einer anderen Einheit.

[0023] In alternativen Ausführungsformen können eines oder alle der folgenden Systeme des AR-Systems 40 und des KI-Assistenten 30 in den Systemcomputer 102 integriert sein. Zudem kann es sich bei der AR-Brille 46 um eine separate IoT-Einheit 50 handeln, die mit dem AR-System 40 verbunden ist.

[0024] Fig. 3 zeigt ein Prozessablauf-/Arbeitsablauf-Schaubild einer Ausführungsform der vorliegenden Erfindung. Wie in Fig. 3 unter Bezugnahme auf Fig. 1 und Fig. 2 dargestellt, wählt der Hauptbenutzer (z.B. der Benutzer A 10 aus Fig. 1 und Fig. 2) im ersten Abschnitt für das Verbinden von Einheiten und die Nutzeridentifikation 110 ein Modul und wird identifiziert 112. Der Benutzer 10 entscheidet sich für die Erfindung und ein historisches Korpus 60, 62 der Stimme des Benutzers 10 wird gestartet. Das Korpus 60, 62 sammelt bei 130 über den Pfad 134 den Tonfall von Benutzerbefehlen, die Häufigkeit von Benutzerbefehlen und die Kontextanalyse von Benutzerbefehlen vom System 100.

[0025] Dann wird das System 100 bei 114 mit einem AR- (erweiterte Realität) System 40, einem KI- (künstliche Intelligenz) Assistenten 30 und IoT- (Internet der Dinge) Einheiten 50 verbunden. Das AI-Sprachassistenzsystem 30 verbindet sich mit der AR-Brille 46. Das KI-Sprachassistenzsystem 30 erkennt den Benutzer 10 anhand einer StimmSignatur, und das AR-System 40 erkennt den Benutzer anhand einer auf einem Retina-Scan beruhenden IoT-Authentifizierung.

[0026] Wenn ein Benutzer (10, 12, 14, 16) eine andere IoT-Einheit 50 trägt, z.B. eine Smartwatch, können vom Benutzer gesammelte biometrische Daten den Benutzer (10, 12, 14, 16) anhand einer Nähe zum AR-System (40), des KI-Sprachassistenten-

ten (30) und anderer biometrischer Daten identifizieren.

[0027] Die Verbindung kann über jede Art von Kommunikationsprotokoll wie BLUETOOTH, WLAN, proprietäres Protokoll usw. erfolgen.

[0028] Dann gibt der Hauptbenutzer 10 einen Befehl, auf den Mitglieder in der Umgebung 116 folgen. Wenn ein Weckbefehl übermittelt wird, wacht das KI-Sprachassistenzsystem 30 auf und empfängt den Sprachbefehl.

[0029] Während der Nutzer 10 den Sprachbefehl abgibt, bleibt das KI-Sprachassistenzsystem 30 wach und empfängt den Sprachbefehl.

[0030] Die Aktivierung der AR-Brille 46 findet wie folgt statt. Sobald der KI-Sprachassistent 30 wach ist, veranlasst er die AR-Brille 46, „aufzuwachen“ und auf andere Geräusche zu hören (wenn die AR-Brille 46 ein Mikrofon 48 enthält).

[0031] Dann bestätigt das System 100 den Sprachbefehl und erfasst den genauen Befehl zusammen mit dem Standort im Raum 118. Während der Benutzer einen Sprachbefehl gibt, erfasst das KI-Sprachassistenzsystem 30, welche anderen Geräusche in der Umgebung erzeugt werden.

[0032] Fig. 4 veranschaulicht die Bestätigung von Befehlen durch das System aus Fig. 3.

[0033] Unter Bezugnahme auf Fig. 4, Fig. 3 und Fig. 2 erfolgt die Bestätigung der Befehle 118a wie folgt. Das AR-System 40 (sofern es über ein Mikrofon 48 verfügt) empfängt auch die Geräusche aus der Umgebung und bestätigt, dass die sekundären Befehle aus der Umgebung kommen. Hierfür wird der Unterschied zwischen Schallwellen analysiert und sichergestellt, dass sie innerhalb desselben Schwellenwertes liegen.

[0034] Mithilfe einer Fourier-Transformation 154 kann die Audiodatei skaliert werden, um ähnliche Dateien zu identifizieren:

$$f(t) = a_0 + \sum_{m=1}^{\infty} a_m \cos\left(\frac{2\pi mt}{T}\right) + \sum_{n=1}^{\infty} b_n \cos\left(\frac{2\pi nt}{T}\right)$$

[0035] Wenn die skalierten Gleichungen innerhalb eines definierten $\varepsilon > 0$ liegen, werden die beiden Audiowellen 156 bestätigt.

[0036] Eine weitere Bestätigung kann durch eine Textanalyse 158 erfolgen. Der Befehl, den das KI-Assistenzsystem 30 registriert, wird mit dem Befehl des AR-Systems 40 verglichen. Dies kann durch den Vergleich der Wörter geschehen, die jedes System

30, 40 erkennt, z.B. mit einem BOW- (Bag of Words) Algorithmus, um Ähnlichkeiten in den Befehlen zu erkennen.

[0037] Die Bestätigung der IoT-Einheit 50 findet wie folgt statt. Andere IoT-Einheiten 50, die mit dem System 100 verbunden sind, können die Informationen bestätigen.

[0038] Fig. 5 veranschaulicht das Erkennen eines Standortes eines Geräusches durch das System 100 am Punkt 118 aus Fig. 3. Wie in Fig. 5, Fig. 2 und Fig. 3 dargestellt, wird der Standort des Geräusches 118b wie folgt erkannt. Die AR-Brille 46 erkennt auch das Geräusch und wer (z.B. die Benutzer 10, 12, 14, 16 usw.) das Geräusch erzeugt.

[0039] Die Identifizierung kann über Einbindung sozialer Medien 204 erfolgen. Wenn die Benutzer (z.B. die Benutzer 10, 12, 14 und 16) auf Social-Media-Plattformen oder über die Kontaktliste eines Benutzers als „Freunde“ identifiziert werden, können sie auf diese Weise identifiziert werden.

[0040] Eine Objekterkennung 206 eines Menschen im Raum kann vom AR-System 40 durchgeführt werden, um zu identifizieren, wem (z.B. dem Benutzer 10, 12, 14, 16 usw.) das Geräusch zuzuordnen ist.

[0041] Eine Richtung des Geräusches 208 kann durch eine Richtung, in der die Schallwelle die höchste Amplitude enthält, näherungsweise bestimmt werden.

[0042] Die IoT-Bestätigung 210 findet wie folgt statt. Andere, mit Mikrofonen ausgestattete, verbundene IoT-Einheiten 50 können helfen, den Standort des Befehls zu triangulieren. Zu den IoT-Einheiten 50 können auch solche zählen, die in der Social-Media-Integration 204 vorhanden sind (d.h. Einheiten von „Freunden“ im selben Netzwerk können zur Standortbestimmung genutzt werden).

[0043] Durch Kombinieren dieser Parameter 204, 206, 208 und 210 lässt sich feststellen, wer welchen Befehl gegeben hat.

[0044] Unter Rückbezug auf Fig. 3 werden jedoch die Nichtbefehle 120 im zweiten Abschnitt 122 ignoriert. Das Ignorieren willkürlicher Geräusche umfasst, dass Geräusche, die nicht als Befehle erkannt werden, ignoriert werden. Dies kann z.B. schwache Geräusche oder andere Arten von Geräuschen usw. umfassen.

[0045] Die Befehle werden dann auf einem AR-Bildschirm im AR-System 40 angezeigt, und der Hauptbenutzer 10 wählt den endgültigen Befehl 124 aus.

[0046] Fig. 6 veranschaulicht das Anzeigen von Befehlen auf der AR-Brille 124a aus Fig. 3. Die erkannten Befehle werden auf der AR-Brille 46 angezeigt.

[0047] Welches Objekt den jeweiligen Befehl oder das jeweilige Geräusch erzeugt hat, wird auf der AR-Brille 46 angezeigt 220.

[0048] Die AR-Brille 46 zeigt die tatsächliche Position und Richtung der Umgebungsgeräusche zusammen mit dem Sprachbefehl in der AR-Umgebung 222 an.

[0049] Während der Abgabe des Sprachbefehls zeigt die AR-Brille das von der Umgebung erzeugte Geräusch in Echtzeit an und wird in der AR-Brille 224 angezeigt. Der Benutzer kann in der AR-Brille auch seinen Sprachbefehl in Text übersetzt und das von der Umgebung erzeugte Geräusch sehen 226.

[0050] Fig. 7 veranschaulicht das Auswählen von Befehlen auf der AR-Brille 124b in einer Ausführungsform der vorliegenden Erfindung aus Fig. 3.

[0051] Der Benutzer 10 kann Fingergesten, Augenkontakt, Sprachbefehle oder Gesichtsausdrücke mit beliebigen Umgebungsgeräuschen verwenden, um zu gestatten, dass diese Geräusche in den Sprachbefehl aufgenommen werden. Auf Grundlage der selektiven Aktionen des Benutzers werden ein oder mehrere Geräusche aus der Umgebung berücksichtigt, und es wird ein Hinzufügen dieser Geräusche zu dem ursprünglichen Sprachbefehl 230 gestattet.

[0052] Der Benutzer 10 kann die Geräusche aus den Geräuschen selektiv auswählen und diese Geräusche dem vom Benutzer 10 gegebenen Sprachbefehl entsprechend hinzufügen 232.

[0053] Die selektiven Geräusche können in den Sprachbefehl eingebunden werden, und der Sprachbefehl kann während der Eingabe des Sprachbefehls durch den Benutzer oder nach der Eingabe des Sprachbefehls eingebunden werden 234.

[0054] Das Auswählen eines Teils eines Geräusches 236 findet wie folgt statt. Während des selektiven Auswählens eines oder mehrerer Geräusche aus der Umgebung kann der Benutzer ein Sprachfragment aus einer beliebigen Quelle in der Umgebung auswählen und dementsprechend diese Geräusche dem ursprünglichen Sprachbefehl 236 hinzufügen.

[0055] Die Auswahl anhand einer Position anstelle von Text 238 findet wie folgt statt. Während des Auswählens eines beliebigen Geräusches aus der Umgebung kann der Benutzer 10 die Position des ursprünglichen Sprachbefehls auswählen, an der

die Geräusche dem Sprachbefehl 238 hinzugefügt werden sollen.

[0056] Fig. 8 veranschaulicht, wie ein KI-Sprachassistent den Befehl aus der AR-Auswahl aus Fig. 3 ausführt.

[0057] Gemäß Fig. 2, Fig. 3 und Fig. 8 führt der KI-Sprachassistent dann den Befehl aus der AR-Auswahl 126 aus. Der aggregierende endgültige Sprachbefehl und die Ausführung 242 finden wie folgt statt. Das vorgeschlagene System 100 aggregiert die aus der Umgebung ausgewählten Stimmen und aggregiert diese mit dem ursprünglichen Sprachbefehl, der vom Benutzer 10 abgegeben 244 wurde. Die AR-Anzeige des AR-Systems 40 zeigt dem Benutzer 10 den endgültigen Sprachbefehl an oder der KI-Assistent 30 wiederholt den endgültigen Sprachbefehl zur Ausführung 246. Dann führt das System 100 nach einer Validierung 248 den Sprachbefehl 250 aus.

[0058] Fig. 9 zeigt eine Rückmeldungsschleife zum historischen Korpus aus Fig. 3.

[0059] Gemäß Fig. 2, Fig. 3 und Fig. 9 wird das Szenario dann erfasst und wieder in das historische Korpus 132 eingespeist. Die Rückmeldungsschleife zum historischen Korpus sieht aus wie folgt. Das KI-Sprachassistenzsystem 30 und das AR-Glas 46 lernen historisch, wie die Umgebungsgeräusche mit dem übermittelten Sprachbefehl kontextuell zusammenhängen, und identifizieren, welche Geräusche geeignet sind und hinzugefügt werden können.

[0060] Wenn bei der Übermittlung eines Sprachbefehls ein oder mehrere Geräusche aus der Umgebung erzeugt werden, bestimmt das vorgeschlagene System 100, welche Geräusche in den Sprachbefehl 252 eingebunden werden sollen.

[0061] Auf Grundlage von historischem Lernen zeigt das vorgeschlagene System dann ein visuelles Schaubild, um zu zeigen, wo das erzeugte Geräusch in den Sprachbefehl aufgenommen werden soll, so dass der Benutzer nachvollziehen kann, wo das Geräusch hinzugefügt werden soll 254.

[0062] Die Interaktion mit sozialen Medien 256 ist wie folgt. Wenn ein sekundärer Benutzer 12, 14 oder 16 viele nützliche Befehlszusätze äußert, dann werden diese mit größerer Wahrscheinlichkeit hinzugefügt, wenn der Hauptbenutzer einen Sprachbefehl gibt und umgekehrt (wenn ein sekundärer Benutzer viele nichtnützliche Befehle hinzufügt, dann werden diese künftig nicht zu Sprachbefehlen hinzugefügt) 256.

[0063] Gemäß Fig. 3 wird das historische Korpus 130 auf Grundlage von Benutzereingaben und

Benutzerinteraktionen mit dem System 100 aus Pfad 134 und dem erfassten Szenario 132 erstellt.

[0064] Eine weitere Ausführungsform kann einen Familienmitglied-Modus umfassen. Die Stimmen von Familienmitgliedern werden dann erkannt und habe eine niedrigere Schwelle für die Zulassung eines Sprachbefehls. Alternativ kann auch der Sprachassistent 30 in das AR-System 40 integriert sein.

[0065] Gemäß **Fig. 2** kann ein Anwendungsfall im Bereich Fertigung folgendermaßen aussehen. Wenn jemand in einer Fertigungsanlage arbeitet und mit einem AR-System 40 ausgestattet ist, werden verschiedene Befehle gegeben, welche Arbeiten zu erledigen sind. Diese Befehle erscheinen auf dem Bildschirm, und der Benutzer kann auswählen, welche Befehle befolgt werden müssen und welche nur Hintergrundgeräusche sind. Das Modul des AR-Systems 40 aggregiert dann diese Befehle zu einem einzigen und stellt diesen als endgültigen Befehl für den Benutzer 10 bereit.

[0066] **Fig. 10** veranschaulicht eine zusammenfassende Übersicht einer Ausführungsform der vorliegenden Erfindung.

[0067] Einige der Merkmale der vorliegenden Erfindung können zusammengefasst werden wie folgt. Zunächst erfolgt in Schritt 302 Übersetzen von Sprachbefehlen an ein AR-System 40 durch das System 100. Während der Abgabe eines Sprachbefehls erkennt das KI-Sprachassistentensystem 40 einzelne Geräusche auf Grundlage von Stimmsignaturen. Dementsprechend wird es erkennen, welche Befehle vom Benutzer 10, 12, 14, 16 kommen und welche in der Umgebung erzeugt werden. Ferner werden die abgegebenen Sprachbefehle auf der AR-Brille 46 angezeigt, um ein oder mehrere Geräusche in den ursprünglichen Sprachbefehl einzubinden.

[0068] Dann wählt das System 100 externe Befehle im AR-Raum 304 aus. Während der Abgabe des Sprachbefehls oder nach der Abgabe des Sprachbefehls kann der Benutzer 10 selektiv einen oder mehrere gesprochene Inhalte, die im AR-System 40 visualisiert werden, einbeziehen, und dementsprechend werden diese im Sprachbefehl berücksichtigt und ausgeführt.

[0069] Anschließend sorgt das System 100 für die Ausführung von Sprachbefehlen in einem AR-Raum 306. Während der Abgabe eines Sprachbefehls durch einen Benutzer 10, 12, 14, 16 zeigt die AR-Brille 46 die Geräuscherzeugung in der Umgebung in Echtzeit an, und dementsprechend kann der Benutzer 10 auf Grundlage von Finger-, Gesichts- und Augenkontaktinteraktion selektiv ein oder meh-

rere Geräusche in den Sprachbefehl aufnehmen, und das KI-Sprachassistentensystem 40 führt dementsprechend den vollständigen Sprachbefehl aus.

[0070] Der Benutzer 10 kann bei der Abgabe eines Sprachbefehls einen gesamten oder einen Teil eines gesprochenen Inhalts aus einer beliebigen Geräuschquelle auswählen, woraufhin das ausgewählte Geräuschfragment in den Sprachbefehl eingebunden und der gesamte Sprachbefehl ausgeführt wird. Beim selektiven Einbinden eines oder mehrerer Geräusche aus der Umgebung zusammen mit dem Sprachbefehl kann der Benutzer eine Position in dem Sprachbefehl auswählen, an der die externen Geräusche selektiv eingebunden werden können, um den Sprachbefehl zu vervollständigen.

[0071] Anschließend stellt das System 100 ein historisches Korpus zum Auswählen von Befehlen 308 bereit. Auf Grundlage von historischem Lernen über die selektive Einbindung eines oder mehrerer Geräusche in den abgegebenen Sprachbefehl empfiehlt das vorgeschlagene System 100, welche Geräusche in den Sprachbefehl aufgenommen werden sollten, und zeigt zudem ein visuelles Schaubild an, um zu empfehlen, wo das Geräusch in den ursprünglichen Sprachbefehl aufgenommen werden sollte.

[0072] Die folgenden **Fig. 11** bis **Fig. 15** zeigen zusätzliche Hardwarekonfigurationen, die implementiert werden können. Unterschiedliche Merkmale, die in unterschiedlichen Figuren der **Fig. 1** bis **Fig. 15** dargestellt sind, können kombiniert, geändert oder zwischen den verschiedenen Beispielen ausgetauscht werden.

[0073] **Fig. 11** veranschaulicht eine andere Hardwarekonfiguration des Systems 100, in der ein Informationsverarbeitungs-/Computersystem 1100 gemäß der vorliegenden Erfindung vorhanden ist, das bevorzugt über mindestens einen Prozessor oder eine Zentraleinheit (CPU) 1110 verfügt, die die erfindungsgemäßen Methoden implementieren kann.

[0074] Die CPUs 1110 sind über einen Systembus 1112 mit einem Direktzugriffsspeicher (RAM) 1114, einem Nur-Lese-Speicher (ROM) 1116, einem Eingabe/Ausgabe-Adapter (E/A-Adapter) 1118 (zum Anschluss von Peripherieeinheiten wie Platteneinheiten 1121 und Bandlaufwerken 1140 an den Bus 1112), einem Benutzerschnittstellenadapter 1122 (zum Anschluss einer Tastatur 1124, einer Maus 1126, eines Lautsprechers 1128, eines Mikrofons 1132 und/oder anderer Benutzerschnittstelleneinheiten an den Bus 1112), einem Datenübertragungsadapter 1134 zum Anschluss eines Informationsverarbeitungssystems an ein Datenverarbeitungsnetzwerk, das Internet, ein Intranet, ein privates Netzwerk (personal area network, PAN) usw., und einem Anzeigeadapter 1136

zum Anschluss des Busses 1112 an eine Anzeigeeinheit 1138 und/oder einen Drucker 1139 (z.B. einen Digitaldrucker oder dergleichen) verbunden.

[0075] Zusätzlich zu der vorstehend beschriebenen Hardware/Software-Umgebung umfasst ein anderer Aspekt der Erfindung ein durch einen Computer implementiertes Verfahren zum Durchführen des vorstehenden Verfahrens. Dieses Verfahren kann zum Beispiel in der vorstehend beschriebenen konkreten Umgebung implementiert werden.

[0076] Ein solches Verfahren kann beispielsweise durch Betreiben eines Computers in Form einer digitalen Datenverarbeitungsvorrichtung zur Ausführung einer Folge durch eine Maschine lesbarer Anweisungen durchgeführt werden. Diese Anweisungen können sich in verschiedenen Arten von signaltragenden Medien befinden.

[0077] Daher ist dieser Aspekt der vorliegenden Erfindung auf ein programmiertes Produkt gerichtet, das signaltragende Speichermedien enthält, die ein Programm mit durch eine Maschine lesbaren Anweisungen materiell enthalten, die durch einen digitalen Datenprozessor ausführbar sind, der die vorstehend genannte CPU 1110 und Hardware enthält, um das erfindungsgemäße Verfahren durchzuführen.

[0078] Diese signaltragenden Speichermedien können z.B. einen in der CPU 1110 enthaltenen RAM enthalten, wie er beispielsweise durch den Schnellzugriffsspeicher repräsentiert wird.

[0079] Alternativ können die Anweisungen auch in anderen signaltragenden Speichermedien 1200 enthalten sein, z.B. einer magnetischen Datendiskette 1210 oder einer optischen Speicherdiskette 1220 (**Fig. 12**), auf die die CPU 1210 direkt oder indirekt zugreifen kann.

[0080] Unabhängig davon, ob sie auf der Diskette 1210, der optischen Platte 1220, dem Computer/der CPU 1210 oder anderswo enthalten sind, können die Anweisungen auf einer Vielzahl durch eine Maschine lesbarer Datenspeichermedien gespeichert sein.

[0081] Bei der vorliegenden Erfindung kann es daher sich um ein System, ein Verfahren und/oder ein Computerprogrammprodukt handeln. Das Computerprogrammprodukt kann (ein) durch einen Computer lesbare(s) Speichermedium (oder -medien) umfassen, auf dem/denen durch einen Computer lesbare Programmanweisungen gespeichert ist/sind, um einen Prozessor dazu zu veranlassen, Aspekte der vorliegenden Erfindung auszuführen.

[0082] Bei dem durch einen Computer lesbaren Speichermedium kann es sich um eine physische Einheit handeln, die Anweisungen zur Verwendung

durch eine Einheit zur Ausführung von Anweisungen behalten und speichern kann. Bei dem durch einen Computer lesbaren Speichermedium kann es sich zum Beispiel um eine elektronische Speichereinheit, eine magnetische Speichereinheit, eine optische Speichereinheit, eine elektromagnetische Speichereinheit, eine Halbleiterspeichereinheit oder jede geeignete Kombination daraus handeln, ohne auf diese beschränkt zu sein. Zu einer nicht erschöpfenden Liste spezifischerer Beispiele des durch einen Computer lesbaren Speichermediums gehören die Folgenden: eine tragbare Computerdiskette, eine Festplatte, ein Direktzugriffsspeicher (RAM), ein Nur-Lese-Speicher (ROM), ein löschbarer programmierbarer Nur-Lese-Speicher (EPROM bzw. Flash-Speicher), ein statischer Direktzugriffsspeicher (SRAM), ein tragbarer Kompaktspeicherplatte-Nur-Lese-Speicher (CD-ROM), eine DVD (digital versatile disc), ein Speicher-Stick, eine Diskette, eine mechanisch codierte Einheit wie zum Beispiel Lochkarten oder gehobene Strukturen in einer Rille, auf denen Anweisungen gespeichert sind, und jede geeignete Kombination daraus. Ein durch einen Computer lesbares Speichermedium soll in der Verwendung hierin nicht als flüchtige Signale an sich aufgefasst werden, wie zum Beispiel Funkwellen oder andere sich frei ausbreitende elektromagnetische Wellen, elektromagnetische Wellen, die sich durch einen Wellenleiter oder ein anderes Übertragungsmedium ausbreiten (z.B. ein Lichtwellenleiterkabel durchlaufende Lichtimpulse) oder durch einen Draht übertragene elektrische Signale.

[0083] Hierin beschriebene, durch einen Computer lesbare Programmanweisungen können von einem durch einen Computer lesbaren Speichermedium auf jeweilige Datenverarbeitungs-/Verarbeitungseinheiten oder über ein Netzwerk wie zum Beispiel das Internet, ein lokales Netzwerk, ein Weitverkehrsnetz und/oder ein drahtloses Netzwerk auf einen externen Computer oder eine externe Speichereinheit heruntergeladen werden. Das Netzwerk kann Kupferübertragungskabel, Lichtwellenübertragungsleiter, drahtlose Übertragung, Leitwegrechner, Firewalls, Vermittlungseinheiten, Gateway-Computer und/oder Edge-Server umfassen. Eine Netzwerkadapterkarte oder Netzwerkschnittstelle in jeder Datenverarbeitungs-/Verarbeitungseinheit empfängt durch einen Computer lesbare Programmanweisungen aus dem Netzwerk und leitet die durch einen Computer lesbaren Programmanweisungen zur Speicherung in einem durch einen Computer lesbaren Speichermedium innerhalb der entsprechenden Datenverarbeitungs-/Verarbeitungseinheit weiter.

[0084] Bei dem durch einen Computer lesbaren Programmanweisungen zum Ausführen von Arbeitsschritten der vorliegenden Erfindung kann es sich um Assembler-Anweisungen, ISA-Anweisungen (Instruction-Set-Architecture), Maschinenanweisungen

gen, maschinenabhängige Anweisungen, Mikrocode, Firmware-Anweisungen, zustandssetzende Daten oder entweder Quellcode oder Objektcode handeln, die in einer beliebigen Kombination aus einer oder mehreren Programmiersprachen geschrieben werden, darunter objektorientierte Programmiersprachen wie Smalltalk, C++ o.ä. sowie herkömmliche prozedurale Programmiersprachen wie die Programmiersprache „C“ oder ähnliche Programmiersprachen. Die durch einen Computer lesbaren Programmanweisungen können vollständig auf dem Computer des Benutzers, teilweise auf dem Computer des Benutzers, als eigenständiges Software-Paket, teilweise auf dem Computer des Benutzers und teilweise auf einem fernen Computer oder vollständig auf dem fernen Computer oder Server ausgeführt werden. In letzterem Fall kann der entfernt angeordnete Computer mit dem Computer des Nutzers durch eine beliebige Art Netzwerk verbunden sein, darunter ein lokales Netzwerk (LAN) oder ein Weitverkehrsnetz (WAN), oder die Verbindung kann mit einem externen Computer hergestellt werden (zum Beispiel über das Internet unter Verwendung eines Internet-Diensteanbieters). In einigen Ausführungsformen können elektronische Schaltungen, darunter zum Beispiel programmierbare Logikschaltungen, vor Ort programmierbare Gatter-Anordnungen (FPGA, field programmable gate arrays) oder programmierbare Logikanordnungen (PLA, programmable logic arrays) die durch einen Computer lesbaren Programmanweisungen ausführen, indem sie Zustandsinformationen der durch einen Computer lesbaren Programmanweisungen nutzen, um die elektronischen Schaltungen zu personalisieren, um Aspekte der vorliegenden Erfindung durchzuführen.

[0085] Aspekte der vorliegenden Erfindung sind hierin unter Bezugnahme auf Ablaufpläne und/oder Blockschaubilder von Verfahren, Vorrichtungen (Systemen) und Computerprogrammprodukten gemäß Ausführungsformen der Erfindung beschrieben. Es versteht sich, dass jeder Block der Ablaufpläne und/oder der Blockschaubilder sowie Kombinationen von Blöcken in den Ablaufplänen und/oder den Blockschaubildern mittels durch einen Computer lesbarer Programmanweisungen ausgeführt werden können.

[0086] Diese durch einen Computer lesbaren Programmanweisungen können einem Prozessor eines Universalcomputers, eines Spezialcomputers oder einer anderen programmierbaren Datenverarbeitungsvorrichtung bereitgestellt werden, um eine Maschine zu erzeugen, so dass die über den Prozessor des Computers bzw. der anderen programmierbaren Datenverarbeitungsvorrichtung ausgeführten Anweisungen ein Mittel zur Umsetzung der in dem Block bzw. den Blöcken der Ablaufpläne und/oder der Blockschaubilder bzw. Schaubilder festgelegten Funktionen/Schritte erzeugen.

[0087] Diese durch einen Computer lesbaren Programmanweisungen können auch auf einem durch einen Computer lesbaren Speichermedium gespeichert sein, das einen Computer, eine programmierbare Datenverarbeitungsvorrichtung und/oder andere Einheiten so steuern kann, dass sie auf eine bestimmte Art funktionieren, so dass das durch einen Computer lesbare Speichermedium, auf dem Anweisungen gespeichert sind, ein Herstellungsprodukt enthält, darunter Anweisungen, welche Aspekte der/des in dem Block bzw. den Blöcken des Ablaufplans und/oder der Blockschaubilder bzw. Schaubilder angegebenen Funktion/Schritts umsetzen.

[0088] Die durch einen Computer lesbaren Programmanweisungen können auch auf einen Computer, eine andere programmierbare Datenverarbeitungsvorrichtung oder eine andere Einheit geladen werden, um das Ausführen einer Reihe von Prozessschritten auf dem Computer bzw. der anderen programmierbaren Vorrichtung oder anderen Einheit zu verursachen, um einen auf einem Computer ausgeführten Prozess zu erzeugen, so dass die auf dem Computer, einer anderen programmierbaren Vorrichtung oder einer anderen Einheit ausgeführten Anweisungen die in dem Block bzw. den Blöcken der Ablaufpläne und/oder der Blockschaubilder festgelegten Funktionen/Schritte umsetzen.

[0089] Die Ablaufpläne und die Blockschaubilder in den Figuren veranschaulichen die Architektur, die Funktionalität und den Betrieb möglicher Ausführungen von Systemen, Verfahren und Computerprogrammprodukten gemäß verschiedenen Ausführungsformen der vorliegenden Erfindung. In diesem Zusammenhang kann jeder Block in den Ablaufplänen oder Blockschaubildern bzw. Schaubildern ein Modul, ein Segment oder einen Teil von Anweisungen darstellen, die eine oder mehrere ausführbare Anweisungen zur Ausführung der bestimmten logischen Funktion(en) enthalten. In einigen alternativen Ausführungen können die in dem Block angegebenen Funktionen in einer anderen Reihenfolge als in den Figuren gezeigt stattfinden. Zwei nacheinander gezeigte Blöcke können zum Beispiel in Wirklichkeit im Wesentlichen gleichzeitig ausgeführt werden, oder die Blöcke können manchmal je nach entsprechender Funktionalität in umgekehrter Reihenfolge ausgeführt werden. Es ist ferner anzumerken, dass jeder Block der Blockschaubilder und/oder der Ablaufpläne sowie Kombinationen aus Blöcken in den Blockschaubildern und/oder den Ablaufplänen durch spezielle auf Hardware beruhende Systeme umgesetzt werden können, welche die festgelegten Funktionen oder Schritte durchführen, oder Kombinationen aus Spezial-Hardware und Computeranweisungen ausführen.

[0090] In Fig. 13 ist ein Blockschaubild 1400 eines Beispiels eines Cloud-Computing-Knotens gezeigt.

Bei dem Cloud-Computing-Knoten 1400 handelt es sich lediglich um ein Beispiel eines geeigneten Cloud-Computing-Knotens, und dieses soll in keiner Weise eine Eingrenzung bezüglich des Verwendungs- oder Funktionalitätsumfangs von Ausführungsformen der hier beschriebenen Erfindung andeuten. Unabhängig davon kann der Cloud-Computing-Knoten 1400 für jedwede vorstehend angegebene Funktionalität realisiert werden und/oder diese ausführen.

[0091] Im Cloud-Computing-Knoten 1400 ist ein Computersystem/Server 1412 vorhanden, welches/r zusammen mit zahlreichen weiteren Universal- oder Spezial-Datenverarbeitungssystem-Umgebungen oder-Konfigurationen betriebsfähig ist. Zu Beispielen bekannter Datenverarbeitungssysteme, -umgebungen und/oder -konfigurationen, die sich für die Verwendung mit dem Computersystem/Server 1412 eignen können, zählen, ohne hierauf beschränkt zu sein, Personal-Computer-Systeme, Servercomputersysteme, Thin Clients, Thick Clients, tragbare oder Laptop-Einheiten, Multiprozessorsysteme, Systeme aus Mikroprozessoren, Set-Top-Boxen, programmierbare Consumer-Elektronik, Netzwerk-PCs, Minicomputer-Systeme, Großrechnersysteme und verteilte Cloud-Computing-Umgebungen, die jedwede der vorstehend genannten Systeme oder Einheiten umfassen, und dergleichen.

[0092] Computersystem/Server 1412 kann im allgemeinen Kontext durch ein Computersystem ausführbarer Anweisungen beschrieben werden, beispielsweise Programmmodule, die durch ein Computersystem ausgeführt werden. Allgemein enthalten Programmmodule Routinen, Programme, Objekte, Komponenten, Logik, Datenstrukturen und so weiter, die bestimmte Aufgaben erfüllen oder bestimmte abstrakte Datentypen implementieren. Computersystem/Server 1412 können in verteilten Cloud-Computing-Umgebungen betrieben werden, in denen Aufgaben durch entfernt angeordnete Verarbeitungseinheiten ausgeführt werden, die durch ein Kommunikationsnetzwerk verbunden sind. In einer verteilten Cloud-Computing-Umgebung können sich Programmmodule sowohl in lokalen als auch in entfernt angeordneten Speichermedien des Computersystems einschließlich Kurzzeit-Speichereinheiten befinden.

[0093] Wie in **Fig. 13** gezeigt, wird Computersystem/Server 1412 in dem Cloud-Computing-Knoten 1400 in Form einer Universal-Datenverarbeitungseinheit gezeigt. Zu den Komponenten des Computersystems/Servers 1412 können, ohne jedoch hierauf eingeschränkt zu sein, ein oder mehrere Prozessoren oder Verarbeitungseinheiten 1416, ein Systemspeicher 1428 und ein Bus 1418 zählen, der verschiedene Systemkomponenten einschließlich des

Systemspeichers 1428 mit einem Prozessor 1416 verbindet.

[0094] Der Bus 1418 stellt eine oder mehrere von verschiedenen Arten von Bus-Strukturen dar, darunter ein Speicherbus oder eine Speichersteuereinheit, ein Peripherie-Bus, ein AGP-Bus und ein Prozessor- oder lokaler Bus, der jedwede einer Vielfalt von Busarchitekturen verwendet. Als nicht einschränkende Beispiele gehören zu solchen Architekturen der ISA- (Industry-Standard-Architecture-) Bus, der Mikrokanal- (Micro-Channel-Architecture, MCA-) Bus, der EISA- (Enhanced-ISA-) Bus, VESA- (Video-Electronics-Standards-Association-) Localbus und der PCI- (Peripheral-Component-Interconnect-) Bus.

[0095] Computersystem/Server 1412 umfasst üblicherweise eine Vielfalt durch ein Computersystem lesbarer Medien. Bei solchen Medien kann es sich um jedwedes verfügbare Medium handeln, auf das von dem Computersystem/Server 1412 zugegriffen werden kann, und dazu zählen sowohl flüchtige als auch nichtflüchtige Medien, entfernbare sowie nicht entfernbare Medien.

[0096] Der Systemspeicher 1428 kann durch ein Computersystem lesbare Medien in Form eines flüchtigen Speichers wie beispielsweise Arbeitsspeicher (RAM) 1430 und/oder Cache 1432 umfassen. Computersystem/Server 1412 kann ferner weitere auswechselbare/nicht auswechselbare, flüchtige/nichtflüchtige Speichermedien des Computersystems umfassen. Lediglich beispielhaft kann das Speichersystem 1434 zum Lesen von einem sowie Schreiben auf ein nicht entfernbare nichtflüchtiges magnetisches Medium (nicht gezeigt und üblicherweise als „Festplattenspeicher“ bezeichnet) bereitgestellt sein. Auch wenn dies nicht gezeigt ist, können ein Magnetplattenlaufwerk zum Lesen von einer und Schreiben auf eine auswechselbare, nichtflüchtige Magnetplatte (z.B. eine „Floppy Disk“) und ein optisches Plattenlaufwerk zum Lesen von einer oder Schreiben auf eine auswechselbare, nichtflüchtige optische Platte wie beispielsweise einen CD-ROM, DVD-ROM oder andere optische Medien bereitgestellt sein. In solchen Fällen kann jedes der Genannten über eine oder mehrere Daten-Medien-Schnittstellen mit dem Bus 1418 verbunden sein. Wie nachstehend weiter dargestellt und beschrieben wird, kann der Speicher 1428 mindestens ein Programmprodukt mit einer Gruppe (z.B. mindestens einem) von Programmmodulen umfassen, die so konfiguriert sind, dass sie die Funktionen von Ausführungsformen der Erfindung ausführen.

[0097] Als nicht einschränkendes Beispiel können in dem Speicher 1428 ein Programm/Dienstprogramm 1440 mit einer Gruppe (mindestens einem) von Programmmodulen 1442 sowie ein Betriebssystem, ein

oder mehrere Anwendungsprogramme, weitere Programmmodule und Programmdateien gespeichert sein. Das Betriebssystem, ein oder mehrere Anwendungsprogramme, andere Programmmodule und Programmdateien oder eine Kombination aus diesen können jeweils eine Implementierung einer Netzwerkumgebung enthalten. Allgemein führen die Programmmodule 1442 die Funktionen und/oder Methoden aus Ausführungsformen der Erfindung wie vorliegend beschrieben aus.

[0098] Computersystem/Server 1412 kann zudem mit einer oder mehreren externen Einheiten 1414 wie beispielsweise einer Tastatur, einer Zeigeeinheit, einer Anzeige 1424 usw., einer oder mehreren Einheiten, die einem Nutzer eine Interaktion mit dem Computersystem/Server 1412 ermöglichen, und/oder jedweden Einheiten (z.B. Netzwerkkarte, Modem usw.) in Datenaustausch stehen, die dem Computersystem/Server 1412 einen Datenaustausch mit einer oder mehreren anderen Datenverarbeitungseinheiten ermöglichen. Ein solcher Datenaustausch kann über Ein/Ausgabe- (E/A-) Schnittstellen 1422 erfolgen. Dennoch kann Computersystem/Server 1412 über einen Netzwerkadapter 1420 in Datenaustausch mit einem oder mehreren Netzwerken wie beispielsweise einem lokalen Netz (LAN), einem allgemeinen Weitverkehrsnetz (WAN) und/oder einem öffentlichen Netz (z.B. dem Internet) stehen. Wie dargestellt, tauscht der Netzwerkadapter 1420 über den Bus 1418 Daten mit den anderen Komponenten des Computersystems/Servers 1412 aus. Es sei angemerkt, dass, auch wenn dies nicht gezeigt ist, andere Hardware- und/oder Software-Komponenten zusammen mit dem Computersystem/Server 1412 verwendet werden könnten. Zu Beispielen zählen, ohne auf diese beschränkt zu sein, Mikrocode, Gerätetreiber, redundante Verarbeitungseinheiten, externe Plattenlaufwerk-Arrays, RAID-Systeme, Bandlaufwerke und Speichersysteme zur Datenarchivierung usw.

[0099] Gemäß Fig. 14 ist eine veranschaulichende Cloud-Computing-Umgebung 1550 abgebildet. Wie gezeigt ist, weist die Cloud-Computing-Umgebung 1550 einen oder mehrere Cloud-Computing-Knoten 1400 auf, mit denen von Cloud-Nutzern verwendete lokale Datenverarbeitungseinheiten wie der elektronische Assistent (PDA, personal digital assistant) oder das Mobiltelefon 1554A, der Desktop-Computer 1554B, der Laptop-Computer 1554C und/oder das Automobil-Computer-System 1554N Daten austauschen können. Die Knoten 1400 können miteinander Daten austauschen. Sie können physisch oder virtuell in ein oder mehrere Netzwerke wie private, Benutzergemeinschafts-, öffentliche oder hybride Clouds gruppiert werden (nicht gezeigt), wie vorstehend beschrieben wurde, oder in eine Kombination daraus. Dies ermöglicht es der Cloud-Computing-Umgebung 1550, Infrastruktur, Plattformen und/oder

Software als Dienste anzubieten, für die ein Cloud-Nutzer keine Ressourcen auf einer lokalen Datenverarbeitungseinheit vorhalten muss. Es sei darauf hingewiesen, dass die Arten von in Fig. 14 gezeigten Datenverarbeitungseinheiten 1554A-N lediglich veranschaulichend sein sollen und dass die Datenverarbeitungsknoten 1400 und die Cloud-Computing-Umgebung 1550 über eine beliebige Art Netzwerk und/oder über eine beliebige Art von über ein Netzwerk aufrufbarer Verbindung (z.B. unter Verwendung eines Web-Browsers) mit einer beliebigen Art von computergestützter Einheit Daten austauschen können.

[0100] Unter Bezugnahme auf Fig. 15 wird ein Satz von funktionalen Abstraktionsschichten gezeigt, die durch die Cloud-Computing-Umgebung 1550 (Fig. 14) bereitgestellt werden. Es sollte vorab klar sein, dass die in Fig. 15 gezeigten Komponenten, Schichten und Funktionen lediglich veranschaulichend sein sollen und Ausführungsformen der Erfindung nicht darauf beschränkt sind. Wie abgebildet ist, werden die folgenden Schichten und entsprechenden Funktionen bereitgestellt:

[0101] Eine Hardware- und Software-Schicht 1660 enthält Hardware- und Software-Komponenten. Zu Beispielen für Hardware-Komponenten zählen Großrechner, in einem Beispiel zSeries®-Systeme von IBM®, Server mit RISC- (Reduced Instruction Set Computer, Computer mit reduziertem Befehlssatz) Architektur, in einem Beispiel pSeries®-Systeme von IBM, xSeries®-Systeme von IBM, BladeCenter®-Systeme von IBM, Speichereinheiten, Netzwerke und Vernetzungskomponenten. Zu Beispielen für Software-Komponenten zählen Netzwerk-Anwendungsserversoftware, in einem Beispiel die Anwendungsserversoftware WebSphere® von IBM, und Datenbank-Software, in einem Beispiel die Datenbank-Software DB2® von IBM. (IBM, zSeries, pSeries, xSeries, BladeCenter, WebSphere und DB2 sind in vielen Rechtssystemen weltweit eingetragene Marken der International Business Machines Corporation.)

[0102] Die Virtualisierungsschicht 1662 stellt eine Abstraktionsschicht bereit, von der aus folgende Beispiele virtueller Einheiten bereitgestellt werden können: virtuelle Server, virtueller Speicher, virtuelle Netzwerke, einschließlich virtueller privater Netzwerke, virtuelle Anwendungen und Betriebssysteme und virtuelle Clients.

[0103] In einem Beispiel kann die Verwaltungsschicht 1664 die nachstehend beschriebenen Funktionen bereitstellen. Ressourcenbereitstellung stellt dynamische Beschaffung von Datenverarbeitungsressourcen und anderen Ressourcen bereit, die verwendet werden, um innerhalb der Cloud-Computing-Umgebung Aufgaben auszuführen. Messung und

Preisermittlung stellen Kostenüberwachung für innerhalb der Cloud-Computing-Umgebung genutzte Ressourcen sowie Abrechnung und Rechnungsstellung für die Inanspruchnahme dieser Ressourcen bereit. In einem Beispiel können diese Ressourcen Anwendungs-Software-Lizenzen aufweisen. Die Sicherheit stellt die Identitätsüberprüfung für Cloud-Kunden und Aufgaben sowie Schutz für Daten und andere Ressourcen bereit. Ein Nutzerportal stellt Zugang zur Cloud-Computing-Umgebung für Kunden und Systemadministratoren bereit. Dienstgüterverwaltung stellt eine Zuteilung und Verwaltung von Cloud-Computing-Ressourcen derart bereit, dass erforderliche Dienstgüter erfüllt werden. Planung und Umsetzung von Dienstgüte-Agreements (SLA) stellen eine Vorbereitung und Beschaffung von Cloud-Computing-Ressourcen, für die mit einem künftigen Bedarf gerechnet wird, gemäß einem SLA bereit.

[0104] Eine Workload-Schicht 1666 stellt Beispiele für die Funktionalität bereit, für welche die Cloud-Computing-Umgebung verwendet werden kann. Zu Beispielen für Workloads und Funktionen, die von dieser Schicht aus bereitgestellt werden können, gehören Funktionen wie Kartierung und Navigation, Softwareentwicklung und Lebenszyklusverwaltung, Bereitstellung von Ausbildung in virtuellen Unterrichtsräumen, Datenanalyseverarbeitung, Transaktionsverarbeitung und, insbesondere im Hinblick auf die vorliegende Erfindung, die APIs und Laufzeitsystemkomponenten zum Erzeugen von Vorschlägen für automatische Suchervollständigung auf Grundlage von Kontexteingaben.

[0105] Die vielen Merkmale und Vorteile der Erfindung sind aus der ausführlichen Beschreibung ersichtlich, und daher sollen die beiliegenden Ansprüche alle Merkmale und Vorteile der Erfindung abdecken, die in den tatsächlichen Umfang der Erfindung fallen. Da für einen Fachmann zahlreiche Modifikationen und Variationen ohne Weiteres ersichtlich sind, ist es nicht erwünscht, die Erfindung auf die exakte Konstruktion und Funktionsweise wie dargestellt und beschrieben einzuschränken, und dementsprechend kann auf alle geeigneten Modifikationen und Äquivalente zurückgegriffen werden, die in den Umfang der Erfindung fallen.

[0106] Es versteht sich, dass die Erfindung in ihrer Anwendung nicht auf die Konstruktionsdetails und die Anordnungen der in der folgenden Beschreibung aufgeführten oder in den Zeichnungen dargestellten Komponenten beschränkt ist. Die Erfindung kann zusätzlich zu den beschriebenen Ausführungsformen weitere Ausführungsformen enthalten und auf verschiedene Art und Weise umgesetzt und ausgeführt werden. Zudem versteht es sich, dass die vorliegend verwendete Phraseologie und Terminologie sowie die Zusammenfassung Beschreibungszwe-

cken dienen und nicht als einschränkend angesehen werden sollten.

[0107] Ein Fachmann versteht somit, dass das Konzept, auf dem diese Offenbarung beruht, ohne Weiteres als Grundlage für die Gestaltung anderer Strukturen, Verfahren und Systeme zur Durchführung der verschiedenen Zwecke der vorliegenden Erfindung verwendet werden kann. Es ist daher wichtig, die Ansprüche als solche äquivalenten Konstruktionen umfassend zu betrachten, soweit sie nicht vom Umfang der vorliegenden Erfindung abweichen.

Patentansprüche

1. Verfahren zum Erzeugen eines erweiterten Sprachbefehls, umfassend:
Identifizieren einer Mehrzahl von Geräuschen von einer jeweiligen Mehrzahl von Wandlern für eine intelligente Einheit;
Erzeugen einer Visualisierung der Geräusche unter Verwendung einer AR-Einheit, wobei eines oder mehrere der Geräusche unter Verwendung der Visualisierung auswählbar sind; und
Erzeugen des erweiterten Sprachbefehls für die intelligente Lautsprechereinheit, wobei der erweiterte Sprachbefehl das eine oder die mehreren Geräusche aufweist, die unter Verwendung der Visualisierung der AR-Einheit ausgewählt werden.
2. Verfahren nach Anspruch 1, ferner umfassend automatisches Auswählen des einen oder der mehreren Geräusche auf Grundlage einer Auswahl- und Abwahlhistorie historischer Geräusche von einem oder mehreren der Wandler.
3. Verfahren nach Anspruch 1, wobei jeder der Mehrzahl von Wandlern einen Lautsprecher aufweist und die intelligente Einheit eine intelligente Lautsprechereinheit aufweist, und wobei die Geräusche gesprochene oder nicht gesprochene Inhalte enthalten, die in einem AR-Raum ausgewählt werden.
4. Verfahren nach Anspruch 1, ferner umfassend Auswählen des einen oder der mehreren Geräusche auf Grundlage einer Auswahl- und Abwahlhistorie historischer Informationen, die in einem Speicher gespeichert sind.
5. Verfahren nach Anspruch 1, ferner umfassend Übersetzen der Sprachbefehle in die Erzeugung der Visualisierung der Geräusche unter Verwendung einer AR-Einheit.
6. Verfahren nach Anspruch 1, ferner umfassend Ausführen des erweiterten Sprachbefehls durch eine Auswahl eines oder mehrerer Geräusche durch ein AR-System.

7. Verfahren nach Anspruch 1, ferner umfassend:

während des Abgebens des Sprachbefehls oder nach dem Abgeben des Sprachbefehls erfolgendes Auswählen eines oder mehrerer gesprochener Inhalte, die in der AR-Einheit visualisiert werden, die entsprechend in dem erweiterten Sprachbefehl zur Ausführung berücksichtigt werden; und Auswählen des einen oder der mehreren Geräusche auf Grundlage einer Auswahl- und Abwahlhistorie historischer Geräusche von einem oder mehreren der Wandler, die in einem historischen Korpus gespeichert sind.

8. Verfahren nach Anspruch 1, das in einer Cloud-Infrastruktur implementiert wird.

9. System zum Erzeugen eines erweiterten Sprachbefehls, wobei das System aufweist: einen Speicher, der Computeranweisungen speichert; und einen Prozessor, der so konfiguriert ist, dass er die Computeranweisungen ausführt, um: eine Mehrzahl von Geräuschen von einer jeweiligen Mehrzahl von Wandlern für eine intelligente Einheit zu identifizieren; eine Visualisierung der Geräusche unter Verwendung einer AR-Einheit zu erzeugen, wobei eines oder mehrere der Geräusche unter Verwendung der Visualisierung auswählbar sind; und den erweiterten Sprachbefehl für die intelligente Lautsprechereinheit zu erzeugen, wobei der erweiterte Sprachbefehl das eine oder die mehreren Geräusche aufweist, die unter Verwendung der Visualisierung der AR-Einheit ausgewählt werden.

10. System nach Anspruch 8, ferner umfassend automatisches Auswählen des einen oder der mehreren Geräusche auf Grundlage einer Auswahl- und Abwahlhistorie historischer Geräusche von einem oder mehreren der Wandler.

11. System nach Anspruch 8, wobei jeder der Mehrzahl von Wandlern einen Lautsprecher aufweist und die intelligente Einheit eine intelligente Lautsprechereinheit aufweist, und wobei die Geräusche gesprochene oder nicht gesprochene Inhalte enthalten, die in einem AR-Raum ausgewählt werden.

12. System nach Anspruch 8, ferner umfassend Auswählen des einen oder der mehreren Geräusche auf Grundlage einer Auswahl- und Abwahlhistorie historischer Informationen, die in dem Speicher gespeichert sind.

13. System nach Anspruch 8, ferner umfassend Übersetzen der Sprachbefehle in die Erzeugung der Visualisierung der Geräusche unter Verwendung einer AR-Einheit.

14. System nach Anspruch 8, ferner umfassend: Ausführen des erweiterten Sprachbefehls durch eine Auswahl eines oder mehrerer Geräusche durch ein AR-System;

während des Abgebens des Sprachbefehls oder nach dem Abgeben des Sprachbefehls erfolgendes Auswählen eines oder mehrerer gesprochener Inhalte, die in der AR-Einheit visualisiert werden, die entsprechend in dem erweiterten Sprachbefehl zur Ausführung berücksichtigt werden; und Auswählen des einen oder der mehreren Geräusche auf Grundlage einer Auswahl- und Abwahlhistorie historischer Geräusche von einem oder mehreren der Wandler, die in einem historischen Korpus gespeichert sind.

15. Computerprogrammprodukt, das ein durch einen Computer lesbares Speichermedium mit darauf enthaltenen Programmanweisungen aufweist, wobei die Programmanweisungen durch einen Computer lesbar und ausführbar sind, um den Prozessor zum Durchführen eines Verfahrens zu veranlassen, das Folgendes umfasst: Identifizieren einer Mehrzahl von Geräuschen von einer jeweiligen Mehrzahl von Wandlern für eine intelligente Lautsprechereinheit; Erzeugen einer Visualisierung der Geräusche unter Verwendung einer AR-Einheit, wobei eines oder mehrere der Geräusche unter Verwendung der Visualisierung auswählbar sind; und Erzeugen des erweiterten Sprachbefehls für die intelligente Lautsprechereinheit, wobei der erweiterte Sprachbefehl das eine oder die mehreren Geräusche aufweist, die unter Verwendung der Visualisierung der AR-Einheit ausgewählt werden.

16. Computerprogrammprodukt nach Anspruch 15, ferner umfassend automatisches Auswählen des einen oder der mehreren Geräusche auf Grundlage einer Auswahl- und Abwahlhistorie historischer Geräusche von einem oder mehreren der Wandler.

17. Computerprogrammprodukt nach Anspruch 15, wobei jeder der Mehrzahl von Wandlern einen Lautsprecher aufweist und die intelligente Einheit eine intelligente Lautsprechereinheit aufweist, und wobei die Geräusche gesprochene oder nicht gesprochene Inhalte enthalten, die in einem AR-Raum ausgewählt werden.

18. Computerprogrammprodukt nach Anspruch 15, ferner umfassend Auswählen des einen oder der mehreren Geräusche auf Grundlage einer Auswahl- und Abwahlhistorie historischer Informationen, die in einem Speicher gespeichert sind.

19. Computerprogrammprodukt nach Anspruch 15, ferner umfassend Übersetzen der Sprachbefehle in die Erzeugung der Visualisierung der Geräusche unter Verwendung einer AR-Einheit.

20. Computerprogrammprodukt nach Anspruch 15, ferner umfassend:
Ausführen des erweiterten Sprachbefehls durch eine Auswahl eines oder mehrerer Geräusche durch ein AR-System;
während des Abgebens des Sprachbefehls oder nach dem Abgeben des Sprachbefehls erfolgendes Auswählen eines oder mehrerer gesprochener Inhalte, die in der AR-Einheit visualisiert werden, die entsprechend in dem erweiterten Sprachbefehl zur Ausführung berücksichtigt werden; und
Auswählen des einen oder der mehreren Geräusche auf Grundlage einer Auswahl- und Abwahlhistorie historischer Geräusche von einem oder mehreren der Wandler, die in einem historischen Korpus gespeichert sind.

Es folgen 15 Seiten Zeichnungen

Anhängende Zeichnungen

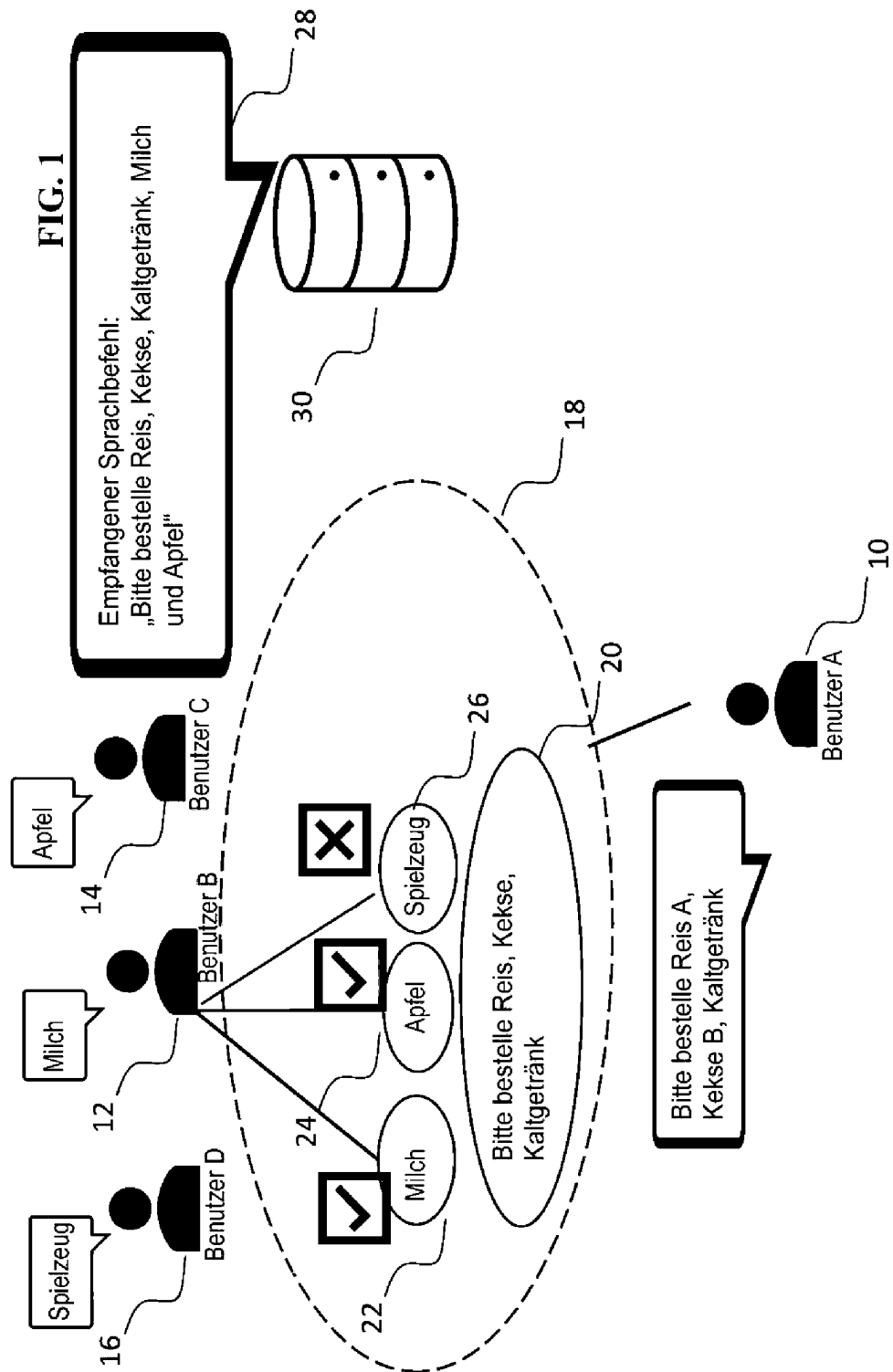


FIG. 2

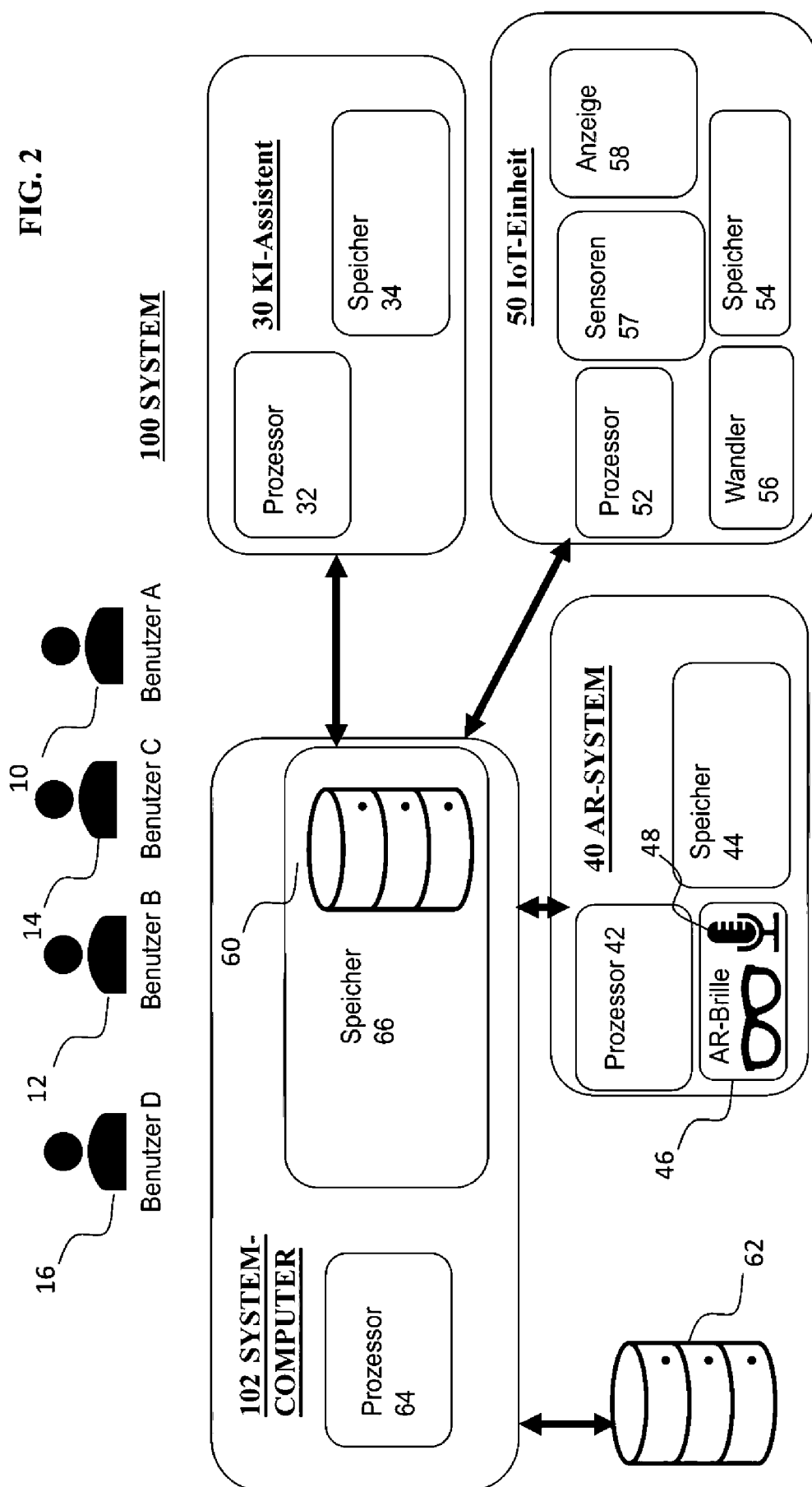


FIG. 3

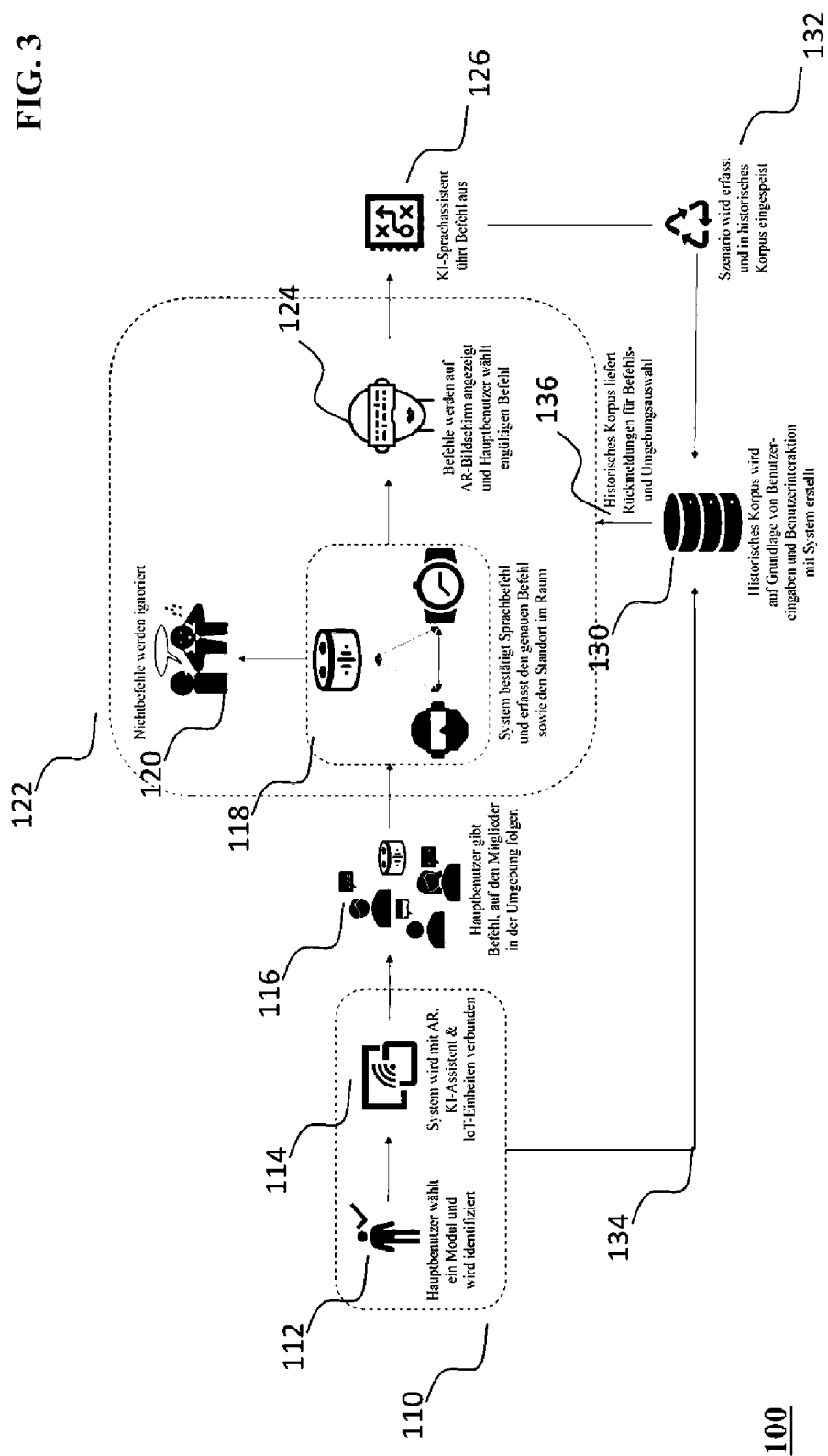


FIG. 4

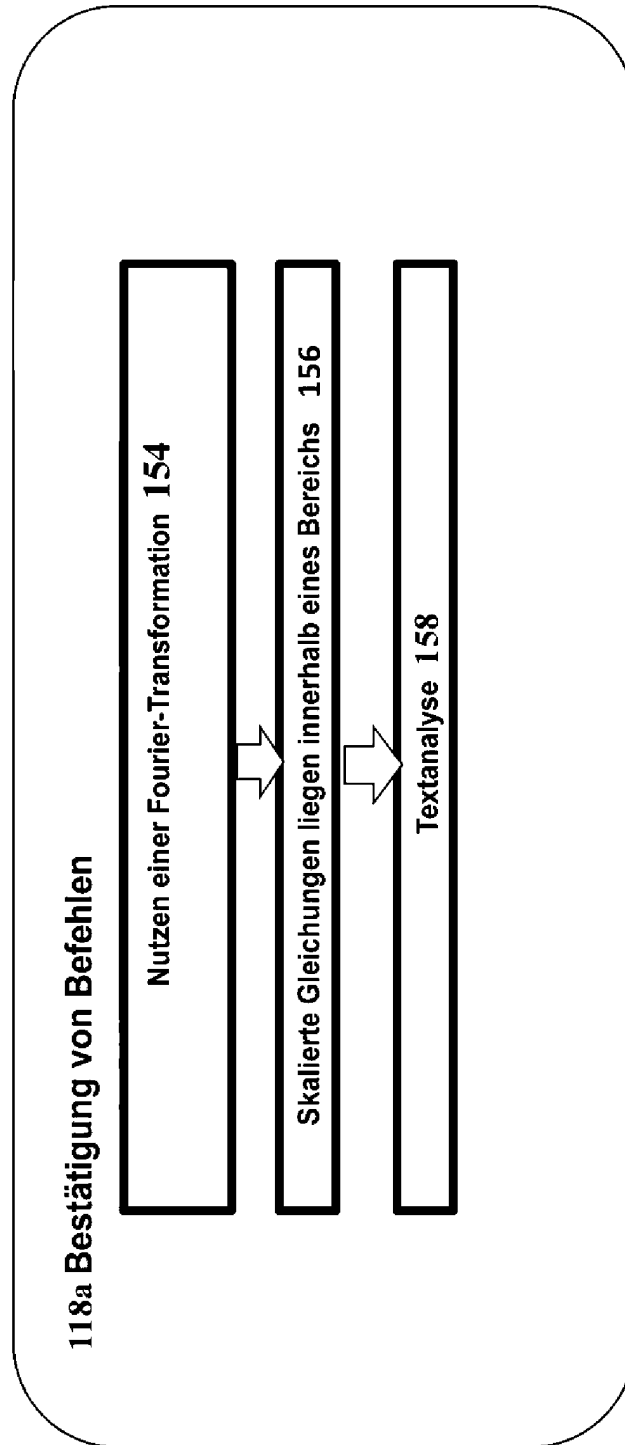


FIG. 5

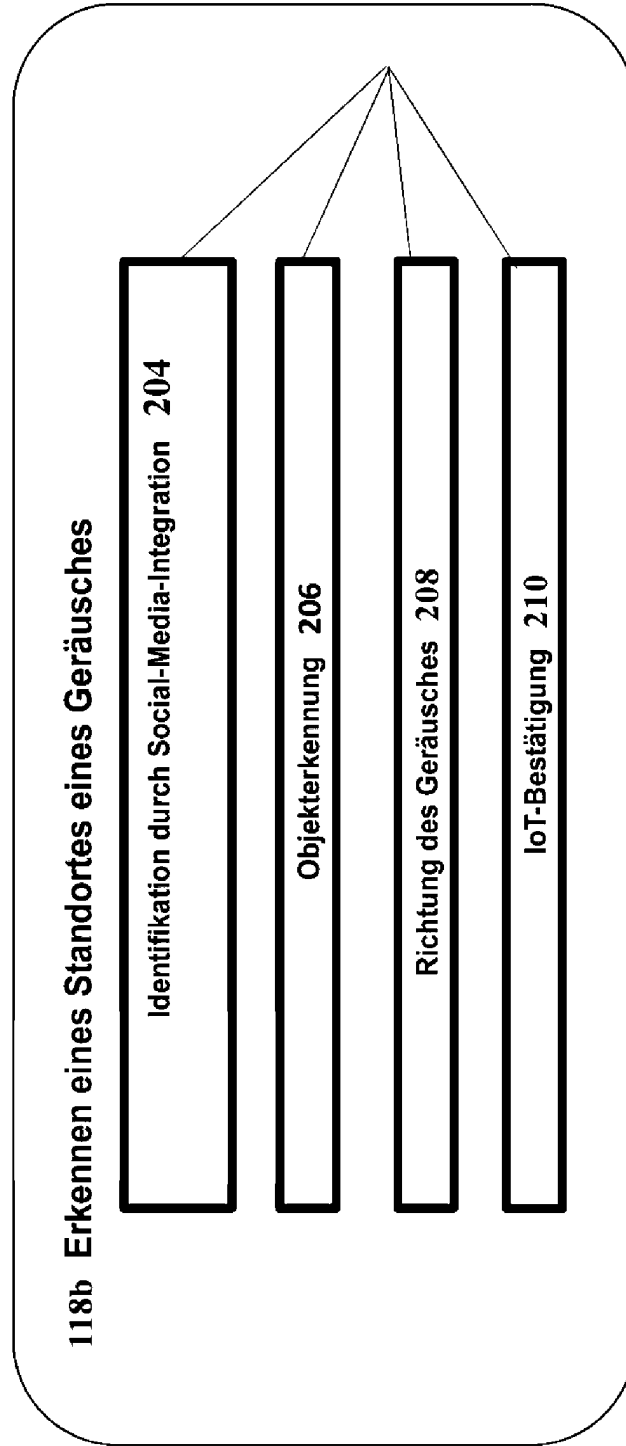


FIG. 6

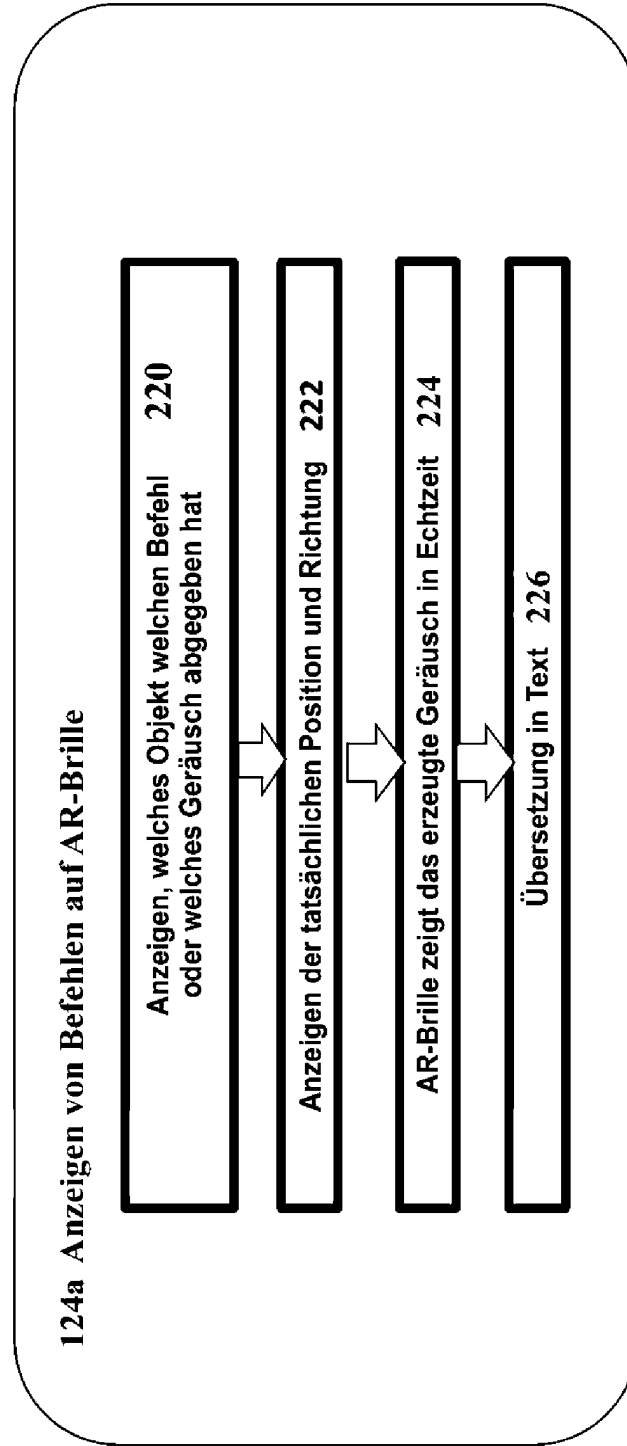


FIG. 7

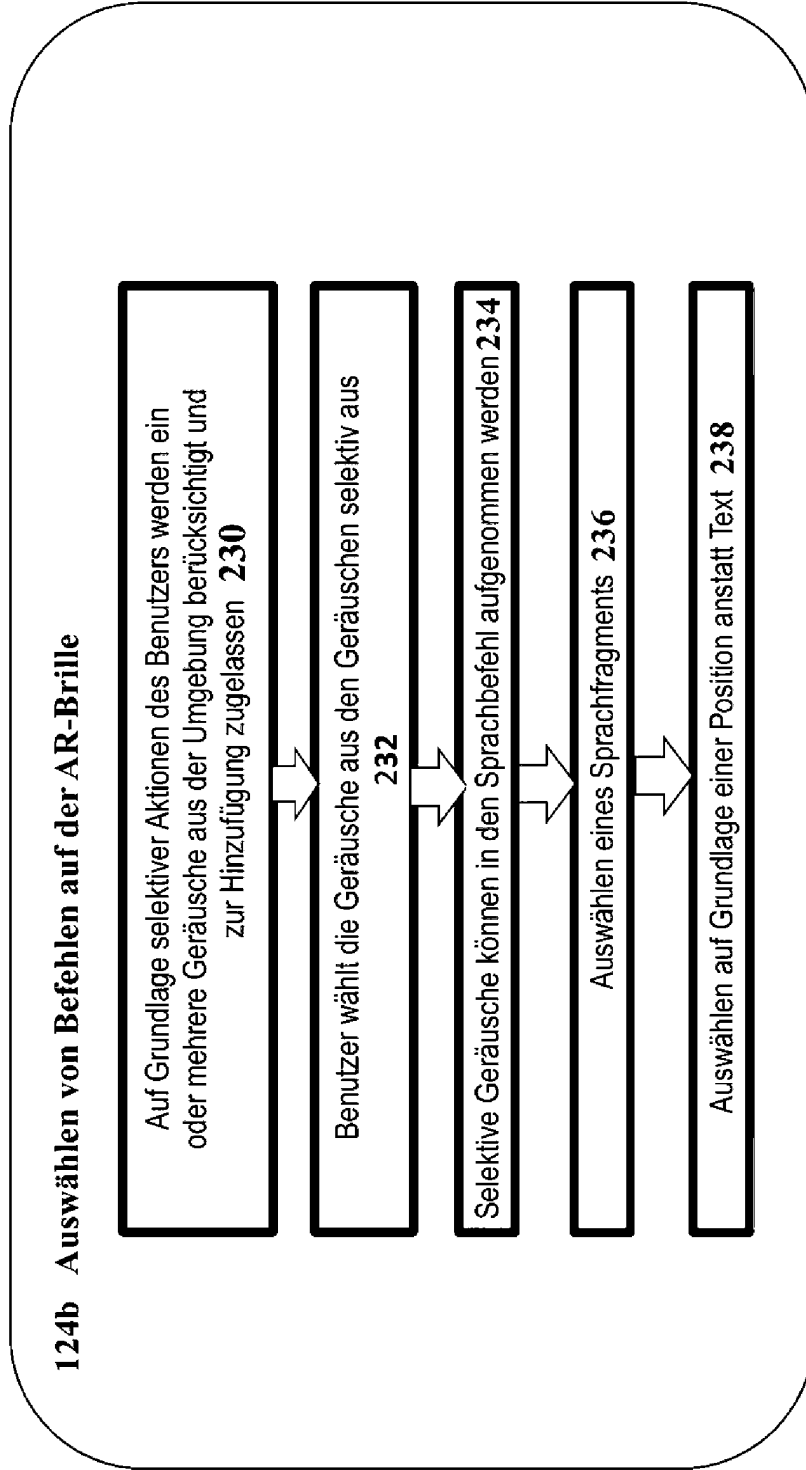


FIG. 8

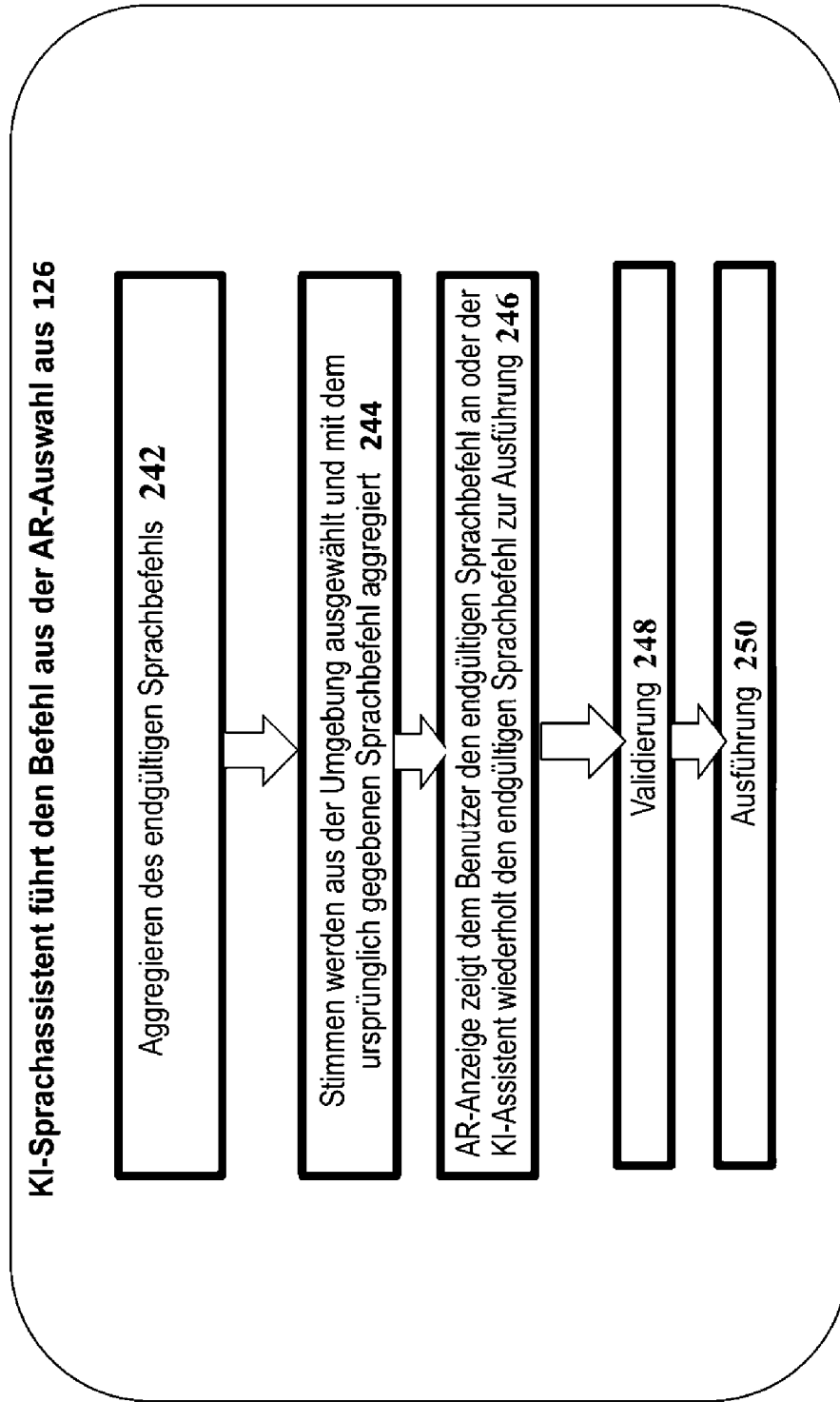


FIG. 9

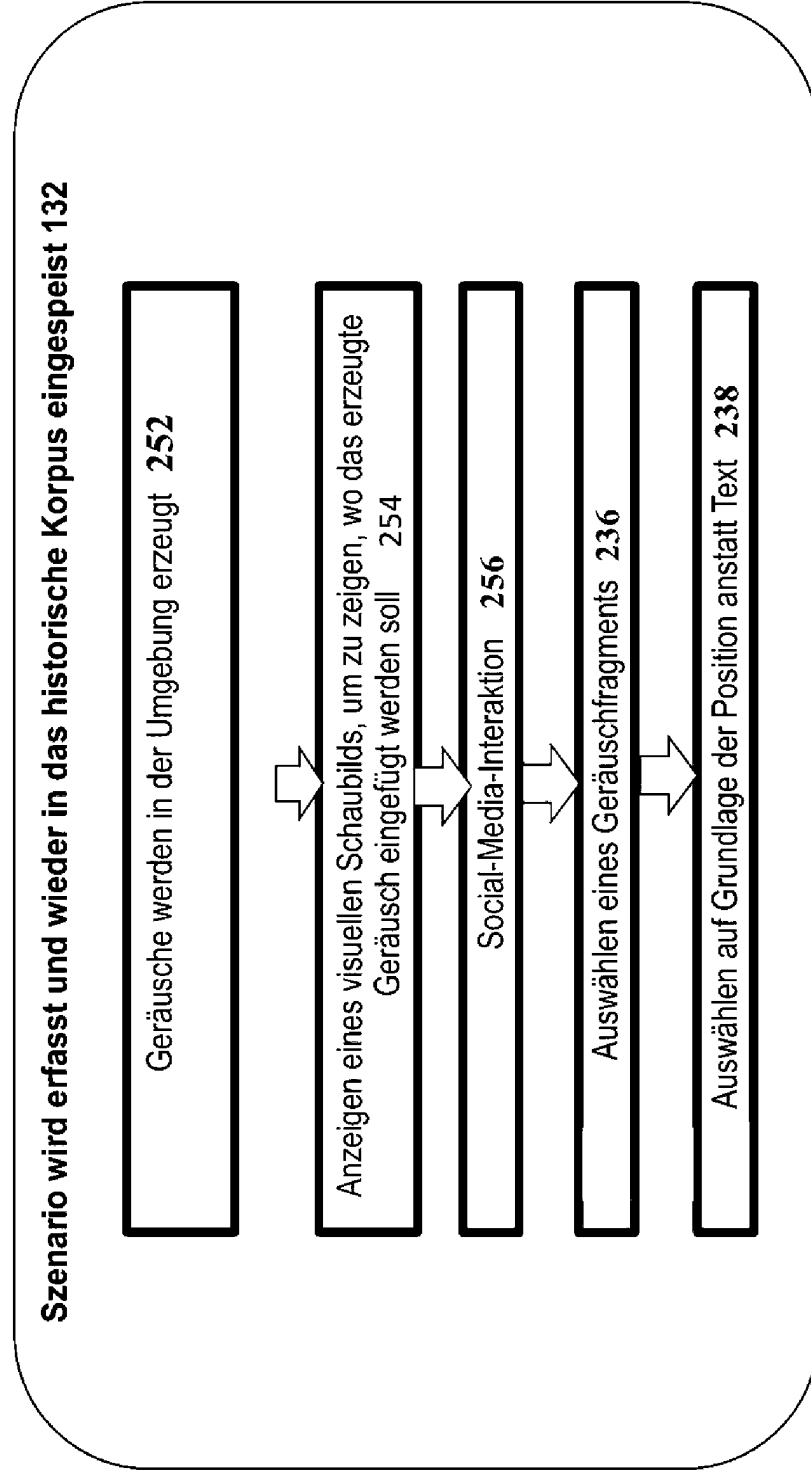
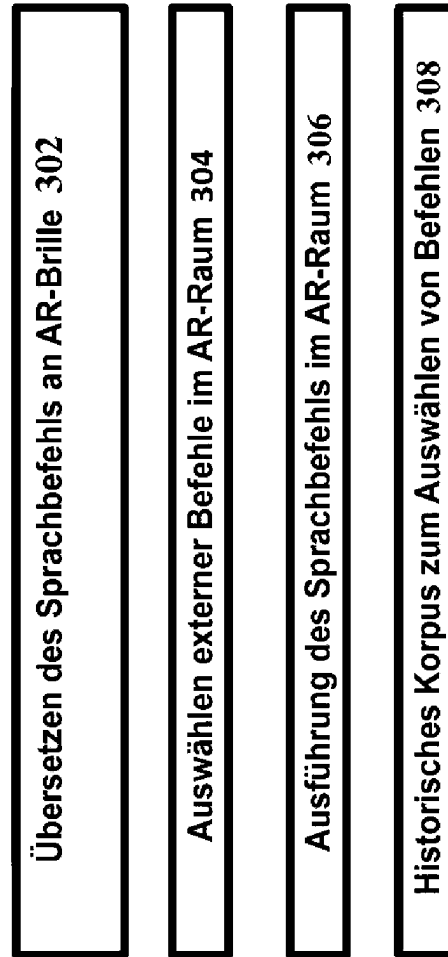


FIG. 10



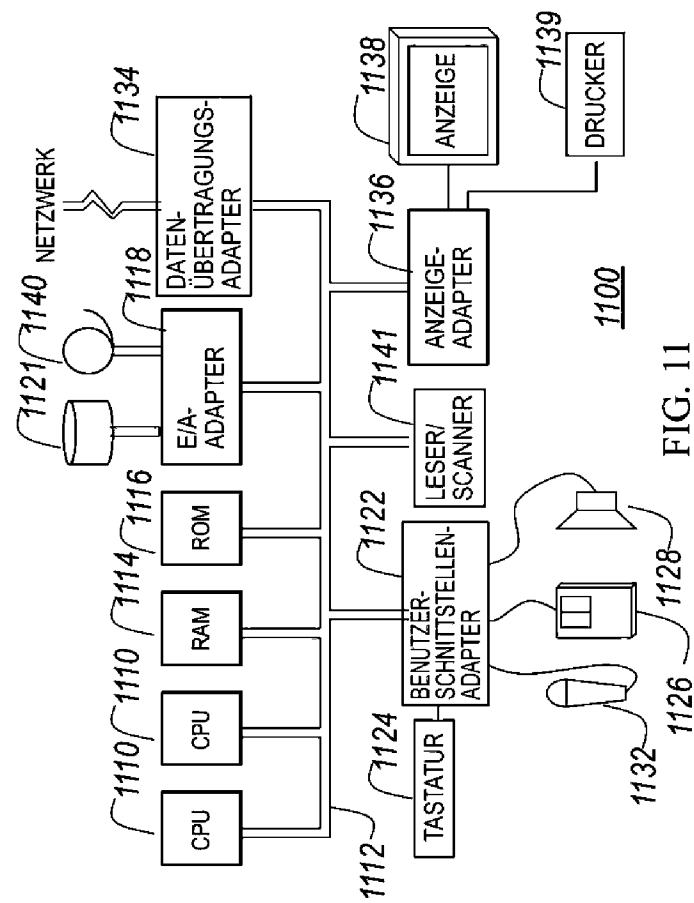
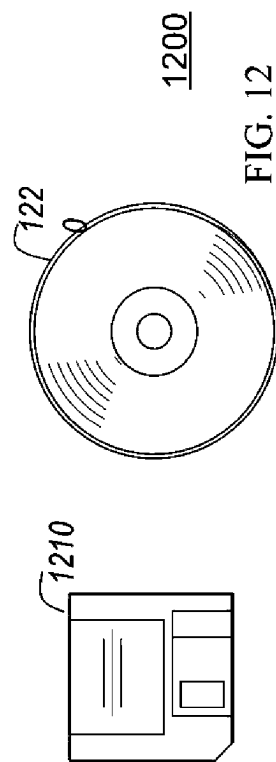


FIG. 11



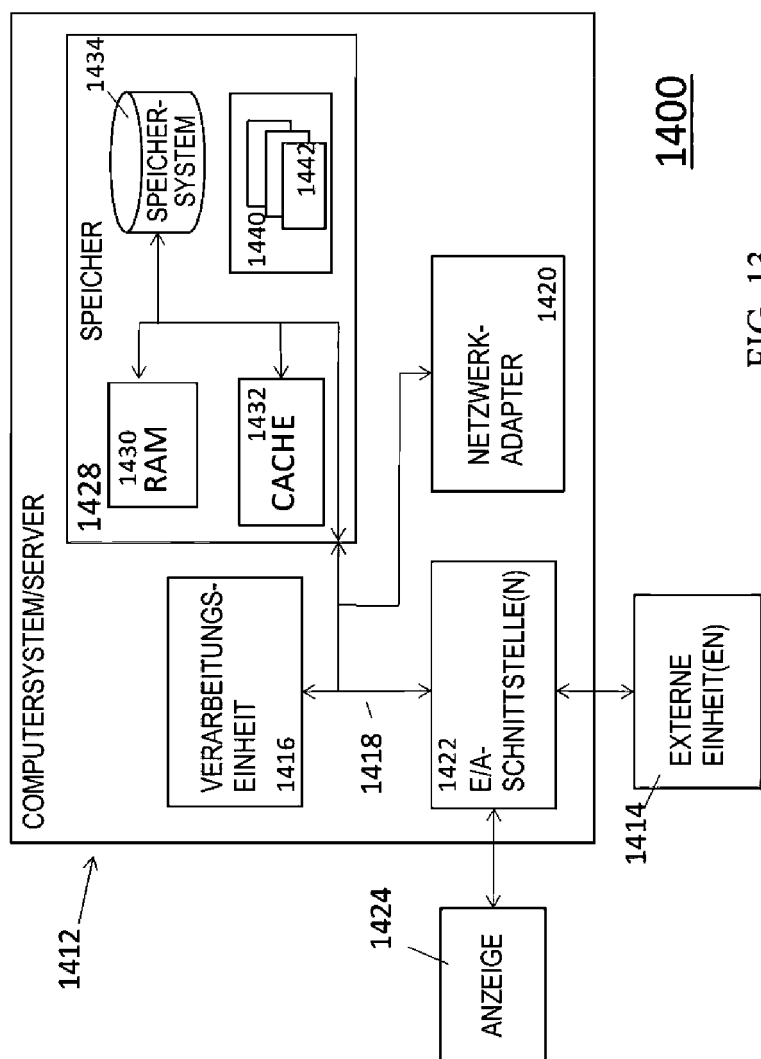
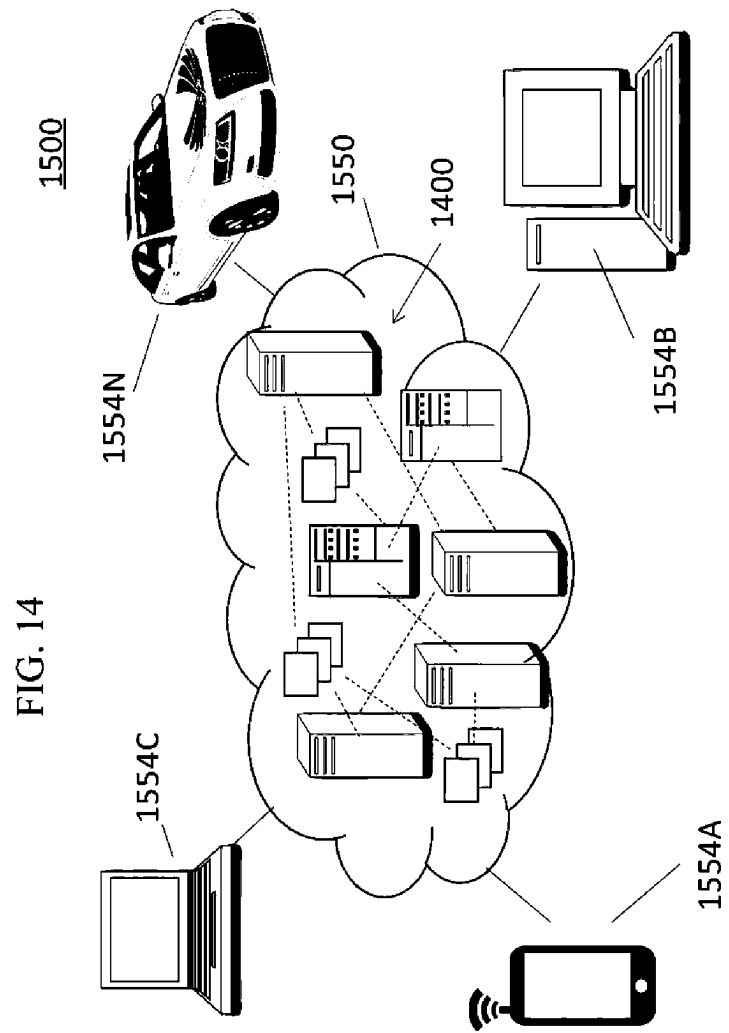


FIG. 13



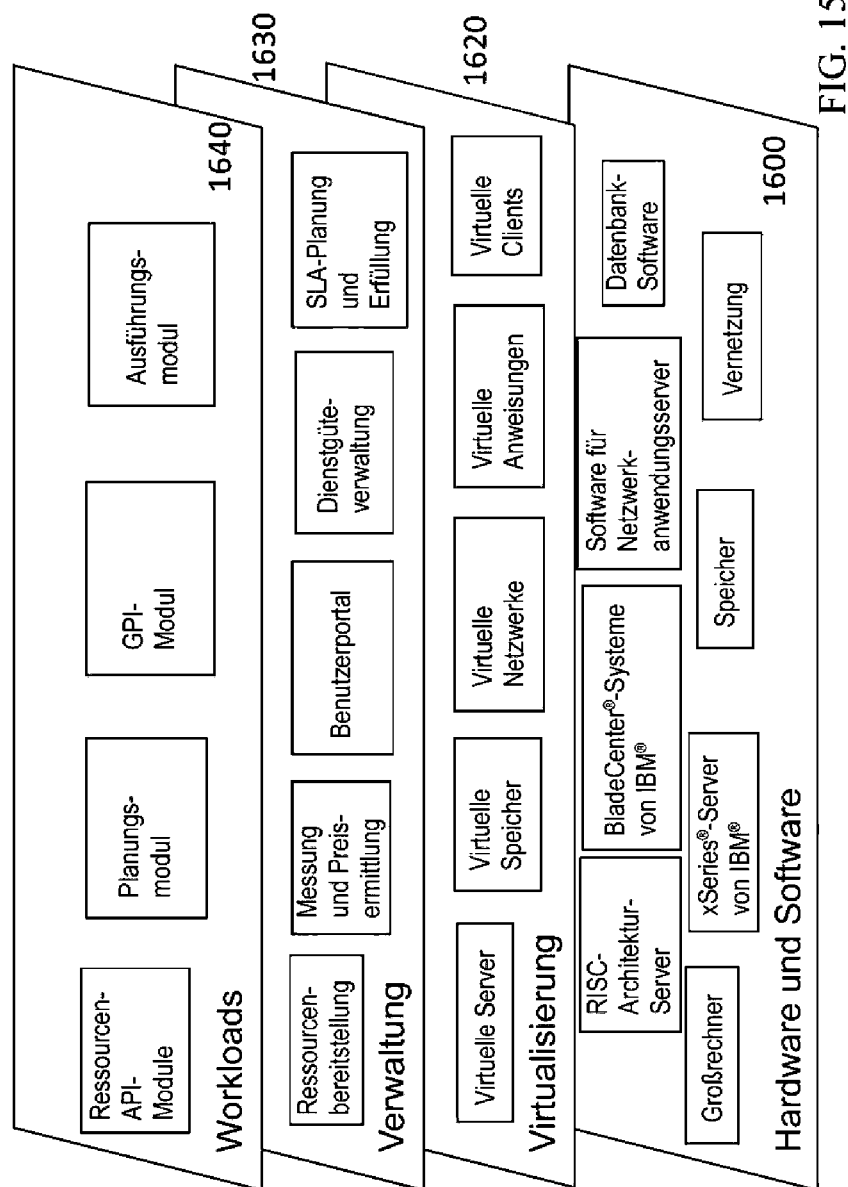


FIG. 15