

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7443559号
(P7443559)

(45)発行日 令和6年3月5日(2024.3.5)

(24)登録日 令和6年2月26日(2024.2.26)

(51)国際特許分類

F I

H O 4 N 19/85 (2014.01)

H O 4 N 19/85

H O 4 N 19/46 (2014.01)

H O 4 N 19/46

請求項の数 12 (全37頁)

(21)出願番号	特願2022-558402(P2022-558402)	(73)特許権者	520353802
(86)(22)出願日	令和3年9月8日(2021.9.8)		テンセント・アメリカ・エルエルシー
(65)公表番号	特表2023-515890(P2023-515890 A)		アメリカ合衆国 9 4 3 0 6 カリフォルニア州 パロアルト パーク・ブールヴァード 2 7 4 7
(43)公表日	令和5年4月14日(2023.4.14)	(74)代理人	100107766
(86)国際出願番号	PCT/US2021/049516		弁理士 伊東 忠重
(87)国際公開番号	WO2022/146503	(74)代理人	100070150
(87)国際公開日	令和4年7月7日(2022.7.7)		弁理士 伊東 忠彦
審査請求日	令和4年10月24日(2022.10.24)	(74)代理人	100135079
(31)優先権主張番号	63/131,656		弁理士 宮崎 修
(32)優先日	令和2年12月29日(2020.12.29)	(72)発明者	アウヨン, チュン
(33)優先権主張国・地域又は機関	米国(US)		アメリカ合衆国 9 4 3 0 6 カリフォルニア州 パロアルト パーク・ブールバード 2 7 4 7 テンセント アメリカ エル
(31)優先権主張番号	17/463,352		最終頁に続く
(32)優先日	令和3年8月31日(2021.8.31)		
	最終頁に続く		

(54)【発明の名称】 ビデオコーティングのための方法及び装置

(57)【特許請求の範囲】

【請求項1】

ビデオ処理のための方法であって、
処理回路が、色空間におけるサブサンプリングフォーマットのピクチャを前記色空間における非サブサンプリングフォーマットに変換するステップと、
前記非サブサンプリングフォーマットのピクチャを入力としてニューラルネットワークに基づくフィルタに提供する前に、前記処理回路が、前記ピクチャを含むビットストリームからのデコーディングされた情報に基づいて、クリッピングのための範囲を決定し、前記非サブサンプリングフォーマットのピクチャの色成分の値を、決定された範囲にクリッピングするステップと、を含む、
方法。

【請求項2】

前記ビットストリームにおけるシーケンスパラメータセット、ピクチャパラメータセット、スライスヘッダー、及びタイルヘッダーのうち少なくとも1つから、前記範囲を示す信号をデコーディングするステップをさらに含む、
請求項1に記載の方法。

【請求項3】

ビットストリームからのデコーディングされた情報に基づいて、前記サブサンプリングフォーマットのピクチャを再構成するステップをさらに含む、
請求項1に記載の方法。

【請求項 4】

クリッピングされた値を有する前記非サブサンプリングフォーマットのピクチャに、ニューラルネットワークに基づくフィルタを適用することで、前記非サブサンプリングフォーマットのフィルタリングされたピクチャを生成するステップと、

前記非サブサンプリングフォーマットのフィルタリングされたピクチャを、前記サブサンプリングフォーマットのフィルタリングされたピクチャに変換するステップとをさらに含む、

請求項 1 に記載の方法。

【請求項 5】

クリッピングされた値を有する前記非サブサンプリングフォーマットのピクチャを保存するステップをさらに含む、

請求項 1 に記載の方法。

【請求項 6】

前記クリッピングされた値を有する前記非サブサンプリングフォーマットで保存されたピクチャをトレーニング入力として提供し、前記ニューラルネットワークに基づくフィルタにおけるニューラルネットワークをトレーニングするステップをさらに含む、

請求項 5 に記載の方法。

【請求項 7】

ビデオ処理のための装置であって、

色空間におけるサブサンプリングフォーマットのピクチャを前記色空間における非サブサンプリングフォーマットに変換し、

前記非サブサンプリングフォーマットのピクチャを入力としてニューラルネットワークに基づくフィルタに提供する前に、前記ピクチャを含むビットストリームからのデコーディングされた情報に基づいて、クリッピングのための範囲を決定し、前記非サブサンプリングフォーマットのピクチャの色成分の値を、決定された範囲にクリッピングする、

ように構成される処理回路を含む、

装置。

【請求項 8】

前記処理回路は、

前記ビットストリームにおけるシーケンスパラメータセット、ピクチャパラメータセット、スライスヘッダー、及びタイルヘッダーのうち少なくとも 1 つから、前記範囲を示す信号をデコーディングするように構成される、

請求項 7 に記載の装置。

【請求項 9】

前記処理回路は、

ビットストリームからのデコーディングされた情報に基づいて、前記サブサンプリングフォーマットのピクチャを再構成するように構成される、

請求項 7 に記載の装置。

【請求項 10】

前記処理回路は、

クリッピングされた値を有する前記非サブサンプリングフォーマットのピクチャに、ニューラルネットワークに基づくフィルタを適用することで、前記非サブサンプリングフォーマットのフィルタリングされたピクチャを生成し、

前記非サブサンプリングフォーマットのフィルタリングされたピクチャを、前記サブサンプリングフォーマットのフィルタリングされたピクチャに変換する、ように構成される、

請求項 7 に記載の装置。

【請求項 11】

クリッピングされた値を有する前記非サブサンプリングフォーマットのピクチャを保存するように構成されるストレージをさらに含む、

請求項 7 に記載の装置。

10

20

30

40

50

【請求項 1 2】

前記処理回路は、

前記クリッピングされた値を有する前記非サブサンプリングフォーマットで保存されたピクチャをトレーニング入力として提供し、前記ニューラルネットワークに基づくフィルタにおけるニューラルネットワークをトレーニングするように構成される、

請求項 1 1 に記載の装置。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本出願は、2020年12月29日に出願された米国仮出願第63／131,656号の「APPLICATION OF CLIPPING TO IMPROVE PRE-PROCESSING IN A NEURAL NETWORK BASED IN-LOOP FILTER IN A VIDEO

CODEC」の優先権を主張する、2021年8月31日に出願された米国特許出願第17／463,352号の「METHOD AND APPARATUS FOR VIDEO CODING」の優先権を主張する。その開示を全て援用により本明細書に組み込むものとする。

【0 0 0 2】

本開示は、全体として、ビデオコーディングの実施形態に関する。より具体的には、本開示は、ニューラルネットワークに基づくループ内フィルタを改善するための技術を提供する。

【背景技術】

【0 0 0 3】

本明細書で提供される背景技術の説明は、本開示の文脈を一般的に提示することを目的としている。本発明者らの研究は、この背景技術の項に記載されている限りにおいて、ならびに出願時に先行技術として認められないことがある説明の態様は、本開示に対する先行技術として明示的にも暗示的にも認められない。

【0 0 0 4】

ビデオコーディングおよびデコーディングは、動き補償を有するピクチャ間予測を使用し実行されることができる。非圧縮デジタルビデオは一連のピクチャを含むことができ、各ピクチャは、例えば、1920×1080ルマサンプルと関連するクロマサンプルの空間次元を持っている。該一連のピクチャは、例えば毎秒60ピクチャまたは60Hzの固定または可変ピクチャレート（非公式には「フレームレート」とも呼ばれる）を持つことができる。非圧縮ビデオには、特定のビットレート要件がある。例えば、サンプルあたり8ビットでの1080p60 4:2:0ビデオ（60Hzフレームレートでの1920×1080ルマサンプル解像度）には、1.5Gbit/sに近い帯域幅が必要となる。このようなビデオを1時間使用するには、600GBytesを超えた記憶空間が必要である。

【0 0 0 5】

ビデオコーディングおよびデコーディングの目的の1つは、圧縮による入力ビデオ信号の冗長性の低減であり得る。圧縮は、前述の帯域幅または記憶空間要件を、場合によっては2桁以上減らすのに役立ち得る。可逆圧縮および非可逆圧縮の両方、ならびにそれらの組み合わせを使用することができる。可逆圧縮とは、原信号の正確な複製を圧縮された原信号から再構成することができる技術を指す。非可逆圧縮を使用する場合、再構成された信号は原信号と同一ではないことがあるが、原信号と再構成された信号との間の歪みは、再構成された信号を意図したアプリケーションに役立てるのに十分小さい。ビデオの場合、非可逆圧縮が広く採用されている。許容される歪みの量はアプリケーションに依存し、例

例えば、特定の消費者向けストリーミングアプリケーションのユーザは、テレビ配信アプリケーションのユーザよりも高い歪みを許容することがある。達成可能な圧縮比は、より高い許容可能/容認可能な歪みがより高い圧縮比をもたらすことができることを反映することができる。

【0006】

ビデオエンコーダおよびデコーダは、例えば、動き補償、変換、量子化、およびエントロピーコーディングを含む、いくつかの広範なカテゴリからの技術を利用し得る。

【0007】

ビデオコーデック技術は、イントラコーディングとして知られる技術を含み得る。イントラコーディングでは、サンプル値は、予め再構成された参照ピクチャからのサンプルまたは他のデータを参照せずに表される。一部のビデオコーデックでは、ピクチャは空間的にサンプルのブロックに細分される。サンプルのすべてのブロックがイントラモードでコーディングされると、そのピクチャはイントラピクチャになる可能性がある。イントラピクチャとその独立したデコーダリフレッシュピクチャなどの派生物は、デコーダの状態のリセットするために使用できるため、コーディングされたビデオビットストリームとビデオセッションの最初のピクチャとして、または静止ピクチャとして使用されることができる。イントラブロックのサンプルは変換にさらされることができ、変換係数はエントロピーコーディングの前に量子化されることができる。イントラ予測は、事前変換領域でサンプル値を最小化する技術であり得る。場合によっては、変換後のDC値が小さく、AC係数が小さいほど、エントロピーコーディング後のブロックを表すために所定の量子化ステップサイズで必要とされるビットが少なくなる。

【0008】

例えばMPEG-2世代コーディング技術から知られているような従来のイントラコーディングは、イントラ予測を使用しない。しかしながら、いくつかのより新しいビデオ圧縮技術は、例えば、空間的に隣接し、デコーディング順で先に位置する、エンコーディング／デコーディング中に得られた周囲のサンプルデータおよび／またはメタデータから試みる技術を含む。そのような技術は、以降、「イントラ予測」技術と呼ばれる。少なくともいくつかのケースでは、イントラ予測は、再構成中の現在のピクチャからの参照データのみを使用し、参照ピクチャからの参照データを使用しないことに留意されたい。

【0009】

イントラ予測はさまざまな形態で存在し得る。そのような技術の複数が所定のビデオコーディング技術に使用できる場合、使用中の技術はイントラ予測モードでコーディングされることができる。場合によっては、モードはサブモードおよび／またはパラメータを有することができる、それらを個別にコーディングするか、もしくはモードコードワードに含めることができる。どのコードワードが所定のモード／サブモード／パラメータの組み合わせに使用されるかは、イントラ予測によるコーディング効率の向上に影響を与える可能性があり、コードワードをビットストリームに変換するために使用されるエントロピーコーディング技術も同様である。

【0010】

イントラ予測の特定のモードはH.264で提出され、H.265で改良され、さらに共同探索モデル(JEM)、多用途ビデオコーディング(VVC)、およびベンチマークセット(BMS)などのより新しいコーディング技術で改良された。予測器ブロックは、すでに利用可能なサンプルに属する隣接サンプル値を使用して形成されることができる。隣接サンプルのサンプル値は、方向に従って予測器ブロックに複製される。使用中の方向への参照は、ビットストリーム内でコーディングされ得るか、またはそのものが予測され得る。

【0011】

図1Aを参照し、右下に示されているのは、H.265の33通りの可能な予測器方向(35通りのイントラモードの33通りの角度モードに対応する)からの9通りの予測器方向のサブセットである。矢印が収束する点(101)は、予測されるサンプルを表す。矢

10

20

30

40

50

印は、サンプルが予測される方向を表す。例えば、矢印（102）は、サンプル（101）が、水平から45度の角度で右上への1つまたは複数のサンプルから予測されることを示す。同様に、矢印（103）は、サンプル（101）が、サンプル（101）の水平から22.5度の角度で左下への1つまたは複数のサンプルから予測されることを示す。

【0012】

引き続き図1Aを参照し、左上には、4×4サンプルの正方形ブロック（104）（太い破線で示される）が示されている。正方形ブロック（104）は、16個のサンプルを含み、それぞれ「S」、Y次元におけるその位置（例えば、行インデックス）、およびX次元におけるその位置（例えば、列インデックス）でラベル付けされている。例えば、サンプルS21は、Y次元（上から）における2番目のサンプルであり、X次元（左から）における1番目のサンプルである。同様に、サンプルS44は、YおよびX次元の両方でのブロック（104）における4番目のサンプルである。ブロックはサイズが4×4サンプルであるので、S44は、右下にある。さらに、同様な番号付け体系に従う参照サンプルが示されている。参照サンプルは、R、ブロック（104）に対するそのY位置（例えば、行インデックス）およびX位置（列インデックス）でラベル付けされる。H.264とH.265の両方で、予測サンプルは再構成中のブロックに隣接している。したがって、負の値を使用する必要はない。

【0013】

イントラピクチャ予測は、シグナリングされた予測方向によって割り当てられた隣接サンプルからの参照サンプル値をコピーすることで行われることができる。例えば、仮に、コーディングされたビデオビットストリームは、このブロックについて矢印（102）（すなわち、水平から45度の角度で右上への1つまたは複数の予測サンプルから予測される）と一致する予測方向を示すシグナリングを含むとする。この場合、サンプルS41、S32、S23、およびS14は同一の参照サンプルR05から予測される。その後、サンプルS44は、参照サンプルR08から予測される。

【0014】

幾つかの場合において、複数の参照サンプルの値は、参照サンプルを算出するために、特に方向が45度で均等に分割されていない場合に、例えば補間によって組み合わせられることができる。

【0015】

ビデオコーディング技術が発展するにつれて、予測可能な方向の数も増えつつある。H.264（2003年）では、9通りの異なる方向を表すことができる。H.265（2013年）で33通りに増加し、JEM/VVC/BMSは開示時に最大65通りの方向をサポートできる。最も見込みのある方向を識別するための実験が行われ、エンロピーコーディングにおける特定の技術は、少ないビットで見込みの高い方向を表すために使用され、見込みの低い方向に対する特定のペナルティを受け入れる。さらに、方向自体は、隣接する、すでにデコーディングされたブロックで使用される隣接方向から予測される場合がある。

【0016】

図1Bは、予測方向の経時的な増加数を示すために、JEMによる65通りのイントラ予測方向を示す概略図（180）を示す。

【0017】

方向を表すコーディングされたビデオビットストリームにおけるイントラ予測方向ビットのマッピングは、ビデオコーディング技術ごとに異なる可能性があり、また、例えば、イントラ予測モードへの予測方向の単純な直接マッピングから、コードワード、最も可能性の高いモードを含む複雑な適応スキーム、および同様の技術までの範囲とすることができる。しかしながら、すべての場合において、他の特定の方向よりも統計的にビデオコンテンツで発生する可能性が低い特定の方向が存在する可能性がある。ビデオ圧縮の目的は冗長性の削減であるため、適切に機能するビデオコーディング技術では、これらの見込みの低い方向は、見込みの高い方向よりも多くのビット数で表される。

10

20

30

40

50

【0018】

動き補償は、非可逆圧縮技術とすることができ、且つ、以前から再構成されたピクチャ又はその一部（参照ピクチャ）からのサンプルデータのブロックが、動きベクトル（以下、MVとも称する）によって示される方向に空間的にシフトされた後に、新たな再構成されるピクチャ又はその一部の予測のために使用される技術に関連付けられることができる。場合によっては、参照ピクチャは現在再構成中のピクチャと同一であり得る。MVは、X及びYの2次元、又は3次元を有することができ、第3次元は、使用中の参照ピクチャの表示である（後者は、間接的に、時間次元とすることができる）。

【0019】

いくつかのビデオ圧縮技術では、サンプルデータのあるエリアに適用可能なMVは、他のMVから、例えば、再構成中の領域に空間的に隣接し、デコーディングの順序でそのMVに先行するサンプルデータの別の領域に関連するMVから予測され得る。このようにすれば、MVのコーディングに必要なデータ量を大幅に削減し、それによって冗長性を除去し、圧縮量を増大させることができる。MV予測は効率的に機能し得る。なぜならば、例えば、カメラから導出された入力ビデオ信号（ナチュラルビデオと呼ばれる）をコーディングする場合、単一のMVを適用可能な領域よりも大きな領域は、類似の方向に沿って移動するという統計的可能性があり、したがって、場合によっては、隣接領域のMVから導出された類似の動きベクトルを用いて予測され得るからである。その結果、所定の領域について見出されたMVは、周囲のMVから予測されたMVと類似又は同一になり、そしてエントロピーコーディングされた後、MVを直接コーディングする場合に使用されるものよりも、より少ない数のビットで表され得る。場合によっては、MV予測は、元の信号（すなわち、サンプルストリーム）から導出された信号（すなわち、MV）の可逆圧縮の例であり得る。他の場合には、MV予測そのものは、例えば、いくつかの周囲MVから予測器を計算する際の丸め誤差のために、非可逆的であり得る。

【0020】

様々なMV予測メカニズムがH. 265/HEVC (ITU-T Rec. H. 265, High Efficiency Video Coding, December 2016) に、記述されている。H. 265が提供する多くのMV予測メカニズムのうち、ここでは、以後「空間マージ」と称されるテクニックについて説明する。

【0021】

図2を参照すると、現在のブロック(201)は、空間的にシフトされた同サイズの前のブロックから予測可能であることが、動き探索プロセス中にエンコーダによって発見されたサンプルを含む。MVを直接コーディングする代わりに、複数の参照ピクチャに関連付けられたメタデータから、例えば、A0、A1、及びB0、B1、B2（それぞれ102から106）と示される5つの周囲サンプルのいずれかに関連付けられたMVを使用して、（デコーディングする順から）最新の参照ピクチャから、MVを導出することができる。H. 265では、MV予測は、隣接ブロックが使用しているのと同じ参照ピクチャからの予測器を使用することができる。

【発明の概要】

【課題を解決するための手段】

【0022】

本開示の態様は、ビデオ処理のための方法及び装置を提供する。いくつかの例では、ビデオ処理のための装置は、処理回路を含む。処理回路は、色空間におけるサブサンプリングフォーマットのピクチャを色空間における非サブサンプリングフォーマットに変換する。その後、処理回路は、非サブサンプリングフォーマットのピクチャを入力としてニューラルネットワークに基づくフィルタに提供する前に、非サブサンプリングフォーマットのピクチャの色成分の値をクリッピングする。

【0023】

いくつかの例では、処理回路は、非サブサンプリングフォーマットのピクチャの色成分の値を、色成分の有効範囲にクリッピングする。一例では、処理回路は、非サブサンプリン

10

20

30

40

50

グフォーマットのピクチャの色成分の値を、ビット深度に基づいて決定される範囲にクリッピングする。別の一例では、処理回路は、非サブサンプリングフォーマットのピクチャの色成分の値を、所定の範囲にクリッピングする。

【0024】

いくつかの例では、処理回路は、ピクチャを含むビットストリームからのデコーディングされた情報に基づいて、値をクリッピングするための範囲を決定し、その後、非サブサンプリングフォーマットのピクチャの色成分の値を、決定された範囲にクリッピングする。一例では、処理回路は、ビットストリームにおけるシーケンスパラメータセット、ピクチャパラメータセット、スライスヘッダー、及びタイルヘッダーのうち少なくとも1つから、範囲を示す信号をデコーディングする。

10

【0025】

いくつかの例では、処理回路は、ビットストリームからのデコーディングされた情報に基づいて、サブサンプリングフォーマットのピクチャを再構成し、サブサンプリングフォーマットのピクチャにデブロッキングフィルタを適用する。いくつかの例では、処理回路は、ニューラルネットワークに基づくフィルタを、クリッピングされた値を有する非サブサンプリングフォーマットのピクチャに適用することで非サブサンプリングフォーマットのフィルタリングされたピクチャを生成し、非サブサンプリングフォーマットのフィルタリングされたピクチャを、サブサンプリングフォーマットのフィルタリングされたピクチャに変換する。

【0026】

いくつかの例では、クリッピングされた値を有する非サブサンプリングフォーマットのピクチャをストレージに保存させる。そして、クリッピングされた値を有する非サブサンプリングフォーマットで保存されたピクチャをトレーニング入力として提供し、前記ニューラルネットワークに基づくフィルタにおけるニューラルネットワークをトレーニングすることができる。

20

【0027】

本開示の態様は、ビデオデコーディングのためにコンピュータによって実行される場合、ビデオ処理のための方法をコンピュータに実行させる命令を記憶した非一時的なコンピュータ読取可能な媒体をさらに提供する。

【図面の簡単な説明】

30

【0028】

開示された主題のさらなる特徴、本質、および様々な利点は、以下の詳細な説明および添付の図面からより明らかになるであろう。

【0029】

【図1A】イントラ予測モードの例示的なサブセットの概略図である。

【0030】

【図1B】例示的なイントラ予測方向の図である。

【0031】

【図2】一例における現在のブロックおよびその周囲の空間マージ候補の概略図である。

【0032】

【図3】実施形態による通信システム（300）の簡略ブロック図の概略図である。

40

【0033】

【図4】実施形態による通信システム（400）の簡略ブロック図の概略図である。

【0034】

【図5】実施形態によるデコーダの簡略ブロック図の概略図である。

【0035】

【図6】実施形態によるエンコーダの簡略ブロック図の概略図である。

【0036】

【図7】別の実施形態によるエンコーダのブロック図を示す。

【0037】

50

【図 8】別の実施形態によるデコーダのブロック図を示す。

【0038】

【図 9】いくつかの例におけるループフィルタユニットのブロック図を示す。

【0039】

【図 10】いくつかの例における別のループフィルタユニットのブロック図を示す。

【0040】

【図 11】いくつかの例におけるニューラルネットワークに基づくフィルタのブロック図を示す。

【0041】

【図 12】いくつかの例における前処理モジュールのブロック図を示す。

10

【0042】

【図 13】いくつかの例におけるニューラルネットワーク構造のブロック図を示す。

【0043】

【図 14】密な残差ユニットのブロック図を示す。

【0044】

【図 15】いくつかの例における後処理モジュールのブロック図を示す。

【0045】

【図 16】いくつかの例における前処理モジュールのブロック図を示す。

【0046】

【図 17】プロセスを概説するフローチャートを示す。

20

【0047】

【図 18】実施形態によるコンピュータシステムの概略図である。

【発明を実施するための形態】

【0048】

図 3 は、本開示の実施形態による通信システム（300）の簡略ブロック図の概略図を示している。通信システム（300）は、例えばネットワーク（350）を介して互いに通信可能な複数の端末装置を含む。例えば、通信システム（300）は、ネットワーク（350）を介して相互接続された第 1 の対の端末装置（310）および（320）を含む。図 3 の例では、第 1 の対の端末装置（310）および（320）は、データの単方向伝送を実行する。例えば、端末装置（310）は、ネットワーク（350）を介して他方の端末装置（320）へ伝送するためにビデオデータ（例えば、端末装置（310）によってキャプチャされたビデオピクチャのストリーム）をコーディングし得る。コーディングされたビデオデータは、1 つまたは複数のコーディングされたビデオビットストリームの形で送信され得る。端末装置（320）は、ネットワーク（350）からコーディングされたビデオデータを受信し、コーディングされたビデオデータをデコーディングしてビデオピクチャを復元し、復元されたビデオデータに従ってビデオピクチャを表示することができる。単方向のデータ伝送は、メディア配信アプリケーションなどで一般的である。

30

【0049】

別の一例では、通信システム（300）は、例えば、ビデオ会議中に発生し得るコーディングされたビデオデータの双方向伝送を実行する第 2 の対の端末装置（330）および（340）を含む。データの双方向伝送の場合、一例では、端末装置（330）および（340）のそれぞれは、ネットワーク（350）を介して端末装置（330）および（340）のうちの他方の端末装置に伝送するためにビデオデータ（例えば、端末装置によってキャプチャされたビデオピクチャのストリーム）をコーディングし得る。端末装置（330）および（340）のそれぞれは、端末装置（330）および（340）のうちの他方の端末装置で伝送されたコーディングされたビデオデータを受信し、コーディングされたビデオデータをデコーディングしてビデオピクチャを復元し、復元されたビデオデータに従ってビデオピクチャにアクセス可能な表示装置に表示させることができる。

40

【0050】

図 3 の例では、端末装置（310）、（320）、（330）および（340）は、サー

50

バ、パーソナルコンピュータおよびスマートフォンとして示され得るが、しかし、本開示の原理はこれに制限されない。本開示の実施形態は、ラップトップコンピュータ、タブレットコンピュータ、メディアプレーヤー、および／または専用のビデオ会議機器に適用できる。ネットワーク（３５０）は、例えば有線および／または無線通信ネットワークを含む、端末装置（３１０）、（３２０）、（３３０）および（３４０）間でコーディングされたビデオデータを伝送する任意の数のネットワークを表す。通信ネットワーク（３５０）は、回線交換および／またはパケット交換チャネルでデータを交換することができる。代表的なネットワークは、電気通信ネットワーク、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および／またはインターネットを含む。本議論の目的のために、ネットワーク（３５０）のアーキテクチャおよびトポロジーは、以下で説明されない限り、本開示の動作にとって重要でない可能性がある。

10

【００５１】

図４は、開示された主題のアプリケーションの例として、ストリーミング環境におけるビデオエンコーダおよびビデオデコーダの配置を示している。開示された主題は、例えば、ビデオ会議、デジタルＴＶ、および、ＣＤ、ＤＶＤ、メモリスティックなどを含むデジタルメディアへの圧縮ビデオの記憶など、を含む他のビデオ機能アプリケーションに同様に適用可能である。

【００５２】

ストリーミングシステムは、例えば非圧縮のビデオピクチャストリーム（４０２）を作成するデジタルカメラなどのビデオソース（４０１）を含み得るキャプチャサブシステム（４１３）を含んでもよい。一例では、ビデオピクチャストリーム（４０２）は、デジタルカメラによって撮像されたサンプルを含む。コーディングされたビデオデータ（４０４）（又はコーディングされたビデオビットストリーム）と比較した際の高いデータ量を強調するために太線で示されたビデオピクチャストリーム（４０２）は、ビデオソース（４０１）に結合されたビデオエンコーダ（４０３）を含む電子機器（４２０）によって処理され得る。ビデオエンコーダ（４０３）は、以下でより詳細に説明されるように、開示された主題の態様を可能にするか或いは実施するためのハードウェア、ソフトウェア、またはそれらの組み合わせを含むことができる。ビデオピクチャ（４０２）のストリームと比較した際の低いデータ量を強調するために細い線で示された、コーディングされたビデオデータ（４０４）（またはコーディングされたビデオビットストリーム（４０４））は、将来に使用するためにストリーミングサーバ（４０５）に記憶され得る。図４におけるクライアントサブシステム（４０６）および（４０８）のような１つまたは複数のストリーミングクライアントサブシステムは、ストリーミングサーバ（４０５）にアクセスして、コーディングされたビデオデータ（４０４）のコピー（４０７）および（４０９）を検索することができる。クライアントサブシステム（４０６）は、例えば、電子機器（４３０）におけるビデオデコーダ（４１０）を含むことができる。ビデオデコーダ（４１０）は、コーディングされたビデオデータの入力コピー（４０７）をデコーディングし、ディスプレイ（４１２）（例えば、表示画面）または他のレンダリングデバイス（図示せず）に表示できるビデオピクチャ（４１１）の出力ストリームを作成する。一部のストリーミングシステムにおいて、コーディングされたビデオデータ（４０４）、（４０７）、および（４０９）（例えば、ビデオビットストリーム）は、特定のビデオコーディング／圧縮規格に従ってコーディングされ得る。これらの規格の例は、ＩＴＵ－Ｔ勧告Ｈ．２６５を含む。一例では、開発中のビデオコーディング規格は、非公式的に多用途ビデオコーディング（ＶＶＣ）として知られている。開示された主題は、ＶＶＣの文脈において使用され得る。

20

30

40

【００５３】

なお、電子機器（４２０）および（４３０）は、他の構成要素（図示せず）を含み得る。例えば、電子機器（４２０）は、ビデオデコーダ（図示せず）を含むことができ、電子機器（４３０）は、ビデオエンコーダ（図示せず）を含むこともできる。

【００５４】

図５は、本開示の実施形態によるビデオデコーダ（５１０）のブロック図を示す。ビデオ

50

デコーダ（５１０）は、電子機器（５３０）に含まれることができる。電子機器（５３０）は、受信機（５３１）（例えば、受信回路）を含み得る。ビデオデコーダ（５１０）は、図４の例におけるビデオデコーダ（４１０）の代わりに使用され得る。

【００５５】

受信機（５３１）は、ビデオデコーダ（５１０）によってデコーディングされた１つまたは複数のコーディングされたビデオシーケンスを受信でき、同一または別の実施形態では、一度に１つのコーディングされたビデオシーケンスであり、各コーディングされたビデオシーケンスのデコーディングは、他のコーディングされたビデオシーケンスから独立している。コーディングされたビデオシーケンスは、コーディングされたビデオデータを記憶する記憶装置へのハードウェア／ソフトウェアリンクであり得るチャネル（５０１）から受信されることができる。受信機（５３１）は、それぞれの使用エンティティ（図示せず）に転送され得る他のデータ、例えば、コーディングされたオーディオデータおよび／または補助データストリームを有する、コーディングされたビデオデータを受信し得る。受信機（５３１）は、コーディングされたビデオシーケンスを他のデータから分離することができる。ネットワークジッタを防止するために、バッファメモリ（５１５）は、受信機（５３１）とエントロピーデコーダ／パーサ（５２０）（以降、「パーサ（５２０）」と称される）の間に結合されてもよい。特定のアプリケーションでは、バッファメモリ（５１５）は、ビデオデコーダ（５１０）の一部である。他の実施形態では、バッファメモリはビデオデコーダ（５１０）（図示せず）の外部に存在し得る。さらに別の実施形態では、例えば、ネットワークジッタを防止するためにビデオデコーダ（５１０）の外部にバッファメモリ（図示せず）が存在し、且つ、例えば、再生タイミングを取り扱うためにビデオデコーダ（５１０）の内部に別のバッファメモリ（５１５）が存在し得る。受信機（５３１）が十分な帯域幅および制御性を有する記憶／転送装置から、または同期ネットワークからデータを受信する際に、バッファメモリ（５１５）は必要とされず、または小さくされる場合がある。インターネットなどのベストエフォートパケットネットワークで使用するために、バッファメモリ（５１５）が必要になる場合があり、バッファメモリ（５１５）は、比較的大きくされ、有利には適応サイズとすることができ、ビデオデコーダ（５１０）の外部のオペレーティングシステムまたは類似の要素（図示せず）に少なくとも部分的に実現され得る。

【００５６】

ビデオデコーダ（５１０）は、コーディングされたビデオシーケンスからシンボル（５２１）を再構成するパーサ（５２０）を含んでもよい。これらのシンボルのカテゴリは、ビデオデコーダ（５１０）の操作を管理するために使用される情報、および、電子機器（５３０）の構成部分ではないが、図５に示されるように電子機器（５３０）に結合され得るレンダリングデバイス（５１２）（例えば、表示画面）のようなレンダリングデバイスを制御する潜在的情報を含む。レンダリングデバイスのための制御情報は、補足拡張情報（ＳＥＩメッセージ）またはビデオユーザビリティ情報（ＶＵＩ）パラメータセットフラグメント（図示せず）の形式でもよい。パーサ（５２０）は、受信されたコーディングされたビデオシーケンスを解析／エントロピーデコーディングすることができる。コーディングされたビデオシーケンスのコーディングは、ビデオコーディング技術または規格に従うことができ、可変長コーディング、ハフマンコーディング、文脈依存を有するまたは有しない算術コーディングなどを含む様々な原理に従うことができる。パーサ（５２０）は、グループに対応する少なくとも１つのパラメータに基づいて、コーディングされたビデオシーケンスから、ビデオデコーダにおける画素の少なくとも１つのサブグループのためのサブグループパラメータのセットを抽出することができる。サブグループは、グループオブピクチャ（ＧＯＰ）、ピクチャ、タイル、スライス、マクロブロック、コーディングユニット（ＣＵ）、ブロック、変換ユニット（ＴＵ）、予測ユニット（ＰＵ）などを含むことができる。パーサ（５２０）は、コーディングされたビデオシーケンスから変換係数、量子化パラメータ値、動きベクトルなどのような情報を抽出することもできる。

【００５７】

パーサ（５２０）は、シンボル（５２１）を作成するために、バッファメモリ（５１５）から受信されたビデオシーケンスに対してエントロピーデコーディング／解析操作を実行することができる。

【００５８】

シンボル（５２１）の再構成には、コーディングされたビデオピクチャまたはその一部（例えば、インターおよびイントラピクチャ、インターおよびイントラブロック）のタイプ、および他の要因に応じて、複数の異なるユニットが関与できる。どのユニットがどのように関与するかは、パーサ（５２０）によってコーディングされたビデオシーケンスから解析されたサブグループ制御情報によって制御され得る。明確化するために、パーサ（５２０）と以下の複数のユニットとの間のそのようなサブグループ制御情報の流れは示されていない。

10

【００５９】

すでに述べた機能ブロックに加えて、ビデオデコーダ（５１０）は、以下で説明されるようにいくつかの機能ユニットに概念的に細分され得る。商業的な制約の下で実際の実現にあたっては、これらのユニットの多くは互いに密接に相互作用し、少なくとも一部に互いに統合することができる。しかしながら、開示された主題の説明の目的で、以下の機能ユニットへの概念的な細分化が適切である。

【００６０】

第１のユニットは、スケーラ／逆変換ユニット（５５１）である。スケーラ／逆変換ユニット（５５１）は、使用する変換、ブロックサイズ、量子化因子、量子化スケーリング行列などを含む制御情報と、量子化された変換係数をシンボル（５２１）としてパーサ（５２０）から受信する。スケーラ／逆変換ユニット（５５１）は、アグリゲータ（５５５）に入力可能なサンプル値を含むブロックを出力することができる。

20

【００６１】

場合によっては、スケーラ／逆変換（５５１）の出力サンプルは、イントラコーディングブロック、すなわち、以前に再構成されたピクチャからの予測情報を使用せず、現在のピクチャの以前に再構成された部分からの予測情報を使用できるブロックに関係し得る。このような予測情報は、イントラピクチャ予測ユニット（５５２）によって提供され得る。場合によっては、イントラピクチャ予測ユニット（５５２）は、現在のピクチャバッファ（５５８）から取り出された周囲の既に再構成された情報を用いて、再構成中のブロックの同サイズおよび形状のブロックを生成する。現在のピクチャバッファ（５５８）は、例えば、部分的に再構成された現在のピクチャおよび／または完全に再構成された現在のピクチャをバッファリングする。アグリゲータ（５５５）は、場合によっては、サンプルごとに、イントラ予測ユニット（５５２）が生成した予測情報を、スケーラ／逆変換ユニット（５５１）によって提供された出力サンプル情報に追加する。

30

【００６２】

他の場合では、スケーラ／逆変換ユニット（５５１）の出力サンプルは、インターコーディングされた、潜在的に動き補償されたブロックに関係することがある。このような場合に、動き補償予測ユニット（５５３）は、参照ピクチャメモリ（５５７）にアクセスして、予測すべきサンプルを取り出すことができる。取り出されたサンプルをブロックに関係するシンボル（５２１）に従って動き補償した後、出力サンプル情報を生成するように、これらのサンプルは、アグリゲータ（５５５）によってスケーラ／逆変換ユニット（５５１）の出力に追加され得る（この場合、残差サンプルまたは残差信号と呼ばれる）。動き補償予測ユニット（５５３）が予測サンプルを取り出す参照ピクチャメモリ（５５７）内のアドレスは、例えば、X、Y、および参照ピクチャ成分を有し得るシンボル（５２１）の形態で動き補償予測ユニット（５５３）に利用可能な動きベクトルによって制御され得る。動き補償は、サブサンプルの正確な動きベクトルが使用中であるときに参照ピクチャメモリ（５５７）から取り出されたサンプル値の補間、動きベクトル予測メカニズムなどを含むこともできる。

40

【００６３】

50

アグリゲータ（５５５）の出力サンプルは、ループフィルタユニット（５５６）において様々なループフィルタリング技術を受けることができる。ビデオ圧縮技術は、コーディングされたビデオシーケンス（コーディングされたビデオビットストリームとも呼ばれる）に含まれる、パーサ（５２０）からのシンボル（５２１）としてループフィルタユニット（５５６）に利用可能とされたパラメータによって制御され、それに、コーディングされたピクチャまたはコーディングされたビデオシーケンスの前の部分（デコーディング順で）のデコーディング進行中に取得されたメタ情報に応じるとともに、以前に再構成されループフィルタリングされたサンプル値に応答することもできるループフィルタ技術を含み得る。

【００６４】

ループフィルタユニット（５５６）の出力は、レンダリングデバイス（５１２）へ出力されるとともに、将来のインターピクチャ予測で使用するために参照ピクチャメモリ（５５７）に記憶されることができるサンプルストリームであり得る。

【００６５】

特定のコーディングされたピクチャは、完全に再構成されると、将来の予測のために参照ピクチャとして使用されることができる。例えば、現在のピクチャに対応するコーディングされたピクチャが完全に再構成され、コーディングされたピクチャが（例えば、パーサ（５２０）によって）参照ピクチャとして識別されると、現在のピクチャバッファ（５５８）は、参照ピクチャメモリ（５５７）の一部になることができ、且つ次のコーディングされたピクチャの再構成を開始する前に新しい現在のピクチャバッファは、再び割り当てられることができる。

【００６６】

ビデオデコーダ（５１０）は、ITU-T Rec. H. 265のような規格での所定のビデオ圧縮技術に従ってデコーディング操作を実行することができる。コーディングされたビデオシーケンスが、ビデオ圧縮技術または規格の構文と、ビデオ圧縮技術または規格で文書化されたプロファイルとの両方に準拠しているという意味で、コーディングされたビデオシーケンスは、使用されているビデオ圧縮技術または規格によって指定されるシンタックスに準拠し得る。具体的には、プロファイルは、ビデオ圧縮技術または規格で使用可能なすべてのツールから、特定のツールをそのプロファイルで使用できる一意のツールとして選択することができる。コーディングされたビデオシーケンスの複雑さがビデオ圧縮技術または規格のレベルで定義される範囲内にあることも、コンプライアンスに必要である。場合によっては、最大ピクチャサイズ、最大フレームレート、最大再構成サンプルレート（例えば、１秒あたりのメガサンプルで測定される）、最大参照ピクチャサイズなどがレベルによって制限される。レベルによって設定された制限は、場合によっては、仮想参照デコーダ（HRD）仕様およびコーディングされたビデオシーケンスでシグナリングされたHRDバッファ管理のためのメタデータによってさらに制限され得る。

【００６７】

実施形態では、受信機（５３１）は、コーディングされたビデオとともに追加の（冗長な）データを受信することができる。追加のデータは、コーディングされたビデオシーケンスの一部として含まれてもよい。追加のデータは、データを適切にデコーディングし、および／または、元のビデオデータをより正確に再構成するためにビデオデコーダ（５１０）によって使用され得る。追加のデータは、例えば、時間的、空間的、または信号対雑音比（SNR）エンハンスメントレイヤ、冗長スライス、冗長ピクチャ、前方向誤り訂正コードなどの形式であり得る。

【００６８】

図６は、本開示の実施形態によるビデオエンコーダ（６０３）のブロック図を示す。ビデオエンコーダ（６０３）は、電子機器（６２０）に含まれる。電子機器（６２０）は、送信機（６４０）（例えば、送信回路）を含む。ビデオエンコーダ（６０３）は、図４の例におけるビデオエンコーダ（４０３）の代わりに使用され得る。

【００６９】

10

20

30

40

50

ビデオエンコーダ（603）は、ビデオエンコーダ（603）によってコーディングされるビデオピクチャをキャプチャし得るビデオソース（601）（図6の例における電子機器（620）の一部ではない）からビデオサンプルを受信することができる。別の一例では、ビデオソース（601）は、電子機器（620）の一部である。

【0070】

ビデオソース（601）は、ビデオエンコーダ（603）によってコーディングされるべきソースビデオシーケンスを、任意の適切なビット深度（例えば、8ビット、10ビット、12ビット、・・・）、任意の色空間（例えば、BT.601 YCrCb、RGB、・・・）および任意の適切なサンプリング構造（例えば、YCrCb 4:2:0、YCrCb 4:4:4）であり得るデジタルビデオサンプルストリームの形式で提供し得る。メディア配信システムでは、ビデオソース（601）は、予め準備されたビデオを記憶する記憶装置であり得る。ビデオ会議システムでは、ビデオソース（601）は、ローカルピクチャ情報をビデオシーケンスとしてキャプチャするカメラであり得る。ビデオデータは、連続して見た際に動きが与えられる複数の個別のピクチャとして提供されてもよい。ピクチャ自体は、画素の空間アレイとして編成されてもよく、各画素は、使用中のサンプリング構造、色空間などに応じて1つまたは複数のサンプルを含むことができる。当業者は、画素とサンプルとの関係を容易に理解することができる。以下の説明ではサンプルを中心に説明する。

【0071】

実施形態によれば、ビデオエンコーダ（603）は、リアルタイムでまたはアプリケーションが要求する任意の他の時間制約の下でソースビデオシーケンスのピクチャをコーディングし、コーディングされたビデオシーケンス（643）に圧縮することができる。適切なコーディング速度を強制することは、コントローラ（650）の機能の1つである。いくつかの実施形態では、コントローラ（650）は、以下で説明される他の機能ユニットを制御し、他の機能ユニットに機能的に結合される。明確にするために、カップリングは描かれていない。コントローラ（650）によって設定されるパラメータは、レート制御関連パラメータ（ピクチャスキップ、量子化、レート歪み最適化技術のラムダ値、・・・）、ピクチャサイズ、グループオブピクチャ（GOP）レイアウト、最大動きベクトル検索範囲などを含むことができる。コントローラ（650）は、特定のシステム設計に対して最適化されたビデオエンコーダ（603）に関する他の適切な機能を有するように構成され得る。

【0072】

いくつかの実施形態では、ビデオエンコーダ（603）は、コーディンググループで動作するように構成される。過度に簡略化した説明として、一例では、コーディンググループは、ソースコーダー（630）（例えば、コーディングしようとする入力ピクチャおよび参照ピクチャに基づくシンボルストリームなどのシンボルの作成を担当する）、およびビデオエンコーダ（603）に埋め込まれた（ローカル）デコーダ（633）を含み得る。デコーダ（633）は、（リモート）デコーダが作成するのと同様な方法でサンプルデータを作成するためにシンボルを再構成する（シンボルとコーディングされたビデオビットストリーム間の如何なる圧縮は、開示された主題で考慮されるビデオ圧縮技術では可逆であるため）。再構成されたサンプルストリーム（サンプルデータ）は参照ピクチャメモリ（634）に入力される。シンボルストリームのデコーディングは、デコーダの位置（ローカルまたはリモート）に関係なくビット正確な結果につながるため、参照ピクチャメモリ（634）のコンテンツもローカルエンコーダとリモートエンコーダの間でビット正確である。言い換えれば、エンコーダの予測部分は、参照ピクチャサンプルとして、デコーディング中に予測を使用する際にデコーダが「見る」と全く同じサンプル値を「見る」。参照ピクチャの同期性（および例えばチャネルエラーに起因して同期性を維持できない場合に生じるドリフト）のかような基本原理は、いくつかの関連分野にも使用されている。

【0073】

「ローカル」デコーダ（633）の動作は、前文で図5を参照して既に詳細に説明された

10

20

30

40

50

、ビデオデコーダ（５１０）などの「リモート」デコーダの動作と同様であり得る。しかしながら、図５を簡単に参照すると、シンボルが使用可能であり、エントロピーコーダー（６４５）およびパーサ（５２０）によるコーディングビデオシーケンスへのシンボルのエンコーディング／デコーディングは可逆であり得るので、バッファメモリ（５１５）、およびパーサ（５２０）を含むビデオデコーダ（５１０）のエントロピーデコーディング部分は、ローカルデコーダ（６３３）では完全に実現されない場合がある。

【００７４】

この点から分かるように、デコーダに存在する解析／エントロピーデコーディング以外の如何なるデコーダ技術も、対応するエンコーダに実質的に同一の機能的形態で必ず存在する必要がある。このため、開示された主題は、デコーダ操作を中心とする。エンコーダ技術の説明は、包括的に説明されたデコーダ技術と逆なものであるため、省略できる。特定の領域でのみ、より詳細な説明が必要となり、以下に提供される。

【００７５】

動作中、いくつかの例では、ソースコーダー（６３０）は、「参照ピクチャ」として指定されたビデオシーケンスからの１つまたは複数の以前にコーディングされたピクチャを参照して入力ピクチャを予測的にコーディングする動き補償予測コーディングを実行してもよい。このようにして、コーディングエンジン（６３２）は、入力ピクチャの画素ブロックと、入力ピクチャへの予測基準として選択され得る参照ピクチャの画素ブロックとの差異をコーディングする。

【００７６】

ローカルビデオデコーダ（６３３）は、ソースコーダー（６３０）で作成されたシンボルに基づいて、参照ピクチャとして指定され得るピクチャのコーディングされたビデオデータをデコーディングすることができる。コーディングエンジン（６３２）の操作は、有利にはロッシェプロセスであり得る。コーディングされたビデオデータをビデオデコーダ（図６に示せず）でデコーディングできる際、再構成されたビデオシーケンスは、通常、いくつかのエラーを伴うソースビデオシーケンスのレプリカであってもよい。ローカルビデオデコーダ（６３３）は、ビデオデコーダによって参照ピクチャを実行し得るデコーディングプロセスを複製し、再構成された参照ピクチャを参照ピクチャキャッシュ（６３４）に記憶させることができる。このようにして、ビデオエンコーダ（６０３）は、遠端ビデオデコーダによって取得される再構成された参照ピクチャと共通するコンテンツ（送信エラー無し）を有する再構成された参照ピクチャのコピーをローカルに記憶し得る。

【００７７】

予測器（６３５）は、コーディングエンジン（６３２）に対する予測検索を実行できる。つまり、コーディング対象となる新しいピクチャについて、予測器（６３５）は、（候補の参照画素ブロックとしての）サンプルデータ、または、参照ピクチャの動きベクトル、ブロック形状など、新しいピクチャの適切な予測基準として機能し得る特定のメタデータを参照ピクチャメモリ（６３４）から検索することができる。予測器（６３５）は、適切な予測基準を見つけるために、サンプルブロック／画素ブロックごとに操作することができる。場合によっては、予測器（６３５）で取得された検索結果によって決定されるように、入力ピクチャは、参照ピクチャメモリ（６３４）に記憶された複数の参照ピクチャから引き出された予測基準を有してもよい。

【００７８】

コントローラ（６５０）は、例えば、ビデオデータをコーディングするために使用されるパラメータおよびサブグループパラメータの設定を含む、ソースコーダー（６３０）のコーディング操作を管理することができる。

【００７９】

前述のすべての機能ユニットの出力は、エントロピーコーダー（６４５）においてエントロピーコーディングを受け得る。エントロピーコーダー（６４５）は、例えば、ハフマンコーディング、可変長コーディング、算術コーディングなどの技術に従ってシンボルを可逆圧縮することにより、様々な機能ユニットによって生成されたシンボルをコーディング

10

20

30

40

50

ビデオシーケンスに変換する。

【0080】

送信機（640）は、コーディングされたビデオデータを記憶する記憶装置へのハードウェア／ソフトウェアリンクであり得る通信チャネル（660）を介した送信の準備のために、エントローピーコーダー（645）によって作成されたコーディングされたビデオシーケンスをバッファリングすることができる。送信機（640）は、ビデオコーダ（603）からのコーディングされたビデオデータを、送信されるべき他のデータ、例えば、コーディングされたオーディオデータおよび／または補助データストリーム（ソースは示されていない）とマージすることができる。

【0081】

コントローラ（650）は、ビデオエンコーダ（603）の操作を管理し得る。コーディングの際に、コントローラ（650）は、各コーディングされたピクチャに特定のコーディングピクチャタイプを割り当てることができ、これは、それぞれのピクチャに適用され得るコーディング技術に影響を及ぼし得る。例えば、ピクチャは、通常、次のピクチャタイプのいずれかに割り当てられ得る。

【0082】

イントラピクチャ（Iピクチャ）は、予測のソースとしてシーケンス内の他のいかなるピクチャを使用せずにコーディングおよびデコーディングされ得るものであり得る。一部のビデオコーデックは、例えば、インディペンデントデコーダリフレッシュ（Independent Decoder Refresh、「IDR」）ピクチャを含む、異なるタイプのイントラピクチャを許容する。当業者は、Iピクチャのそれらの変形およびそれらのそれぞれの用途および特徴を知っている。

【0083】

予測ピクチャ（Pピクチャ）は、各ブロックのサンプル値を予測するために最大1つの動きベクトルおよび参照インデックスを使用したイントラ予測またはインター予測によりコーディングおよびデコーディングされ得るものであり得る。

【0084】

双方向予測ピクチャ（Bピクチャ）は、各ブロックのサンプル値を予測するために最大2つの動きベクトルおよび参照インデックスを使用したイントラ予測またはインター予測によりコーディングおよびデコーディングされ得るものであり得る。同様に、複数の予測ピクチャは、単一のブロックを再構成するために2つを超える参照ピクチャおよび関連メタデータを使用することができる。

【0085】

ソースピクチャは、一般に、複数のサンプルブロック（例えば、それぞれ、 4×4 、 8×8 、 4×8 、または 16×16 サンプルのブロック）に空間的に細分され、ブロックごとにコーディングされ得る。ブロックは、ブロックのそれぞれのピクチャに適用されるコーディング割り当てによって決定された他の（既にコーディングされた）ブロックを参照して予測的にコーディングされ得る。例えば、Iピクチャのブロックは、非予測的にコーディングされてもよく、或いは、同一のピクチャの既にコーディングされたブロック（空間予測またはイントラ予測）を参照して予測的にコーディングされてもよい。Pピクチャの画素ブロックは、1つの予めコーディングされた参照ピクチャを参照して、空間予測を介してまたは時間予測を介して予測的にコーディングされ得る。Bピクチャのブロックは、1つまたは2つの以前にコーディングされた参照ピクチャを参照して、空間予測を介してまたは時間予測を介して予測的にコーディングされ得る。

【0086】

ビデオエンコーダ（603）は、ITU-T Rec. H. 265などの予め設定されたビデオコーディング技術または規格に従って、コーディング操作を実行することができる。動作中、ビデオエンコーダ（603）は、入力ビデオシーケンスの時間的および空間的冗長性を利用する予測コーディング操作を含む、様々な圧縮操作を実行することができる。したがって、コーディングされたビデオデータは、使用されるビデオコーディング技術ま

10

20

30

40

50

たは規格によって指定されたシンタックスに準拠する場合がある。

【0087】

実施形態では、送信機（640）は、コーディングされたビデオとともに追加のデータを送信することができる。ソースコーダー（630）は、このようなデータをコーディングされたビデオシーケンスの一部として含み得る。追加のデータは、時間的／空間的／SNRエンハンスメントレイヤ、冗長なピクチャやスライスなどの他の形式での冗長データ、SEIメッセージ、VUIパラメータセットフラグメントなどを含み得る。

【0088】

ビデオは、時系列の複数のソースピクチャ（ビデオピクチャ）としてキャプチャされ得る。イントラピクチャ予測（「イントラ予測」と略されることが多い）は、所定のピクチャにおける空間的相関性を利用し、インターピクチャ予測は、ピクチャ間の（時間的または他の）相関性を利用する。一例では、現在のピクチャと呼ばれるエンコーディング／デコーディング中の特定のピクチャは、ブロックに分割される。現在のピクチャにおけるブロックが、ビデオにおける以前にコーディングされ、まだバッファリングされている参照ピクチャの参照ブロックに類似している場合、現在のピクチャにおけるブロックは、動きベクトルと呼ばれるベクトルによってコーディングされ得る。動きベクトルは、参照ピクチャの参照ブロックを指し、複数の参照ピクチャが使用されている場合、参照ピクチャを識別する第3次元を有することができる。

【0089】

いくつかの実施形態では、インターピクチャ予測において双方向予測法を使用することができる。双方向予測法によれば、ビデオにおける現在のピクチャよりもデコーディング順序がそれぞれ前である（ただし、表示順序でそれぞれ過去および未来にあり得る）第1の参照ピクチャおよび第2の参照ピクチャのような2つの参照ピクチャを使用する。現在のピクチャにおけるブロックは、第1の参照ピクチャ内の第1の参照ブロックを指す第1の動きベクトル、および第2の参照ピクチャ内の第2の参照ブロックを指す第2の動きベクトルによってコーディングされ得る。ブロックは、第1の参照ブロックと第2の参照ブロックとの組み合わせによって予測され得る。

【0090】

さらに、マージモード技術をインターピクチャ予測に適用して、コーディング効率を向上させることができる。

【0091】

本開示のいくつかの実施形態によれば、インターピクチャ予測およびイントラピクチャ予測などの予測は、ブロック単位で実行される。例えば、HEVC規格によれば、ビデオピクチャのシーケンス内のピクチャは、圧縮のためにコーディングツリーユニット（CTU）に分割され、ピクチャ内のCTUは、 64×64 画素、 32×32 画素、または 16×16 画素など、同一のサイズを有する。一般に、CTUは、1つのルマCTBと2つのクロマCTBである3つのコーディングツリーブロック（CTB）を含む。各CTUは、1つまたは複数のコーディングユニット（CU）に再帰的に四分木分割され得る。例えば、 64×64 画素のCTUは、 64×64 画素の1つのCU、 32×32 画素の4つのCU、または 16×16 画素の16個のCUに分割され得る。一例では、各CUを解析して、インター予測タイプまたはイントラ予測タイプなど、CUの予測タイプを決定する。CUは、時間的および／または空間的予測可能性に応じて、1つまたは複数の予測ユニット（PU）に分割される。通常、各PUは、ルマ予測ブロック（PB）と2つのクロマPBを含む。実施形態では、コーディング（エンコーディング／デコーディング）における予測操作は、予測ブロックの単位で実行される。ルマ予測ブロックを予測ブロックの例として用いて、予測ブロックは、 8×8 画素、 16×16 画素、 8×16 画素、 16×8 画素など画素の値（例えば、ルマ値）の行列を含む。

【0092】

図7は、本開示の他の実施形態によるビデオエンコーダ（703）の図を示す。ビデオエンコーダ（703）は、ビデオピクチャシーケンスにおける現在のビデオピクチャ内のサ

10

20

30

40

50

ンプル値の処理ブロック（例えば、予測ブロック）を受信し、処理ブロックを、コーディングされたビデオシーケンスの一部であるコーディングされたピクチャにコーディングするように構成される。一例では、図4の例におけるビデオエンコーダ（403）の代わりにビデオエンコーダ（703）を使用する。

【0093】

HEVCの例では、ビデオエンコーダ（703）は、8×8サンプルのような予測ブロックなどの処理ブロックのサンプル値の行列を受信する。ビデオエンコーダ（703）は、例えばレート歪み最適化を用いて、処理ブロックをイントラモード、インターモード、または双方向予測モードにより最良にコーディングするか否かを決定する。処理ブロックがイントラモードでコーディングされようとする場合、ビデオエンコーダ（703）は、イントラ予測法を用いて処理ブロックをコーディングされたピクチャにコーディングすることができる。また、処理ブロックがインターモードまたは双予測モードでコーディングされようとする場合、ビデオエンコーダ（703）は、それぞれインター予測または双方向予測法を用いて、処理ブロックをコーディングされたピクチャにコーディングすることができる。特定のビデオコーディング技術では、マージモードは、予測器外のコーディングされた動きベクトル成分を介することなく、1つまたは複数の動きベクトル予測器から動きベクトルを導出するインターピクチャ予測サブモードであり得る。特定の他のビデオコーディング技術では、対象ブロックに適用可能な動きベクトル成分が存在し得る。一例では、ビデオエンコーダ（703）は、処理ブロックのモードを決定するためのモード決定モジュール（図示せず）などの他の構成要素を含む。

【0094】

図7の例では、ビデオエンコーダ（703）は、図7に示すように互いに結合されたインターエンコーダ（730）、イントラエンコーダ（722）、残差算出部（723）、スイッチ（726）、残差エンコーダ（724）、汎用コントローラ（721）およびエントロピーエンコーダ（725）を含む。

【0095】

インターエンコーダ（730）は、現在のブロック（例えば、処理ブロック）のサンプルを受信し、ブロックを参照ピクチャ内の1つまたは複数の参照ブロック（例えば、前の先行ピクチャおよび後のピクチャ内のブロック）と比較し、インター予測情報（例えば、インターコーディング法による冗長情報の記述、動きベクトル、マージモード情報）を生成し、インター予測情報に基づいて任意の適切な技術を用いてインター予測結果（例えば、予測されたブロック）を算出するように構成される。いくつかの例では、参照ピクチャは、コーディングされたビデオ情報に基づいてデコーディングされた参照ピクチャである。

【0096】

イントラエンコーダ（722）は、現在のブロック（例えば、処理ブロック）のサンプルを受信し、場合によっては該ブロックを同一のピクチャで既にコーディングされたブロックと比較し、変換後に、量子化された係数を生成し、場合によってはイントラ予測情報（例えば、1つまたは複数のイントラコーディング技術によるイントラ予測方向情報）をも生成するように構成される。一例では、イントラエンコーダ（722）は、同一のピクチャ内の参照ブロックおよびイントラ予測情報に基づいてイントラ予測結果（例えば、予測されたブロック）も算出する。

【0097】

汎用コントローラ（721）は、汎用制御データを決定し、汎用括制御データに基づいてビデオエンコーダ（703）の他の構成要素を制御するように構成される。一例では、汎用コントローラ（721）は、ブロックのモードを決定し、このモードに基づいて制御信号をスイッチ（726）に提供する。例えば、モードがイントラモードである場合、汎用コントローラ（721）は、残差算出部（723）が使用するためのイントラモード結果を選択するようにスイッチ（726）を制御するとともに、イントラ予測情報を選択してイントラ予測情報をビットストリームに含ませるようにエントロピーエンコーダ（725）を制御する。また、モードがインターモードである場合、汎用コントローラ（721）

は、残差算出部（７２３）が使用するためのインター予測結果を選択するようにスイッチ（７２６）を制御するとともに、インター予測情報を選択してインター予測情報をビットストリームに含ませるようにエントロピーエンコーダ（７２５）を制御する。

【００９８】

残差算出部（７２３）は、受信されたブロックとイントラエンコーダ（７２２）またはインターエンコーダ（７３０）から選択された予測結果との差（残差データ）を算出するように構成される。残差エンコーダ（７２４）は、残差データに基づいて動作し、残差データをコーディングして変換係数を生成するように構成される。一例では、残差エンコーダ（７２４）は、残差データを空間領域から周波数領域へ変換し、変換係数を生成するように構成される。その後、変換係数に対して量子化処理を行い、量子化された変換係数を得る。様々な実施形態では、ビデオエンコーダ（７０３）は、残差デコーダ（７２８）をさらに含む。残差デコーダ（７２８）は、逆変換を実行し、デコーディングされた残差データを生成するように構成される。デコーディングされた残差データは、イントラエンコーダ（７２２）およびインターエンコーダ（７３０）によって適切に使用され得る。例えば、インターエンコーダ（７３０）は、デコーディング残差データおよびインター予測情報に基づいて、デコーディングブロックを生成することができ、且つイントラエンコーダ（７２２）は、デコーディングされた残差データおよびイントラ予測情報に基づいて、デコーディングブロックを生成することができる。いくつかの例では、デコーディングブロックは、デコーディングピクチャを生成するように適切に処理され、デコーディングピクチャは、メモリ回路（図示せず）にバッファリングされ、参照ピクチャとして使用され得る。

【００９９】

エントロピーエンコーダ（７２５）は、コーディングされたブロックを含むようにビットストリームをフォーマットするように構成される。エントロピーエンコーダ（７２５）は、H E V C規格などの適切な規格に従って様々な情報を含むように構成される。一例では、エントロピーエンコーダ（７２５）は、汎用制御データ、選択された予測情報（例えば、イントラ予測情報またはインター予測情報）、残差情報、およびビットストリームにおける他の適切な情報を含むように構成される。開示された主題によれば、インターモードまたは双方向予測モードのマージサブモードでブロックをコーディングする場合、残差情報はないことに留意されたい。

【０１００】

図８は、本開示の他の実施形態によるビデオデコーダ（８１０）の図を示す。ビデオデコーダ（８１０）は、コーディングされたビデオシーケンスの一部であるコーディングされたピクチャを受信し、コーディングされたピクチャをデコーディングして、再構成ピクチャを生成するように構成される。一例では、図４の例におけるビデオデコーダ（４１０）の代わりにビデオデコーダ（８１０）を使用する。

【０１０１】

図８の例では、ビデオデコーダ（８１０）は、図８に示されるように互いに結合されたエントロピーデコーダ（８７１）、インターデコーダ（８８０）、残差デコーダ（８７３）、再構成モジュール（８７４）、およびイントラデコーダ（８７２）を含む。

【０１０２】

エントロピーデコーダ（８７１）は、コーディングされたピクチャから、コーディングされたピクチャを構成する構文要素を表す特定のシンボルを再構成するように構成されることができる。このようなシンボルは、例えば、ブロックがコーディングされるモード（例えば、イントラモード、インターモード、双方向予測モード、後の両者のマージサブモードまたは他のサブモード）、それぞれイントラデコーダ（８７２）またはインターデコーダ（８８０）による予測に使用される特定のサンプルまたはメタデータを識別できる予測情報（例えば、イントラ予測情報またはインター予測情報）、例えば、量子化された変換係数の形式の残差情報などを含むことができる。一例では、予測モードがインターまたは双方向予測モードであれば、インター予測情報は、インターデコーダ（８８０）に提供される。また、予測タイプがイントラ予測タイプであれば、イントラ予測情報は、イントラ

10

20

30

40

50

デコーダ（８７２）に提供される。残差情報は、逆量子化され、残差デコーダ（８７３）に提供されることができる。

【０１０３】

インターデコーダ（８８０）は、インター予測情報を受信し、インター予測情報に基づいてインター予測結果を生成するように構成される。

【０１０４】

イントラデコーダ（８７２）は、イントラ予測情報を受信し、イントラ予測情報に基づいて予測結果を生成するように構成される。

【０１０５】

残差デコーダ（８７３）は、逆量子化を実行して、逆量子化された変換係数を抽出し、逆量子化された変換係数を処理して残差を周波数領域から空間領域に変換するように構成される。残差デコーダ（８７３）は、特定の制御情報（量子化器パラメータ（ＱＰ）を含めるように）を必要とする場合があり、この情報は、エントロピーデコーダ（８７１）によって提供されてもよい（データパスは、少量の制御情報のみであり得るため、示されていない）。

【０１０６】

再構成モジュール（８７４）は、空間領域において、残差デコーダ（８７３）によって出力された残差と、（場合によってはインターまたはイントラ予測モジュールによって出力される）予測結果とを組み合わせ、再構成ビデオの一部となり得る再構成ピクチャの一部であり得る再構成ブロックを形成するように構成される。なお、視覚的品質を改善するために、デブロッキング操作などの他の適切な操作を実行することができる。

【０１０７】

なお、ビデオエンコーダ（４０３）、（６０３）および（７０３）とビデオデコーダ（４１０）、（５１０）および（８１０）は、任意の適切な技術を用いて実施されることができる。実施形態では、ビデオエンコーダ（４０３）、（６０３）および（７０３）とビデオデコーダ（４１０）、（５１０）および（８１０）は、１つまたは複数の集積回路を用いて実現され得る。別の実施形態では、ビデオエンコーダ（４０３）、（６０３）および（７０３）とビデオデコーダ（４１０）、（５１０）および（８１０）は、ソフトウェア命令を実行する１つまたは複数のプロセッサを用いて実現され得る。

【０１０８】

ニューラルネットワーク技術はビデオコーディング技術と併用でき、ニューラルネットワークを利用したビデオコーディング技術はハイブリッドビデオコーディング技術と呼ぶことができる。たとえば、ループフィルタユニット（５５６）などのループフィルタユニットは、サンプルフィルタリングのために、さまざまなループフィルタを適用できる。１つまたは複数のループフィルタはニューラルネットワークによって実現され得る。本開示の各態様は、ニューラルネットワークを使用して画質を改善するためのハイブリッドビデオコーディング技術におけるループ内フィルタリング技術を提供する。具体的には、本開示の一態様によれば、データをニューラルネットワークに基づくループフィルタのカーネルに供給する前に、データをクリッピングする技術を使用することができる。

【０１０９】

本開示の一態様によれば、ループフィルタは、参照データに影響を与えるフィルタである。例えば、ループフィルタユニット（５５６）によってフィルタリングされた画像は、参照ピクチャメモリ（５５７）などのバッファに、さらなる予測のために保存されている。ループ内フィルタは、ビデオコーデックのビデオ品質を向上させることができる。

【０１１０】

図９は、いくつかの例におけるループフィルタユニット（９００）のブロック図を示している。一例では、ループフィルタユニット（９００）は、ループフィルタユニット（５５６）の代わりに使用され得る。図９の例では、ループフィルタユニット（９００）は、デブロッキングフィルタ（９０１）、サンプル適応オフセット（ＳＡＯ）フィルタ（９０２）、および適応ループフィルタ（ＡＬＦ）フィルタ（９０３）を含む。いくつかの例では

10

20

30

40

50

、A L Fフィルタ（9 0 3）は、クロス成分適応ループフィルタ（C C A L F）を含み得る。

【0 1 1 1】

一例では、動作中、ループフィルタユニット（9 0 0）は、再構成されたピクチャを受信し、再構成されたピクチャに様々なフィルタを適用し、再構成されたピクチャに応答して出力ピクチャを生成する。

【0 1 1 2】

いくつかの例では、デブロッキングフィルタ（9 0 1）及びS A Oフィルタ（9 0 2）は、ブロックコーディング技術を使用する際に導入されたブロッキングアーチファクトを除去するように構成される。デブロッキングフィルタ（9 0 1）は、ブロックコーディング技術を使用する際に形成された形状エッジを平滑化することができる。S A Oフィルタ（9 0 2）は、ビデオフレーム内の他のサンプルに対する歪みを減らすために、サンプルに特定のオフセットを適用できる。A L F（9 0 3）は、たとえば、サンプルのブロックに分類を適用し、そして分類に関連付けられたフィルタをサンプルのブロックに適用することができる。幾つかの例では、フィルタのフィルタ係数は、エンコーダによって決定され、デコーダにシグナリングされる。

【0 1 1 3】

いくつかの例（例えば、J V E T - T 0 0 5 7）では、デブロッキングフィルタ（9 0 1）とS A Oフィルタ（9 0 2）の間に密な残差畳み込みニューラルネットワークに基づくループ内フィルタ（Dense residual convolutional neural network based in-loop filter, D R N L F）と呼ばれる付加のフィルタを挿入することができる。D R N L Fは、画質をさらに向上させることが期待できる。

【0 1 1 4】

図10は、いくつかの例におけるループフィルタユニット（1 0 0 0）のブロック図を示す。一例では、ループフィルタユニット（1 0 0 0）は、ループフィルタユニット（5 5 6）の代わりに使用され得る。図10の例では、ループフィルタユニット（1 0 0 0）は、デブロッキングフィルタ（1 0 0 1）、S A Oフィルタ（1 0 0 2）、A L Fフィルタ（1 0 0 3）、およびデブロッキングフィルタ（1 0 0 1）とS A Oとの間に配置されたD R N L Fフィルタ（1 0 1 0）を含む。

【0 1 1 5】

デブロッキングフィルタ（1 0 0 1）は、デブロッキングフィルタ（9 0 1）と同様に構成され、S A Oフィルタ（1 0 0 2）は、S A Oフィルタ（9 0 2）と同様に構成され、A L Fフィルタ（1 0 0 3）は、A L Fフィルタ（9 0 3）と同様に構成される。

【0 1 1 6】

D R N L Fフィルタ（1 0 1 0）は、デブロッキングされたピクチャ（1 0 1 1）によって示されるデブロッキングフィルタ（1 0 0 1）の出力を受信するとともに、再構成されたピクチャの量子化パラメータ（Q P）マッピングも受信する。Q Pマップには、再構成されたピクチャにおけるブロックの量子化パラメータが含まれている。D R N L Fフィルタ（1 0 1 0）は、画質が向上されたフィルタリングされたピクチャ（1 0 1 9）によって示されるピクチャを出力することができ、フィルタリングされたピクチャ（1 0 1 9）は、さらなるフィルタ処理のためにS A Oフィルタ（1 0 0 2）に供給される。

【0 1 1 7】

本開示の一態様によれば、ビデオ処理のためのニューラルネットワークは、色空間における色成分を処理するための複数のチャネルを含み得る。一例では、色空間は、Y C b C rモデルを使用して定義されることができる。Y C b C rモデルにおいて、Yは、ルマ成分（明るさ）を表し、C bおよびC rはクロマ成分を表す。以下の説明では、Y U Vは、Y C b C rモデルを使用してエンコードされたフォーマットを表すために使用されることに留意されたい。

【0 1 1 8】

本開示の一態様によれば、ニューラルネットワークにおける複数のチャネルは、同じサイ

10

20

30

40

50

ズの色成分を操作するように構成される。いくつかの例では、ピクチャは、異なるサイズの色成分によって表されることができる。例えば、人間の視覚システムは、色よりも明るさの変化にはるかに敏感であるため、ビデオシステムは、人間の目で視認されるような視覚的な差異が生じないように、クロマ成分を圧縮してファイルサイズを縮小し伝送時間を短縮することができる。いくつかの例では、クロマサブサンプリング手法は、ルマよりも色差に対する人間の視覚システムの鋭敏さを利用して、ルマ情報よりもクロマ情報の解像度を低くするために使用される。

【0119】

いくつかの例では、サブサンプリングは、たとえば、4:4:4、4:2:0、4:2:2、4:1:1などの3部分の比率として表すことができる。例えば、4:4:4 (YUV444とも呼ばれる) は、各YCbCr成分がサブサンプリングなしで同じサンプルレートを有することを示している。4:2:0 (YUV420とも呼ばれます) は、クロマ成分がサブサンプリングされたことを示し、4ピクセルごとに (またはY成分) がCb成分とCr成分に対応することができる。以下の説明では、サブサンプリングフォーマットとしてYUV420を使用することを例に、本開示に開示された技術について説明する。説明の便宜上、サブサンプリングなしで同じサンプルレートを持つ色成分のフォーマット (例えば、YUV444) は、非サブサンプリングフォーマットと呼ばれ、サブサンプリングされた少なくとも1つの色成分のフォーマット (例えば、YUV420、YUV422、YUV411など) は、サブサンプリングフォーマットと呼ばれる。

【0120】

通常、ニューラルネットワークは非サブサンプリングフォーマットのピクチャ (例えば、YUV444) を操作することができる。したがって、サブサンプリングフォーマットのピクチャの場合、ピクチャは入力としてニューラルネットワークネットワークに提供される前に非サブサンプリングフォーマットに変換される。

【0121】

図11は、いくつかの例におけるDRNFLFフィルタ (1100) のブロック図を示す。一例では、DRNFLFフィルタ (1010) の代わりにDRNFLFフィルタ (1100) を使用することができる。DRNFLFフィルタ (1100) は、図11に示すように互いに結合されたQPマップ量子化器 (1110)、前処理モジュール (1120)、メイン処理モジュール (1130) および後処理モジュール (1140) を含む。メイン処理モジュール (1130) には、図11に示されるように互いに結合されたパッチフェッチャー (1131)、パッチに基づくDRNFLFカーネル処理モジュール (1132) およびパッチリアセンブラ (1133) を含む。

【0122】

いくつかの例では、QPマップには、現在の再構成ピクチャ内の各ブロックを再構成するために適用されるQP値のマップが含まれている。QPマップ量子化器 (1110) は、値を一連の所定の値に量子化することができる。一例 (JVET-T0057など) では、QP値は、QPマップ量子化器 (1110) によって22、27、32および37のうちの1つに量子化することができる。

【0123】

前処理モジュール (1120) は、第1のフォーマットのデブロッキングされたピクチャを受け取り、メイン処理モジュール (1130) によって使用される第2のフォーマットに変換することができる。例えば、メイン処理モジュール (1130) は、YUV444フォーマットを有するピクチャを処理するように構成される。前処理モジュール (1120) がYUV444フォーマットとは異なるフォーマットのデブロッキングされたピクチャを受信すると、前処理モジュール (1120) は、異なるフォーマットのデブロッキングされたピクチャを処理し、YUV444フォーマットのデブロッキングされたピクチャを出力することができる。たとえば、前処理モジュール (1120) は、YUV420フォーマットのデブロッキングされたピクチャを受信し、次に、UおよびVクロマチャネルを水平方向および垂直方向に係数2補間して、YUV444フォーマットのデブロッキン

10

20

30

40

50

グされたピクチャを生成する。

【0124】

メイン処理モジュール(1130)は、YUV444フォーマットのデブロッキングされたピクチャを受信し、量子化されたQPマップを入力として受信できる。パッチフェッチャー(1131)は、入力をパッチに分解する。DRNLFカーネル処理モジュール(1132)は、DRNLFカーネルに基づいて、それぞれのパッチを処理することができる。パッチリアセンブラ(1133)は、DRNLFカーネル処理モジュール(1132)によって処理されたパッチを、YUV444フォーマットのフィルタリングされたピクチャに組み立てることができる。

【0125】

後処理モジュール(1140)は、第2のフォーマットのフィルタリングされたピクチャを第1のフォーマットに戻す。例えば、後処理モジュール(1140)は、(メイン処理モジュール(1130)から出力された)YUV444フォーマットのフィルタリングされたピクチャを受信し、YUV420フォーマットのフィルタリングされたピクチャを出力する。

【0126】

図12は、いくつかの例における前処理モジュール(1220)のブロック図を示す。一例では、前処理モジュール(1120)の代わりに前処理モジュール(1220)が使用されている。

【0127】

前処理モジュール(1220)は、YUV420フォーマットのデブロッキングされたピクチャを受信し、デブロッキングされたピクチャをYUV444フォーマットに変換し、YUV444フォーマットのデブロッキングされたピクチャを出力することができる。具体的には、前処理モジュール(1220)は、Y成分用のルマ入力チャネルと、それぞれU(Cb)成分用及びV(Cr)成分用の2つのクロマ入力チャネルを含む3つの入力チャネルによって、デブロッキングされたピクチャを受信する。また、前処理モジュール(1220)は、Y成分用のルマ出力チャネルと、それぞれU(Cb)成分用及びV(Cr)成分用の2つのクロマ出力チャネルを含む3つの出力チャネルによって、デブロッキングされたピクチャを出力する。

【0128】

一例では、デブロッキングされたピクチャがYUV420フォーマットの場合、Y成分のサイズは(H、W)、U成分のサイズは(H/2、W/2)、V成分のサイズは(H/2、W/2)である。ここで、Hは、デブロッキングされたピクチャの高さ(例えば、サンプル単位で)を表し、Wは、デブロッキングされたピクチャの幅(例えば、サンプル単位で)を表す。

【0129】

図12の例では、前処理モジュール(1220)は、Y成分のサイズを変更しない。前処理モジュール(1220)は、ルマ入力チャネルから、サイズ(H、W)のY成分を受け取り、サイズ(H、W)のY成分をルマ出力チャネルへ出力する。

【0130】

前処理モジュール(1220)は、それぞれU成分およびV成分のサイズを変更する。前処理モジュール(1220)は、U成分およびV成分をそれぞれ処理するための第1のサイズ変更ユニット(1221)および第2のサイズ変更ユニット(1222)を含む。たとえば、第1のサイズ変更ユニット(1221)は、サイズ(H/2、W/2)のU成分を受信し、U成分のサイズをサイズ(H、W)に変更し、サイズ(H、W)のU成分をU成分用のクロマ出力チャネルへ出力する。第2のサイズ変更ユニット(1222)は、サイズ(H/2、W/2)のV成分を受信し、V成分のサイズをサイズ(H、W)に変更し、サイズ(H、W)のV成分をV成分用のクロマ出力チャネルへ出力する。幾つかの例では、第1のサイズ変更ユニット(1221)は補間に基づいて、例えばLanczos補間フィルタを使用してU成分のサイズを変更する。同様に、いくつかの例では、第2のサイズ

10

20

30

40

50

変更ユニット（１２２２）は補間に基づいて、例えばLanczos補間フィルタを使用してV成分のサイズを変更する。

【０１３１】

いくつかの例では、Lanczos補間フィルタなどを使用する補間演算は、補間演算の出力が、意味のあるU（Cb）成分およびV（Cr）成分に対して非負であるなど、意味のある値であることを保証できない。いくつかの例では、前処理されたYUV444フォーマットのデブロッキングされたピクチャを保存し、その後、保存されたYUV444フォーマットのピクチャをニューラルネットワークのトレーニングプロセスで利用できる。U（Cb）成分及びV（Cr）成分の負の値は、ニューラルネットワークのトレーニングプロセスの結果に悪影響を与える恐れがある。

10

【０１３２】

図１３は、ニューラルネットワーク構造（１３００）のブロック図を示す。いくつかの例では、ニューラルネットワーク構造（１３００）は、密な残差畳み込みニューラルネットワークに基づくループ内フィルタ（DRNLF）に使用され、パッチによるDRNLFカーネル処理モジュール（１１３２）に代わって使用され得る。ニューラルネットワーク構造（１３００）には、DRU（１３０１）～

DRU（１３０４）などの一連の密な残差ユニット（Dense Residual Unit, DRU）が含まれ、DRUの数はNで表される。図１３では、畳み込みカーネルの数はMで表され、Mは畳み込み用の出力チャンネルの数でもある。たとえば、「CONV3×3×M」は、カーネルサイズが3×3のM個の畳み込みカーネルによる標準畳み込みを示し、「DSC3×3×M」は、カーネルサイズが3×3のM個の畳み込みカーネルによる深さ方向の分離可能な畳み込みを示す。NとMは、計算効率と性能の両立から設定され得る。一例（例えばJ-VET-T0057）では、Nは4、Mは32に設定される。

20

【０１３３】

動作中、ニューラルネットワーク構造（１３００）は、デブロッキングされたピクチャをパッチごとに処理する。YUV444フォーマットのデブロッキングされたピクチャの各パッチについて、パッチは正規化され（たとえば、図１３の例では1023で除算される）、デブロッキングされたピクチャの平均値が正規化されたパッチから削除され、内部入力（１３１３）の第１の部分（１３１１）を取得する。内部入力（１３１３）の第２の部分はQPマップからのものである。たとえば、第１の部分（１３１１）を形成するパッチに対応するQPマップのパッチ（QPマップパッチと呼ばれる）は、QPマップから取得される。QPマップパッチは正規化され（たとえば、図１３では51で除算される）、正規化されたQPマップパッチは、内部入力（１３１３）の第２の部分（１３１２）となる。第２の部分（１３１２）は、内部入力（１３１３）を取得するために、第１の部分（１３１１）に連結されている。内部入力（１３１３）は、第１の標準畳み込みブロック（１３５１）（CONV3×3×Mで示される）に提供される。そして、第１の標準畳み込みブロック（１３５１）の出力はN個のDRUによって処理される。

30

【０１３４】

DRUごとに、中間入力を受信され処理される。DRUの出力は中間入力と連結されて、次のDRUのための中間入力を形成する。例として、DRU（１３０２）を使用する場合、DRU（１３０２）は中間入力（１３２１）を受信し、中間入力（１３２１）を処理し、出力（１３２２）を生成する。出力（１３２２）は中間入力（１３２１）と連結されて、DRU（１３０３）のための中間入力（１３２３）が形成される。

40

【０１３５】

なお、中間入力（１３２１）はM個以上のチャンネルを有するため、中間入力（１３２１）に「CONV1×1×M」の畳み込み演算を適用して、DRU（１３０２）でさらに処理するためのM個のチャンネルを生成することができる。また、第１の標準畳み込みブロック（１３５１）の出力にはM個のチャンネルが含まれているため、「CONV1×1×M」の畳み込み演算を使用せずに、この出力をDRU（１３０１）で処理することができること

50

にも留意されたい。

【0136】

最後のDRUの出力は、最後の標準畳み込みブロック(1359)に提供される。最後の標準畳み込みブロック(1359)の出力は、例えば、図13に示すように、デブロッキングされたピクチャの平均値を加算し1023を乗算することによって、標準ピクチャパッチ値に変換される。

【0137】

図14は、密な残差ユニット(DRU)(1400)のブロック図を示す。いくつかの例では、DRU(1301) DRU(1302)、DRU(1303)及びDRU(1304)などの図13における各DRUの代わりにDRU(1400)を使用することができる。

10

【0138】

図14の例では、DRU(1400)は、中間入力xを受信するとともに、ショートカット(1401)を介して中間入力xを後続のDRUに直接伝播する。DRU(1400)はまた、標準処理パス(1402)を含む。いくつかの例では、標準処理パス(1402)は、標準畳み込み層(1411)、深さ方向の分離可能な畳み込み(DSC)層(1412)および(1414)、ならびに正規化線形ユニット(ReLU)層(1413)を含む。例えば、中間入力xは、標準処理パス(1402)の出力と連結されて、後続のDRUのための中間入力を形成する。

【0139】

いくつかの例では、DSC層(1412)および(1414)は、計算コストを削減するために使用される。

20

【0140】

本開示の一態様によれば、ニューラルネットワーク構造(1300)は、Y、U(Cb)、V(Cr)成分にそれぞれ対応する3つのチャンネルを含む。幾つかの例では、この3つのチャンネルは、Yチャンネル、Uチャンネル、およびUチャンネルと呼ぶことができる。DRNFLフィルタ(1100)は、イントラピクチャとインターピクチャの両方に適用できる。いくつかの例では、ピクチャレベルとCTUレベルでDRNFLフィルタ(1100)のオン/オフを示す追加のフラグがシグナリングされる。

【0141】

図15は、いくつかの例における後処理モジュール(1540)のブロック図を示す。一例では、後処理モジュール(1540)は、後処理モジュール(1140)の代わりに使用され得る。後処理モジュール(1540)には、Y成分、U成分、V成分の値をそれぞれ予め定められた非負の範囲[a、b]にクリッピングするクリッピングユニット(1541)～(1543)を含む。一例では、非負の範囲の下限aと上限bは、 $a = 16 \times 4$ および $b = 234 \times 4$ として設定され得る。さらに、後処理モジュール(1540)は、それぞれ、クリッピングされたU成分及びV成分をサイズ(H、W)からサイズ(H/2、W/2)にクリッピングするサイズ変更ユニット(1545)及びサイズ変更ユニット(1546)を含む。ここで、Hは、元のピクチャ(デブロッキングされたピクチャなど)の高さ、Wは元のピクチャの幅である。

30

【0142】

本開示の態様は、前処理の技術を提供する。前処理されたデータを保存してニューラルネットワークのトレーニングに使用し、より良いトレーニングと推論の結果を得ることができる。

【0143】

図16は、いくつかの例における前処理モジュール(1620)のブロック図を示す。一例では、前処理モジュール(1120)の代わりに前処理モジュール(1620)が使用されている。

40

【0144】

前処理モジュール(1620)は、YUV420フォーマットのデブロッキングされたピ

50

クチャを受信し、デブロッキングされたピクチャをYUV444フォーマットに変換し、YUV444フォーマットのデブロッキングされたピクチャを出力することができる。具体的には、前処理モジュール(1620)は、Y成分用のルマ入力チャンネルと、それぞれU(Cb)成分用及びV(Cr)成分用の2つのクロマ入力チャンネルを含む3つの入力チャンネルによって、デブロッキングされたピクチャを受信する。また、前処理モジュール(1620)は、Y成分用のルマ出力チャンネルと、それぞれU(Cb)成分用及びV(Cr)成分用の2つのクロマ出力チャンネルを含む3つの出力チャンネルによって、デブロッキングされたピクチャを出力する。

【0145】

一例では、デブロッキングされたピクチャがYUV420フォーマットの場合、Y成分のサイズは(H、W)、U成分のサイズは(H/2、W/2)、V成分のサイズは(H/2、W/2)である。ここで、Hは、デブロッキングされたピクチャの高さ(例えば、サンプル単位で)を表し、Wは、デブロッキングされたピクチャの幅(例えば、サンプル単位で)を表す。

【0146】

図16の例では、前処理モジュール(1620)は、Y成分のサイズを変更しない。前処理モジュール(1620)は、ルマ入力チャンネルから、サイズ(H、W)のY成分を受け取り、サイズ(H、W)のY成分をルマ出力チャンネルへ出力する。

【0147】

前処理モジュール(1620)は、U成分およびV成分のサイズをそれぞれ変更する。前処理モジュール(1620)は、U成分およびV成分をそれぞれ処理するための第1のサイズ変更ユニット(1621)および第2のサイズ変更ユニット(1622)を含む。たとえば、第1のサイズ変更ユニット(1621)は、サイズ(H/2、W/2)のU成分を受信し、U成分のサイズをサイズ(H、W)に変更し、サイズ(H、W)のU成分をU成分用のクロマ出力チャンネルへ出力する。第2のサイズ変更ユニット(1622)は、サイズ(H/2、W/2)のV成分を受信し、V成分のサイズをサイズ(H、W)に変更し、サイズ(H、W)のV成分をV成分用のクロマ出力チャンネルへ出力する。幾つかの例では、第1のサイズ変更ユニット(1621)は補間に基づいて、例えばLanczos補間フィルタを使用してU成分のサイズを変更する。同様に、いくつかの例では、第2のサイズ変更ユニット(1622)は補間に基づいて、例えばLanczos補間フィルタを使用してV成分のサイズを変更する。

【0148】

いくつかの例では、Lanczos補間フィルタなどを使用する補間演算は、補間演算の出力が、意味のあるU(Cb)成分およびV(Cr)成分に対して非負であるなど、意味のある値であることを保証できない。

【0149】

図16の例では、前処理モジュール(1620)は、補間された後のU成分およびV成分の値を[c、d]の範囲にそれぞれクリッピングするためのクリッピングユニット(1625)および(1626)を含む。いくつかの例では、前処理用のY成分、U成分及びV成分の値は、ビット深度が10であるため、cとdは、 $c=0$ 且つ $d=2^{\text{bit depth}-1}=1023$ として設定できる。

【0150】

一例では、c値とd値が予め定義され使用される。別の一例では、c値とd値の複数のペアが予め定義され、クリッピングで使用されるc値とd値のペアのインデックスは、シーケンスパラメータセット(SPS)、ピクチャパラメータセット(PPS)、スライスまたはタイルヘッダーなどのビットストリームでシグナリングされ得る。

【0151】

いくつかの例では、U成分とV成分のクリッピングされた値とY成分の値は、デブロッキングされたピクチャとしてYUV444フォーマットで保存されることができる。いくつかの実施形態では、YUV444フォーマットの保存されたピクチャは、メイン処理モジ

10

20

30

40

50

ルール（１１３０）のニューラルネットワークなどのニューラルネットワークのトレーニングプロセスで入力として使用され得る。いくつかの例では、UおよびV成分の値は、ニューラルネットワークのトレーニングプロセスに悪影響を及ぼさない範囲にクリッピングされる。一例では、U成分とV成分の値が非負になるようにクリッピングされている。

【０１５２】

いくつかの例では、YUV444フォーマットで保存されたクリッピングされた値をもつピクチャを使用する場合、トレーニング中の前処理（例えば、サイズ変更、クリッピング）ステップを回避することにより、時間が節約されているため、ニューラルネットワークのトレーニングを高速化することができる。また、ニューラルネットワークは、圧縮効率および／または画質を改善できる、より優れたモデルパラメータを使用してトレーニング

10

【０１５３】

いくつかの例では、前処理モジュール（１６２０）にクリッピングユニット（１６２５）および（１６２６）を追加することにより、例えば、より低いピオンテガードデルタレート（BDレート）で、圧縮効率および／または画質を改善することができる。

【０１５４】

図１７は、本開示の実施形態によるプロセス（１７００）を概説するフローチャートを示す。プロセス（１７００）は、ビデオを処理するために使用され得る。様々な実施形態では、プロセス（１７００）は、端末装置（３１０）、（３２０）、（３３０）および（３４０）内の処理回路、ビデオエンコーダ（４０３）の機能を実行する処理回路、ビデオデコーダ（４１０）の機能を実行する処理回路、ビデオエンコーダ（６０３）の機能を実行する処理回路などの処理回路によって実行される。いくつかの実施形態では、プロセス（１７００）はソフトウェア命令で実現されるため、処理回路がソフトウェア命令を実行すると、処理回路はプロセス（１７００）を実行する。該プロセスは、（Ｓ１７０１）で開始し、（Ｓ１７１０）に進む。

20

【０１５５】

（Ｓ１７１０）において、色空間におけるサブサンプリングフォーマットのピクチャを色空間における非サブサンプリングフォーマットに変換する。いくつかの例では、変換は補間に基づいて実行され、無効値が生じる可能性がある。一例では、変換によってYCbCrモデルに無効な負の値が生じる可能性がある。

30

【０１５６】

（Ｓ１７２０）において、非サブサンプリングフォーマットのピクチャを入力としてニューラルネットワークに基づくフィルタに提供する前に、非サブサンプリングフォーマットのピクチャの１つまたは複数の色成分の値をクリッピングする。いくつかの例では、１つまたは複数の色成分はクロマ成分であり得る。次に、プロセスは（Ｓ１７９９）に進む。

【０１５７】

一例では、非サブサンプリングフォーマットのピクチャの色成分の値を、該色成分の有効範囲にクリッピングする。一例では、非サブサンプリングフォーマットのピクチャの色成分の値は非負になるようにクリッピングされる。別の一例では、ビット深度に基づいて範囲を決定する。たとえば、この範囲の下限は０、且つこの範囲の上限は（ $2^{\text{bit depth}}$ ）-１に設定される。

40

【０１５８】

いくつかの例では、この範囲は予め定められている。いくつかの例では、ピクチャを含むビットストリームからのデコーディングされた情報に基づいて、範囲を決定する。いくつかの例では、ビットストリームにおけるシーケンスパラメータセット、ピクチャパラメータセット、スライスヘッダー及びタイルヘッダーのうち少なくとも１つから、範囲を示す信号をデコーディングする。

【０１５９】

一例では、複数の範囲を予め定めることができる。そして、ビットストリームにおけるシーケンスパラメータセット、ピクチャパラメータセット、スライスヘッダー及びタイルヘ

50

ッダーのうち少なくとも1つに、複数の範囲のうちの1つを示すインデックスを含むことができる。

【0160】

いくつかの例では、デコーダにおいてプロセス(1700)を使用する。例えば、ビットストリームからのデコーディングされた情報に基づいて、サブサンプリングフォーマットのピクチャを再構成し、且つサブサンプリングフォーマットのピクチャをサブサンプリングフォーマットから非サブサンプリングフォーマットに変換する前に、サブサンプリングフォーマットのピクチャにデブロッキングフィルタを適用する。別の一例では、ニューラルネットワークに基づくフィルタを、クリッピングされた値を有する非サブサンプリングフォーマットのピクチャに適用することで、非サブサンプリングフォーマットのフィルタリングされたピクチャを生成し、その後、非サブサンプリングフォーマットのフィルタリングされたピクチャを、サブサンプリングフォーマットのフィルタリングされたピクチャに変換する。

10

【0161】

いくつかの例では、クリッピングされた値を有する非サブサンプリングフォーマットのピクチャをストレージに保存させる。次に、クリッピングされた値を有する非サブサンプリングフォーマットのピクチャおよび他のピクチャをトレーニング入力として提供し、ニューラルネットワークに基づくフィルターにおけるニューラルネットワークをトレーニングすることができる。

20

【0162】

上記の説明における様々なユニット、ブロックおよびモジュールは、処理回路、ソフトウェア命令を実行するプロセッサ、ハードウェアとソフトウェアの組み合わせなどの様々な技術によって実現され得ることに留意されたい。

【0163】

以上で説明された技術は、コンピュータ読取可能な命令を使用するコンピュータソフトウェアとして実現され、1つまたは複数のコンピュータ読取可能な媒体に物理的に記憶されることができる。例えば、図18は、開示された主題の特定の実施形態を実行することに適したコンピュータシステム(1800)を示す。

【0164】

コンピュータソフトウェアは、アセンブリ、コンパイル、リンク、またはそのようなメカニズムを施されて、1つまたは複数のコンピュータ中央処理装置(CPU)、グラフィックスプロセッシングユニット(GPU)などによって直接、または解釈、マイクロコード実行などによって実行され得る命令を含むコードを作成する任意の適切な機械コードまたはコンピュータ言語を用いてコーディングされることができる。

30

【0165】

命令は、例えば、パーソナルコンピュータ、タブレットコンピュータ、サーバ、スマートフォン、ゲームデバイス、モノのインターネットデバイスなどを含む、様々なタイプのコンピュータまたはそのコンポーネント上で実行されることができる。

【0166】

コンピュータシステム(1800)について、図18に示される例示的なコンポーネントは、本質的に例示的なものであり、本開示の実施形態を実施するコンピュータソフトウェアの使用または機能の範囲に関していかなる限定を示唆することも意図されない。またコンポーネントの構成は、コンピュータシステム(1800)の例示的な実施形態で示されるコンポーネントのうちのいずれか1つ又は組み合わせに関する任意の依存性又は必要性を有するとして解釈されるべきでもない。

40

【0167】

コンピュータシステム(1800)は、特定のヒューマンインターフェース入力デバイスを含み得る。このようなヒューマンインターフェース入力デバイスは、例えば、触覚入力(キーストローク、スワイプ、データグローブの動きなど)、オーディオ入力(音声、拍手など)、視覚入力(ジェスチャーなど)、嗅覚入力(示されていない)によって、1人

50

以上のユーザによる入力に応答することができる。ヒューマンインターフェースデバイスは、オーディオ（音声、音楽、環境音など）、ピクチャ（走査画像、静止画像カメラから取得される写真画像など）、ビデオ（2次元ビデオ、立体ビデオを含む3次元ビデオなど）など、人間による意識的な入力に必ずしも直接関係しない特定のメディアをキャプチャすることにも使用され得る。

【0168】

入力ヒューマンインターフェースデバイスは、キーボード（1801）、マウス（1802）、トラックパッド（1803）、タッチスクリーン（1810）、データグローブ（図示せず）、ジョイスティック（1805）、マイクフォン（1806）、スキャナ（1807）、カメラ（1808）（各種につき1つのみ示されている）のうちの1つまたは複数を含み得る。

10

【0169】

コンピュータシステム（1800）は、特定のヒューマンインターフェース出力デバイスをも含み得る。このようなヒューマンインターフェース出力デバイスは、例えば、触覚出力、音声、光、および嗅覚／味覚を介して1人以上のユーザの感覚を刺激し得る。このようなヒューマンインターフェース出力デバイスは、触覚出力デバイス（例えば、タッチスクリーン（1810）、データグローブ（図示せず）、またはジョイスティック（1805）による触覚フィードバックがあるが、入力デバイスとして機能しない触覚フィードバックデバイスであってもよい）、オーディオ出力デバイス（スピーカ（1809）、ヘッドホン（図示せず）など）、視覚出力デバイス（CRTスクリーン、LCDスクリーン、プラズマスクリーン、OLEDスクリーンを含むスクリーン（1810）（それぞれタッチスクリーン入力能力を有するかもしくは有せず、それぞれ触覚フィードバック能力を有するかもしくは有しない。それらの一部は、ステレオグラフィック出力などの手段を介して、2次元の視覚出力または3次元以上の出力を出力することができる）、仮想現実眼鏡（図示せず）、ホログラフィックディスプレイおよびスモークタンク（図示せず）など）、およびプリンタ（図示せず）を含み得る。

20

【0170】

コンピュータシステム（1800）は、人間がアクセス可能な記憶装置およびそれらの関連する媒体、例えば、CD／DVDなどの媒体（1821）付きのCD／DVD ROM／RW（1820）を含む光学媒体、サムドライブ（1822）、リムーバブルハードドライブまたはソリッドステートドライブ（1823）、テープやフロッピー（登録商標）ディスクなどの従来の磁気媒体（図示せず）、セキュリティ dongle などの専用のROM／ASIC／PLDベースのデバイス（図示せず）などをも含み得る。

30

【0171】

ここで、開示された主題に合わせて使用される「コンピュータ読取可能な媒体」という用語は、送信媒体、搬送波、または他の一時的な信号を含まないことは、当業者に理解されるべきであろう。

【0172】

コンピュータシステム（1800）は、1つまたは複数の通信ネットワーク（1855）へのインターフェース（1854）をさらに含み得る。ネットワークは、例えば、無線、有線、光学的であり得る。ネットワークは、さらに、ローカル、広域、大都市圏、車両用および産業用、リアルタイム、遅延耐性などであり得る。ネットワークの例としては、イーサネット（登録商標）、無線LANなどのローカルエリアネットワーク、GSM、3G、4G、5G、LTEなどを含むセルラーネットワーク、ケーブルTV、衛星TV、および地上放送TVを含むTV有線または無線広域デジタルネットワーク、CANBusを含む車両用や産業用などが含まれる。特定のネットワークは、一般に、特定の汎用データポートまたは周辺バス（1849）（例えば、コンピュータシステム（1800）のUSBポートなど）に接続された外部ネットワークインターフェースアダプターを必要とする。他のものは一般に、以下で説明するようにシステムバスに接続することにより、コンピュータシステム（1800）のコアに統合される（例えば、PCコンピュータシステム

40

50

へのイーサネット（登録商標）インターフェースまたはスマートフォンコンピュータシステムへのセルラーネットワークインターフェース）。これらのネットワークのいずれかを用いて、コンピュータシステム（１８００）は、他のエンティティと通信することができる。このような通信は、単方向の受信のみ（例えば、放送ＴＶ）、単方向の送信のみ（例えば、特定のＣＡＮｂｕｓデバイスへのＣＡＮｂｕｓ）、または双方向、例えばローカルまたはワイドエリアデジタルネットワークを用いる他のコンピュータシステムへの送信であり得る。特定のプロトコルおよびプロトコルスタックを上述したこれらのネットワークおよびネットワークインターフェースのそれぞれでを使用することができる。

【０１７３】

前述のヒューマンインターフェースデバイス、人間がアクセス可能な記憶装置、およびネットワークインターフェースは、コンピュータシステム（１８００）のコア（１８４０）に接続されることができる。

【０１７４】

コア（１８４０）は、１つまたは複数の中央処理装置（ＣＰＵ）（１８４１）、グラフィックスプロセッシングユニット（ＧＰＵ）（１８４２）、フィールドプログラマブルゲートアレイ（ＦＰＧＡ）（１８４３）の形式の専用プログラマブル処理ユニット、特定のタスク向けのハードウェアアクセラレータ（１８４４）、グラフィックアダプタ（１８５０）などを含むことができる。これらのデバイスは、リードオンリーメモリ（ＲＯＭ）（１８４５）、ランダムアクセスメモリ（１８４６）、非ユーザアクセス可能な内部ハードドライブ、ＳＳＤなどの内部大容量記憶装置（１８４７）とともに、システムバス（１８４８）を介して接続されてもよい。一部のコンピュータシステムでは、システムバス（１８４８）は、１つまたは複数の物理プラグの形態でアクセスでき、追加のＣＰＵ、ＧＰＵなどによる拡張を可能にする。周辺機器は、コアのシステムバス（１８４８）に直接、または周辺バス（１８４９）を介して接続され得る。一例では、ディスプレイ（１８１０）は、グラフィックアダプタ（１８５０）に接続され得る。周辺バスのアーキテクチャは、ＰＣＩ、ＵＳＢなどを含む。

【０１７５】

ＣＰＵ（１８４１）、ＧＰＵ（１８４２）、ＦＰＧＡ（１８４３）、およびアクセラレータ（１８４４）は、組み合わせて、前述のコンピュータコードを構成することができる特定の命令を実行することができる。そのコンピュータコードは、ＲＯＭ（１８４５）またはＲＡＭ（１８４６）に記憶され得る。移行データはＲＡＭ（１８４６）にも記憶できるが、永続データは、例えば、内部大容量ストレージ（１８４７）に記憶され得る。１つまたは複数のＣＰＵ（１８４１）、ＧＰＵ（１８４２）、大容量ストレージ（１８４７）、ＲＯＭ（１８４５）、ＲＡＭ（１８４６）などと密接に関連付けることができるキャッシュメモリを使用することにより、任意のメモリデバイスへの高速保存および検索が可能になる。

【０１７６】

コンピュータ読取可能な媒体は、様々なコンピュータ実現操作を実行するためのコンピュータコードを備えることができる。媒体およびコンピュータコードは、本開示の目的のために特別に設計および構築されたものであり得るか、もしくは、それらは、コンピュータソフトウェア技術の当業者に周知であって利用可能な種類のものであり得る。

【０１７７】

限定ではなく、一例として、アーキテクチャを有するコンピュータシステム（１８００）、特にコア（１８４０）は、１つまたは複数の有形のコンピュータ読取可能な媒体に組み込まれたソフトウェアを実行するプロセッサ（ＣＰＵ、ＧＰＵ、ＦＰＧＡ、アクセラレータなどを含む）の結果としての機能を提供することができる。このようなコンピュータ読取可能な媒体は、以上で説明されたようにユーザがアクセス可能な大容量ストレージ、および、コア内部大容量ストレージ（１８４７）またはＲＯＭ（１８４５）などの非一時的な性質を持つコア（１８４０）の特定のストレージに関連付けられた媒体であり得る。本開示の様々な実施形態を実行するソフトウェアは、このようなデバイスに記憶され、コア

10

20

30

40

50

(1840)によって実行されることができる。コンピュータ読取可能な媒体は、特定の必要に応じて、1つまたは複数のメモリデバイスまたはチップを含むことができる。ソフトウェアは、コア(1840)、具体的にはその中のプロセッサ(CPU、GPU、FPGAなどを含む)に、RAM(1846)に記憶されたデータ構造を定義すること、および、ソフトウェアで定義されたプロセスに従ってこのようなデータ構造を変更することを含む、ここで説明する特定のプロセスまたは特定のプロセスの特定の部分を実行させることができる。加えて、または、代替として、コンピュータシステムは、本明細書に記載された特定のプロセスまたは特定のプロセスの特定の部分を実行するためにソフトウェアの代わりにまたは一緒に動作することができる回路(例えば、アクセラレータ(1844))に有線接続されたまたは組み込まれたロジックの結果としての機能を提供することができる。ソフトウェアへの引用は、必要に応じて、ロジックを含むことができ、その逆も同様である。コンピュータ読取可能な媒体への引用は、必要に応じて、実行のためのソフトウェアを記憶する回路(集積回路(IC)など)、実行のためのロジックを具現化する回路、またはその両方を含むことができる。本開示は、ハードウェアとソフトウェアの任意の適切な組み合わせを含む。

【0178】

付録A：頭字語

JEM: Joint Exploration Model (共同探索モデル)
VVC: Versatile Video Coding (多用途ビデオコーディング)
BMS: Benchmark Set (ベンチマークセット)
MV: Motion Vector (動きベクトル)
HEVC: High Efficiency Video Coding (高効率ビデオコーディング)
SEI: Supplementary Enhancement Information (補足拡張情報)
VUI: Video Usability Information (ビデオユーザビリティ情報)
GOP: Groups of Pictures (グループオブピクチャ)
TU: Transform Units (変換ユニット)
PU: Prediction Units (予測ユニット)
CTU: Coding Tree Units (コーディングツリーユニット)
CTB: Coding Tree Blocks (コーディングツリーブロック)
PB: Prediction Blocks (予測ブロック)
HRD: Hypothetical Reference Decoder (仮想参照デコーダ)
SNR: Signal Noise Ratio (信号対雑音比)
CPU: Central Processing Units (中央処理装置)
GPU: Graphics Processing Units (グラフィックスプロセッシングユニット)
CRT: Cathode Ray Tube (陰極線管)
LCD: Liquid-Crystal Display (液晶ディスプレイ)
OLED: Organic Light-Emitting Diode (有機発光ダイオード)
CD: Compact Disc (コンパクトディスク)
DVD: Digital Video Disc (デジタルビデオディスク)
ROM: Read-Only Memory (リードオンリーメモリ)
RAM: Random Access Memory (ランダムアクセスメモリ)
ASIC: Application-Specific Integrated Circuit (特定用途向け集積回路)
PLD: Programmable Logic Device (プログラマブルロジック)

10

20

30

40

50

クデバイス)

LAN: Local Area Network (ローカルエリアネットワーク)

GSM: Global System for Mobile Communications (グローバルモバイル通信システム)

LTE: Long-Term Evolution (長期的な進化)

CANBus: Controller Area Network Bus (コントローラエリアネットワークバス)

USB: Universal Serial Bus (ユニバーサルシリアルバス)

PCI: Peripheral Component Interconnect (ペリフェラルコンポーネントインターコネクト)

FPGA: Field Programmable Gate Array (フィールドプログラマブルゲートアレイ)

SSD: Solid-State Drive (ソリッドステートドライブ)

IC: Integrated Circuit (集積回路)

CU: Coding Unit (コーディングユニット)

【0179】

本開示はいくつかの例示的な実施形態について説明してきたが、本開示の範囲内に含まれる変更、置換、および様々な代替の均等物が存在する。したがって、当業者は、本明細書では明記または記載されていないが、本開示の原理を具現化し、その思想および範囲内に含まれる様々なシステムおよび方法を考案できることは、理解されるべきであろう。

10

20

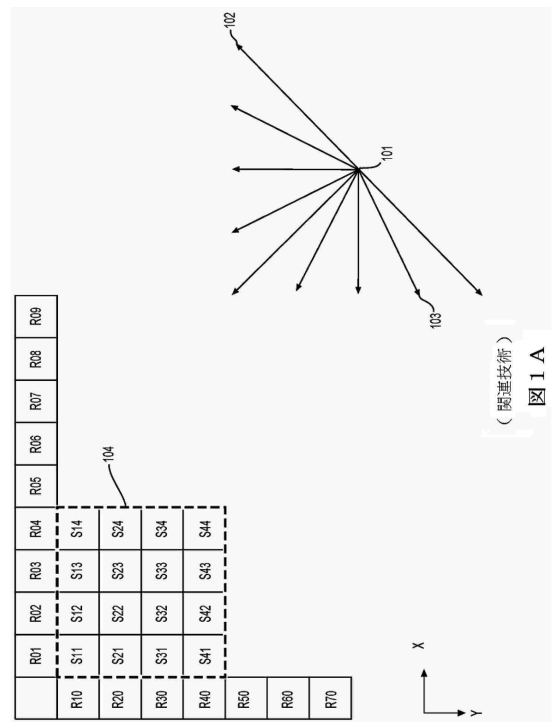
30

40

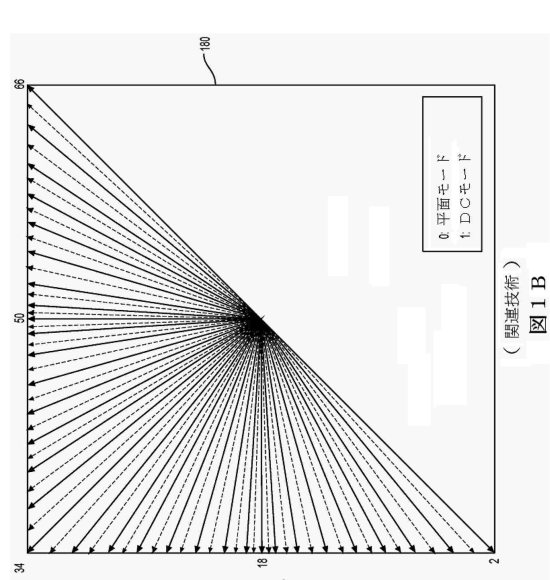
50

【図面】

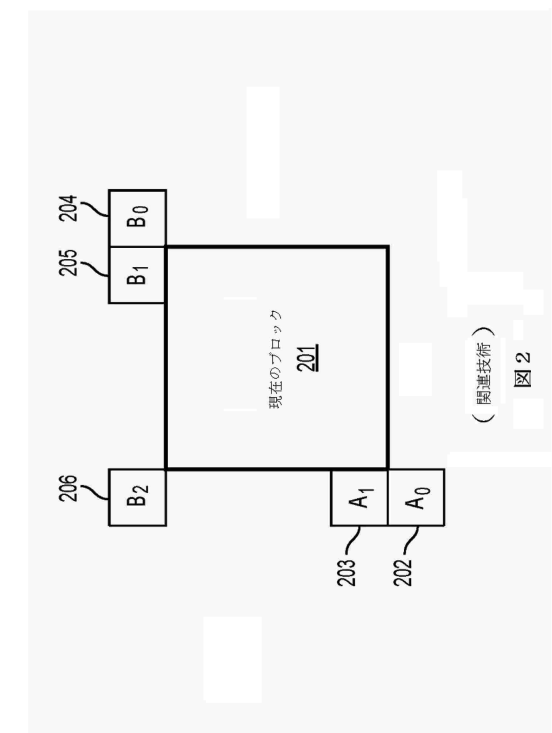
【図 1 A】



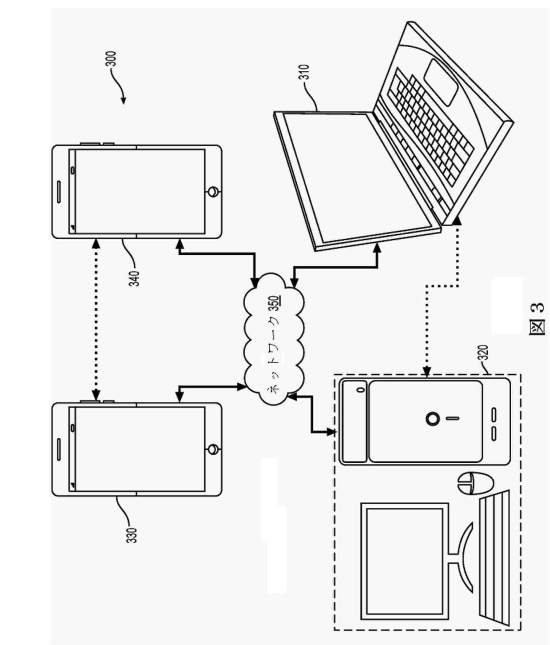
【図 1 B】



【図 2】



【図 3】



10

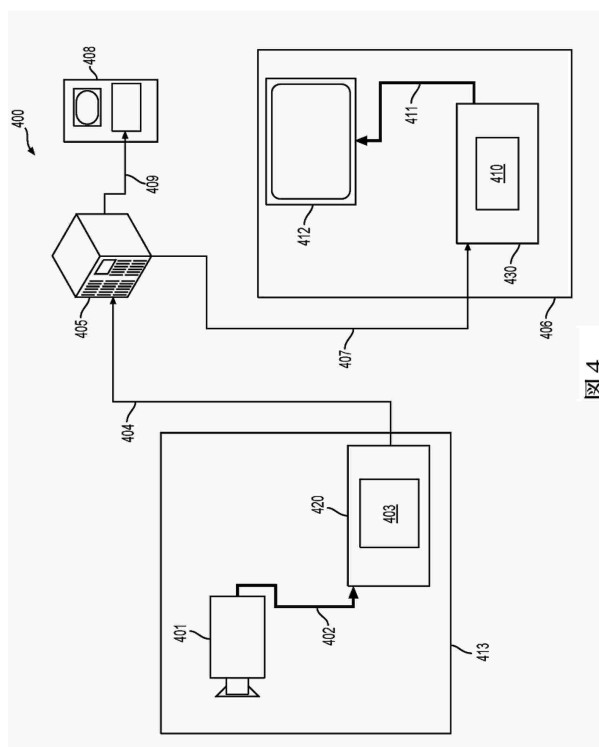
20

30

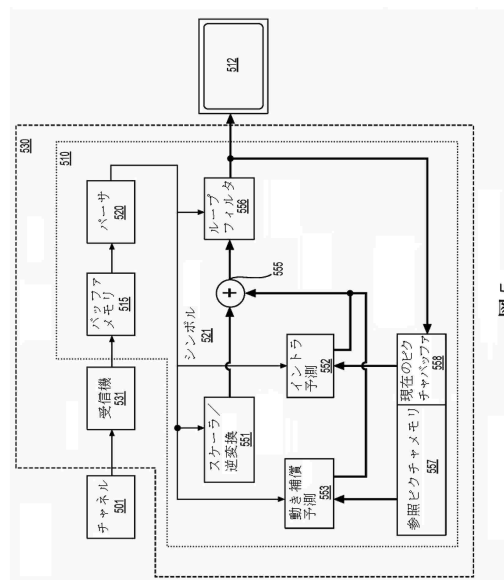
40

50

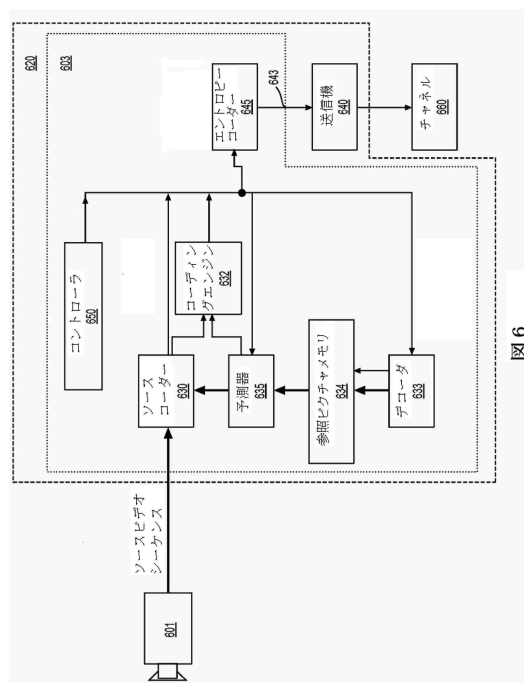
【図 4】



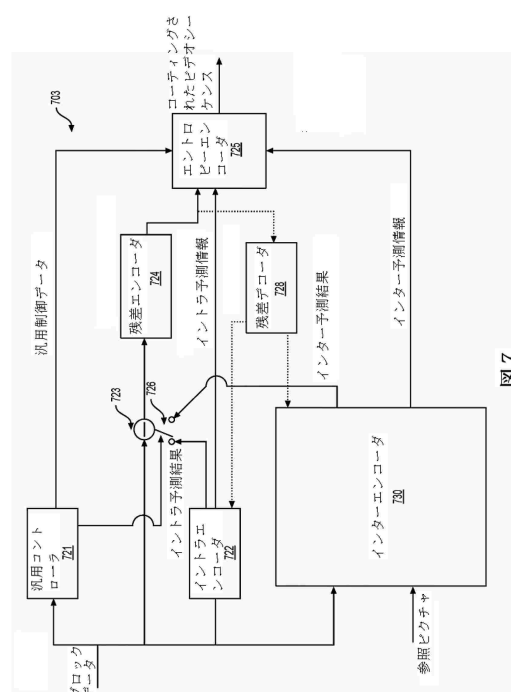
【図 5】



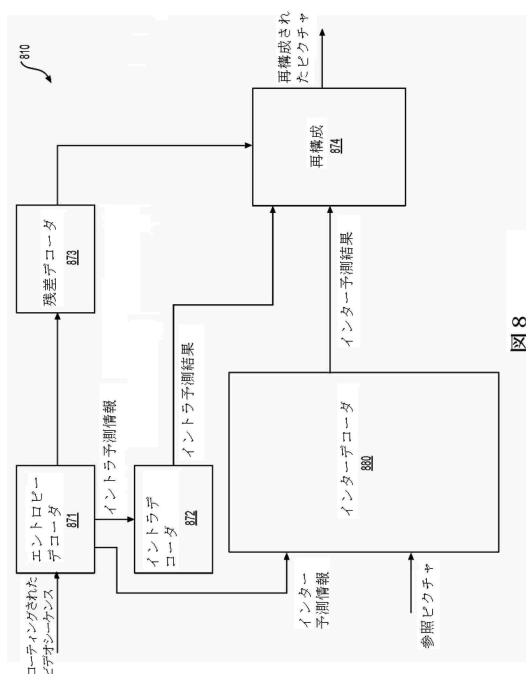
【圖 6】



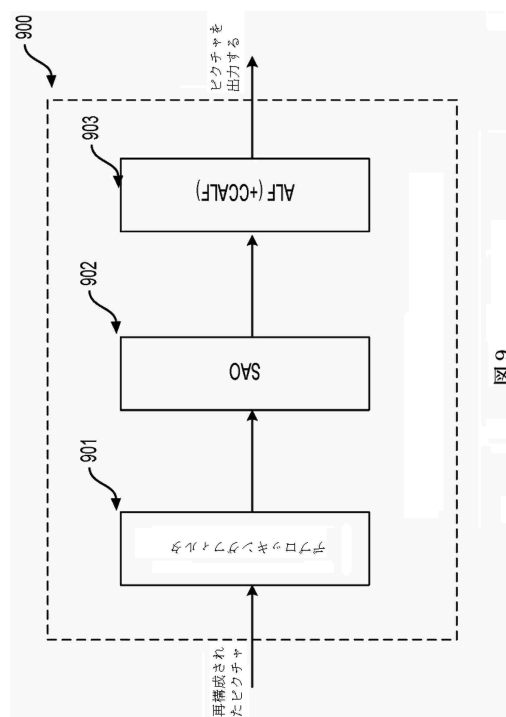
【圖 7】



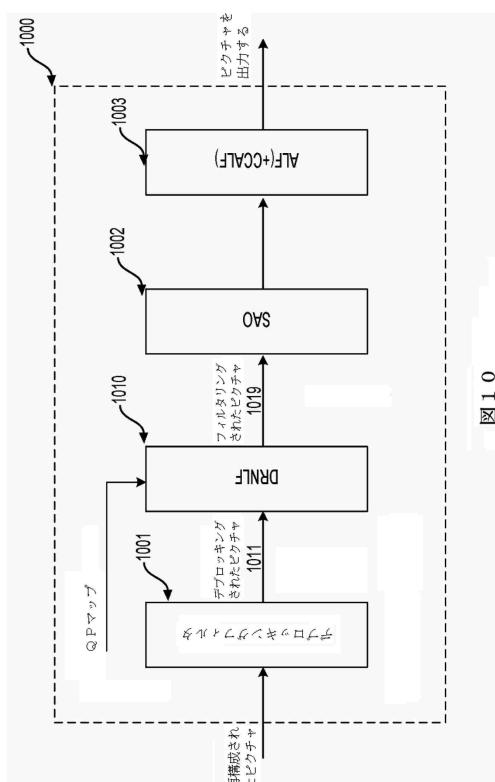
【圖 8】



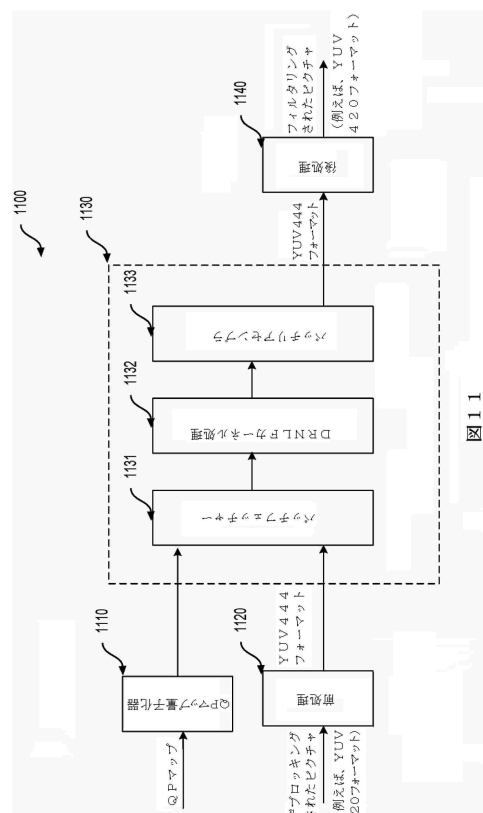
【圖 9】



【図 10】



【図 1 1】



【図 12】

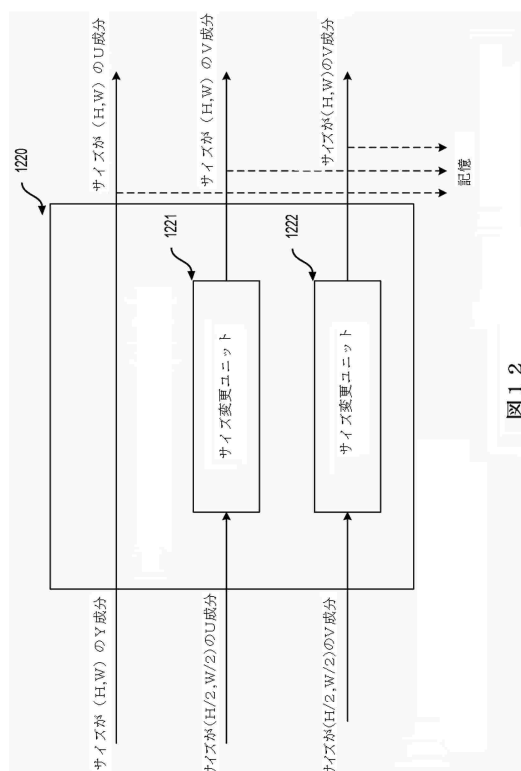


图 12

【図 13】

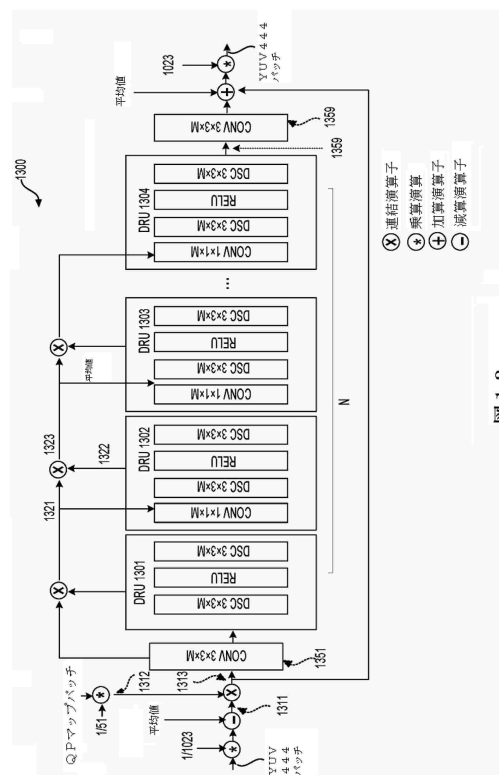


图 13

【図 14】

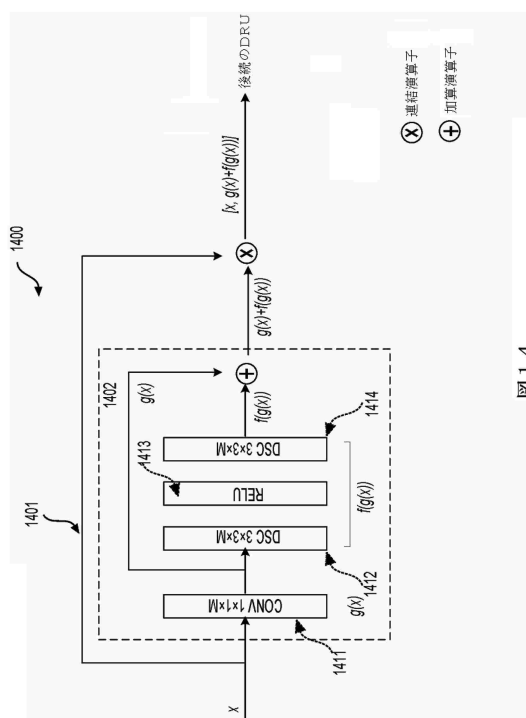


图 14

【図 15】

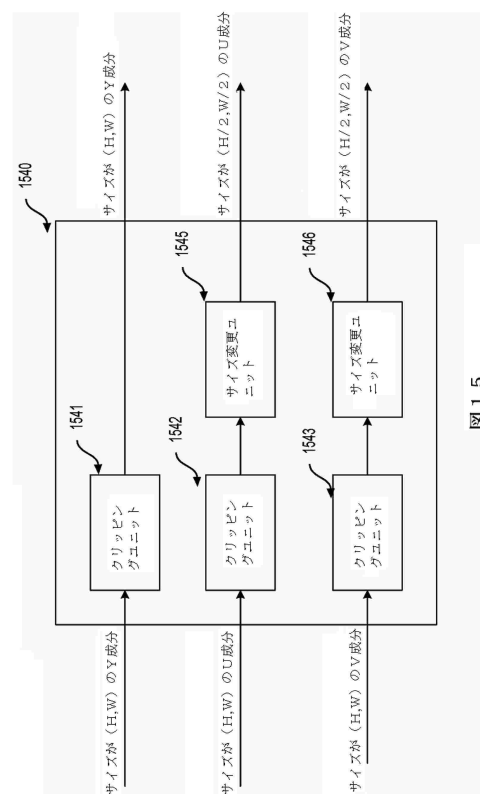


图 15

【図 16】

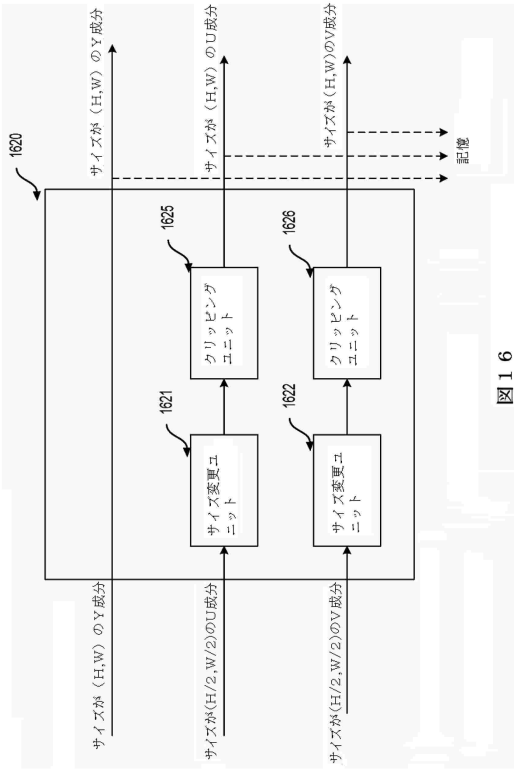


図 16

【図 17】

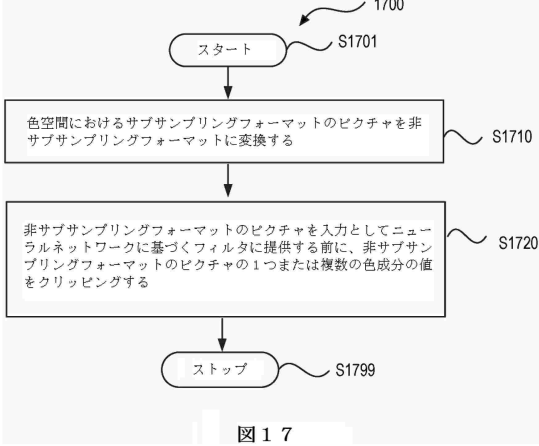


図 17

【図 18】

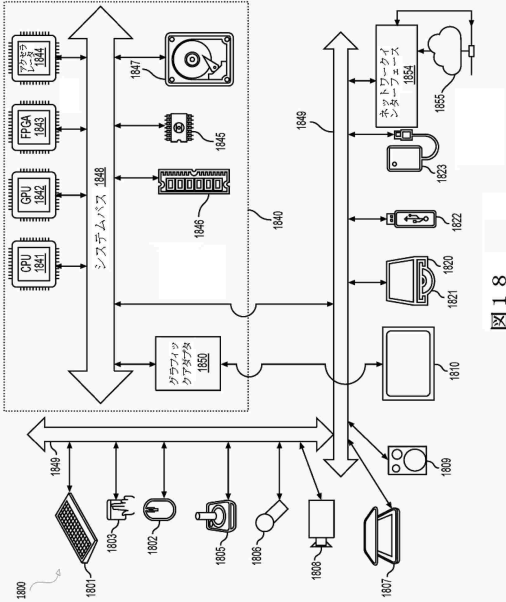


図 18

10

20

30

40

50

フロントページの続き

(33)優先権主張国・地域又は機関

米国(US)

エルシー内

(72)発明者 リー, シャン

アメリカ合衆国 9 4 3 0 6 カリフォルニア州 パロアルト パーク・ブールバード 2 7 4 7 テンセント アメリカ エルエルシー内

(72)発明者 リウ, シャン

アメリカ合衆国 9 4 3 0 6 カリフォルニア州 パロアルト パーク・ブールバード 2 7 4 7 テンセント アメリカ エルエルシー内

審査官 岩井 健二

(56)参考文献

米国特許出願公開第2020/0404335 (US, A1)

米国特許出願公開第2020/0252654 (US, A1)

Cheung Auyeung, et al., AHG11: A Case Study to Reduce Computation of Neural Network based In-Loop Filter by Pruning, Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29, VET-T0057_r3, 20th Meeting, by teleconference, 2020年10月, pp.1-9

Tsung-Chuan Ma, et al., AHG11: In-loop filtering based on neural network, Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29, JVET-T0094_r1, 20th Meeting, by teleconference, 2020年10月, pp.1-3

Zhao Wang, et al., [DNNVC] Preliminary results of Neural Network Loop Filter, Joint Video Experts Team (JVET) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29, JVET-T0128, 20th Meeting, by teleconference, 2020年10月, pp.1-3

(58)調査した分野 (Int.Cl., DB名)

H04N 19/00 - 19/98