

(19)日本国特許庁(JP)

(12)公表特許公報(A)

(11)公表番号
特表2024-520682
(P2024-520682A)

(43)公表日 令和6年5月24日(2024.5.24)

(51)国際特許分類
H 0 4 N 19/103 (2014.01)
H 0 4 N 19/172 (2014.01)

F I
H 0 4 N 19/103
H 0 4 N 19/172

テーマコード (参考)
5 C 1 5 9

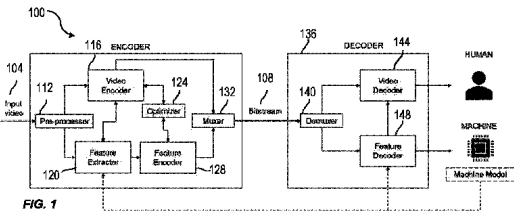
審査請求 未請求 予備審査請求 未請求 (全28頁)

(21)出願番号	特願2023-574502(P2023-574502)	(71)出願人	521328755 オーピー ソリューションズ, エルエル シー アメリカ合衆国 マサチューセッツ 0 1 0 0 2, アマースト, ミドル ストリ ート 3 6 8
(86)(22)出願日	令和4年6月3日(2022.6.3)	(74)代理人	100105957 弁理士 恩田 誠
(85)翻訳文提出日	令和6年1月30日(2024.1.30)	(74)代理人	100068755 弁理士 恩田 博宣
(86)国際出願番号	PCT/US2022/032048	(74)代理人	100142907 弁理士 本田 淳
(87)国際公開番号	WO2022/260934	(72)発明者	カルバ、ハリ アメリカ合衆国 3 3 4 3 1 フロリダ州 ボカ ラトン グレイズ ロード 7 7 7
(87)国際公開日	令和4年12月15日(2022.12.15)		最終頁に続く
(31)優先権主張番号	63/197,834		
(32)優先日	令和3年6月7日(2021.6.7)		
(33)優先権主張国・地域又は機関	米国(US)		
(81)指定国・地域	AP(BW,GH,GM,KE,LR,LS,MW,MZ,NA ,RW,SD,SL,ST,SZ,TZ,UG,ZM,ZW),EA(AM,AZ,BY,KG,KZ,RU,TJ,TM),EP(AL,A T,BE,BG,CH,CY,CZ,DE,DK,EE,ES,FI,FR ,GB,GR,HR,HU,IE,IS,IT,LT,LU,LV,MC, 最終頁に続く		

(54)【発明の名称】 機械向け映像符号化（VCM）のためのエンコーダ及びデコーダ

(57)【要約】

機械向け映像符号化（VCM）エンコーダは、入力映像をビットストリームに符号化するように構成されている第1の映像エンコーダを含む。VCMエンコーダは、入力映像における1つ以上の特徴を検出するように構成されている特徴抽出器を含む。VCMエンコーダは、入力映像及び1つ以上の特徴に応じて特徴ビットストリームを符号化するように構成されている第2のエンコーダを含む。



【特許請求の範囲】

【請求項 1】

機械向け映像符号化（VCM）エンコーダであって、
入力映像をビットストリームに符号化するように構成されている第 1 の映像エンコーダと、
前記入力映像における 1 つ以上の特徴を検出するように構成されている特徴抽出器と、
前記入力映像及び 1 つ以上の特徴に応じて特徴ビットストリームを符号化するように構成されている第 2 のエンコーダと、を備える VCM エンコーダ。

【請求項 2】

前記特徴抽出器は、1 つ以上の特徴マップを出力するように構成されている機械学習モデルを更に含む、請求項 1 に記載の VCM エンコーダ。

10

【請求項 3】

前記機械学習モデルは、畳み込みニューラルネットワークを更に含む、請求項 2 に記載の VCM エンコーダ。

【請求項 4】

前記畳み込みニューラルネットワークは、
複数の畳み込み層と、
複数のプーリング層と、を備える、請求項 3 に記載の VCM エンコーダ。

【請求項 5】

前記特徴抽出器は分類器を更に備え、前記分類器は、前記機械学習モデルの出力を 1 つ以上の特徴に分類するように構成される、請求項 2 に記載の VCM エンコーダ。

20

【請求項 6】

前記分類器は、深層ニューラルネットワークを更に含む、請求項 5 に記載の VCM エンコーダ。

【請求項 7】

前記第 2 のエンコーダは、1 つ以上の特徴への分類に従って前記 1 つ以上の特徴マップの特徴マップをグループ化するように更に構成される、請求項 5 に記載の VCM エンコーダ。

【請求項 8】

前記第 2 のエンコーダは特徴エンコーダを更に含む、請求項 1 に記載の VCM エンコーダ。

30

【請求項 9】

前記第 2 のエンコーダは映像エンコーダを更に含む、請求項 1 に記載の VCM エンコーダ。

【請求項 10】

前記第 1 の映像エンコーダは、前記特徴抽出器に結合され、前記特徴抽出器から特徴信号を受信する、請求項 1 に記載の VCM エンコーダ。

【請求項 11】

マルチプレクサを更に備え、前記マルチプレクサは、前記映像ビットストリーム及び前記特徴ビットストリームを結合するように構成される、請求項 1 に記載の VCM エンコーダ。

40

【請求項 12】

前記特徴抽出器は、複数の特徴マップを生成するように構成され、前記特徴マップは、符号化に先立って空間的に配置される、請求項 1 に記載の VCM エンコーダ。

【請求項 13】

前記特徴マップは、少なくとも部分的に前記特徴マップのテクスチャ成分に基づいて空間的に配置される、請求項 12 に記載の VCM エンコーダ。

【請求項 14】

符号化されたハイブリッドビットストリームを受信するように構成されている VCM デコーダであって、

50

前記ハイブリッドビットストリームを受信するデマルチプレクサと、
符号化された特徴ビットストリームを前記デマルチプレクサから受信し、機械処理に向けて特徴の復号されたセットを提供する特徴デコーダと、
前記特徴デコーダに結合された機械モデルと、
符号化された映像ビットストリームを前記デマルチプレクサから受信し、復号された映像信号を人による消費に向けて提供する映像デコーダと、を備えるVCMデコーダ。

【請求項15】

前記特徴デコーダは、複数の空間的に配置された特徴マップを含むビットストリームを受信し、前記空間的に配置された特徴マップを復号し、特徴マップの元のシーケンスを再構築するように構成される、請求項14に記載のVCMデコーダ。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、一般的には映像（ビデオ）の符号化及び復号の分野に関する。特に、本発明は、機械向け映像符号化（VCM）のためのエンコーダ及びデコーダに関する。

【背景技術】

【0002】

映像コーデックは、デジタル映像を圧縮又は解凍する電子回路又はソフトウェアを含むことができる。電子回路又はソフトウェアは、圧縮されていない映像を圧縮形式に又は圧縮された映像を圧縮されていない形式に変換することができる。映像圧縮の場合、映像を圧縮する（及び／又はそのうちの何らかの機能を実行する）デバイスは典型的にはエンコーダと呼ぶことができ、映像を解凍する（及び／又はそのうちの何らかの機能を実行する）デバイスはデコーダと呼ぶことができる。

20

【0003】

圧縮データの形式は、標準映像圧縮仕様に準拠することができる。圧縮は、圧縮された映像が元の映像に存在した幾らかの情報を欠くという点で情報の損失を伴い得る。この結果の1つとして、解凍された映像は、元の映像を正確に再構築するには情報が不十分であるため、元の圧縮されていない映像よりも低い品質を有し得る。

【0004】

映像の品質、映像を表すために使用されるデータ量（例えば、ビットレートによって決まる）、符号化アルゴリズム及び復号アルゴリズムの複雑さ、データの損失及び誤差の影響の受けやすさ、編集の容易さ、ランダムアクセス、エンドツーエンド遅延（例えば、待ち時間）等の間には複雑な関係が存在し得る。

30

【0005】

動き補償は、映像中のカメラ及び／又は物体の動きを考慮することにより、前のフレーム及び／又は将来のフレーム等の参照フレームを所与として、映像フレーム又はその一部分を予測する手法を含むことができる。動き補償は、映像圧縮に向けた映像データの符号化及び復号、例えば、動画像専門家集団（MPEG）の高度動画像符号化（AVC）規格（H.264とも呼ばれる）を使用した符号化及び復号に採用することができる。動き補償は、現在ピクチャへの参照ピクチャの変換に関してピクチャを記述することができる。参照ピクチャは、現在ピクチャと比べて時間的に前のピクチャ、現在ピクチャと比べて将来からのピクチャであることができる。画像が前に送信され及び／又は記憶された画像から正確に合成することができる場合、圧縮効率を改善することができる。

40

【0006】

映像は従来、人が消費するメディアであり、映像圧縮方法論は、解凍後に人である視聴者により知覚される映像の忠実度を維持することにフォーカスしていた。しかしながら、今日、膨大な量の映像は機械により解析されている。したがって、機械解析に向けて最適化された映像圧縮方法を開発し最適化する必要性が増している。用途に応じて、機械は、解析及び機能を実行するために、映像内容から同じ情報を必要としないことがある。代わりに、映像信号中の特定の特徴で十分であり得る。機械向け映像符号化（「VCM」）は

50

、従来の映像ストリームと、その映像ストリームから抽出された機械解析に適した特徴との両方を圧縮することにより圧縮ビットストリームを生成する手法である。

【発明の概要】

【0007】

エンコーダ及びデコーダを含む機械向け映像符号化（VCM）のシステムが提供される。VCMエンコーダは、好ましくは、入力映像信号をビットストリームに符号化するように構成されている第1の映像エンコーダを含む。VCMエンコーダは、入力映像中の1つ以上の特徴を検出するように構成されている特徴抽出器を更に含む。第2のエンコーダは、入力映像及び1つ以上の特徴に応じて特徴ビットストリームを符号化するように構成される。

10

【0008】

幾つかの実施形態では、映像デコーダは特徴抽出器に結合されて、特徴抽出器から特徴信号を受信する。好ましくは、機械モデルが特徴抽出器に含まれ又は特徴抽出器に提供されることができる。マルチプレクサが提供されて、符号化された映像及び特徴信号を結合して、デコーダに送信されるビットストリームにすることができる。

【0009】

幾つかの好ましくは実施形態では、特徴抽出器は、1つ以上の特徴マップを出力するように構成されている機械学習モデルを更に含む。なお好ましくは、機械学習モデルは畳み込みニューラルネットワークを更に含むことができる。幾つかの実施形態では、畳み込みニューラルネットワークは、複数の畳み込み層及び複数のプーリング層を含む。

20

【0010】

特徴抽出器は、機械学習モデルの出力を1つ以上の特徴に分類するように構成されている分類器を更に含む。特定の実施形態では、分類器は深層ニューラルネットワークである。

【0011】

特徴抽出器は、複数の特徴マップを生成し、符号化に先立って複数の特徴マップの少なくとも一部分を空間的に配置するように構成し得る。特徴マップは、テクスチャ等の特徴マップのパラメータに基づいて空間的に配置し得る。

【0012】

更に別の実施形態では、第2のエンコーダは、1つ以上の特徴の分類に従って特徴マップをグループ化するように更に構成することができる。

30

VCMデコーダは、符号化されたハイブリッドビットストリームを受信するように構成される。VCMデコーダは、ハイブリッドビットストリームを受信し、映像ビットストリーム及びそこからの特徴ビットストリームを提供するデマルチプレクサを含む。特徴デコーダが提供される。特徴デコーダは、デマルチプレクサから符号化された特徴ビットストリームを受信し、機械処理に向けて復号された特徴セットを提供する。機械モデルは、好ましくは、特徴デコーダに結合される。映像デコーダは、デマルチプレクサから符号化された映像ビットストリームを受信し、人が消費するのに適した、復号された映像信号を提供するために提供される。

【0013】

幾つかの実施形態では、VCMデコーダは、複数の空間的に配置された特徴マップを含むビットストリームを受信し、空間的に配置された特徴マップを復号し、特徴マップの元のシーケンスを再構築するように構成することができる。

40

【0014】

本発明の非限定的な実施形態のこれら及び他の態様及び特徴は、添付図面と併せて本発明の特定の非限定的な実施形態の以下の説明を考察することで当業者には明らかになる。

【0015】

本発明を説明するために、図面に本発明の1つ又は複数の実施形態の態様を示す。しかしながら、本発明が、図面に示されている精密な配置及び手段に限定されないことを理解

50

されたい。

【図面の簡単な説明】

【0016】

【図1】VCMエンコーダ及びデコーダシステムの例示的な一実施形態を示す簡易ブロック図。

【図2】VCMエンコーダ及びデコーダシステムの例示的な一実施形態を示す簡易ブロック図。

【図3】機械モデルの例示的な一実施形態を絵で示した模式図。

【図4】畳み込み単位の組み立てプロセスの例示的な一実施形態を示す図。

【図5】符号化単位表現及び畳み込み単位表現における映像の例示的な一実施形態を示すブロック図。 10

【図6】インター予測プロセスの例示的な一実施形態を示すブロック図。

【図7】イントラ予測プロセスの例示的な一実施形態を示すブロック図。

【図8】映像入力ピクチャ及び対応する畳み込み層の例示的な一実施形態を示す模式図。

【図9】映像デコーダの例示的な一実施形態を示す簡易ブロック図。

【図10】映像エンコーダの例示的な一実施形態を示すブロック図。

【図11】機械学習モジュールの例示的な一実施形態を示すブロック図。

【図12】本明細書に開示される方法論の任意の1つ又は複数及びそれらの1つ又は複数の部分の実施に使用することができる計算システムのブロック図。

【発明を実施するための形態】 20

【0017】

図面は必ずしも一定の縮尺ではなく、幻像線、線図表現、及び断片図により示され得る。特定の場合、実施形態の理解に必要な細部又は他の細部を理解しにくくする細部は省かれていることがある。

【0018】

実施形態において、VCMエンコーダは、映像モード又はハイブリッドモードで動作可能であり得る。いずれのモードでも、VCMエンコーダは、人により閲覧される出力映像として、復号される映像ビットストリームを提供し得る。

【0019】

これより図1を参照して、機械向け映像符号化(VCM)のエンコーダの例示的な一実施形態を示す。VCMエンコーダ100は、限定ではなく、デジタル回路及び／又はアナログ回路を含む任意の回路を使用して実施し得る。VCMエンコーダ100は、ハードウェア構成、ソフトウェア構成、ファームウェア構成、及び／又はそれらの任意の組合せを使用して構成し得る。VCMエンコーダ100は、計算デバイスとして及び／又は計算デバイスの構成要素として実施し得、計算デバイスは、限定ではなく、後述の任意の計算デバイスを含み得る。一実施形態において、VCMエンコーダ100は、入力映像104を受信し、出力ビットストリーム108を生成するように構成し得る。入力映像104の受信は、後述の任意の様式で達成し得る。ビットストリームは、限定ではなく、後述の任意のビットストリームを含み得る。VCMエンコーダ100は、限定ではなく、プリプロセッサ112、映像エンコーダ116、特徴抽出器120、オプティマイザ124、特徴エンコーダ128、及び／又はマルチプレクサ132を含み得る。プリプロセッサ112は、入力映像104のストリームを受信し、それを分析してストリームの映像サブストリーム、オーディオサブストリーム、及びメタデータサブストリームに分け得る。プリプロセッサ112は、更に詳細に後述するようにデコーダを含み得及び／又はデコーダと通信し得る。換言すれば、プリプロセッサ112は入力ストリームを復号する能力を有し得る。これにより、非限定的な例において、入力映像104の復号が可能になり得、下流のピクセルドメイン解析を促進し得る。

【0020】

図1を更に参照すると、VCMエンコーダ100はハイブリッドモード及び／又は映像モードで動作し得る。ハイブリッドモードの場合、VCMエンコーダ100は、人である 50

消費者を意図した視覚信号を符号化するとともに、機械である消費者を意図した特徴信号を符号化するように構成し得る。機械である消費者は、限定ではなく更に詳細に後述する計算デバイスを含め、限定ではなく任意のデバイス及び／又は構成要素を含み得る。入力信号は、例えば、ハイブリッドモードである場合にプリプロセッサ 112 を通して渡され得る。

【0021】

なお図 1 を参照すると、人による消費に向けて映像信号を符号化している映像エンコーダ 116 は、限定ではなく、更に詳細に後述するように任意の映像エンコーダ 116 を含み得る。VCM エンコーダ 100 がハイブリッドモードである場合、VCM エンコーダ 100 は、変更されていない入力映像 104 を映像エンコーダ 116 に送信し得、同じ入力映像 104 のコピー及び／又は何らかの様式で変更された入力映像 104 を特徴抽出器 120 に送信し得る。入力映像 104 への変更は、任意のスケール、変換、又は本開示全体を考慮した上で当業者が思い至り得る他の変更を含み得る。限定ではなく例として、入力映像 104 はより小さな解像度にリサイズされてよく、入力映像 104 中の一連のピクチャにおける特定数のピクチャが破棄されて、入力映像 104 のフレームレートを下げよく、限定ではなく例として、RGB を変換することにより色情報が変更されてよく、映像がグレースケール映像に変換等されよい。

【0022】

なお図 1 を参照すると、映像エンコーダ 116 及び特徴抽出器 120 は接続され、有用情報を両方向で交換し得る。限定ではなく例として、映像エンコーダ 116 は、運動推定情報を特徴抽出器 120 に、又は特徴抽出器 120 は、運動推定情報を映像エンコーダ 116 に転送し得る。映像エンコーダ 116 は、映像エンコーダ 116 及び／又は特徴抽出器 120 が識別し得る関心領域 (ROI) に基づいて、量子化マッピング及び／又はそのデータ記述を特徴抽出器 120 に提供し得、又は特徴抽出器 120 は、映像エンコーダ 116 及び／又は特徴抽出器 120 が識別し得る関心領域 (ROI) に基づいて、量子化マッピング及び／又はそのデータ記述を映像エンコーダ 116 に提供し得る。映像エンコーダ 116 は、入力映像 104、入力信号、及び／又は任意のフレーム及び／又はそのサブフレームに存在及び／又は識別された特徴に基づいて、1 つ又は複数のパーティション決定を記述するデータを特徴抽出器 120 に提供し得る；特徴抽出器 120 は、入力映像 104、入力信号、並びに／或いは任意のフレーム及び／又はそのサブフレームに存在及び／又は識別された特徴に基づいて、1 つ又は複数のパーティション決定を記述するデータを映像エンコーダ 116 に提供し得る。映像エンコーダ 116、特徴抽出器 120 は、最適なピクチャ群 (GOP) 決定のために時間的情報を互いと共有及び／又は伝送し得る。これらの技法及び／又はプロセスの各々は、限定ではなく、更に詳細に後述されるように実行し得る。

【0023】

引き続き図 1 を参照すると、特徴抽出器 120 はオフラインモード又はオンラインモードで動作し得る。特徴抽出器 120 は、特徴を識別並びに／或いは他の方法で特徴に対して作用及び／又は特徴を操作し得る。「特徴」は、本開示で 사용되는場合、データの特定の構造属性及び／又は内容属性である。特徴の例は、スケール不変特徴変換 (SIFT)、オーディオ特徴、色ヒストグラム、動きヒストグラム、発話レベル、音の大きさレベル等を含み得る。特徴はタイムスタンプを有し得る。各特徴に、フレーム群のうちの単一のフレームを関連付け得る。特徴は、タイムスタンプ、映像中の人物及び物体のラベル、関心のある物体及び／又は領域の座標、領域ベースの量子化のフレームマスク、及び／又は本開示全体を考慮した上で当業者が思い至り得る任意の他の特徴等の高レベル内容特徴を含み得る。更なる非限定的な例として、特徴は、フレーム又はフレーム群の空間的特性及び／又は時間的特性を記述する特徴を含み得る。空間的特性及び／又は時間的特性を記述する特徴の例には、動き、テクスチャ、色、明るさ、エッジカウント、ブラー、ブロックノイズ等があり得る。オフラインモードの場合、更に詳細に後述する全ての機械モデルは、エンコーダ及び／又はエンコーダのメモリに記憶され得並びに／或いはエンコーダに

10

20

30

40

50

よりアクセス可能であり得る。そのようなモデルの例には、限定ではなく、全体若しくは部分畳み込みニューラルネットワーク、キーポイント抽出器、エッジ検出器、特徴マップコンストラクタ等があり得る。オンラインモードの場合、1つ又は複数のモデルは、リモート機械によりリアルタイムで又は抽出前の何らかの時点で特徴抽出器 120 に通信し得る。

【0024】

なお図 1 を参照すると、特徴エンコーダ 128 は、限定ではなく例として、特徴抽出器 120 により生成された特徴信号を符号化するように構成される。一実施形態では、特徴を抽出した後、特徴抽出器 120 は、抽出された特徴を特徴エンコーダ 128 に渡し得る。

10

【0025】

特徴エンコーダ 128 は、限定ではなく例として、エントロピー符号化及び／又は同様の技法を使用して特徴ストリームを生成し得、特徴ストリームはマルチプレクサ 132 に渡され得る。

【0026】

映像エンコーダ 116 及び／又は特徴エンコーダ 128 は、オブティマイザ 124 を介して接続され得、オブティマイザ 124 は、それらの映像エンコーダ 116 と特徴エンコーダ 128 との間で有用な情報を交換し得る。限定ではなく例として、コードワード構造及び／又はエントロピー符号化の長さに関連する情報が、最適な圧縮に向けてオブティマイザ 124 を介して交換され再使用され得る。

20

【0027】

一実施形態では、図 1 を参照すると、映像エンコーダ 116 は映像ストリームを生成し得、映像ストリームはマルチプレクサ 132 に渡され得る。マルチプレクサ 132 は、映像ストリームを特徴エンコーダ 128 により生成された特徴ストリームと多重化し得る。それに代えて又は加えて、映像ビットストリーム及び特徴ビットストリームは、別個のチャネル、別個のネットワークを経由して別個のデバイスに及び／又は別個の時間又は時間間隔で（時間多重化）伝送されてよい。映像ストリーム及び特徴ストリームの各々は、本開示に記載の任意のビットストリームの実施に適した任意の様式で実施されてよい。一実施形態では、多重化された映像ストリーム及び特徴ストリームはハイブリッドビットストリームを生成し得、ハイブリッドビットストリームは更に詳細に後述するように伝送され得る。

30

【0028】

なお図 1 を参照すると、VCM エンコーダ 100 が映像モードである場合、VCM エンコーダ 100 は映像エンコーダ 116 を映像及び特徴の両方の符号化に使用し得る。特徴抽出器 120 は特徴を映像エンコーダ 116 に送信し得、映像エンコーダ 116 は特徴を符号化して映像ストリームにし得、映像ストリームは対応する映像デコーダ 144 により復号し得る。VCM エンコーダ 100 は単一の映像エンコーダ 116 を映像符号化及び特徴符号化の両方に使用し得、その場合、映像及び特徴に異なるパラメータセットを使用し得、代替的には、VCM エンコーダ 100 が、並列に動作し得る 2 つの別個の映像エンコーダ 116 を使用してもよいことに留意されたい。

40

【0029】

なお図 1 を参照すると、システム 100 は、VCM デコーダ 136 を含み得及び／又は VCM デコーダ 136 と通信し得る。VCM デコーダ 136 及び／又はその要素は、上述した VCM エンコーダ 100 の構成に適した任意の回路及び／又は任意の種類の構成を使用して実施し得る。VCM デコーダ 136 は、限定ではなく、デマルチプレクサ 140 を含み得る。デマルチプレクサ 140 は、上述したように多重化された場合、ビットストリームを逆多重化するように動作し得る。限定ではなく例として、デマルチプレクサ 140 は、1つ又は複数の映像ビットストリーム及び1つ又は複数の特徴ビットストリームを含む多重化されたビットストリームを別個の映像ビットストリーム及び特徴ビットストリームに分け得る。

50

【0030】

図1を引き続き参照すると、VCMデコーダ136は映像デコーダ144を含み得る。映像デコーダ144は、限定ではなく、更に詳細に後述するように、デコーダに適した任意の様式で実施し得る。一実施形態では、限定ではなく、映像デコーダ144は出力映像を生成し得、出力映像は、人又は他の生物及び／又は視覚能力を有するデバイスにより閲覧し得る。

【0031】

なお図1を参照すると、VCMデコーダ136は特徴デコーダ148を含み得る。一実施形態では、限定ではなく、特徴デコーダ148は、1つ又は複数の復号されたデータを機械に提供するように構成し得る。機械は、限定ではなく任意のマイクロコントローラ、プロセッサ、組み込みシステム、システムオンチップ、ネットワークノード等を含む、限定ではなく任意の計算デバイスを含み得る。機械は、更に詳細に後述するように、機械モデルの操作、記憶、トレーニング、機械モデルからの入力を受信、機械モデルの出力の生成、及び／又は機械モデルとの他の方法での対話を行い得る。機械は、幾つかはデスクトップコンピュータ、ラップトップコンピュータ、及び／又はモバイルデバイス等の従来の計算デバイスでなくてよい処理構成要素及び通信構成要素を有する物体のネットワークとして定義されるモノのインターネット（IoT）に含まれてよい。IoT内の物体は、限定ではなく、マイクロプロセッサ及び／又はマイクロコントローラ及びローカルエリアネットワーク（LAN）及び／又はワイドエリアネットワーク（WAN）と対話するための1つ又は複数の構成要素が組み込まれた任意のデバイスを含み得る；1つ又は複数の構成要素は、限定ではなく例として、ワシントン州カークランドに所在のブルトゥースSIG社（Bluetooth SIG, Inc.）により公表されたプロトコルに従うブルトゥース（登録商標）送受信機のような2.4～2.485GHzで通信する無線送受信機及び／又は仏国リュエイマルメゾン（Rueil-Malmaison）に所在のシュナイダーエレクトリックSE（Schneider Electric SE）により公表されたMODBUSプロトコル及び／又は電気電子技術者協会（IEEE）により公表されたIEEE 802.15.4規格のZIGBEE（登録商標）仕様に従って動作するネットワーク通信構成要素を含み得る。当業者ならば、本開示全体を考慮した上で、本開示と一貫して採用し得る代替又は追加の通信プロトコル及びそのようなプロトコルをサポートするデバイスを認識し、それらの各々は本開示の範囲内であると企図される。

【0032】

引き続き図1を参照すると、VCMエンコーダ100及び／又はVCMデコーダ136の各々は、本開示に記載の任意の実施形態における任意の方法、方法ステップ、又は方法ステップのシーケンスを任意の順序且つ任意の反復程度で実行するように設計及び／又は構成し得る。例えば、VCMエンコーダ100及び／又はVCMデコーダ136の各々は、所望又は命令された結果が達成されるまで、単一のステップ又はステップのシーケンスを繰り返し実行するように構成し得、ステップ又はステップのシーケンスの反復は、前の反復の出力を続く反復の入力として使用し、反復の入力及び／又は出力を集計して集計結果を生成し、大域変数等の1つ又は複数の変数の低減又はデクリメントを使用し、及び／又はより大きな処理タスクの、繰り返しアドレス指定されるより小さな処理タスクのセットへの分割を使用して、反復的及び／又は再帰的に実行し得る。VCMエンコーダ100及び／又はVCMデコーダ136の各々は、本開示に記載の任意のステップ又はステップのシーケンスを並列に実行し得、例えば、2つ以上の並列スレッド、プロセッサコア等を使用してステップを2回以上同時に及び／又は略同時に実行し得；並列スレッド及び／又はプロセッサ間のタスクの分割は、反復間のタスクの分割に適した任意のプロトコルに従って実行し得る。当業者ならば、本開示全体を考慮した上で、ステップ、ステップのシーケンス、処理タスク、及び／又はデータを細分、共有、並びに／或いは反復、再帰、及び／又は並列処理の使用に他の方法で対処し得る種々の方法を認識するであろう。

【0033】

これより図2を参照して、映像モードで動作しているVCMエンコーダ100の例示的

な一実施形態 200 を示す。VCM エンコーダ 100 は、ユーザ、プログラム、メモリロケーション、及び／又は VCM エンコーダ 100 と対話している 1 つ又は複数の追加のデバイス（図示せず）からの命令を受信すると、映像モードとハイブリッドモードとを切り替えるように構成し得る。映像モードにおいて、VCM エンコーダ 100 は第 2 の映像エンコーダ 204 を動作させるように構成し得る。

【0034】

第 2 の映像エンコーダ 204 は、上述した特徴エンコーダ 128 及び／又は映像エンコーダ 116 に適した任意の様式で特徴及び／又は第 2 の映像ビットストリームを符号化し得る。第 2 の映像エンコーダ 204 は、特徴抽出器 120 及び／又はオプティマイザ 124 からデータを受信し得るか、特徴抽出器 120 及び／又はオプティマイザ 124 にデータを送信し得るか、又はその両方を行い得る。図 2 を引き続き参照すると、映像モードは、視覚情報として表現し得るあらゆる又は全ての種類の特徴の符号化に使用し得る。非限定的な例として、映像モードは、特徴マップ、エッジ、ライン等を表す画像等のフィルタリングされた画像、更に詳細に後述する畳み込みニューラルネットワークの特徴マップ等の符号化に使用し得る。

【0035】

これより図 3 を参照すると、VCM エンコーダ 100 と併用される畳み込みニューラルネットワーク（CNN）の例示的な一実施形態 300 を示す。本開示で使用される場合、人工ニューラルネットワークとしても知られる「ニューラルネットワーク」は、「ノード」又は 1 つ若しくは複数の入力、1 つ若しくは複数の出力、及び入力に基づいて出力を決定する関数を有するデータ構造である。そのようなノードは、限定ではなく、ノードの入力層、1 つ又は複数の中間層、及びノードの出力層を含む畳み込みニューラルネットワーク等のネットワークに編成し得る。ノード間の接続は、ネットワークの「トレーニング」プロセスを介して作成し得、ネットワークのトレーニングプロセスでは、トレーニングデータセットからの要素が入力ノードに適用され、次いで、適したトレーニングアルゴリズム（レーベンバーグ・マルカート（Levenberg-Marquardt）、共役勾配、模擬アニーリング、又は他のアルゴリズム等）が使用されてニューラルネットワークの隣接層におけるノード間の接続及び重みを調整し、出力ノードにおいて所望の値を生成し得る。このプロセスは深層学習と呼ばれることがある。

【0036】

なお図 3 を参照すると、ノードは、限定ではなく、当該ノードを含むニューラルネットワークへの入力から及び／又は他のノードから数値を受信し得る複数の入力 x_i を含む得る。ノードは、各入力 x_i により乗算される重み w_i を使用して入力の加重和を実行し得る。これに加えて又は代えて、層の入力から独立したニューラルネットワーク層中の各ユニットにオフセットが追加されるようなバイアス b を入力の加重和に追加し得る。次いで、加重和は関数 ψ に入力され得、関数 ψ は 1 つ又は複数の出力 y を生成し得る。入力 x_i に適用される重み w_i は、入力が、例えば、大きな数値を有する対応する重みにより、1 つ又は複数の出力 y に対して強い影響を有することを示す「興奮性」であるか及び／又は例えば、小さな数値を有する対応する重みにより、1 つ又は複数の入力 y に対して弱い影響を有することを示す「抑制性」であるかを示し得る。重み w_i の値は、トレーニングデータを使用してニューラルネットワークをトレーニングすることにより決定し得、トレーニングは、上述した任意の適したプロセスを使用して実行し得る。一実施形態では、限定ではなく、ニューラルネットワークは、入力としてセマンティックユニットを受信し、本開示に記載の機械学習プロセスを使用して導出される重み w_i に従ってそのようなセマンティックユニットを表すベクトルを出力し得る。「畳み込みニューラルネットワーク」は、本開示で使用される場合、1 つ以上の隠れ層が、プーリング層、全結合層等の 1 つ又は複数の追加の層と共に、その層への入力を「カーネル」として知られる入力のサブセットで畳み込む畳み込み層であるニューラルネットワークである。CNN は、限定ではなく、深層ニューラルネットワーク（DNN）を含み得、DNN は、2 つ以上の隠れ層を有するニューラルネットワークとして定義される。

10

20

30

40

50

【0037】

一実施形態では、引き続き図3を参照すると、入力ピクチャの標準映像エンコーダ116要件に合うように調整された特徴マップを符号化し及び／又は合うように特徴マップを再アレンジするために、CNN及び／又は他のモデルを使用し得る。例示的な一実施形態では、幅W及び高さHの入力画像がCNN及び／又は他のモデルに入力され得る。畳み込み及びプーリングとして知られる一連の動作を通して、入力ピクチャ中の画像を1つ又は複数の特徴マップに変換し得る。CNNの各層は出力として幾つの特徴マップを含み得る。合計でn個の畳み込み層(C)及びプーリング層(P)があり得る。各層は、幅w及び高さhとして表現される同じ又は異なる寸法の畳み込み及びプーリングカーネルを有し得る；例えば、畳み込み層1は幅C1_w及び高さC1_hを有し得、プーリング層1は幅P1_w及び高さP1_hを有し得、畳み込み層2は幅C2_w及び高さC2_hを有し得、プーリング層2は幅P2_w及び高さP2_hを有し得、以下同様である。

10

【0038】

幾つの実施形態では、なお図3を参照すると、最後のプーリング演算後、出力は、入力ベクトルの入力に使用し得る、限定ではなく深層ニューラルネットワーク等の分類器に渡すことができる。本開示で使用される場合、「分類アルゴリズム」は、入力をデータのカテゴリ又はビンにソートし、データのカテゴリ又はビン及び／又はそれらと関連付けられたラベルを出力する、更に詳細に後述する機械学習アルゴリズム及び／又はプロセスであり；「分類器」は、本開示で使用される場合、更に詳細に後述する機械学習プロセスにより生成された数学モデル、ニューラルネット、又はプログラム等の機械学習モデルである。分類器は、後述の距離メトリック下で近いことがわかった、一緒にクラスタリングされたデータのセット等をラベリング又は他の方法で識別する1つ以上のデータを出力するように構成し得る。計算デバイス及び／又は別のデバイスは、プロセスとして定義される分類アルゴリズムを使用して分類器を生成し得、それにより、計算デバイスは、トレーニングデータから分類器を導出する。DNN分類器等のニューラルネットワーク分類器に加えて、分類は、限定ではなく、限定ではなくロジスティック回帰及び／又はナイーブベイズ分類器等の線形分類器、k最近傍分類器等の最近傍分類器、サポートベクターマシン、最小二乗サポートベクターマシン、フィッシャーの線形判別、二次分類器、決定木、ブーストツリー、ランダムフォレスト分類器、学習ベクトル量子化、及び／又はニューラルネットワークベースの分類器を使用して実行し得る。

20

30

【0039】

一実施形態では、なお図3を参照すると、出力ニューロンの値の組合せがVCMエンコーダ100により使用されて、関心のある特定のクラスの有無を判断し得る。例えば、人が検出される場合、[0, 1]の組合せを出力し、車が検出される場合、[1, 0]を出力する等のように、2つ以上の出力を有するニューラルネットワークをトレーニングし得る。畳み込み及びプーリングは出力として2次元行列の値を有するため、そのような2次元行列の値は視覚ピクチャとして表現し得る。このピクチャは「特徴マップ」と呼ぶこともできる。幾つの実施形態では、限定ではなく、VCMエンコーダ100は完全なCNNを含み得、その場合、特徴抽出器120の出力は最終的に、プーリングされたベクトル又はDNNの中間層若しくは最後の層の出力のいずれかを含み得る。他の実施形態では、VCMエンコーダ100はCNNの部分のみを含み得る。例えば、VCMエンコーダ100は、C1、P1、C2、P2、・・・、Ck、Pkとして示され得るCNNの最初のk層を含み得、ここで、kは、1及びn-1を含め、1~n-1の任意の整数を含み得る。この場合、特徴抽出器120の出力は、最終層(k)からの特徴マップのセットを含み得る。幾つの実施形態では、特徴抽出器120の出力が畳み込み層及び／又はプーリング層からの特徴マップのセットを含む場合、その出力は映像エンコーダ116を使用して符号化し得る。特徴マップセットを映像エンコーダ116に送信する前、特徴抽出器120は各特徴マップを適合及び／又は再アレンジして、映像エンコーダ116への入力として適したものにし得る。特徴マップは通常、映像エンコーダ116により使用される典型的なピクチャサイズよりもはるかに小さいため、特徴マップのセットを再アレンジする一方

40

50

法は、映像符号化に適したより大きな矩形単位に組み立てることを含み得る。

【0040】

なお図3を参照すると、CNN、DNN、並びに／又は特徴の再アレンジ及び／若しくはマッピングに使用される任意の他のモデル及び／又はプロセスは、更に詳細に後述する任意の機械学習プロセスを使用してトレーニングし得る。

【0041】

これより図4を参照して、空間的再組み立てのプロセス400の例示的な一実施形態を示す。一実施形態では、畳み込み層並びに／又は本明細書に記載の他の層及び／若しくは要素からの特徴マップのセットは矩形単位にアレンジし得、それに代えて又は加えて、単位は、スライス及び／若しくはタイルへの矩形形態の組合せ並びに／又は多角形及び／若しくは曲線の周縁を有する任意の形状を含め、任意の適した形状を有してよい。1つのそのような形状は、ドーナツのような又は他のエンクロージャを形成する別のそのような形状を囲み得る。個々のマップの空間的位置は、畳み込み演算における順番に対応する象限にマップが位置する簡易な順次配置により決定し得る。単位におけるマップの空間的位置はこれに代えて又は加えて、結果として生成される映像符号化が最適化されるように配置してよい。そのような最適化に適用し得る演算の一例は、同様のテクスチャを有するマップを互いの隣に配置することであり、したがって、限定ではなく例として後述するような映像エンコーダ116によるイントラ予測の効率を上げる。テクスチャの尺度は、マップにおけるピクセル値の分散として表現し得る。さらに、単一の畳み込み層から組み立てられる特徴単位は、空間的又は時間的のいずれかで他の畳み込み層からの特徴単位と組み合わせられ得る。

【0042】

これより図5を参照して、図4において生成されるように配置された出力として定義される、単位から組み立てられたピクチャマップのセットの例示的な一実施形態を示す。非限定的な例として、畳み込み単位は、映像エンコーダ116により符号化されるピクチャ内部の近隣ブロックとして組み立て得る。なお、幾つかの場合、単位の境界を映像エンコーダ116の符号化単位の境界と並べることが有益であり得る。これは、限定ではなく、符号化単位と一致したサイズ（例えば、 64×64 ピクセル若しくは 128×128 ピクセル）の畳み込みカーネルを使用することにより又は符号化単位サイズに合うように畳み込み単位を再スケーリングすることにより達成し得る。再スケーリングは、畳み込み単位のピクセルに適用される簡易な線形再スケーリング又はバイキュービック再スケーリング技法を使用して達成し得る。

【0043】

なお図5を参照すると、ピクチャの単一の符号化単位は、例えば図5の右下に示されるように、1つ又は複数の畳み込み単位を含み得る。特定の配置における畳み込み単位の空間的位置は、限定ではなく、畳み込み単位の順序に基づくか又はテクスチャ等の同様の特性を有する単位を近傍に配置する最適化アルゴリズムを使用することにより計算し得る。映像エンコーダ116は時間的予測及び／又はインター予測も同様に単位で適用し得るため、映像符号化のイントラ予測の改善の他に、空間位置への時間的マッピングも同様に考慮し得る。

【0044】

これより図6を参照して、動き推定を使用して計算されるインター予測の例示的な一実施形態及び動きベクトル(MV)における結果図示される。畳み込みマップの均質性、ひいては連続ピクチャにおける単位の類似性は入力映像104の変化に依存するため、映像エンコーダ116は、最適な圧縮を達成するために、動的解像度変更を適用し得る。例えば、映像エンコーダ116は、関連する畳み込みマップ数が減る場合、ピクチャの解像度を下げ得る。これは、入力映像104が、多くの物体及び関心領域を含むシーンから1つの物体及び／又はより小さな関心領域を有するシーンに切り替わる場合に生じ得る。そのような場合、畳み込み単位の数は下がり得、その結果として、符号化ピクチャの解像度を下げることが有益であり得る。解像度の変更は、限定ではなく、ヘッダ情報の使用を含め

、本開示全体を考慮した上で当業者が思い付き得る任意の様式でデコーダに伝え得る。畳み込みマップに関連する他の適切な情報及び／又はパラメータは、補足強化情報（S E I）ストリーム又は他の同様のメタデータストリームを使用して符号化し得る。S E I等のメタデータストリームは、機械モデルを記述する関連情報の符号化にも同様に使用し得る。

【0045】

なお図6を参照すると、大半の標準エンコーダに固有であり得るレート歪み最適化（R D O）により所与の入力ピクチャでの映像エンコーダ116の最適な性能は保証し得るが、幾つかのパラメータは特徴抽出器120により調整、更新され得る。この調整の理由は、映像エンコーダ116は符号化された映像の視覚的品質に向けて最適化することができ、一方、機械の効用関数は、特定の場合、その尺度からずれ得ることを含み得る。限定ではなく例として、C N Nにおいて大きな値の重みと関連付けられた特徴マップは、より重要であるものとして伝えられ得、その意味において、映像エンコーダ116によるそれらの圧縮を調整する必要があるし得る。そのような機能の非限定的な一例は、所与の畳み込み単位と関連付けられた重みの大きさに反比例するようにエンコーダの量子化レベルを更新することであり得る。重みの大きさは、C N Nにおける全ての重みと比較した相対値として表現し得る。

【0046】

これより図7を参照して、動きベクトルマッピングプロセスの例示的な一実施形態を示す。一実施形態において、動きベクトルマッピングは、特徴抽出器120と映像エンコーダ116との間の対話を含み得及び／又は当該対話により生じ得る；この対話はハイブリッドモード及び／又は映像モードで行われ得る。一実施形態において、動き推定情報は、特徴抽出器120から映像エンコーダ116に及び／又は映像エンコーダ116から特徴抽出器120に転送し得る。時間依存性を有する入力映像104の場合、連続ピクチャが何らかの現実世界の動きを表す場合、動きは、動き推定を利用してインター予測を使用して推定し得る。結果として生成される動きベクトル（M V）は、連続フレーム間の予測単位の変位を表し得る。

【0047】

一例として、引き続き図7を参照すると、単純な並進運動は、現在フレームにおける及び／又はあるフレームから次のフレームへのブロック、符号化単位、符号化ツリー単位、畳み込み単位、及び／又はピクセルの変位を記述する2つの成分M V x、M V yを有する動きベクトル（M V）を使用して記述し得る。回転、ズーム、及び／又はワープ等のより複雑な動きは、アフィン動きベクトルを使用して記述し得、「アフィン動きベクトル」は、本開示で使用される場合、映像ピクチャ及び／又はピクチャで表現されるピクセル又は点のセット、例えば、運動中の見掛けの形状を変えことなく、映像におけるビューにわたって動いている物体を示すピクセルのセットの一樣変位を記述するベクトルである。映像符号化及び／又は復号の幾つかの手法は、ピクチャ間符号化における動き補償に対して4パラメータ又は6パラメータアフィンモデルを使用し得る。

【0048】

例えば、なお図7を参照すると、6パラメータアフィン運動は、
$$x' = ax + by + c \quad y' = dx + ey + f$$
として記述することができる。

【0049】

4パラメータアフィン運動は、
$$x' = ax + by + c \quad y' = -bx + ay + f$$
として記述することができ、式中、 (x, y) 及び (x', y') はそれぞれ、現在ピクチャ及び参照ピクチャにおけるピクセルロケーションであり、 a 、 b 、 c 、 d 、 e 、及び f はアフィン運動モデルのパラメータである。

【0050】

なお図7を参照すると、上述したように、動き推定は畳み込み単位に対して及び／又は

より一般的には特徴マップに対して行うことができる。映像モードの場合、この推定は映像エンコーダ116により実行し得る。しかしながら、幾つかの場合、高速で単純な動き推定は特徴抽出器120自体で実施することができ、したがって、特徴エンコーダ128に送信する前、特徴マップ間の時間依存性を除去することができる。このようにして得られた動き情報は、入力映像104を符号化するために映像エンコーダ116により再使用し得る。解像度の低い特徴マップに対する動き推定ほど、はるかに効率的であり得るため、これは映像エンコーダ116の複雑性を大きく下げ得る。換言すれば、特徴マップは上述したように生成し得、動きベクトルは特徴マップから導出し得、動きベクトルは映像エンコーダ116に伝えられ得る。映像エンコーダ116は、そうして伝えられた動きベクトルを使用して、限定ではなく例として、後述するように映像ビットストリームを符号化し得る。

10

【0051】

幾つかの場合、なお図7を参照すると、動きベクトルマッピングは、映像エンコーダ116により計算され得、特徴抽出器120に転送され得る。これは、特徴モダリティにおける動きベクトルの精度を上げるために行うことができる。動きベクトルは、畳み込み単位とピクチャ予測単位との解像度差に比例する適切なスケーリング定数を適用することにより、モダリティ間、例えば特徴モダリティと映像モダリティとの間で転送し得る。これより図8を参照して、特徴から映像にマッピングし得る、ROIに基づく量子化マッピングプロセスの例示的な一実施形態を示す。一実施形態において、映像エンコーダ116は、RDOに基づいて符号化単位に量子化を適用して、所与のビット配分での符号化単位の視覚的品質を最大化し得る。幾つかの場合、ピクチャの特定の部分は他の部分よりも知覚的に重要であり得、より高品質で符号化され得、一方、ピクチャの残りの部分は幾らか低い品質で符号化され得る。知覚的に重要な部分の例は、人の顔、物体、低テクスチャ等を含む部分である。ピクチャの部分の重要度は、効用関数によっても同様に特定し得る。例えば、監視映像では、関心のある顔及び小さな物体の細部を保持することが最も重要であり得る。特徴抽出器120により取得された情報を使用して、ピクチャのそのような部分を示すことが可能であり得る。これは、境界ボックスとして表現し得る関心領域を使用し得る。境界ボックスは、限定ではなく、ピクチャ及び／又は特徴マップの左上隅等のピクチャ及び／又は特徴マップ内のロケーションの座標 x 、 y 及びピクセル値で表現される幅及び高さ w 、 h を含む任意の適した様式で画定し得る。この空間情報は映像エンコーダ116に渡すことができ、映像エンコーダ116は、ROI内の全ての符号化単位のRDO計算により低い歪み及びより高いレートを割り当て得る。

20

30

【0052】

図8を更に参照すると、重要度は、重要度係数 S_N に従って特定、記憶、及び／又は伝えられ得、重要度係数 S_N は、外部専門家により供給されてよく及び／又は限定ではなく畳み込み単位、符号化単位、符号化ツリー単位等として定義される等のピクチャ内のエリアの特性に基づいて計算されてよい。エリアの「特性」は、本明細書で使用される場合、その内容に基づいて特定されるエリアの測定可能な属性であり、特性は、第1のエリアに対して実行される1つ又は複数の計算の出力を使用して数値で表現し得る。1つ又は複数の計算は、第1のエリアにより表される任意の信号の任意の解析を含み得る。非限定的な一例には、品質モデリング用途において平滑な背景を有するエリアにはより高い S_N を割り当て、平滑性の低いエリアにはより低い S_N を割り当てることがあり得る；非限定的な例として、平滑性はキャニーエッジ検出を使用して特定されて、エッジの数を特定し得、その場合、数が少ないことは、平滑性が高いことを示す。自動平滑性検出の更なる例には、エリアにわたる空間変数における信号にわたる高速フーリエ変換（FFT）の使用があり得、信号は、任意の2次元座標系にわたり及び赤緑青の色値を表すチャンネル等にわたり解析し得、FFTを使用して計算される周波数ドメインにおいて低周波成分が相対的に優位なことは、平滑性がより高いことを示し得、一方、高周波成分が相対的に優位なことは、背景エリアにわたる色値及び／又は陰影値の遷移が頻繁で急速であることを示し得、これは平滑性スコアの低下に繋がり得、セマンティック的に重要な物体はユーザ入力により

40

50

識別し得る。これに代えて又は加えて、セマンティック重要性はエッジ構成及び／又はテクスチャパターンに従って検出されてよい。背景は、限定ではなく、限定ではなく意味論的に重要な物体を含む顔又は他の物品等の重要な又は「前景」物体を表すエリアの一部分を受信及び／又は検出することにより識別し得る。別の例には、人の顔等の意味論的に重要な物体を含むエリアにより高い S_N を割り当てることがあり得る。

【0053】

例示的な一実施形態において、なお図8を参照すると、CNN又は本開示に記載の他の要素は、最初の特徴マップにおいて、層 C_n において顔を検出し得る。対応する境界ボックスが、映像エンコーダ116により符号化されるピクチャにマッピングされ得る。指定された符号化単位により高い優先度を割り当てることができ、適切なRDO更新でもって符号化し得る。他の例では、キーポイント抽出器、エッジ検出器、特徴マップコンストラクタ等の他の特徴モデルからの導出を実行することができる。

【0054】

引き続き図8を参照すると、VCMエンコーダ100は、限定ではなく例として、特徴抽出器120から映像エンコーダ116に提供される情報を使用して、特徴に基づいてパーティション決定を実行し得る。一実施形態において、映像エンコーダ116は特徴抽出器120から受信した関連情報を使用して、パーティション等の他の符号化パラメータを更新し得る。限定ではなく例として、符号化単位ツリーの深度は、限定ではなく境界ボックス等の関心領域内の単位に対して増大し得る。別の例では、映像エンコーダ116は、最小符号化単位のサイズに関連する境界ボックスのサイズと合わせて、可能な限り多くの細部を保持し、予測単位間のブロック境界に導入される歪みを回避し得る。

【0055】

なお図8を参照すると、特徴抽出器120から映像エンコーダ116に又は映像エンコーダ116から特徴抽出器120に伝送される情報は時間情報を含み得る。非限定的な例では、特徴抽出器120を使用して、入力映像104における大きな変化、限定ではなく例として、シーン変化を検出し、タイムスタンプを映像エンコーダ116に伝え得、映像エンコーダ116はこの情報を使用して、1つ又は複数のピクチャ群(GOP)の構造及び長さを最適に決定し得る。キーフレーム又は換言すればイントラ符号化されるフレームは、非限定的な例では、シーンの最初のピクチャに対応し得る。一実施形態において、連続ピクチャの変動性により、GOPで使用されるフレームの最適な種類：イントラ(I)又はインター(P、B)及びそのようなフレームの数及び／又は順序が決まり得る。他方、例えば、特徴抽出器120が動き推定を含まない場合、映像エンコーダ116により得られる情報は特徴抽出器120により使用され得る。そのような場合、映像エンコーダ116の動き推定が使用されて、特徴抽出器120による動き追跡及び活動検出を改善し得る。

【0056】

図9は、映像デコーダ144及び／又は特徴デコーダ148を実施して、圧縮されたハイブリッドビットストリームから映像及び特徴を復号するのに適し得る一例のデコーダを示すシステムブロック図である。デコーダ900は、エン트로ピーデコーダプロセッサ904、逆量子化及び逆変換プロセッサ908、デブロッキングフィルタ912、フレームバッファ916、動き補償プロセッサ920、及び／又はイントラ予測プロセッサ924を含み得る。

【0057】

動作において、なお図9を参照すると、ビットストリーム928がデコーダ900により受信され、エン트로ピーデコーダプロセッサ904に入力され得、エン트로ピーデコーダプロセッサ904はビットストリームの部分を復号して量子化係数にし得る。量子化係数は逆量子化及び逆変換プロセッサ908に提供され得、逆量子化及び逆変換プロセッサ908は逆量子化及び逆変換を実行して残差信号を作成し得、これは処理モードに従って動き補償プロセッサ920又はイントラ予測プロセッサ924の出力に追加し得る。動き補償プロセッサ920及びイントラ予測プロセッサ924の出力は、前に復号されたプロ

ックに基づいてブロック予測を含み得る。予測と残差との和がデブロッキングフィルタ 912 により処理され、フレームバッファ 916 に記憶し得る。

【0058】

一実施形態において、なお図 9 を参照すると、デコーダ 900 は、任意の実施形態において上述された任意の動作を任意の順に任意の反復程度で実施するように構成されている回路を含み得る。例えば、デコーダ 900 は、所望又は命令された結果が達成されるまで、単一のステップ又はステップのシーケンスを繰り返し実行するように構成し得、ステップ又はステップのシーケンスの反復は、前の反復の出力を続く反復の入力として使用し、反復の入力及び／又は出力を集計して集計結果を生成し、大域変数等の 1 つ又は複数の変数の低減又はデクリメントを使用し、及び／又はより大きな処理タスクの、繰り返しアドレス指定されるより小さな処理タスクのセットへの分割を使用して、反復的及び／又は再帰的に実行し得る。デコーダは、本開示に記載の任意のステップ又はステップのシーケンスを並列に実行し得、例えば、2 つ以上の並列スレッド、プロセッサコア等を使用してステップを 2 回以上同時に及び／又は略同時に実行し得；並列スレッド及び／又はプロセッサ間のタスクの分割は、反復間のタスクの分割に適した任意のプロトコルに従って実行し得る。当業者ならば、本開示全体を考慮した上で、ステップ、ステップのシーケンス、処理タスク、及び／又はデータを細分、共有、並びに／或いは反復、再帰、及び／又は並列処理の使用に他の方法で対処し得る種々の方法を認識するであろう。

10

【0059】

図 10 は、適応クロッピングが可能な一例の映像エンコーダ 116、1000 を示すシステムブロック図である。一例の映像エンコーダ 116、1000 は入力映像 104、1004 を受信し得、これはまず、ツリー構造マクロブロックパーティション方式（例えば、四分木二分木）等の処理方式に従ってセグメント化又は分割し得る。ツリー構造のマクロブロックパーティション方式の一例には、ピクチャフレームを符号化ツリー単位（CTU: coding tree unit）と呼ばれる大きなブロック要素にパーティションすることがあり得る。幾つかの実施態様では、各 CTU は符号化単位（CU）と呼ばれる幾つかのサブブロックに更にパーティションされ得る。このパーティションの最終結果は、予測単位（PU）と呼ばれ得るサブブロック群を含み得る。変換単位（TU）が利用されてもよい。

20

【0060】

なお図 10 を参照すると、一例の映像エンコーダ 116、1000 は、イントラ予測プロセッサ 1008、グローバル動きベクトル候補を動きベクトル候補リストに追加することを含む動きベクトル候補を構築可能な、インター予測プロセッサと呼ばれることもある動き推定／補償プロセッサ 1012、変換／量子化プロセッサ 1016、逆量子化／逆変換プロセッサ 1020、インループフィルタ 1024、復号ピクチャバッファ 1028、及び／又はエントロピー符号化プロセッサ 1032 を含み得る。ビットストリームパラメータがエントロピー符号化プロセッサ 1032 に入力されて、出力ビットストリーム 1036 に含まれてもよい。

30

【0061】

動作において、引き続き図 10 を参照すると、入力映像 104、1004 のフレームの各ブロックについて、ピクチャ内予測を介してブロックを処理すべきかそれとも動き推定／補償を使用してブロック処理すべきかを判断し得る。ブロックは、イントラ予測プロセッサ 1008 又は動き推定／補償プロセッサ 1012 に提供され得る。ブロックがイントラ予測を介して処理されるべき場合、イントラ予測プロセッサ 1008 が処理を実行して予測子を出力し得る。ブロックが動き推定／補償を介して処理されるべき場合、動き推定／補償プロセッサ 1012 が、該当する場合、グローバル動きベクトル候補を動きベクトル候補リストに追加することを含む動きベクトル候補リストを構築することを含む処理を実行し得る。

40

【0062】

更に図 10 を参照すると、予測子を入力映像 104 から減算することにより残差を形成

50

し得る。残差は変換／量子化プロセッサ1016により受信され得、変換／量子化プロセッサ1016は変換処理（例えば、離散コサイン変換（DCT））を実行して係数を生成し得、係数は量子化し得る。量子化された係数及び関連する任意のシグナリング情報はエントロピー符号化プロセッサ1032に提供されて、エントロピー符号化され、出力ビットストリーム1036に含まれ得る。エントロピー符号化プロセッサ1032は、現在ブロックの符号化に関連するシグナリング情報の符号化をサポートし得る。加えて、量子化された係数は逆量子化／逆変換プロセッサ1020に提供され得、逆量子化／逆変換プロセッサ1020はピクセルを再現し得、これは予測子と結合され、インループフィルタ1024により処理され得、インループフィルタ1024の出力は、グローバル動きベクトル候補を動きベクトル候補リストに追加することを含む動きベクトル候補リストを構築することが可能な動き推定／補償プロセッサ1012による使用に向けて復号ピクチャバッファ1028に記憶され得る。

10

【0063】

引き続き図10を参照すると、少数の変形形態について詳細に上述したが、他の変更又は追加も可能である。例えば、幾つかの実施態様において、現在ブロックは任意の対称ブロック（ 8×8 、 16×16 、 32×32 、 64×64 、 128×128 等）及び任意の非対称ブロック（ 8×4 、 16×8 等）を含み得る。

【0064】

幾つかの実施態様では、なお図10を参照すると、四分木二分決定木（QTBT: quadtree plus binary decision tree）を実施し得る。QTBTでは、符号化ツリー単位レベルにおいて、QTBTのパーティションパラメータは、いかなるオーバーヘッドも送信せずにローカル特性に適合するように動的に導出し得る。続けて、符号化単位レベルにおいて、共同分類器決定木構造が不必要な反復をなくし、予測誤りリスクを制御し得る。幾つかの実施態様では、LTRフレームブロック更新モードが、QTBTのあらゆるリーフノードで利用可能な追加選択肢として利用可能であり得る。

20

【0065】

幾つかの実施態様において、なお図10を参照すると、追加のシンタックス要素がビットストリームの異なる階層レベルで伝えられ得る。例えば、有シーケンスパラメータセット（SPS）に符号化された有効化フラグを含むことにより、シーケンス全体でフラグを有効化し得る。さらに、CTUフラグが符号化ツリー単位（CTU）レベルに符号化されてよい。

30

【0066】

幾つかの実施形態は、1つ又は複数の計算システムの1つ又は複数のデータプロセッサにより実行されると、1つ以上のデータプロセッサに本明細書における動作を実行させる命令を記憶した非一時的コンピュータプログラム製品（即ち、物理的に具体化されるコンピュータプログラム製品）を含み得る。

【0067】

なお図10を参照すると、エンコーダ1000は、任意の実施形態において上述された任意の動作を任意の順に任意の反復程度で実施するように構成されている回路を含み得る。例えば、エンコーダ1000は、所望又は命令された結果が達成されるまで、単一のステップ又はステップのシーケンスを繰り返し実行するように構成し得、ステップ又はステップのシーケンスの反復は、前の反復の出力を続く反復の入力として使用し、反復の入力及び／又は出力を集計して集計結果を生成し、大域変数等の1つ又は複数の変数の低減又はデクリメントを使用し、及び／又はより大きな処理タスクの、繰り返しアドレス指定されるより小さな処理タスクのセットへの分割を使用して、反復的及び／又は再帰的に実行し得る。エンコーダ1000は、本開示に記載の任意のステップ又はステップのシーケンスを並列に実行し得、例えば、2つ以上の並列スレッド、プロセッサコア等を使用してステップを2回以上同時に及び／又は略同時に実行し得；並列スレッド及び／又はプロセッサ間のタスクの分割は、反復間のタスクの分割に適した任意のプロトコルに従って実行し

40

50

得る。当業者ならば、本開示全体を考慮した上で、ステップ、ステップのシーケンス、処理タスク、及び／又はデータを細分、共有、並びに／或いは反復、再帰、及び／又は並列処理の使用に他の方法で対処し得る種々の方法を認識するであろう。

【0068】

引き続き図10を参照すると、非一時的コンピュータプログラム製品（即ち、物理的に具体化されるコンピュータプログラム製品）は、1つ又は複数の計算システムの1つ又は複数のデータプロセッサにより実行されると、1つ以上のデータプロセッサに、限定ではなく上述の任意の動作並びに／或いはデコーダ900及び／又はエンコーダ1000が実行するように構成し得る任意の動作を含め、本開示に記載の動作及び／又はそのステップを実行させる命令を記憶し得る。同様に、1つ又は複数のデータプロセッサと、1つ又は複数のデータプロセッサに結合されたメモリとを含み得るコンピュータシステムも記載される。メモリは、1つ以上のプロセッサに本明細書に記載の動作の1つ又は複数を実行させる命令を一時的又は永久的に記憶し得る。加えて、方法は、単一の計算システム内の又は2つ以上の計算システムに分散した1つ又は複数のデータプロセッサにより実施することができる。そのような計算システムは、ネットワーク（例えば、インターネット、無線ワイドエリアネットワーク、ローカルエリアネットワーク、ワイドエリアネットワーク、有線ネットワーク等）を経由した接続、複数の計算システムの1つ又は複数間の直接接続等を含め、1つ又は複数の接続を介して接続することができ及びデータ及び／又はコマンド又は他の命令等を交換することができる。

【0069】

これより図11を参照して、本開示に記載の1つ又は複数の機械学習プロセスを実行し得る機械学習モジュール1100の例示的な一実施形態を示す。機械学習モジュールは、機械学習プロセスを使用して、本開示に記載の判断、分類、及び／又は分析ステップ、方法、プロセス等を実行し得る。「機械学習プロセス」は、本開示で使用される場合、トレーニングデータ1104を自動的に使用して、計算デバイス／モジュールにより実行されて、入力1112として提供されたデータを所与として出力1108を生成するアルゴリズムを生成するプロセスである。これは、実行されるコマンドが一般に、ユーザにより事前に決定され、プログラミング言語で書かれる非機械学習ソフトウェアプログラムとは対照的である。

【0070】

なお図11を参照すると、「トレーニングデータ」は、本明細書で使用される場合、機械学習プロセスがデータ要素の2つ以上のカテゴリ間の関係をモデリングするために使用し得る相関を含むデータである。例えば、限定ではなく、トレーニングデータ1104は複数のデータエントリを含み得、各エントリは、一緒に記録、受信、及び／又は生成されたデータ要素のセットを表し、データ要素は、所与のデータエントリにおける近接性等により所与のデータエントリ内での共有される存在により相関し得る。トレーニングデータ1104中の複数のデータエントリは、データ要素のカテゴリ間の相関における1つ又は複数の傾向を明らかにし得る；例えば、限定ではなく、データ要素の第1のカテゴリに属する第1のデータ要素の値がより高いことは、データ要素の第2のカテゴリに属する第2のデータ要素の値がより高いことと相関する傾向があり得、これらの2つのカテゴリに属する値をリンクする比例又は他の数学的関係の可能性を示す。データ要素の複数のカテゴリは種々の相関に従ってトレーニングデータ1104において関連し得る；相関は、データ要素のカテゴリ間の因果的リンク及び／又は予測的リンクを示し得、これは、更に詳細に後述するように機械学習プロセスにより数学的関係等の関係としてモデリングすることができる。トレーニングデータ1104は、データ要素のカテゴリにより、例えばデータ要素に、データ要素のカテゴリに対応する1つ又は複数の記述子を関連付けることにより形式化及び／又は編成し得る。非限定的な例として、トレーニングデータ1104は、ある形態での所与のフィールド内の所与のデータ要素のエントリをカテゴリの1つ又は複数の記述子にマッピングすることができるよう、人又はプロセスにより標準化された形態で入力されるデータを含み得る。トレーニングデータ1104内の要素は、タグ、トーク

ン、又は他のデータ要素によりカテゴリの記述子にリンクすることができる；例えば、限定ではなく、トレーニングデータ1104は、固定長形式、データの位置をカンマ区切り値（CSV）形式等のカテゴリにリンクする形式、及び／又はプロセッサ又はデバイスがデータのカテゴリを検出できるようにする拡張マークアップ言語（XML）、J A V A（登録商標）S c r i p tオブジェクト表記（J S O N）等の自己記述形式で提供し得る。

【0071】

これに代えて又は加えて、引き続き図11を参照すると、トレーニングデータ1104は分類されていない1つ又は複数の要素を含み得る；即ち、トレーニングデータ1104は、データの幾つかの要素に関して形式化されていない又はデータの幾つかの要素の記述子を含まないことがある。機械学習アルゴリズム及び／又は他のプロセスは、例えば、自然言語処理アルゴリズム、トークン化、生データにおける相関値の検出等を使用して、1つ又は複数の分類を使用してトレーニングデータ1104をソートし得る；カテゴリは、相関及び／又は他の処理アルゴリズムを使用して生成し得る。非限定的な例として、テキストのコーパスにおいて、他の名詞で修飾された名詞等の幾つか「n」個の複合語を構成する句は、特定の順でそのような語を含むnグラムの統計学的に有意な出現率に従って識別し得る；そのようなnグラムは、単一の語と同様に追跡されるべき「語」等の言語の要素として分類し得、統計解析の結果として新しいカテゴリを生成する。同様に、幾らかのテキストデータを含むデータエントリにおいて、人の名前は、リスト、辞書、又は他の用語一覧を参照して識別し得、機械学習アルゴリズムによるアドホック分類及び／又は記述子を有するデータエントリにおけるデータの自動関連付け又は所与の形式への分類を可能にする。更に詳細に後述するように、データエントリを自動的に分類する能力により、同じトレーニングデータ1104を2つ以上の別個の機械学習アルゴリズムに適用できるようになり得る。機械学習モジュール1100により使用されるトレーニングデータ1104は、本開示に記載の任意の入力データを本開示に記載の任意の出力データと相関付け得る。

【0072】

図11を更に参照すると、トレーニングデータは、更に後述する1つ又は複数の教師あり及び／又は教師なし機械学習プロセス及び／又はモデルを使用してフィルタリング、ソート、及び／又は選択し得、そのようなモデルは、限定ではなく、トレーニングデータ分類器1116を含み得る。トレーニングデータ分類器1116は「分類器」を含み得、分類器は、本開示で使用される場合、入力をデータのカテゴリ又はピンにソートし、データのカテゴリ又はピン及び／又はそれと関連付けられたラベルを出力する数学モデル、ニューラルネット、又は更に後述する「分類アルゴリズム」として知られる機械学習アルゴリズムにより生成されるプログラム等の以下定義される機械学習モデルである。分類器は、一緒にクラスタリングされた又は後述される距離メトリック下で近いことがわかった等のデータのセットをラベリング又は他の方法で識別する1つ以上のデータを出力するように構成し得る。機械学習モジュール1100は、計算デバイス並びに／或いは計算デバイスで動作している任意のモジュール及び／又はコンポーネントがトレーニングデータ1104から分類器を導出するプロセスとして定義される分類アルゴリズムを使用して分類器を生成し得る。

【0073】

分類は、限定ではなく、限定ではなくロジスティック回帰及び／又はナイーブベイズ分類器等の線形分類器、k最近傍分類器等の最近傍分類器、サポートベクターマシン、最小二乗サポートベクターマシン、フィッシャーの線形判別、二次分類器、決定木、ブーストツリー、ランダムフォレスト分類器、学習ベクトル量子化、及び／又はニューラルネットワークベースの分類器を使用して実行し得る。

【0074】

なお図11を参照すると、機械学習モジュール1100は遅延学習プロセス1120及び／又はプロトコルを実行するように構成し得、遅延学習プロセス1120及び／又はプロトコルは、代替的には「遅延読み込み」又は「必要時呼び出し（call-when-

10

20

30

40

50

needed)」プロセス及び／又はプロトコルと呼ばれ得、入力とトレーニングセットとを組み合わせ、オンデマンドで出力を生成するのに使用されるアルゴリズムを導出することにより、出力に変換すべき入力を受信したとき、機械学習を行うプロセスであり得る。例えば、シミュレーションの初期セットを実行して、出力及び／又は関係において初期ヒューリスティック及び／又は「最初の推測」をカバーし得る。非限定的な例として、初期ヒューリスティックは入力とトレーニングデータ1104の要素との間の関連のランクを含み得る。ヒューリスティックは、幾つかの数の最高ランク及び／又はトレーニングデータ1104の要素を選択することを含み得る。遅延学習は、限定ではなく、k最近傍アルゴリズム、遅延ナイーブベイズアルゴリズム等を含め、任意の適した遅延学習アルゴリズムを実施し得、当業者ならば、本開示全体を考慮した上で、限定ではなく、更に後述する機械学習アルゴリズムの遅延学習適用を含め、本開示に記載の出力を生成するために適用し得る種々の遅延学習アルゴリズムを認めるであろう。

10

【0075】

これに代えて又は加えて、図11を引き続き参照すると、本開示に記載の機械学習プロセスは、機械学習モデル1124の生成に使用し得る。「機械学習モデル」は、本開示で使用される場合、限定ではなく、上述した任意のプロセスを含む任意の機械学習プロセスを使用して生成され、メモリに記憶される、入力と出力との間の関係の数学的表現及び／又はアルゴリズム表現であり、機械学習モデル1124が作成されると、入力が機械学習モデル1124に提出され、機械学習モデル1124は、導出された関係に基づいて出力を生成する。限定ではなく例として、線形回帰アルゴリズムを使用して生成された線形回帰モデルは、機械学習プロセス中に導出された係数を使用して入力データの一次結合を計算して、出力データを算出し得る。更なる非限定的な例として、機械学習モデル1124は、ノードの入力層と、1つ又は複数の中間層と、ノードの出力層とを含む畳み込みニューラルネットワーク等の人工ニューラルネットワークを作成することにより生成し得る。ノード間の接続は、ネットワークの「トレーニング」プロセスを介して作成し得、トレーニングプロセスにおいて、トレーニングデータ1104のセットからの要素が入力ノードに適用され、次いで、適したトレーニングアルゴリズム（レーベンバーグ・マルカート、共役勾配、模擬アニーリング、又は他のアルゴリズム等）が使用されて、ニューラルネットワークの隣接層内のノード間の接続及び重みを調整して、出力ノードにおいて所望の値を生成する。このプロセスは深層学習と呼ばれることがある。

20

30

【0076】

なお図11を参照すると、機械学習アルゴリズムは1つ以上の教師あり機械学習プロセス1128を含み得る。1つ以上の教師あり機械学習プロセス1128は、本明細書に定義されるように、幾つかの入力を幾つかの出力に関連付けるトレーニングセットを受信し、入力を出力に関連付ける1つ又は複数の数学的関係を見つけようとするアルゴリズムを含み、1つ又は複数の数学的関係の各々は、何らかのスコアリング関数を使用して、アルゴリズムに指定された何らかの基準に従って最適である。例えば、教師あり機械学習アルゴリズムは、本開示において入力として記載される入力、本開示において出力として記載される出力、及び入力と出力との間で検出されるべき関係の所望の形態を表すスコアリング関数を含み得、スコアリング関数は、例えば、入力及び／又は要素の組合せを所与として、所与の入力が所与の出力と関連付けられない確率を最小化するように入力が所与の出力と関連付けられる確率を最大化しようとし得る。スコアリング関数は、入力を出力に関連付けるアルゴリズムの「期待損失」を表すリスク関数として表現し得、損失は、関係により生成された予測が、トレーニングデータ1104において提供される所与の入力-出力対と比べて不正確である程度を表す誤差関数として計算される。当業者ならば、本開示全体を考慮した上で、入力と出力との間の関係を突き止めるために使用し得る1つ以上の教師あり機械学習プロセス1128の種々の可能な変形形態を認めるであろう。教師あり機械学習プロセスは、上記定義された分類アルゴリズムを含み得る。

40

【0077】

図11を更に参照すると、機械学習プロセスは1つ以上の教師なし機械学習プロセス1

50

1 3 2を含み得る。教師なし機械学習プロセスは、本明細書で使用される場合、ラベルを考慮せずにデータセット内の推測を導出するプロセスであり、その結果、教師なし機械学習プロセスは、データにおいて提供される任意の構造、関係、及び／又は相関を自由に発見することができる。教師なしプロセスは応答変数を必要としなくてよく、教師なしプロセスは、変数間の興味深いパターン及び／又は推論を見つけるため、2つ以上の変数間の相関の程度を突き止めるため等に使用し得る。

【0078】

なお図11を参照すると、機械学習モジュール1100は、線形回帰モデルの発展させる技法を使用して機械学習モデル1124を作成するように設計及び構成し得る。線形回帰モデルは、予測された結果と実際の結果との間の差を測定するのに適したノルム（例えば、ベクトル空間距離ノルム）に従ってそのような差の二乗を最小化することを目的とした通常の最小二乗回帰を含み得、その結果生成された一次方程式の係数は最小化を改善するように変更し得る。線形回帰モデルはリッジ回帰法を含み得、リッジ回帰法では、最小化すべき関数は、最小二乗関数に、大きな係数にペナルティを科すスカラー量により各係数の二乗を乗算する項を含む。線形回帰モデルは最小絶対値縮小選択演算子（LASSO：least absolute shrinkage and selection operator）モデルを含み得、このモデルでは、リッジ回帰が、最小二乗項を係数1で乗算し、サンプル数の2倍で除算することと組み合わせられる。線形回帰モデルはマルチタスクLASSOモデルを含み得、このモデルでは、LASSOモデルの最小二乗項に適用されるノルムは、全項の二乗和の平方根になるフロベニウスノルムである。線形回帰モデルは、弾性ネットモデル、マルチタスク弾性ネットモデル、最小角度回帰モデル、LARS LASSOモデル、直交マッチング追跡モデル、ベイズ回帰モデル、ロジスティック回帰モデル、確率勾配降下モデル、パーセプトロンモデル、パッシブ-アグレッシブアルゴリズム、ロバスト回帰モデル、フーバー回帰モデル、又は本開示全体を考慮した上で当業者が思い付き得る任意の他の適したモデルを含み得る。線形回帰モデルは、一実施形態では、多項式回帰モデルに一般化することができ、それにより、多項式（例えば、二次、三次、又はより高次の方程式）は最良予測出力／適合が探される実際の出力を提供し、本開示全体を考慮した上で当業者には明らかになるように、上述した方法と同様の方法を適用して誤差関数を最小化し得る。

【0079】

引き続き図11を参照すると、機械学習アルゴリズムは、限定ではなく、線形判別解析を含み得る。機械学習アルゴリズムは二次判別解析を含み得る。機械学習アルゴリズムはカーネルリッジ回帰を含み得る。機械学習アルゴリズムは、限定ではなく、サポートベクター分類ベースの回帰プロセスを含め、サポートベクターマシンを含み得る。機械学習アルゴリズムは、確率勾配降下に基づく分類及び回帰アルゴリズムを含め、確率勾配降下アルゴリズムを含み得る。機械学習アルゴリズムは最近傍アルゴリズムを含み得る。機械学習アルゴリズムは、変分正則化等の種々の形態の線形空間正則化を含み得る。機械学習アルゴリズムは、ガウス過程回帰等のガウス過程を含み得る。機械学習アルゴリズムは、部分最小二乗及び／又は正準相関解析を含め、クロス分解アルゴリズムを含み得る。機械学習アルゴリズムはナイーブベイズ法を含み得る。機械学習アルゴリズムは、決定木分類又は回帰アルゴリズム等の決定木に基づくアルゴリズムを含み得る。機械学習アルゴリズムは、バギングメタ推定器、ランダムトレスフォレスト（forest of randomized trees）、AdaBoost、勾配ブーストツリー、及び／又は投票分類器等のアンサンブル法を含み得る。機械学習アルゴリズムは、畳み込みニューラルネットワークプロセス等のニューラルネットワークアルゴリズムを含み得る。

【0080】

なお、コンピュータ分野の当業者には明らかになるように、本明細書に記載の態様及び実施形態の任意の1つ又は複数は、本明細書の教示に従ってプログラムされた1つ又は複数の機械（例えば、電子ドキュメントのユーザ計算デバイスとして利用される1つ又は複数の計算デバイス、ドキュメントサーバ等の1つ又は複数のサーバデバイス等）を使用し

て好都合に実施し得る。適切なソフトウェアコーディングは、ソフトウェア分野の当業者には明らかになるように、本開示の教示に基づいて熟練したプログラマにより容易に準備することができる。ソフトウェア及び／又はソフトウェアモジュールを採用した上記態様及び実施態様は、ソフトウェア及び／又はソフトウェアモジュールの機械実行可能命令の実施を支援するのに適したハードウェアを含むこともできる。

【0081】

そのようなソフトウェアは、機械可読記憶媒体を採用するコンピュータプログラム製品であり得る。機械可読記憶媒体は、機械（例えば、計算デバイス）により実行される命令のシーケンスを記憶及び／又は符号化することが可能であり、機械に本明細書に記載の方法論及び／又は実施形態のいずれか1つを実行させる任意の媒体であり得る。機械可読記憶媒体の例には、限定ではないが、磁気ディスク、光ディスク（例えば、CD、CD-R、DVD、DVD-R等）、光磁気ディスク、読取専用メモリ「ROM」デバイス、ランダムアクセスメモリ「RAM」デバイス、磁気カード、光学カード、ソリッドステートメモリデバイス、EPROM、EEPROM、及び／又はそれらの任意の組合せがある。機械可読媒体は、本明細書で使用される場合、単一の媒体及び物理的に別個の媒体の集合、例えば、コンピュータメモリと組み合わせたコンパクトディスク又は1つ若しくは複数のハードディスクドライブの集合体等を含むことを意図する。本明細書で使用される場合、機械可読記憶媒体は一過性形態の信号伝送を含まない。

【0082】

そのようなソフトウェアはまた、搬送波等のデータキャリア上のデータ信号として搬送される情報（例えば、データ）を含んでもよい。例えば、機械実行可能情報はデータキャリアの中に具現化されるデータ搬送信号として含まれてもよく、信号は、機械（例えば、計算デバイス）により実行される命令のシーケンス又はその一部及び本明細書に記載の方法論及び／又は実施形態のいずれか1つを機械に実行させる任意の関連する情報（例えば、データ構造及びデータ）を符号化する。

【0083】

計算デバイスの例には、限定ではないが、電子書籍読取デバイス、コンピュータワークステーション、端末コンピュータ、サーバコンピュータ、ハンドヘルドデバイス（例えば、タブレット型コンピュータ、スマートフォン等）、ウェブ装置、ネットワークルータ、ネットワークスイッチ、ネットワークブリッジ、その機械によってとられるべき動作を規定する命令のシーケンスを実行可能な任意の機械、及びそれらの任意の組合せを含む。一例では、計算デバイスはキオスクを含んでもよく及び／又はその中に含まれてもよい。

【0084】

図12は、制御システムに本開示の態様及び／又は方法論の任意の1つ又は複数を実行させる命令セットを内部で実行し得るコンピュータシステム1200の例示的な形態における計算デバイスの一実施形態の図式表現を示す。本開示の態様及び／又は方法論の任意の1つ又は複数を実行させるデバイスの1つ又は複数に実行させる、特に構成されている命令セットを実施するのに複数の計算デバイスを利用してもよいことも企図される。コンピュータシステム1200は、バス1212を介して互い及び他の構成要素と通信するプロセッサ1204及びメモリ1208を含む。バス1212は、限定ではなく、多様な任意のバスアーキテクチャを使用するメモリバス、メモリコントローラ、周辺機器用バス、ローカルバス、及びそれらの任意の組合せを含む幾つかの種類の任意のバス構造を含み得る。

【0085】

プロセッサ1204は、限定ではなく、状態機械を用いて調整され得、メモリ及び／又はセンサからの動作入力により指示され得る、算術論理ユニット（ALU）等の算術演算及び論理演算を実行する論理回路を組み込んだプロセッサ等の任意の適したプロセッサを含み得、プロセッサ1204は、非限定的な例としてフォンノイマン及び／又はハーバードアーキテクチャに従って編成し得る。プロセッサ1204は、限定ではなく、マイクロコントローラ、マイクロプロセッサ、デジタル信号プロセッサ（DSP）、フィールドプログラマブルゲートアレイ（FPGA）、コンプレックスプログラマブル論理デバイス（

C P L D)、グラフィカル処理ユニット (G P U)、汎用 G P U、テンソル処理ユニット (T P U)、アナログ又は混合信号プロセッサ、トラステッドプラットフォームモジュール (T P M)、浮動小数点ユニット (F P U)、及び／又はシステムオンチップ (S o C) を含んでよく、組み込んでよく、及び／又は組み込まれてよい。

【0086】

メモリ1208は、限定ではないが、ランダムアクセスメモリコンポーネント、読取専用コンポーネント、及びそれらの任意の組合せを含め、種々のコンポーネント（例えば、機械可読媒体）を含んでもよい。一例では、スタートアップ中にコンピュータシステム1200内の要素間で情報を転送するのに役立つ基本ルーチンを含む基本入力／出力システム1216 (B I O S) が、メモリ1208に記憶されてもよい。メモリ1208はまた、本開示の態様及び／又は方法論のうちの任意の1つ又は複数を具現化する命令（例えば、ソフトウェア）1220を含んで（例えば、1つ又は複数の機械可読媒体に記憶されて）もよい。別の例では、メモリ1208は更に、限定ではないが、オペレーティングシステム、1つ又は複数のアプリケーションプログラム、他のプログラムモジュール、プログラムデータ、及びそれらの任意の組合せを含め、任意の数のプログラムモジュールを含んでもよい。

10

【0087】

コンピュータシステム1200はまた、記憶デバイス1224を含んでもよい。記憶デバイス（例えば、記憶デバイス1224）の例には、限定ではないが、ハードディスクドライブ、磁気ディスクドライブ、光学媒体との組み合わせた光ディスクドライブ、ソリッドステートメモリデバイス、及びそれらの任意の組合せがある。記憶デバイス1224は、適切なインタフェース（図示せず）によってバス1212に接続されてよい。インタフェースの例は、限定ではないが、S C S I、アドバンスド・テクノロジー・アタッチメント (A T A)、シリアルA T A、ユニバーサルシリアルバス (U S B)、I E E E 1394 (ファイアワイヤ (F I R E W I R E) (登録商標))、及びそれらの任意の組合せがある。一例では、記憶デバイス1224（又は1つ若しくは複数のそのコンポーネント）は、コンピュータシステム1200と取り外し可能にインタフェースし得る（例えば、外部ポートコネクタ（図示せず）を介して）。特に、記憶デバイス1224及び関連する機械可読媒体1228は、コンピュータシステム1200の機械可読命令、データ構造、プログラムモジュール、及び／又は他のデータの不揮発性及び／又は揮発性記憶装置を提供し得る。一例では、ソフトウェア1220は、完全又は部分的に、機械可読媒体1228内に常駐してよい。別の例では、ソフトウェア1220は、完全又は部分的に、プロセッサ1204内に常駐してよい。

20

30

【0088】

コンピュータシステム1200はまた、入力デバイス1232を含んでもよい。一例では、コンピュータシステム1200のユーザが、入力デバイス1232を介してコンピュータシステム1200にコマンド及び／又は他の情報を入力し得る。入力デバイス1232の例には、限定ではないが、英数字入力デバイス（例えば、キーボード）、ポインティングデバイス、ジョイスティック、ゲームパッド、オーディオ入力デバイス（例えば、マイクロホン、音声応答システム等）、カーソル制御デバイス（例えば、マウス）、タッチパッド、光学スキャナ、映像キャプチャデバイス（例えば、静止カメラ、映像カメラ）、タッチスクリーン、及びそれらの任意の組合せがある。入力デバイス1232は、限定ではないが、シリアルインタフェース、パラレルインタフェース、ゲームポート、U S Bインタフェース、ファイアワイヤ (F I R E W I R E) インタフェース、バス1212への直接インタフェース、及びそれらの任意の組合せを含め、種々の任意のインタフェース（図示せず）を介してバス1212とインタフェースし得る。入力デバイス1232は、更に後述されるディスプレイ1236の一部であってもよく又はそれと別個であってもよい。タッチスクリーンインタフェースを含んでもよい。入力デバイス1232は、上述したようなグラフィカルインタフェース内で1つ又は複数のグラフィカル表現を選択するためのユーザ選択デバイスとして利用されてもよい。

40

50

【0089】

ユーザはまた、記憶デバイス1224（例えば、リムーバブルディスクドライブ、フラッシュドライブ等）及び／又はネットワークインタフェースデバイス1240を介してコマンド及び／又は他の情報をコンピュータシステム1200に入力することもできる。ネットワークインタフェースデバイス1240等のネットワークインタフェースデバイスは、ネットワーク1244等の種々のネットワークの1つ又は複数及びそれに接続される1つ又は複数のリモートデバイス1248にコンピュータシステム1200を接続するために利用し得る。ネットワークインタフェースデバイスの例には、限定ではないが、ネットワークインタフェースカード（例えば、モバイルネットワークインタフェースカード、LANカード）、モデム、及びそれらの任意の組合せがある。ネットワークの例には、限定ではないが、ワイドエリアネットワーク（例えば、インターネット、企業ネットワーク）、ローカルエリアネットワーク（例えば、オフィス、ビル、キャンパス、又は他の比較的小さい地理的空間と関連付けられるネットワーク）、電話ネットワーク、電話／音声プロバイダと関連付けられるデータネットワーク（例えば、モバイル通信プロバイダのデータ及び／又は音声ネットワーク）、2つの計算デバイス間の直接接続、及びそれらの任意の組合せがある。ネットワーク1244等のネットワークは、通信の有線モード及び／又は無線モードを採用し得る。一般に、いかなるネットワークトポロジも使用可能である。情報（例えば、データ、ソフトウェア1220等）が、ネットワークインタフェースデバイス1240を介してコンピュータシステム1200に及び／又はコンピュータシステム1200から通信し得る。コンピュータシステム1200は、ディスプレイデバイス1236等のディスプレイデバイスに表示可能な画像を通信するための映像ディスプレイアダプタ1252を更に含み得る。ディスプレイデバイスの例には、限定ではなく、液晶ディスプレイ（LCD）、陰極線管（CRT）、プラズマディスプレイ、発光ダイオード（LED）ディスプレイ、及びそれらの任意の組合せがある。

10

20

【0090】

ディスプレイアダプタ1252及びディスプレイデバイス1236がプロセッサ1204と組み合わせて利用されて、本開示の態様のグラフィカル表現を提供し得る。ディスプレイデバイスに加えて、コンピュータシステム1200は、限定ではないが、オーディオスピーカ、プリンタ、及びそれらの任意の組合せを含む1つ又は複数の他の周辺機器出力デバイスを含み得る。そのような周辺機器出力デバイスは、周辺機器インタフェース1256を介してバス1212に接続し得る。周辺機器インタフェースの例には、限定ではなく、シリアルポート、USB接続、ファイアワイヤ（FIREWIRE）接続、パラレル接続、及びそれらの任意の組合せがある。

30

【0091】

上記は本発明の例示的な実施形態の詳細な説明であった。本発明の主旨及び範囲から逸脱せずに種々の変更及び追加を行うことができる。上述した種々の実施形態の各々の特徴は、関連する新しい実施形態において多くの特徴組合せを提供するために、記載の他の実施形態の特徴と適宜組み合わせることができる。さらに、上記は幾つかの別個の実施形態について説明しているが、本明細書に記載されたものは単に、本発明の原理の適用の例示である。さらに、本明細書における特定の方法是特定の順序で実行されるものとして示され及び／又は説明されていることがあるが、順序は、本開示による方法、システム、及びソフトウェアを達成するために通常技能内で大きく変更可能である。したがって、この説明は例としてのみ解釈されることが意図され、本発明の範囲の限定を意図しない。

40

【0092】

例示的な実施形態について上に開示し添付図面に示してきた。本発明の主旨及び範囲から逸脱せずに、本明細書に特に開示されたものに種々の変更、省略、及び追加を行い得ることが当業者には理解されよう。

50

【図面】
【図 1】

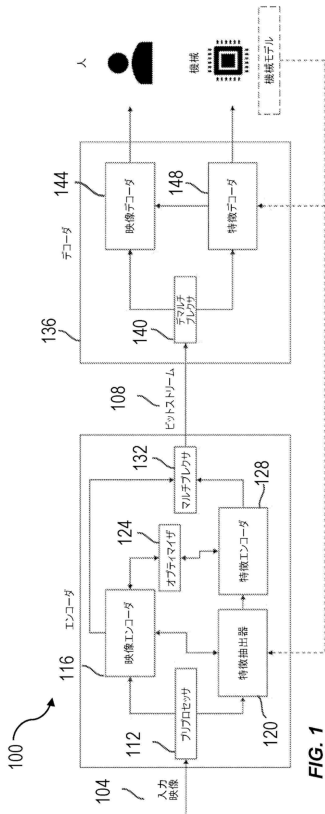


FIG. 1

【図 2】

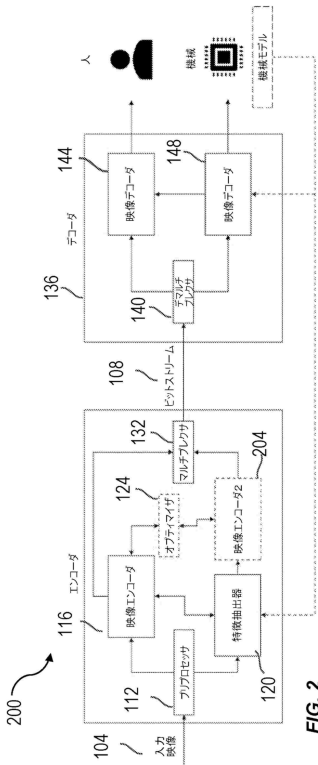


FIG. 2

【図 3】

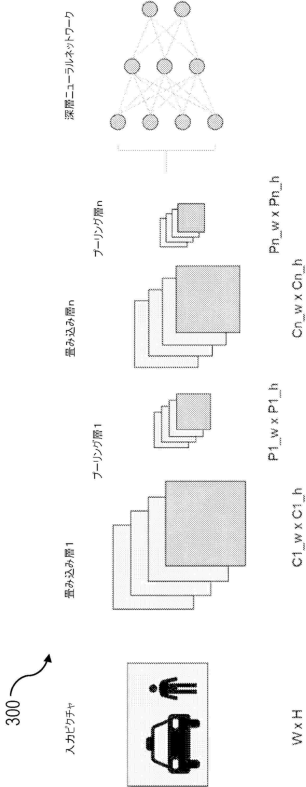


FIG. 3

【図 4】

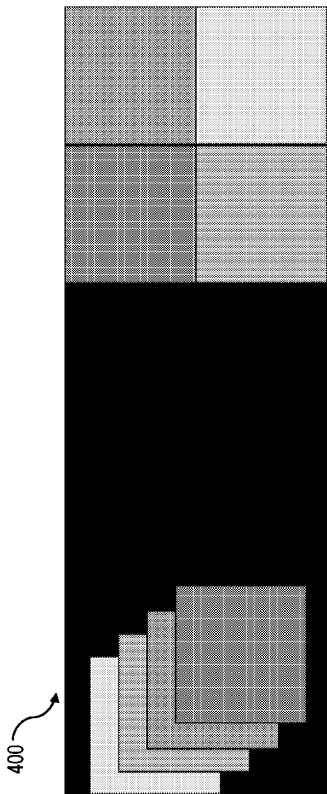


FIG. 4

10

20

30

40

50

【図 5】

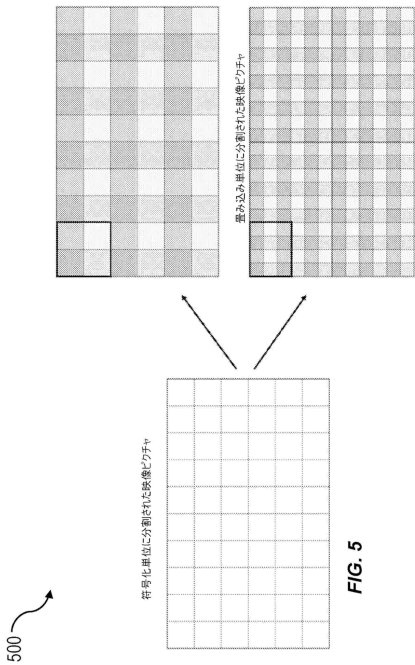


FIG. 5

【図 6】

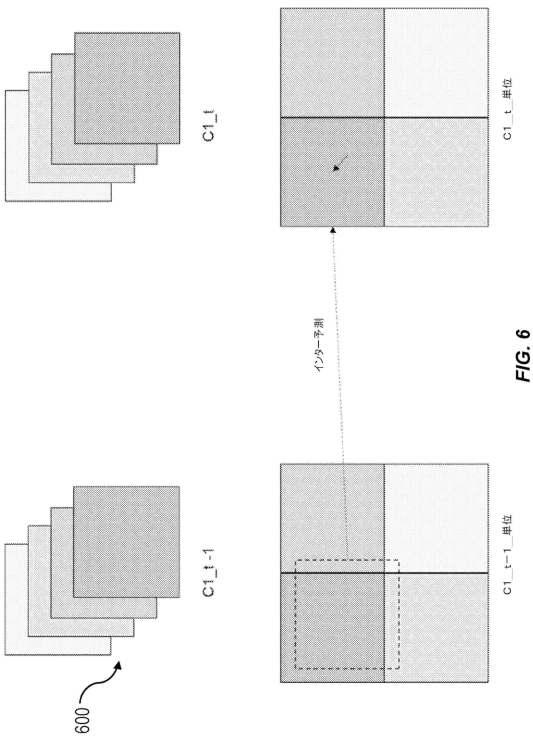


FIG. 6

【図 7】

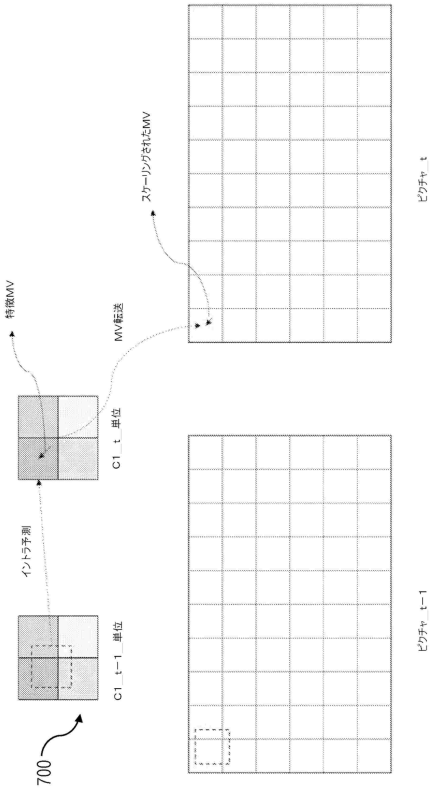


FIG. 7

【図 8】

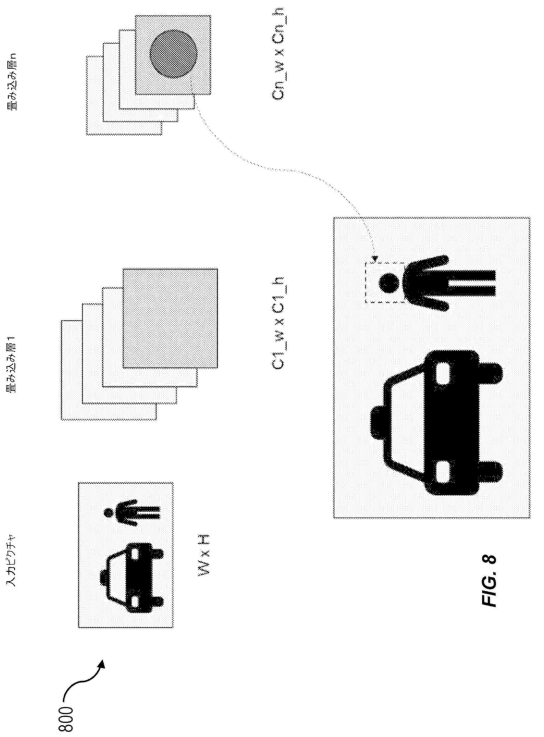


FIG. 8

【図 9】

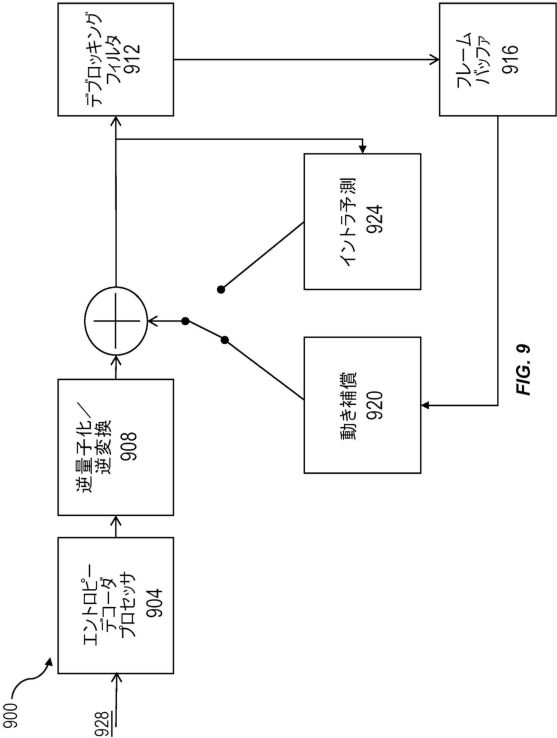


FIG. 9

【図 10】

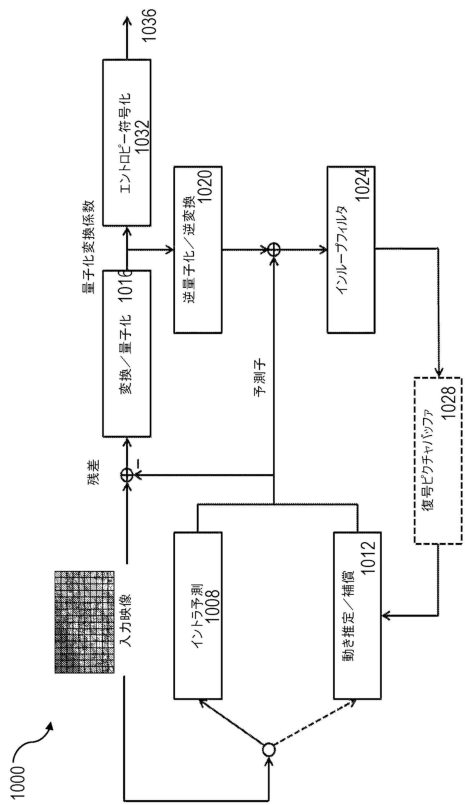


FIG. 10

【図 11】

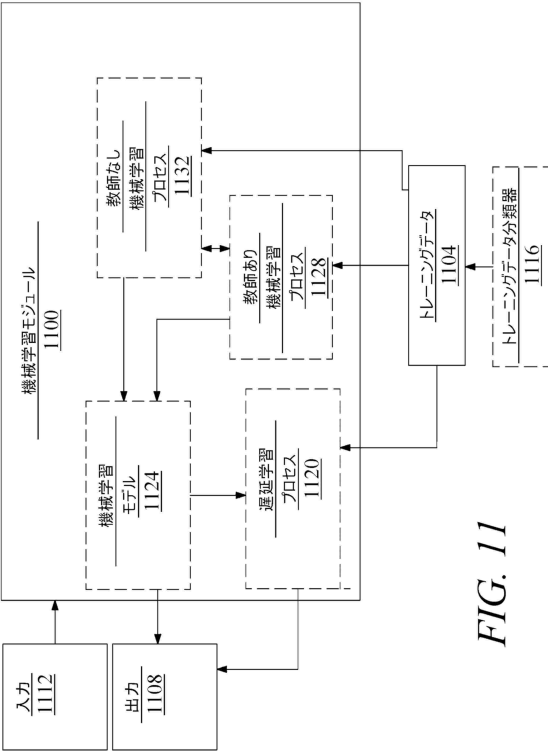


FIG. 11

【図 12】

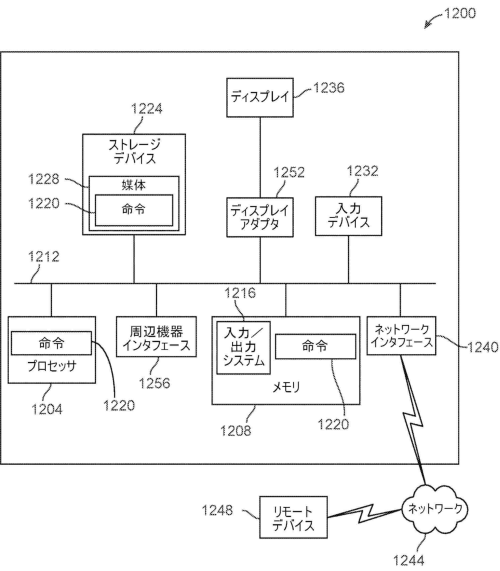


FIG. 12

10

20

30

40

50

【国際調査報告】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US 22/32048
A. CLASSIFICATION OF SUBJECT MATTER IPC - INV. G06V 10/40; ADD. G06V 10/764, G06V 10/82 (2022.01) CPC - INV. G06V 10/40; ADD. G06V 10/764, G06V 10/82		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) See Search History document		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched See Search History document		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) See Search History document		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2020/055279 A1 (Huawei Technologies Co., Ltd.) 19 March 2020 (19.03.2020) entire document (especially pg. 26, ln 26 to pg. 27, ln 21; pg. 28, ln 3; pg. 29, ln 19-29; pg. 31, ln 11-22; pg. 33, ln 1-18 and 29-31)	1-15
A	US 2019/0306526 A1 (Electronics and Telecommunications Research Institute) 03 October 2019 (03.10.2019) entire document	1-15
☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "D" document cited by the applicant in the international application "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 26 September 2022		Date of mailing of the international search report OCT 11 2022
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1460, Alexandria, Virginia 22313-1460 Facsimile No. 571-273-8300		Authorized officer Kari Rodriguez Telephone No. PCT Helpdesk: 571-272-4300

Form PCT/ISA/210 (second sheet) (July 2019)

フロントページの続き

MK,MT,NL,NO,PL,PT,RO,RS,SE,SI,SK,SM,TR),OA(BF,BJ,CF,CG,CI,CM,GA,GN,GQ,GW,KM,ML,MR,N
E,SN,TD,TG),AE,AG,AL,AM,AO,AT,AU,AZ,BA,BB,BG,BH,BN,BR,BW,BY,BZ,CA,CH,CL,CN,CO,CR,CU,
CZ,DE,DJ,DK,DM,DO,DZ,EC,EE,EG,ES,FI,GB,GD,GE,GH,GM,GT,HN,HR,HU,ID,IL,IN,IQ,IR,IS,IT,JM,J
O,JP,KE,KG,KH,KN,KP,KR,KW,KZ,LA,LC,LK,LR,LS,LU,LY,MA,MD,ME,MG,MK,MN,MW,MX,MY,M
Z,NA,NG,NI,NO,NZ,OM,PA,PE,PG,PH,PL,PT,QA,RO,RS,RU,RW,SA,SC,SD,SE,SG,SK,SL,ST,SV,SY,TH,
TJ,TM,TN,TR,TT,TZ,UA,UG,US,UZ,VC,VN,WS,ZA,ZM,ZW

(72)発明者 フルト、ポリヴォジェ
アメリカ合衆国 3 3 4 3 1 フロリダ州 ボカ ラトン グレイズ ロード 7 7 7

(72)発明者 アジッチ、ベリボル
アメリカ合衆国 3 0 1 4 4 ジョージア州 カントン バレー コート 5 0 1

F ターム (参考) 5C159 MA04 MA05 MA21 MC11 ME01 PP04 RC11 TA16 TB04