

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6728074号
(P6728074)

(45) 発行日 令和2年7月22日(2020.7.22)

(24) 登録日 令和2年7月3日(2020.7.3)

(51) Int.Cl.	F 1
HO4N 19/70 (2014.01)	HO4N 19/70
HO4N 19/30 (2014.01)	HO4N 19/30
HO4N 19/597 (2014.01)	HO4N 19/597

請求項の数 15 (全 38 頁)

(21) 出願番号	特願2016-575195 (P2016-575195)
(86) (22) 出願日	平成27年6月25日 (2015.6.25)
(65) 公表番号	特表2017-525256 (P2017-525256A)
(43) 公表日	平成29年8月31日 (2017.8.31)
(86) 國際出願番号	PCT/US2015/037787
(87) 國際公開番号	W02015/200694
(87) 國際公開日	平成27年12月30日 (2015.12.30)
審査請求日	平成30年6月6日 (2018.6.6)
(31) 優先権主張番号	62/017,120
(32) 優先日	平成26年6月25日 (2014.6.25)
(33) 優先権主張国・地域又は機関	米国(US)
(31) 優先権主張番号	14/749,544
(32) 優先日	平成27年6月24日 (2015.6.24)
(33) 優先権主張国・地域又は機関	米国(US)

(73) 特許権者	595020643 クアルコム・インコーポレイテッド QUALCOMM INCORPORATED アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775
(74) 代理人	100108855 弁理士 蔵田 昌俊
(74) 代理人	100158805 弁理士 井関 守三
(74) 代理人	100112807 弁理士 岡田 貴志

最終頁に続く

(54) 【発明の名称】マルチレイヤビデオコーディング

(57) 【特許請求の範囲】

【請求項 1】

マルチレイヤHigh Efficiency Video Coding (HEVC)
)ビデオデータを符号化する方法であって、

アクセスユニットの第1のピクチャのために第1のビデオコーディングレイヤ(VCL)
)ネットワーク抽象化レイヤ(NAL)ユニットを生成すること、前記第1のVCL N
ALユニットは第1のスライスタイルを備える、と、

前記アクセスユニットの第2のピクチャのために第2のVCL NALユニットを生成
すること、前記第2のVCL NALユニットは第2のスライスタイルを備える、と、

前記第1のスライスタイルおよび前記第2のスライスタイルに基づいて前記アクセスユ
ニットに対するアクセスユニットデリミタ(AUD)NALユニットを生成すること、こ
こにおいて前記AUD NALユニットのピクチャタイプシンタックス要素は、前記アク
セスユニット中のすべてのピクチャに対するピクチャタイプを示し、前記ピクチャタイプ
は、複数のVCL NALユニットの各々によって備えられたスライスタイルを備えるス
ライスタイルの関連付けられたセットを有する、と、
を備える、方法。

【請求項 2】

前記複数のVCL NALユニットの各々によって備えられた前記スライスタイルが、
Iスライス、Pスライス、およびBスライスからなる群から選択される、請求項1に記載
の方法。

10

20

【請求項 3】

前記複数の VCL NAL ユニットの各々によって備えられた前記ライスタイルが、
Iスライスおよび Pスライスからなる群から選択される、請求項 1 に記載の方法。

【請求項 4】

前記 AUD NAL ユニットが、前記アクセスユニットと先行するアクセスユニットとの境界を特定する、請求項 1 に記載の方法。

【請求項 5】

ワイヤレス通信デバイス上で実行可能であり、前記ワイヤレス通信デバイスが、

前記マルチレイヤ High Efficiency Video Coding (HEVC) ビデオデータを記憶するように構成されるメモリと、10

前記メモリに記憶されている前記マルチレイヤ High Efficiency Video Coding (HEVC) ビデオデータを処理するための命令を実行するように構成される 1 つまたは複数のプロセッサと、

前記マルチレイヤ High Efficiency Video Coding (HEVC) ビデオデータを送信するように構成される送信機と

を備える、請求項 1 に記載の方法。

【請求項 6】

前記ワイヤレス通信デバイスが携帯電話であり、前記送信機が、セルラー通信規格に従って変調されたデータとして前記マルチレイヤ High Efficiency Video Coding (HEVC) ビデオデータを送信するように構成される、請求項 5 に記載の方法。20

【請求項 7】

High Efficiency Video Coding (HEVC) ビデオデータを処理するためのデバイスであって、

マルチレイヤビデオデータのビットストリームの少なくとも一部分を記憶するように構成されるメモリと、

1 つまたは複数のプロセッサとを備え、前記 1 つまたは複数のプロセッサが、

アクセスユニットの第 1 のピクチャのために第 1 のビデオコーディングレイヤ (VC-L) ネットワーク抽象化レイヤ (NAL) ユニットを生成すること、前記第 1 の VCL NAL ユニットは第 1 のライスタイルを備える、と、30

前記アクセスユニットの第 2 のピクチャのために第 2 の VCL NAL ユニットを生成すること、前記第 2 の VCL NAL ユニットは第 2 のライスタイルを備える、と、

前記第 1 のライスタイルおよび前記第 2 のライスタイルに基づいて前記アクセスユニットに対するアクセスユニットデリミタ (AUD) NAL ユニットを生成すること、ここにおいて前記 AUD NAL ユニットのピクチャタイプシングルタクス要素は、前記アクセスユニット中のすべてのピクチャに対するピクチャタイプを示し、前記ピクチャタイプは、複数の VCL NAL ユニットの各々によって備えられたライスタイルを備えるライスタイルの関連付けられたセットを有する、と、

を行うように構成される、デバイス。

【請求項 8】

前記複数の VCL NAL ユニットの各々によって備えられた前記ライスタイルが、Iスライス、Pスライス、および Bスライスからなる群から選択される、請求項 7 に記載のデバイス。

【請求項 9】

前記複数の VCL NAL ユニットの各々によって備えられた前記ライスタイルが、Iスライスおよび Pスライスからなる群から選択される、または前記 AUD NAL ユニットが、前記アクセスユニットと先行するアクセスユニットとの境界を特定する、請求項 7 に記載のデバイス。

【請求項 10】

ワイヤレス通信デバイスであり、50

前記マルチレイヤビデオデータを送信するように構成される送信機をさらに備える、請求項7に記載のデバイス。

【請求項11】

前記ワイヤレス通信デバイスが携帯電話を備え、前記送信機が、セルラー通信規格に従つて変調されたデータとして前記マルチレイヤビデオデータを送信するように構成される、請求項10に記載のデバイス。

【請求項12】

マルチレイヤHigh Efficiency Video Coding (HEVC)
)ビデオデータを符号化するための装置であつて、

アクセスユニットの第1のピクチャのために第1のビデオコーディングレイヤ (VCL)
)ネットワーク抽象化レイヤ (NAL) ユニットを生成するための手段、前記第1のVCL
NALユニットは第1のスライスタイルを備える、と、

前記アクセスユニットの第2のピクチャのために第2のVCL NALユニットを生成するための手段、前記第2のVCL NALユニットは第2のスライスタイルを備える、と、

前記第1のスライスタイルおよび前記第2のスライスタイルに基づいて前記アクセスユニットに対するアクセスユニットデリミタ (AUD) NALユニットを生成するための手段、ここにおいて前記AUD NALユニットのピクチャタイプシンタックス要素は、前記アクセスユニット中のすべてのピクチャに対するピクチャタイプを示し、前記ピクチャタイプは、複数のVCL NALユニットの各々によって備えられたスライスタイルを備えるスライスタイルの関連付けられたセットを有する、と、

を備える、装置。

【請求項13】

前記複数のVCL NALユニットの各々によって備えられた前記スライスタイルが、
Iスライス、Pスライス、およびBスライスからなる群から選択される、請求項12に記載の装置。

【請求項14】

前記複数のVCL NALユニットの各々によって備えられた前記スライスタイルが、
IスライスおよびPスライスからなる群から選択される、または前記AUD NALユニットが、前記アクセスユニットと先行するアクセスユニットとの境界を特定する、請求項12に記載の装置。

【請求項15】

1つまたは複数のプロセッサによって実行されると、前記1つまたは複数のプロセッサに、

アクセスユニットの第1のピクチャのために第1のビデオコーディングレイヤ (VCL)
)ネットワーク抽象化レイヤ (NAL) ユニットを生成すること、前記第1のVCL NAL
ユニットは第1のスライスタイルを備える、と、

前記アクセスユニットの第2のピクチャのために第2のVCL NALユニットを生成すること、前記第2のVCL NALユニットは第2のスライスタイルを備える、と、

前記第1のスライスタイルおよび前記第2のスライスタイルに基づいて前記アクセスユニットに対するアクセスユニットデリミタ (AUD) NALユニットを生成すること、ここにおいて前記AUD NALユニットのピクチャタイプシンタックス要素は、前記アクセスユニット中のすべてのピクチャに対するピクチャタイプを示し、前記ピクチャタイプは、複数のVCL NALユニットの各々によって備えられたスライスタイルを備えるスライスタイルの関連付けられたセットを有する、と、

を行わせる、High Efficiency Video Coding (HEVC)
)ビデオデータを符号化するための命令を記憶した、コンピュータ可読記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

[0001]本出願は、その内容全体が参照により本明細書に組み込まれる、2014年6月25日に出願された米国仮出願第62/017,120号の利益を主張する。

【0002】

[0002]本開示は、ビデオコーディングと、ビデオ圧縮と、ビットストリーム中の圧縮されたビデオと関連付けられるデータのシグナリングとに関する。

【背景技術】

【0003】

[0003]デジタルビデオ能力は、デジタルテレビジョン、デジタルダイレクトブロードキャストシステム、ワイヤレスブロードキャストシステム、携帯情報端末(PDA)、ラップトップコンピュータもしくはデスクトップコンピュータ、タブレットコンピュータ、電子ブックリーダー、デジタルカメラ、デジタル記録デバイス、デジタルメディアプレーヤ、ビデオゲームデバイス、ビデオゲームコンソール、セルラー電話もしくは衛星無線電話、いわゆる「スマートフォン」、ビデオ会議デバイス、ビデオストリーミングデバイスなどを含む、広範囲のデバイスに組み込まれ得る。デジタルビデオデバイスは、MPEG-2、MPEG-4、ITU-T H.263、ITU-T H.264/MPEG-4、Part 10、Advanced Video Coding (AVC)、High Efficiency Video Coding (HEVC) 規格によって定義された規格、およびそのような規格の拡張に記述されているビデオ圧縮技法などのビデオ圧縮技法を実装する。これらのビデオデバイスは、そのようなビデオ圧縮技法を実装することによって、デジタルビデオ情報を効率的に送信し、受信し、符号化し、復号し、および/または記憶し得る。

【0004】

[0004]ビデオ圧縮技法は、ビデオシーケンスに固有の冗長性を低減または除去するために空間的(イントラピクチャ)予測および/または時間的(インターピクチャ)予測を実行する。ブロックベースのビデオコーディングでは、ビデオスライス(すなわち、ビデオフレームまたはビデオフレームの一部)はビデオブロックに区分されてよく、これらのビデオブロックは、ツリーブロック、コーディングユニット(CU)および/またはコーディングノードと呼ばれることがある。ピクチャのイントラコーディングされた(I)スライス中のビデオブロックは、同じピクチャ中の隣接ブロック中の参照サンプルに対する空間予測を使用して符号化される。ピクチャのインターフォーマットコーディングされた(PまたはB)スライス中のビデオブロックは、同じピクチャ中の隣接ブロック中の参照サンプルに対する空間的予測、または他の参照ピクチャ中の参照サンプルに対する時間的予測を使用し得る。

【0005】

[0005]空間的予測または時間的予測は、コーディングされるべきブロックの予測ブロックをもたらす。残差データは、コーディングされるべき元のブロックと予測ブロックとの間のピクセル差分を表す。インターフォーマットコーディングされたブロックは、予測ブロックを形成する参照サンプルのブロックを指す動きベクトルに従って符号化され、残差データは、コーディングされたブロックと予測ブロックとの間の差分を示す。イントラコーディングされたブロックは、イントラコーディングモードと残差データとに従って符号化される。さらなる圧縮のために、残差データは、ピクセル領域から変換領域に変換されて残差変換係数をもたらすことができ、その残差変換係数が、次いで量子化され得る。最初に2次元アレイで構成される量子化された変換係数は、変換係数の1次元ベクトルを生成するために走査されてよく、なお一層の圧縮を達成するためにエントロピーコーディングが適用されてよい。

【発明の概要】

【0006】

[0006]本開示は、マルチレイヤビデオコーディングに関する技法を紹介し、より具体的には、独立した非ベースレイヤ(INBL: independent non-base layer)のための表現フォーマットを含むマルチレイヤビデオコーディングの態様に関する。本開示は、あるS

10

20

30

40

50

E I メッセージおよびツーアクセスユニットデリミタの解釈のために S P S が使用されるか P P S が使用されるかを決定するための技法を紹介する。

【 0 0 0 7 】

[0007]一例では、マルチレイヤビデオデータを符号化する方法は、アクセスユニットの第1のピクチャのための第1のビデオコーディングレイヤ (VCL) ネットワーク抽象化レイヤ (NAL) ユニットを生成することと、第1のVCL NALユニットは第1のスライスタイルを備える、アクセスユニットの第2のピクチャのための第2のVCL NALユニットを生成することと、第2のVCL NALユニットは第2のスライスタイルを備える、第1のスライスタイルおよび第2のスライスタイルに基づいて、アクセスユニットデリミタ (AUD) NALユニットを生成することとを含む。

10

【 0 0 0 8 】

[0008]別の例では、マルチレイヤビデオデータを符号化するための装置は、アクセスユニットの第1のピクチャのための第1のビデオコーディングレイヤ (VCL) ネットワーク抽象化レイヤ (NAL) ユニットを生成するための手段と、第1のVCL NALユニットは第1のスライスタイルを備える、アクセスユニットの第2のピクチャのための第2のVCL NALユニットを生成するための手段と、第2のVCL NALユニットは第2のスライスタイルを備える、第1のスライスタイルおよび第2のスライスタイルに基づいて、アクセスユニットデリミタ (AUD) NALユニットを生成するための手段とを含む。

20

【 0 0 0 9 】

[0009]別の例では、コンピュータ可読記憶媒体は、1つまたは複数のプロセッサによって実行されると、1つまたは複数のプロセッサに、アクセスユニットの第1のピクチャのための第1のビデオコーディングレイヤ (VCL) ネットワーク抽象化レイヤ (NAL) ユニットを生成することと、第1のVCL NALユニットは第1のスライスタイルを備える、アクセスユニットの第2のピクチャのための第2のVCL NALユニットを生成することと、第2のVCL NALユニットは第2のスライスタイルを備える、第1のスライスタイルおよび第2のスライスタイルに基づいて、アクセスユニットデリミタ (AUD) NALユニットを生成することとを行わせる命令を記憶する。

【 0 0 1 0 】

[0010]一例では、ビデオデータを処理するためのデバイスは、マルチレイヤビデオデータのビットストリームの少なくとも一部分を記憶するように構成されるメモリと、1つまたは複数のプロセッサとを含み、1つまたは複数のプロセッサは、アクセスユニットの第1のピクチャのための第1のビデオコーディングレイヤ (VCL) ネットワーク抽象化レイヤ (NAL) ユニットを生成することと、第1のVCL NALユニットは第1のスライスタイルを備える、アクセスユニットの第2のピクチャのための第2のVCL NALユニットを生成することと、第2のVCL NALユニットは第2のスライスタイルを備える、第1のスライスタイルおよび第2のスライスタイルに基づいて、アクセスユニットデリミタ (AUD) NALユニットを生成することとを行うように構成される。

30

【 0 0 1 1 】

[0011]本開示の1つまたは複数の例の詳細が添付の図面および以下の説明に記載される。他の特徴、目的、および利点は、説明、図面、および特許請求の範囲から明らかになろう。

40

【図面の簡単な説明】

【 0 0 1 2 】

【図1】[0012]本開示で説明される技法を利用し得る例示的なビデオ符号化および復号システムを示すブロック図。

【図2】[0013]本開示で説明される技法を実装し得る例示的なビデオエンコーダを示すブロック図。

【図3】[0014]本開示で説明される技法を実装し得る例示的なビデオデコーダを示すブロック図。

50

【図4】[0015]本開示の1つまたは複数の態様が実装され得る1つの例示的なネットワークを示すブロック図。

【図5】[0016]本開示の技法による例示的な方法を示すフローチャート。

【図6】[0017]本開示の技法による例示的な方法を示すフローチャート。

【図7】[0018]本開示の技法による例示的な方法を示すフローチャート。

【図8】[0019]本開示の技法による例示的な方法を示すフローチャート。

【発明を実施するための形態】

【0013】

[0020]本開示は、マルチレイヤビデオコーディングに関する技法を紹介し、より具体的には、独立した非ベースレイヤ (INBL : independent non-base layer) のための表現フォーマットを含むマルチレイヤビデオコーディングの態様に関する。本開示は、あるSEIメッセージおよびツーアクセスユニットデリミタの解釈のためにSPSが使用されるかPPSが使用されるかを決定するための技法を紹介する。10

【0014】

[0021]本開示は、マルチレイヤの状況においてHigh Efficiency Video Coding (HEVC) 規格において定義される補足強化情報 (SEI) メッセージを適用するための技法を含む。いくつかの事例では、以下で述べられるように、本技法は、HEVCに対するマルチビュービデオコーディング拡張 (MV-HEVC) またはHEVCに対するスケーラブルビデオコーディング (SVC) 拡張 (SHVC) のような、HEVC規格に対するマルチレイヤ拡張とともに実行され得る。本開示の技法は全般にHEVCの用語を使用して説明されるが、本開示の技法は、いかなる特定のビデオコーディング規格にも必ずしも限定されず、加えて、または代替的に、HEVCに対する他の拡張、他のマルチビューコーディング規格および/または他のマルチレイヤビデオコーディング規格とともに使用され得る。加えて、別段述べられない限り、以下で説明されるように、本開示の技法は独立にまたは組み合わせて適用され得ることを想定されたい。20

【0015】

[0022]ビデオデータの「レイヤ」は一般に、ビュー、解像度、忠実度、予備表現などのような少なくとも1つの共通の特性を有するピクチャのシーケンスを指し得る。たとえば、レイヤは、マルチビュービデオデータの特定のビュー（たとえば、視点）と関連付けられるビデオデータを含み得る。別の例として、レイヤは、スケーラブルビデオデータの特定のレイヤと関連付けられるビデオデータを含み得る。したがって、本開示は、ビデオデータのレイヤおよびビューを互換的に指すことがある。すなわち、ビデオデータのビューはビデオデータのレイヤと呼ばれることがあり、またはその逆のこともあり、複数のビューまたは複数のスケーラブルレイヤは、同じように、たとえばマルチレイヤコーディングシステムでは複数のレイヤと呼ばれることがある。加えて、（マルチレイヤビデオコーダまたはマルチレイヤエンコーダ - デコーダとも呼ばれる）マルチレイヤコーデックは、マルチビューコーデックまたはスケーラブルコーデック（たとえば、MV-HEVC、SHVC、または別のマルチレイヤコーディング技法を使用してビデオデータを符号化および/または復号するように構成されたコーデック）を指すことがある。30

【0016】

[0023]ピクチャはフレームまたはフィールドのいずれかであってよく、フィールドはフレームの1つおきの行（たとえば、偶数番目の行または奇数番目の行）を指し、フレームは偶数番目のフィールド（上フィールドとも呼ばれる）および奇数番目のフィールド（下フィールドとも呼ばれる）のような2つのフィールドの合成を指す。本開示は全般にピクチャまたはフレームに対する参照を伴う技法を説明するが、別段明示的に述べられない限り、説明される技法はフィールドにも適用可能であり得ることを理解されたい。40

【0017】

[0024]HEVC規格は一般に、すべてが特定の値のnuh_layer_idを有するネットワーク抽象化レイヤ (NAL) ユニットと、関連する非ビデオコーディングレイヤ (VCL) NALユニットとのセットとして、または階層的関係を有する構文的構造のセ50

ットのうちの1つとして、レイヤを定義する。HEVC規格は一般に、NALユニットに含まれるデータのタイプの指示と、ローバイトシーケンスペイロード(RBSP)の形態でそのデータを含むバイトとを含む、シンタックス構造として、NALユニットを定義する。シンタックス要素「`nuh_layer_id`」は、NALユニットが属するレイヤを識別する。

【0018】

[0025]マルチレイヤビットストリームは、たとえばSHVCではベースレイヤと1つまたは複数の非ベースレイヤを含むことがある、またはたとえばMV-HEVCでは複数のビューを含むことがある。スケーラブルビットストリームでは、ベースレイヤは通常、0に等しいレイヤ識別子(たとえば、`nuh_layer_id`)を有し得る。非ベースレイヤは、0よりも大きなレイヤ識別子を有してよく、ベースレイヤに含まれない追加のビデオデータを提供することができる。たとえば、マルチビュービデオデータの非ベースレイヤは、ビデオデータの追加のビューを含み得る。スケーラブルビデオデータの非ベースレイヤは、スケーラブルビデオデータの追加のレイヤを含み得る。非ベースレイヤは、交換可能にエンハンスメントレイヤと呼ばれ得る。

【0019】

[0026]マルチレイヤビットストリームのアクセスユニット(AUと省略されることがある)は、一般に、共通の時間インスタンスのためのすべてのレイヤ成分(たとえば、すべてのNALユニット)を含むデータのユニットである。アクセスユニットのレイヤ成分は通常、一緒に出力される(すなわち、実質的に同時に出力される)ことが意図されており、ここで、ピクチャを出力することは一般に、復号ピクチャバッファ(DPB)からピクチャを転送すること(たとえば、DPBから外部メモリにピクチャを記憶すること、DPBからディスプレイへピクチャを送ることなど)を伴う。SHVC拡張とMV-HEVC拡張とを含むHEVC規格は一般に、指定された分類規則に従って互いに関連付けられ、復号順序において連続的であり、`nuh_layer_id`の任意の特定の値を伴う多くとも1つのコーディングされたピクチャを含む、NALユニットのセットとして、アクセスユニットを定義する。コーディングされたピクチャのVCL_NALユニットを含むことに加えて、アクセスユニットは、非VCL_NALユニットも含み得る。アクセスユニットの復号は、`nuh_layer_id`の任意の特定の値を伴う多くとも1つの復号されたピクチャをもたらす。アクセスユニットのあるシーケンスは、コーディングされたビデオシーケンス(CVS:coded video sequence)と呼ばれ得る。

【0020】

[0027]ビデオデータの符号化された表現を含むビットストリームは、一連のNALユニットを含み得る。NALユニットは、VCL_NALユニットと非VCL_NALユニットとを含み得る。VCL_NALユニットは、ピクチャのコーディングされたスライスを含み得る。非VCL_NALユニットは、たとえば、ビデオパラメータセット(VPS)、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)、1つまたは複数のSEIメッセージ、または他のタイプのデータのような、他の情報をカプセル化し得る。

【0021】

[0028]ビットストリームのNALユニットは、ビットストリームの様々なレイヤと関連付けられ得る。SHVCでは、上で述べられたように、ベースレイヤ以外のレイヤは、「エンハンスメントレイヤ」と呼ばれてよく、ビデオデータの再生の品質を改善するデータを含み得る。MV-HEVCのようなマルチビューコーディングおよび3次元ビデオ(3DV)コーディングでは、レイヤは、異なるビューと関連付けられるデータを含み得る。ビットストリームの各レイヤは、異なるレイヤ識別子と関連付けられる。

【0022】

[0029]さらに、NALユニットは時間識別子を含み得る。ビットストリームの各動作点は、レイヤ識別子のセットと時間識別子とを有する。NALユニットが、ある動作点のためのレイヤ識別子のセット中のレイヤ識別子を指定し、NALユニットの時間識別子が動

10

20

30

40

50

作点の時間識別子以下である場合、N A L ユニットは、その動作点と関連付けられる。

【 0 0 2 3 】

[0030] H . 2 6 4 / A V C と H E V C の両方でサポートされる S E I 機構により、ビデオエンコーダは、出力ピクチャのサンプル値のビデオデコーダまたは他のデバイスによる正しい復号のために必要とされないが、ピクチャ出力タイミング、表示、ならびに損失の検出および補償のような様々な他の目的で使用され得るようなメタデータをビットストリームに含めることが可能になる。1つまたは複数の S E I メッセージをカプセル化する N A L ユニットは、本明細書では S E I N A L ユニットと呼ばれる。1つのタイプの S E I メッセージは、スケーラブルネスティング S E I メッセージである。スケーラブルネスティング S E I メッセージは、1つまたは複数の追加 S E I メッセージを含む S E I メッセージである。スケーラブルネスティング S E I メッセージは、S E I メッセージがマルチレイヤビットストリームの特定のレイヤに適用されるかまたは時間サブレイヤに適用されるかを示すために使用され得る。スケーラブルネスティング S E I メッセージ中に含まれない S E I メッセージは、本明細書では、ネストされていない S E I メッセージと呼ばれる。

【 0 0 2 4 】

[0031] ある特定のタイプの S E I メッセージは、特定の動作点にのみ適用可能な情報を含む。ビットストリームの動作点は、レイヤ識別子のセットおよび時間識別子と関連付けられる。動作点の表現は、動作点と関連付けられる各 N A L ユニットを含み得る。動作点の表現は、元のビットストリームとは異なるフレームレートおよび / またはビットレートを有することがある。これは、動作点の表現が元のビットストリームのいくつかのピクチャおよび / またはいくつかのデータを含まないことがあるからである。

【 0 0 2 5 】

[0032] 本開示はさらに、アクセスユニットデリミタ (A U D) N A L ユニットに関する技法を含む。H E V C 規格によれば、A U D N A L ユニットは、コーディングされたピクチャ中に存在するスライスのタイプを示すために、および、アクセスユニットの境界の検出を簡単にするために使用され得る。アクセスユニットデリミタと関連付けられる規範的な復号処理はない。

【 0 0 2 6 】

[0033] 以下でより詳細に説明されるように、ビデオデータの表現フォーマットを示すこと、ある S E I メッセージの解釈のために S P S が使用されるか P P S が使用されるかを決定すること、および A U D N A L ユニットの解釈は、シングルレイヤビデオ中に存在しないマルチレイヤビデオにおいてはいくつかの課題をもたらす。本開示は、それらの課題に対処し得る技法を紹介する。

【 0 0 2 7 】

[0034] 図 1 は、本開示で説明される技法を利用し得る例示的なビデオ符号化および復号システム 1 0 を示すブロック図である。システム 1 0 は、マルチレイヤビデオデータを符号化し、カプセル化し、送信し、カプセル化解除し、復号するように構成され得る。図 1 に示されるように、システム 1 0 は、宛先デバイス 1 4 によって後で復号されるべき符号化されたビデオデータを生成するソースデバイス 1 2 を含む。ソースデバイス 1 2 および宛先デバイス 1 4 は、デスクトップコンピュータ、ノートブック (すなわち、ラップトップ) コンピュータ、タブレットコンピュータ、セットトップボックス、いわゆる「スマート」フォンのようなワイヤレス / セルラー電話ハンドセット、いわゆる「スマート」パッド、テレビジョン、カメラ、ディスプレイデバイス、デジタルメディアプレーヤ、ビデオゲームコンソール、ビデオストリーミングデバイスなどを含む、広範囲のデバイスのいずれかを備え得る。いくつかの場合には、ソースデバイス 1 2 および宛先デバイス 1 4 は、ワイヤレス通信に対応し得る。

【 0 0 2 8 】

[0035] 宛先デバイス 1 4 は、リンク 1 6 を介して、復号されるべき符号化されたビデオデータを受信し得る。リンク 1 6 は、ソースデバイス 1 2 から宛先デバイス 1 4 に符号化

10

20

30

40

50

されたビデオデータを移動することが可能な任意のタイプの媒体またはデバイスを備え得る。一例では、リンク 16 は、ソースデバイス 12 が、符号化されたビデオデータをリアルタイムで宛先デバイス 14 に直接送信することを可能にするための通信媒体を備え得る。符号化されたビデオデータは、ワイヤレス通信プロトコルなどの通信規格に従って変調され、宛先デバイス 14 に送信され得る。通信媒体は、高周波 (R F) スペクトルまたは 10 1つもしくは複数の物理伝送線路のような、任意のワイヤレスまたは有線通信媒体を備え得る。通信媒体は、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、またはインターネットなどのグローバルネットワークなどのパケットベースのネットワークの一部を形成し得る。通信媒体は、ソースデバイス 12 から宛先デバイス 14 への通信を容易にするために有用であり得る、ルータ、スイッチ、基地局、または任意の他の機器を含み得る。通信媒体はまた、セルラーネットワークまたはモバイルネットワークの一部を形成することがあり、ソースデバイス 12 および宛先デバイス 14 は、G S M (登録商標) ネットワーク、C D M A ネットワーク、L T E (登録商標) ネットワーク、または他のそのようなネットワークのような、セルラー通信規格を使用して通信するように構成され得る。

【 0 0 2 9 】

[0036]代替的に、符号化されたデータは出力インターフェース 22 から記憶デバイス 32 に出力され得る。同様に、符号化されたデータは、入力インターフェースによって記憶デバイス 32 からアクセスされ得る。記憶デバイス 32 は、ハードドライブ、B l u - r a y (登録商標) ディスク、D V D、C D - R O M、フラッシュメモリ、揮発性メモリもしくは不揮発性メモリ、または、符号化されたビデオデータを記憶するための任意の他の好適なデジタル記憶媒体のような、種々の分散したまたはローカルでアクセスされるデータ記憶媒体のいずれかを含み得る。さらなる例では、記憶デバイス 32 は、ソースデバイス 12 によって生成された符号化されたビデオを保持し得るファイルサーバまたは別の中間記憶デバイスに対応し得る。宛先デバイス 14 は、ストリーミングまたはダウンロードを介して、記憶デバイス 32 から、記憶されたビデオデータにアクセスすることができる。ファイルサーバは、符号化されたビデオデータを記憶することと、その符号化されたビデオデータを宛先デバイス 14 に送信することが可能な任意のタイプのサーバであり得る。例示的なファイルサーバは、(たとえば、ウェブサイトのための) ウェブサーバ、F T P サーバ、ネットワークアタッチストレージ (N A S) デバイス、またはローカルディスクドライブを含む。宛先デバイス 14 は、インターネット接続を含む、任意の標準的なデータ接続を通じて符号化されたビデオデータにアクセスし得る。これは、ワイヤレスチャネル (たとえば、W i - F i (登録商標) 接続)、有線接続 (たとえば、D S L、ケーブルモデムなど)、または、ファイルサーバに記憶された符号化されたビデオデータにアクセスするのに適した、両方の組合せを含み得る。記憶デバイス 32 からの符号化されたビデオデータの送信は、ストリーミング送信、ダウンロード送信、または両方の組合せであり得る。

【 0 0 3 0 】

[0037]本開示の技法は、必ずしもワイヤレスの用途または設定に限定されるとは限らない。本技法は、オーバージエアテレビジョン放送、ケーブルテレビジョン送信、衛星テレビジョン送信、たとえばインターネットを介したストリーミングビデオ送信、データ記憶媒体に記憶するためのデジタルビデオの符号化、データ記憶媒体に記憶されたデジタルビデオの復号、または他の適用例のような、種々のマルチメディア適用例のいずれかをサポートするビデオコーディングに適用され得る。いくつかの例では、システム 10 は、ビデオストリーミング、ビデオ再生、ビデオブロードキャスト、および / またはビデオ電話のような適用例をサポートするために、一方向または双方向のビデオ送信をサポートするように構成され得る。

【 0 0 3 1 】

[0038]図 1 の例では、ソースデバイス 12 は、ビデオソース 18 と、ビデオエンコーダ 20 と、カプセル化ユニット 21 と、出力インターフェース 22 とを含む。いくつかの場合、出力インターフェース 22 は、変調器 / 復調器 (モデム) および / または送信機を含

10

20

30

40

50

み得る。ソースデバイス 12において、ビデオソース 18は、ビデオキャプチャデバイス、たとえばビデオカメラ、以前にキャプチャされたビデオを含んでいるビデオアーカイブ、ビデオコンテンツプロバイダからビデオを受信するためのビデオフィードインターフェース、および／またはソースビデオとしてコンピュータグラフィックスデータを生成するためのコンピュータグラフィックスシステムのようなソース、またはそのようなソースの組合せを含み得る。一例として、ビデオソース 18がビデオカメラである場合、ソースデバイス 12および宛先デバイス 14は、いわゆるカメラ電話またはビデオ電話を形成し得る。しかしながら、本開示で説明される技法は、ビデオコーディング全般に適用可能であることがあり、ワイヤレスおよび／または有線の適用例に適用され得る。

【0032】

10

[0039]キャプチャされたビデオ、以前にキャプチャされたビデオ、またはコンピュータにより生成されたビデオは、ビデオエンコーダ 20によって符号化され得る。カプセル化ユニット 21は、マルチメディアコンテンツの1つまたは複数の表現を形成することができ、ここで、表現の各々は1つまたは複数のレイヤを含み得る。いくつかの例では、ビデオエンコーダ 20は、たとえば、異なるフレームレート、異なるビットレート、異なる解像度、または他のそのような違いを伴う、異なる方法で各レイヤを符号化し得る。したがって、カプセル化ユニット 21は、様々な特性、たとえば、ビットレート、フレームレート、解像度などを有する様々な表現を形成し得る。

【0033】

20

[0040]表現の各々は、宛先デバイス 14によって取り出され得るそれぞれのビットストリームに対応し得る。カプセル化ユニット 21は、たとえば、マルチメディアコンテンツに対する media presentation description (MPD) データ構造内で、各表現に含まれるビューのビュー識別子 (view_id) の範囲の指示を与え得る。たとえば、カプセル化ユニット 21は、表現のビューに対する最大ビュー識別子と最小ビュー識別子との指示を与え得る。MPD はさらに、マルチメディアコンテンツの複数の表現の各々に対する出力を対象とするビューの最大の数の指示を与え得る。MPD またはそのデータは、いくつかの例では、(1つまたは複数の) 表現に対するマニフェストに記憶され得る。

【0034】

30

[0041]符号化されたビデオデータは、ソースデバイス 12の出力インターフェース 22を介して宛先デバイス 14に直接送信され得る。符号化されたビデオデータは、さらに(または代替的に)、復号および／または再生のための宛先デバイス 14または他のデバイスによる後のアクセスのために、記憶デバイス 32に記憶され得る。

【0035】

30

[0042]宛先デバイス 14は、入力インターフェース 28と、カプセル化解除ユニット 29と、ビデオデコーダ 30と、ディスプレイデバイス 31とを含む。いくつかの場合、入力インターフェース 28は、受信機および／またはモデムを含み得る。宛先デバイス 14の入力インターフェース 28は、リンク 16を通じて符号化されたビデオデータを受信する。リンク 16を通じて通信された、または記憶デバイス 32に提供された、符号化されたビデオデータは、ビデオデータを復号する際に、ビデオデコーダ 30のようなビデオデコーダが使用するための、ビデオエンコーダ 20によって生成された種々のシンタックス要素を含み得る。そのようなシンタックス要素は、通信媒体上で送信される、記憶媒体に記憶される、またはファイルサーバに記憶される、符号化されたビデオデータとともに含まれ得る。

40

【0036】

[0043]宛先デバイス 14のカプセル化解除ユニット 29は、ビットストリーム(またはマルチレイヤコーディングの状況では「動作点」と呼ばれるビットストリームのサブセット)からの SEI メッセージをカプセル化解除するユニットを表し得る。カプセル化解除ユニット 29は、カプセル化ユニット 21によって実行される動作とは逆の順序で動作を実行して、SEI メッセージのようなカプセル化された符号化されたビットストリームか

50

らのデータをカプセル化解除することができる。

【0037】

[0044]ディスプレイデバイス31は、宛先デバイス14と一緒にあってよく、またはその外部にあってよい。いくつかの例では、宛先デバイス14は、一体化されたディスプレイデバイスを含んでよく、また、外部ディスプレイデバイスとインターフェースするように構成されてよい。他の例では、宛先デバイス14はディスプレイデバイスであり得る。一般に、ディスプレイデバイス31は、復号されたビデオデータをユーザに表示し、液晶ディスプレイ(LCD)、プラズマディスプレイ、有機発光ダイオード(LED)ディスプレイ、または別のタイプのディスプレイデバイスのような、種々のディスプレイデバイスのいずれかを備え得る。

10

【0038】

[0045]ビデオエンコーダ20およびビデオデコーダ30は各々、1つまたは複数のマイクロプロセッサ、デジタル信号プロセッサ(DSP)、特定用途向け集積回路(ASIC)、フィールドプログラマブルゲートアレイ(FPGA)、ディスクリート論理、ソフトウェア、ハードウェア、ファームウェアまたはそれらの任意の組合せのような、種々の好適なエンコーダ回路のいずれかとして実装され得る。本技法が部分的にソフトウェアで実装されるとき、デバイスは、好適な非一時的コンピュータ可読媒体にソフトウェアの命令を記憶し、1つまたは複数のプロセッサを使用してその命令をハードウェアで実行して、本開示の技法を実行し得る。ビデオエンコーダ20およびビデオデコーダ30の各々は1つまたは複数のエンコーダまたはデコーダ中に含まれてよく、そのいずれもが、それぞれのデバイスにおいて複合エンコーダ/デコーダ(コーデック)の一部として統合され得る。

20

【0039】

[0046]図1には示されていないが、いくつかの態様では、ビデオエンコーダ20およびビデオデコーダ30は各々、オーディオエンコーダおよびデコーダと統合されてよく、共通のデータストリームまたは別個のデータストリーム中のオーディオとビデオの両方の符号化を扱うために、適切なMUX-DEMUXユニット、または他のハードウェアおよびソフトウェアを含み得る。適用可能な場合、いくつかの例では、MUX-DEMUXユニットは、ITU-H.223マルチプレクサプロトコル、またはユーザデータプロトコル(UDP)のような他のプロトコルに準拠し得る。

30

【0040】

[0047]本開示は全般に、ビデオエンコーダ20が、ある情報をビデオデコーダ30のような別のデバイスに「シグナリング」することに言及することがある。「シグナリング」という用語は、全般に、圧縮されたビデオデータを復号するために使用されるシンタックス要素および/または他のデータの通信を指し得る。そのような通信は、リアルタイムまたはほぼリアルタイムで発生し得る。代替的に、そのような通信は、符号化の時に符号化されたビットストリームの中でシンタックス要素をコンピュータ可読記憶媒体に記憶するときに発生し得るなど、ある時間の長さにわたって発生することがあり、これらの要素は次いで、この媒体に記憶された後の任意の時間に復号デバイスによって取り出され得る。

40

【0041】

[0048]いくつかの例では、ビデオエンコーダ20およびビデオデコーダ30は、そのスケーラブルビデオコーディング(SVC)拡張と、マルチビュービデオコーディング(MVC)拡張と、MVCベースの3DV拡張とを含む、ISO/IEC MPEG-4 Visualおよび(ISO/IEC MPEG-4 AVCとしても知られる)ITU-T H.264のようなビデオ圧縮規格に従って動作する。他の例では、ビデオエンコーダ20およびビデオデコーダ30は、Joint Collaboration Team on Video Coding (JCT-VC) of ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Motion Picture Experts Group (MPEG)によって開発されたHEVCに従って動作し得る。

50

【0042】

[0049]さらに、HEVCのためのスケーラブルビデオコーディング拡張と、マルチビューコーディング拡張と、3DV拡張とを作成する作業が進行中である。HEVCのスケーラブルビデオコーディング拡張は、SHVCと呼ばれることがある。SHVCの最近のワーキングドラフト(WD)(以後、SHVC WD5または現在のSHVC WDと呼ばれる)は、Chen他、「High Efficiency Video Coding (HEVC) scalable extension draft 5」、Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11、文書JCTVC-P1008_v4、第16回会合、サンノゼ、2014年1月に記載されている。MV-HEVCの最近のワーキングドラフト(WD)(以後、MV-HEVC WD7または現在のMV-HEVC WDと呼ばれる)は、Tech他、「MV-HEVC Draft Text 7」、Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11、文書JCTVC-G1004_v7、第16回会合、サンノゼ、2014年1月に記載されている。10

【0043】

[0050]HEVCおよび他のビデオコーディング規格では、ビデオシーケンスは一般に、一連のピクチャを含む。ピクチャは「フレーム」と呼ばれることがある。ピクチャは、 S_L 、 S_{Cb} および S_{Cr} と表される3つのサンプルアレイを含み得る。 S_L は、ルーマサンプルの2次元アレイ(すなわち、ブロック)である。 S_{Cb} は、Cbクロミナンスサンプルの2次元アレイである。 S_{Cr} は、Crクロミナンスサンプルの2次元アレイである。クロミナンスサンプルは、本明細書では「クロマ」サンプルと呼ばれることがある。他の事例では、ピクチャは、モノクロームであってよく、ルーマサンプルのアレイのみを含み得る。20

【0044】

[0051]ビデオエンコーダ20は、ピクチャの符号化された表現を生成するために、コーディングツリーユニット(CTU)のセットを生成することができる。CTUの各々は、ルーマサンプルのコーディングツリーブロックと、クロマサンプルの2つの対応するコーディングツリーブロックと、それらのコーディングツリーブロックのサンプルをコーディングするために使用されるシンタックス構造とを備え得る。モノクロームピクチャまたは3つの別々のカラープレーンを有するピクチャでは、CTUは、単一のコーディングツリーブロックと、そのコーディングツリーブロックのサンプルをコーディングするために使用されるシンタックス構造とを備え得る。コーディングツリーブロックは、サンプルの $N \times N$ のブロックであり得る。CTUは、「ツリーブロック」または「最大コーディングユニット」(LCU)と呼ばれることがある。HEVCのCTUは、H.264/AVCのような他の規格のマクロブロックに広い意味で類似し得る。しかしながら、CTUは、必ずしも特定のサイズに限定されるとは限らず、1つまたは複数のコーディングユニット(CU)を含み得る。スライスは、ラスタースキャン順序で連続的に順序付けられた整数個のCTUを含み得る。30

【0045】

[0052]コーディングされたCTUを生成するために、ビデオエンコーダ20は、コーディングツリーブロックをコーディングブロックに分割するために、CTUのコーディングツリーブロックに対して4分木区分を再帰的に実行することができ、したがって「コーディングツリーユニット」という名称である。コーディングブロックは、サンプルの $N \times N$ のブロックであり得る。CUは、ルーマサンプルアレイとCbサンプルアレイとCrサンプルアレイとを有するピクチャのルーマサンプルのコーディングブロックと、そのピクチャのクロマサンプルの2つの対応するコーディングブロックと、それらのコーディングブロックのサンプルをコーディングするために使用されるシンタックス構造とを備え得る。モノクロームピクチャまたは3つの別々のカラープレーンを有するピクチャでは、CUは40

、単一のコーディングブロックと、そのコーディングブロックのサンプルをコーディングするために使用されるシンタックス構造とを備え得る。

【0046】

[0053]ビデオエンコーダ20は、CUのコーディングブロックを1つまたは複数の予測ブロックに区分し得る。予測ブロックは、同じ予測が適用されるサンプルの矩形（すなわち、正方形または非正方形）ブロックである。CUの予測ユニット（PU）は、ルーマサンプルの予測ブロックと、クロマサンプルの2つの対応する予測ブロックと、それらの予測ブロックを予測するために使用されるシンタックス構造とを備え得る。モノクロームピクチャまたは3つの別個のカラーブレーンを有するピクチャでは、PUは、単一の予測ブロックと、その予測ブロックを予測するために使用されるシンタックス構造とを備え得る。ビデオエンコーダ20は、CUの各PUのルーマ予測ブロック、Cb予測ブロック、およびCr予測ブロックのための、予測ルーマブロックと、予測Cbブロックと、予測Crブロックとを生成することができる。10

【0047】

[0054]ビデオエンコーダ20は、PUの予測ブロックを生成するためにイントラ予測またはインター予測を使用することができる。ビデオエンコーダ20がPUの予測ブロックを生成するためにイントラ予測を使用する場合、ビデオエンコーダ20は、PUと関連付けられたピクチャの復号されたサンプルに基づいてPUの予測ブロックを生成することができる。ビデオエンコーダ20がPUの予測ブロックを生成するためにインター予測を使用する場合、ビデオエンコーダ20は、PUと関連付けられるピクチャ以外の1つまたは複数のピクチャの復号されたサンプルに基づいて、PUの予測ブロックを生成することができる。20

【0048】

[0055]ビデオエンコーダ20がCUの1つまたは複数のPUのための予測ルーマブロックと、予測Cbブロックと、予測Crブロックとを生成した後、ビデオエンコーダ20は、CUのためのルーマ残差ブロックを生成することができる。CUのルーマ残差ブロック中の各サンプルは、CUの予測ルーマブロックの1つの中のルーマサンプルとCUの元のルーマコーディングブロック中の対応するサンプルとの間の差分を示す。加えて、ビデオエンコーダ20は、CUのためのCb残差ブロックを生成することができる。CUのCb残差ブロック中の各サンプルは、CUの予測Cbブロックの1つの中のCbサンプルと、CUの元のCbコーディングブロック中の対応するサンプルとの間の差を示し得る。ビデオエンコーダ20は、CUのためのCr残差ブロックを生成することもできる。CUのCr残差ブロック中の各サンプルは、CUの予測Crブロックの1つの中のCrサンプルと、CUの元のCrコーディングブロック中の対応するサンプルとの間の差分を示し得る。30

【0049】

[0056]さらに、ビデオエンコーダ20は、CUのルーマ残差ブロックと、Cb残差ブロックと、Cr残差ブロックとを、1つまたは複数のルーマ変換ブロック、Cb変換ブロック、およびCr変換ブロックに分解するために、4分木区分を使用し得る。変換ブロックは、同じ変換が適用されるサンプルの矩形（たとえば、正方形または非正方形）ブロックである。CUの変換ユニット（TU）は、ルーマサンプルの変換ブロックと、クロマサンプルの2個の対応する変換ブロックと、それらの変換ブロックサンプルを変換するために使用されるシンタックス構造とを備え得る。したがって、CUの各TUは、ルーマ変換ブロック、Cb変換ブロック、およびCr変換ブロックと関連付けられ得る。TUと関連付けられるルーマ変換ブロックは、CUのルーマ残差ブロックのサブブロックであってよい。Cb変換ブロックは、CUのCb残差ブロックのサブブロックであってよい。Cr変換ブロックは、CUのCr残差ブロックのサブブロックであってよい。モノクロームピクチャまたは3つの別個のカラーブレーンを有するピクチャでは、TUは、単一の変換ブロックと、その変換ブロックのサンプルを変換するために使用されるシンタックス構造とを備え得る。40

【0050】

10

20

30

40

50

[0057]ビデオエンコーダ20は、TUのルーマ係数ブロックを生成するために、TUのルーマ変換ブロックに1回または複数回の変換を適用することができる。係数ブロックは、変換係数の2次元アレイであり得る。変換係数は、スカラー量であってよい。ビデオエンコーダ20は、TUのためのCb係数ブロックを生成するために、TUのCb変換ブロックに1回または複数回の変換を適用することができる。ビデオエンコーダ20は、TUのためのCr係数ブロックを生成するために、TUのCr変換ブロックに1回または複数回の変換を適用することができる。

【0051】

[0058]ビデオエンコーダ20は、係数ブロック（たとえば、ルーマ係数ブロック、Cb係数ブロック、またはCr係数ブロック）を生成した後、係数ブロックを量子化することができる。量子化は一般に、変換係数を表すために使用されるデータの量をできるだけ低減するために変換係数が量子化されさらなる圧縮を実現する、処理を指す。ビデオエンコーダ20が係数ブロックを量子化した後、ビデオエンコーダ20は、量子化された変換係数を示すシンタックス要素をエントロピー符号化し得る。たとえば、ビデオエンコーダ20は、量子化された変換係数を示すシンタックス要素に対してコンテキスト適応型バイナリ算術コーディング（CABAC）を実行し得る。

【0052】

[0059]ビデオエンコーダ20は、コーディングされたピクチャおよび関連するデータの表現を形成するビットのシーケンスを含むビットストリームを出力し得る。ビットストリームは、一連のNALユニットを備え得る。NALユニットは、NALユニット中のデータのタイプの指示と、必要に応じてエミュレーション防止ビットが点在しているRBSPの形態でそのデータを含むバイトとを含む、シンタックス構造である。NALユニットの各々は、NALユニットヘッダを含み、RBSPをカプセル化する。NALユニットヘッダは、NALユニットタイプコードを指示するシンタックス要素を含み得る。NALユニットのNALユニットヘッダによって指定されるNALユニットタイプコードは、NALユニットのタイプを示す。RBSPは、NALユニット内にカプセル化された整数個のバイトを含むシンタックス構造であり得る。いくつかの例では、RBSPは0ビットを含む。

【0053】

[0060]異なるタイプのNALユニットは、異なるタイプのRBSPをカプセル化し得る。たとえば、第1のタイプのNALユニットはPPSのためのRBSPをカプセル化することができ、第2のタイプのNALユニットはコーディングされたスライスのためのRBSPをカプセル化することができ、第3のタイプのNALユニットはSEIメッセージのためのRBSPをカプセル化することができ、以下同様である。ビデオコーディングデータのためのRBSPをカプセル化するNALユニットは（パラメータセットおよびSEIメッセージのためのRBSPとは対照的に）、VCL_NALユニットと呼ばれることがある。

【0054】

[0061]ビデオデコーダ30は、ビデオエンコーダ20によって生成されたビットストリームを受信し得る。加えて、ビデオデコーダ30は、ビットストリームからシンタックス要素を取得するために、ビットストリームを構文解析し得る。ビデオデコーダ30は、ビットストリームから取得されたシンタックス要素に少なくとも一部基づいて、ビデオデータのピクチャを再構築することができる。ビデオデータを再構築するための処理は全般に、ビデオエンコーダ20によって実行される処理の逆であり得る。加えて、ビデオデコーダ30は、現在のCUのTUと関連付けられる係数ブロックを逆量子化し得る。ビデオデコーダ30は、現在のCUのTUと関連付けられる変換ブロックを再構築するために、係数ブロックに対して逆変換を実行し得る。ビデオデコーダ30は、現在のCUのPUのための予測ブロックのサンプルを現在のCUのTUの変換ブロックの対応するサンプルに追加することによって、現在のCUのコーディングブロックを再構築することができる。ピクチャの各CUのためのコーディングブロックを再構築することによって、ビデオデコ-

10

20

30

40

50

ダ30はピクチャを再構築することができる。

【0055】

[0062]マルチビューコーディングでは、異なる視点からの同じシーンの複数のビューが存在することがある。上で述べられたように、アクセスユニットは、同じ時間インスタンスに対応するピクチャのセットを含む。したがって、ビデオデータは、時間とともに生じる一連のアクセスユニットとして概念化され得る。「ビュー成分」は、単一のアクセスユニット中のビューのコーディングされた表現であり得る。本開示では、「ビュー」は、同じビュー識別子と関連付けられる一連のビュー成分を指し得る。例示的なタイプのビュー成分は、テクスチャビュー成分と深度ビュー成分とを含み得る。

【0056】

[0063]マルチビューコーディングは、ビュー間予測をサポートする。ビュー間予測は、HEVCにおいて使用されるインター予測と同様であり、同じシンタックス要素を使用し得る。しかしながら、ビデオコーダが現在のビデオユニット(P U のような)に対してビュー間予測を実行するとき、ビデオエンコーダ20は、参照ピクチャとして、その現在のビデオユニットと同じアクセスユニット中にあるが異なるビューの中にあるピクチャを使用し得る。対照的に、従来のインター予測は、参照ピクチャとして異なるアクセスユニット内のピクチャのみを使用する。

【0057】

[0064]マルチビューコーディングでは、ビデオデコーダ(たとえば、ビデオデコーダ30)が、あるビュー中のピクチャを、任意の他のビュー中のピクチャを参照せずに復号することができる場合、そのビューは「ベースビュー」と呼ばれ得る。非ベースビューの1つの中のピクチャをコーディングするとき、あるピクチャが異なるビュー中にあるが、ビデオコーダが現在コーディングしているピクチャと同じ時間インスタンス(すなわち、アクセスユニット)内にある場合、(ビデオエンコーダ20またはビデオデコーダ30のよう)ビデオコーダは、参照ピクチャリストにそのピクチャを追加することができる。他のインター予測参照ピクチャと同様に、ビデオコーダは、参照ピクチャリストの任意の位置にビュー間予測参照ピクチャを挿入することができる。

【0058】

[0065]H.264/AVCとHEVCの両方においてサポートされるSEI機構は、ビデオエンコーダ(たとえば、ビデオエンコーダ20)が、出力ピクチャのサンプル値の正確な復号のためには要求されないが、ピクチャ出力タイミング、表示、ならびに損失の検出および補償のような他の様々な目的で使用され得るようなメタデータをビットストリームに含めることを可能にする。ビデオエンコーダ20は、ピクチャのサンプル値の正確な復号に必要ではないメタデータをビットストリームに含めるために、SEIメッセージを使用することができる。しかしながら、ビデオデコーダ30または他のデバイスは、SEIメッセージに含まれるメタデータを様々な他の目的で使用することができる。たとえば、ビデオデコーダ30または別のデバイスは、ピクチャ出力タイミング、ピクチャ表示、損失検出、および誤り補償のために、SEIメッセージ中のメタデータを使用することができる。

【0059】

[0066]ビデオエンコーダ20は、アクセスユニットに含めるための1つまたは複数のSEI_NALユニットを生成することができる。言い換えれば、任意の数のSEI_NALユニットがアクセスユニットと関連付けられることがある。さらに、各SEI_NALユニットは、1つまたは複数のSEIメッセージを含み得る。つまり、ビデオエンコーダは、任意の数のSEI_NALユニットをアクセスユニットに含めることができ、各SEI_NALユニットは、1つまたは複数のSEIメッセージを含み得る。SEI_NALユニットは、NALユニットヘッダとペイロードとを含み得る。SEI_NALユニットのNALユニットヘッダは、少なくとも第1のシンタックス要素と第2のシンタックス要素とを含む。第1のシンタックス要素は、SEI_NALユニットのレイヤ識別子を指定する。第2のシンタックス要素は、SEI_NALユニットの時間識別子を指定する。

10

20

30

40

50

【 0 0 6 0 】

[0067] ネストされた S E I メッセージは、スケーラブルネスティング S E I メッセージに含まれる S E I メッセージを指す。ネストされていない S E I メッセージは、スケーラブルネスティング S E I メッセージに含まれない S E I メッセージを指す。 S E I N A L ユニットのペイロードは、ネストされた S E I メッセージまたはネストされていない S E I メッセージを備え得る。

【 0 0 6 1 】

[0068] H E V C 規格は、様々なタイプの S E I メッセージのためのシンタックスとセマンティクスとを記述する。しかしながら、 S E I メッセージが規範となる復号処理に影響を及ぼさないので、 H E V C 規格は S E I メッセージの扱いを記述しない。 H E V C 規格に S E I メッセージを有する 1 つの理由は、補足データが H E V C を使用する様々なシステムにおいて同じように解釈されることを可能にするためである。 H E V C を使用する規格およびシステムは、ビデオエンコーダに、いくつかの S E I メッセージを生成するよう requirement 10 に要求することがあり、または特定のタイプの受信された S E I メッセージの固有の扱いを定義することがある。

【 0 0 6 2 】

[0069] 以下の表 1 は、 H E V C において規定される S E I メッセージを列挙し、それらの目的を簡潔に記述している。

【 0 0 6 3 】

【表1】

表1—SEIメッセージの概要

SEIメッセージ	目的
パッファリング期間	仮想参照デコーダ(HRD)動作のための初期遅延
ピクチャタイミング	HRD動作のためのピクチャ出力時間およびピクチャ/サブピクチャ削除時間、ならびにピクチャ構造関連の情報
パンスキヤン方形	出力ピクチャのピクチャアスペクト比(PAR)とは異なるPARでの表示
フィラーペイロード	固有の制約を満たすためのビットレートの調整
ユーザデータ登録 ユーザデータ未登録	外部エンティティによって指定されるべきSEIメッセージ
復元点	クリーンランダムアクセスのための追加の情報。 漸次復号リフレッシュ。
シーン情報	シーンの変更および遷移についての情報
フルフレーム スナップショット	関連する復号されたピクチャをビデオコンテンツの静止画像スナップショットとして標識するための指示
漸進的 精緻化セグメント	動いているシーンではなくピクチャの品質の漸進的な精緻化をいくつかの連続的なピクチャが表すことを示す
フィルムグレイン特性	デコーダがフィルムグレインを合成することを可能にする
デブロッキング フィルタ表示選好	表示されるピクチャがループ内デブロッキングフィルタ処理を経るべきかどうかを推奨する

【0064】

ポストフィルタヒント	ポストフィルタ設計のための提案されるポストフィルタ係数または相関情報を提供する	
トーンマッピング情報	符号化において使用または仮定される別の色空間への再マッピング	
フレームパッキング構成	HEVCビットストリームへの立体視ビデオのパック	10
表示方向	出力ピクチャが表示されるときに出力ピクチャに適用されるべきフリップおよび／または回転を指定する	
ピクチャの構造の記述	ビットストリームの時間的なインター予測構造を記述する	
復号ピクチャハッシュ	誤り検出のために使用され得る復号されるピクチャのチェックサム	
有効パラメータセット	有効VPS、SPSなどについての情報を提供する	
復号ユニット情報	HRD動作のためのサブピクチャ削除時間、ならびに復号ユニットインデックス	20
時間的レベル0インデックス	時間的レベル0インデックスの値を提供する	
スケーラブルネスティング	異なる動作点およびレイヤへの関連付けのためにSEIメッセージをネストするための機構を提供する	
領域リフレッシュ情報	漸次的復号リフレッシュのためのリフレッシュされる領域およびリフレッシュされない領域についての情報を提供する	30

【0065】

[0070]上で紹介されたように、本開示は、マルチレイヤビデオコーディングに関する技法、より具体的には、独立の非ベースレイヤ(INBL)の表現フォーマットと、あるSEIメッセージの解釈のためにSPSが使用されるかPPSが使用されるかということと、AUDINALユニットの処理とを含む、マルチレイヤビデオコーディングの態様に関するツー技法を紹介する。

【0066】

[0071]ビデオエンコーダ20はマルチレイヤビデオデータを生成するように構成されてよく、ビデオデコーダ30はマルチレイヤビデオデータを復号するように構成されてよい。マルチレイヤビデオデータは、ベースレイヤと1つまたは複数の非ベースレイヤとを含み得る。非ベースレイヤは、復号のために他のレイヤに従属する従属ベースレイヤと、復号のために他のレイヤに従属しないINBLの両方を含み得る。マルチレイヤビデオコーディングにおけるINBLの既存の実装形態には、いくつかの潜在的な問題がある。

【0067】

[0072]既存の実装形態についての潜在的な問題の1つの例として、0よりも大きなレイヤ識別子(たとえば、n_u_h_l_a_y_e_r_i_d)を伴うレイヤは、そのレイヤがINBLであるかどうかとは無関係に、VPSにおいてシグナリングされる表現フォーマットを使用する。表現フォーマットは、幅、高さ、ビット深度、およびカラーフォーマットのようなパラメータを含む。1に等しいV1ComptibleSPSFlagまたは0に等しいn_u_h_l_a_y_e_r_i_dを伴うSPSを参照するINBLが、たとえばJCTV

C - R 0 0 1 0 v 3 / J C T 3 V - I 0 0 1 0 v 3 (h t t p : / / p h e n i x . i n t - e v r y . f r / j c t / d o c _ e n d _ u s e r / d o c u m e n t s / 1 8 _ S a p p o r o / w g 1 1 / J C T V C - R 0 0 1 0 - v 3 . z i p) の A H G 1 0 成果テキストにおいて規定されるような、 H E V C バージョン 1 に適合するベースレイヤとなるように書き換えられることになる場合、 V P S からの使用される表現フォーマットが S P S においてシグナリングされる（パラメータのいずれに対する）表現フォーマットとも異なるとき、書換え処理は、表現フォーマットが V P S からの使用される表現フォーマットと同じとなるように、 S P S を変更する必要がある。この要件は、 S P S 全体の書換えを必要とすることがある、このことは、 H E V C バージョン 1 に適合するベースレイヤへの I N B L の書換え処理を極めて複雑にし得る。

10

【 0 0 6 8 】

[0073] V 1 C o m p a t i b l e S P S F l a g が 1 に等しいとき、 S P S は、 H E V C バージョン 1 において規定される S P S シンタックスに適合しており、 H E V C バージョン 1 に従って実装されるレガシーの H E V C デコーダによって構文解析され得る。 V 1 C o m p a t i b l e S P S F l a g が 0 に等しいとき、 S P S は、 H E V C バージョン 1 において規定される S P S シンタックスに適合しておらず、 H E V C バージョン 1 に従って実装されるレガシーの H E V C デコーダによって構文解析され得ない。

【 0 0 6 9 】

[0074] 本開示は、上で説明された問題に対処することができるいくつかの技法を紹介する。一例として、ビデオエンコーダ 2 0 は、 V P S からの使用される表現フォーマットおよび I N B L のために S P S においてシグナリングされる表現フォーマットが同一となるように、マルチレイヤビデオを符号化するように構成され得る。加えて、または代替的に、 I N B L しか含まれないビットストリーム区分に含まれるその I N B L に対して、指定される表現フォーマットは、有効 S P S においてシグナリングされる表現フォーマットである。加えて、または代替的に、任意の I N B L に対して、指定される表現フォーマットは、レイヤのための有効 S P S においてシグナリングされる表現フォーマットであり得る。

20

【 0 0 7 0 】

[0075] マルチレイヤビデオの既存の実装形態についての潜在的な問題の別の例として、 S E I メッセージは、たとえば S E I メッセージがネストされるとき、複数のレイヤ、または、複数のレイヤを含む（出力）レイヤセットと関連付けられる（出力）動作点に適用されることがある。そのような事例では、複数の有効 S P S および複数の有効 P P S が存在することがある。たとえば、各レイヤに対して有効 S P S および有効 P P S が存在することがあり、このことは、いくつかの S E I メッセージがどの P P S または S P S に適用されるかを不明確にすることがある。

30

【 0 0 7 1 】

[0076] たとえば、フラグ「 g e n e r a l _ p r o g r e s s i v e _ s o u r c e _ f l a g 」および「 g e n e r a l _ i n t e r l a c e d _ s o u r c e _ f l a g 」が、フレームフィールド情報 S E I メッセージにおける f f i n f o _ s o u r c e _ s c a n _ t y p e のセマンティクスにおいて参照される。これらの 2 つのフラグは、プロファイル、層、およびレベル（ P T L : profile, tier and level ）のシンタックス構造の中に位置し、このシンタックス構造は有効 S P S の各々の中にあることがあり、フラグの複数のインスタンスは V P S の中にも存在することがある。したがって、フレームフィールド情報 S E I メッセージが複数のレイヤに適用されるとき、 2 つのフラグを含むどのシンタックス構造が適用されるかが明確にされるべきである。フレームフィールド情報 S E I メッセージにおいて f f i n f o _ p i c _ s t r u c t および f f i n f o _ d u p l i c a t e _ f l a g のセマンティクスで使用される S P S フラグ「 f i e l d _ s e q _ f l a g 」について、同様の問題が存在する。

40

【 0 0 7 2 】

[0077] 上で説明された問題をより高い確率で解決するために、本開示は、フレームフィールド情報 S E I メッセージが適用されるレイヤのためのすべての有効 S P S に対して f

50

`field_seq_flag` の値が同じであることを要求し得る、符号化の制約を導入する。加えて、または代替的に、本開示は、`field_seq_flag` の値がレイヤのためのすべてのアクティブ SPS に対して同じではない場合に、レイヤのセットに適用されるフレームフィールド情報 SEI メッセージが存在しないことを要求し得る、符号化の制約を導入する。HEVC では、1 に等しい `field_seq_flag` は、フィールドを表すピクチャを CVS が伝えることを示し、ピクチャタイミング SEI メッセージが現在の CVS の各アクセスユニット中に存在すべきであることを規定する。HEVC では、0 に等しい `field_seq_flag` は、フレームを表すピクチャを CVS が伝えることと、ピクチャタイミング SEI メッセージが現在の CVS の任意のアクセスユニット中に存在することもまたはしないこともあることを示す。

10

【0073】

[0078] 同様に、フラグ「`general_progressive_source_flag`」および「`general_interlaced_source_flag`」に対して、本開示はそれぞれ、フレームフィールド情報 SEI メッセージが適用されるレイヤを含むビットストリーム区分に適用されるすべての PTL シンタックス構造に対して、フラグが同一であることを要求し得る符号化の制約を導入する。HEVC によれば、`general_progressive_source_flag` および `general_interlaced_source_flag` は、次のように解釈される。

- `general_progressive_source_flag` が 1 に等しく、`general_interlaced_source_flag` が 0 に等しい場合、CVS におけるピクチャのソーススキャンタイプは、プログレッシブのみであると解釈されるべきである。

20

- そうではなく、`general_progressive_source_flag` が 0 に等しく、`general_interlaced_source_flag` が 1 に等しい場合、CVS におけるピクチャのソーススキャンタイプは、インターレースのみであると解釈されるべきである。

- そうではなく、`general_progressive_source_flag` が 0 に等しく、`general_interlaced_source_flag` が 0 に等しい場合、CVS におけるピクチャのソーススキャンタイプは、未知または未指定であると解釈されるべきである。

30

【0074】

[0079] それ以外の場合 (`general_progressive_source_flag` が 1 に等しく、`general_interlaced_source_flag` が 1 に等しい)、CVS の中の各ピクチャのソーススキャンタイプは、ピクチャタイミング SEI メッセージにおいてシンタックス要素 `source_scan_type` を使用してピクチャレベルで示される。そのような制約は、他の SEI メッセージのシンタックス要素にも適用され得る。複数のレイヤまたは複数の（出力）レイヤセットに適用される任意の SEI メッセージ `seiA` に対して、`seiA` の一部ではない任意のシンタックス要素の複数のインスタンスが `seiA` の一部である任意のシンタックス要素のセマンティクスに関与するとき、`seiA` の一部ではないシンタックス要素の値はすべてのインスタンスに対して同じであることが要求される。したがって、ビデオエンコーダ 20 が、SEI メッセージがマルチレイヤビデオデータの 2 つ以上のレイヤに適用されることを決定し、マルチレイヤビデオデータの 2 つ以上のレイヤと関連付けられる複数のインスタンスを有するシンタックス要素を SEI メッセージが参照することを決定する場合、ビデオエンコーダ 20 は、複数のインスタンスのすべてを同じ値に設定する。シンタックス要素は、`field_seq_flag` シンタックス要素、`general_progressive_source_flag` シンタックス要素、`general_interlaced_source_flag` シンタックス要素、または本明細書で説明される任意の他のそのようなシンタックス要素のいずれであってもよい。

40

【0075】

50

[0080]上で紹介されたように、本開示はまた、AUD_NALユニットに関するいくつかの技法も紹介し、AUD_NALユニットは、現在のように実装されると、マルチレイヤビデオとともに使用されるときにいくつかの潜在的な欠点を有する。AUD_NALユニットにおけるシンタックス要素「pic_type」のセマンティクスは、マルチレイヤビットストリームのようなマルチレイヤの状況では明確ではないことがある。マルチレイヤの状況におけるアクセスユニットは、1つまたは複数のコーディングされたピクチャを含み得るが、pic_typeのセマンティクスは、「コーディングされたピクチャ」の中に存在するスライスのタイプを示すためにpic_typeが使用され得ることを示している。したがって、アクセスユニットの中に複数のピクチャが存在する可能性がある、マルチレイヤビデオの状況では、現在のセマンティクスは不明確である。

10

【0076】

[0081]この潜在的な欠点に対処するために、本開示は、AUD_NALユニットのセマンティクスを次のように変更することを提案する。アクセスユニットデリミタのRBSのセマンティクスが次のように変更される（下線付きのテキストが追加され、[[プラットテキスト]]が削除される）。

アクセスユニットデリミタは、[[a]]アクセスユニットデリミタNALユニットを含むアクセスユニット中のすべてのコーディングされたピクチャの中に存在するスライスのタイプを示し、アクセスユニットの境界の検出を簡単にするために使用され得る。アクセスユニットデリミタと関連付けられる規範的な復号処理はない。

pic_typeは、アクセスユニットデリミタNALユニットを含むアクセスユニット中のコーディングされたピクチャのすべてのスライスに対するslice_typeの値がpic_typeの所与の値に対して表7-2において列挙されるセットの一員であることを示す。pic_typeの値は、本規格のこのバージョンに適合するビットストリームにおいては0、1、または2に等しいものとする。pic_typeの他の値は、ITU-T|ISO/IECによる将来の使用のために確保されている。本規格のこのバージョンに適合するデコーダは、pic_typeの確保されている値を無視するものとする。pic_typeの値は、本規格のこのバージョンに適合するビットストリームにおいては0、1、または2に等しいものとする。pic_typeの他の値は、ITU-T|ISO/IECによる将来の使用のために確保されている。本規格のこのバージョンに適合するデコーダは、pic_typeの確保されている値を無視するものとする。

20

【0077】

【表2】

表7-2 pic_typeの解釈

pic_type	コーディングされたピクチャの中に存在し得るslice_typeの値
0	I
1	P, I
2	B, P, I

40

【0078】

[0082]代替的に、pic_typeのセマンティクスは、pic_typeの値が、アクセスユニットデリミタNALユニットのnuh_layer_idに等しいnuh_layer_idを伴うアクセスユニット中のコーディングされたピクチャのすべてのスライスに対するslice_typeの値がpic_typeの所与の値に対して表7-2において列挙されるセットの一員であることを示すように、修正され得る。

【0079】

[0083]マルチレイヤビデオコーディングの既存の実装形態には、0以外の数に等しいn

50

`u h _ l a y e r _ i d` を伴う A U D N A L ユニットを処理することに関するいくつかの潜在的な問題もある。たとえば、 H E V C バージョン 1 の規格には、次の制約がある。アクセスユニットデリミタ N A L ユニットが存在するとき、それは第 1 の N A L ユニットであるものとする。どのアクセスユニット中にも多くとも 1 つのアクセスユニットデリミタ N A L ユニットがあるものとする。

【 0 0 8 0 】

[0084]この制約は、 H E V C バージョン 1 の規格に当てはまる。しかしながら、マルチレイヤビットストリームが 0 よりも大きな `n u h _ l a y e r _ i d` を伴う A U D N A L ユニットを含むとき、そのような A U D N A L ユニットは、バージョン 1 のデコーダにより、新たなアクセスユニットの始点であると見なされるべきではない。そうしないと、デコーダは、 0 よりも大きな `n u h _ l a y e r _ i d` を有するいずれの N A L ユニットも認識せず、そのようなマルチレイヤビットストリームのベースレイヤが別様に復号可能であり得るとしても、そのビットストリームを適合しないものとして見なすので、空のアクセスユニットを見ることになる。

【 0 0 8 1 】

[0085]本開示の技法によれば、 A U D N A L ユニットの制約は次のように修正され得る（下線付きのテキストが追加されている）。

0 に等しい `n u h _ l a y e r _ i d` を伴うアクセスユニットデリミタ N A L ユニットが存在するとき、それは第 1 の N A L ユニットであるものとする。どのアクセスユニット中にも 0 に等しい `n u h _ l a y e r _ i d` を伴う多くとも 1 つのアクセスユニットデリミタ N A L ユニットがあるものとする。

【 0 0 8 2 】

[0086]代替的に、この制約は次のように修正される。

【 0 0 8 3 】

[0087] 0 に等しい `n u h _ l a y e r _ i d` を伴うアクセスユニットデリミタ N A L ユニットが存在するとき、それは第 1 の N A L ユニットであるものとする。どのアクセスユニット中にも多くとも 1 つのアクセスユニットデリミタ N A L ユニットがあるものとする。したがって、本開示の技法によれば、ビデオエンコーダ 2 0 は、第 1 のアクセスユニットに対して、 0 に等しいレイヤ識別子の値（たとえば、 `n u h _ l a y e r _ i d` ）の値を有する第 1 の A U D N A L ユニットを生成し、第 1 のアクセスユニットに対して、 0 よりも大きなレイヤ識別子を有する後続の A U D N A L ユニットを生成することができる。ビデオデコーダ 3 0 は、第 1 のアクセスユニットに対して、 0 に等しいレイヤ識別子を伴う第 1 の A U D N A L ユニットを受信することができ、第 1 のアクセスユニットに対して、 0 よりも大きなレイヤ識別子を有する後続の A U D N A L ユニットを受信することができる。第 1 のアクセスユニットにおいて、第 2 の A U D N A L ユニットは、第 1 の A U D N A L ユニットおよび少なくとも 1 つの V C L N A L ユニットに後続する。

【 0 0 8 4 】

[0088] 図 2 は、本開示で説明される技法を実装し得る例示的なビデオエンコーダ 2 0 を示すブロック図である。図 2 は、説明のために与えられており、本開示で広く例示され説明される技法を限定するものと見なされるべきではない。説明のために、本開示は、 H E V C コーディングの状況においてビデオエンコーダ 2 0 を説明する。しかしながら、本開示の技法は、他のコーディング規格または方法にも適用可能であり得る。

【 0 0 8 5 】

[0089] ビデオエンコーダ 2 0 は、本開示で説明される技法を実装し得る別の例示的なデバイスである後処理エンティティ 2 7 にビデオを出力するように構成され得る。後処理工エンティティ 2 7 は、メディア認識ネットワーク要素（ M A N E ）、分割 / 編集デバイスまたは他の中間デバイスのような、ビデオエンコーダ 2 0 からの符号化されたビデオデータを処理し得るビデオエンティティの例を表すものとする。いくつかの事例では、後処理工エンティティ 2 7 はネットワークエンティティの例であってよい。いくつかのビデオ符号化

10

20

30

40

50

システムでは、後処理エンティティ 27 およびビデオエンコーダ 20 は別個のデバイスの部分であってよく、他の事例では、後処理エンティティ 27 に関して説明される機能は、ビデオエンコーダ 20 を備える同じデバイスによって実行されてよい。

【 0 0 8 6 】

[0090] ビデオエンコーダ 20 は、ビデオスライス内のビデオブロックのイントラコーディングとインターフォーマンスとを実行し得る。イントラコーディングは、所与のビデオフレームまたはピクチャ内のビデオの空間冗長性を低減または除去するために空間予測に依拠する。インターフォーマンスは、ビデオシーケンスの隣接するフレームまたはピクチャ内のビデオの時間的冗長性を低減または除去するために時間予測に依拠する。イントラモード (I モード) は、いくつかの空間ベースの圧縮モードのいずれかを指し得る。單方向予測 (P モード) または双予測 (B モード) のようなインターモードは、いくつかの時間ベースの圧縮モードのいずれかを指し得る。 10

【 0 0 8 7 】

[0091] 図 2 の例では、ビデオエンコーダ 20 は、ビデオデータメモリ 33 と、区分ユニット 35 と、予測処理ユニット 41 と、フィルタユニット 63 と、復号ピクチャバッファ (D P B) 64 と、加算器 50 と、変換処理ユニット 52 と、量子化ユニット 54 と、エントロピー符号化ユニット 56 とを含む。予測処理ユニット 41 は、動き推定ユニット 42 と、動き補償ユニット 44 と、イントラ予測処理ユニット 46 とを含む。ビデオブロックの再構築のために、ビデオエンコーダ 20 はまた、逆量子化ユニット 58 と、逆変換処理ユニット 60 と、加算器 62 とを含む。フィルタユニット 63 は、デブロッキングフィルタ、適応ループフィルタ (A L F)、およびサンプル適応オフセット (S A O) フィルタのような、1つまたは複数のループフィルタを表すことが意図されている。図 2 では、フィルタユニット 63 はループ内フィルタであるものとして示されているが、他の構成では、フィルタユニット 63 はループ後フィルタとして実装され得る。 20

【 0 0 8 8 】

[0092] 図 2 に示されるように、ビデオエンコーダ 20 はビデオデータを受信し、受信されたビデオデータをビデオデータメモリ 33 に記憶する。ビデオデータメモリ 33 は、ビデオエンコーダ 20 のコンポーネントによって符号化されるべきビデオデータを記憶し得る。ビデオデータメモリ 33 に記憶されたビデオデータは、たとえば、ビデオソース 18 から取得され得る。D P B 64 は、たとえば、イントラコーディングモードまたはインターフォーマンスモードでビデオエンコーダ 20 によってビデオデータを符号化する際に使用するための、参照ビデオデータを記憶する参照ピクチャメモリであり得る。ビデオデータメモリ 33 および D P B 64 は、同期 D R A M (S D R A M) を含むダイナミックランダムアクセスメモリ (D R A M)、磁気抵抗 R A M (M R A M)、抵抗性 R A M (R R A M (登録商標))、または他のタイプのメモリデバイスのような、様々なメモリデバイスのいずれかによって形成され得る。ビデオデータメモリ 33 および D P B 64 は、同じメモリデバイスまたは別個のメモリデバイスによって与えられ得る。様々な例では、ビデオデータメモリ 33 は、ビデオエンコーダ 20 の他のコンポーネントとともにオンチップであるか、またはそれらのコンポーネントに対してオフチップであり得る。 30

【 0 0 8 9 】

[0093] 区分ユニット 35 は、ビデオデータメモリ 33 からビデオデータを取り出し、ビデオデータをビデオブロックに区分する。この区分は、たとえば、L C U および C U の 4 分割構造に従って、スライス、タイル、または他のより大きいユニットへの区分、ならびにビデオブロックの区分も含み得る。ビデオエンコーダ 20 は一般に、符号化されるべきビデオスライス内のビデオブロックを符号化するコンポーネントを示す。スライスは、複数のビデオブロック (場合によってはタイルと呼ばれるビデオブロックのセット) に分割され得る。予測処理ユニット 41 は、誤差結果 (たとえばコーディングレートおよびひずみレベル) に基づいて現在のビデオブロックについて、複数のイントラコーディングモードの 1 つ、または複数のインターフォーマンスモードの 1 つのような、複数の可能なコーディングモードの 1 つを選択し得る。予測処理ユニット 41 は、得られたイントラコーディングモードの 1 つを選択し得る。 40

イングされたブロックまたはインターフォーマンスされたブロックを、残差ブロックデータを生成するために加算器 50 に与え、参照ピクチャとして使用するための符号化されたブロックを再構築するために加算器 62 に与え得る。

【 0 0 9 0 】

[0094] 予測処理ユニット 41 内のイントラ予測処理ユニット 46 は、空間的圧縮を行うために、コーディングされるべき現在のブロックと同じフレームまたはスライス中の 1 つまたは複数の隣接ブロックに対して現在のビデオブロックのイントラ予測コーディングを実行することができる。予測処理ユニット 41 内の動き推定ユニット 42 および動き補償ユニット 44 は、時間的圧縮を行うために、1 つまたは複数の参照ピクチャ中の 1 つまたは複数の予測ブロックに対して現在のビデオブロックのイントラ予測コーディングを実行する。10

【 0 0 9 1 】

[0095] 動き推定ユニット 42 は、ビデオシーケンスの所定のパターンに従ってビデオスライスのためのイントラ予測モードを決定するように構成され得る。所定のパターンは、シーケンス中のビデオスライスを P スライスまたは B スライスとして指定することができる。動き推定ユニット 42 および動き補償ユニット 44 は、高度に統合され得るが、概念的な目的のために別々に示されている。動き推定ユニット 42 によって実行される動き推定は、ビデオブロックの動きを推定する動きベクトルを生成する処理である。動きベクトルは、たとえば、参照ピクチャ内の予測ブロックに対する現在のビデオフレームまたはピクチャ内のビデオブロックの P U の変位を示し得る。20

【 0 0 9 2 】

[0096] 予測ブロックは、絶対値差分和 (S A D) 、差分 2 乗和 (S S D) 、または他の差分の尺度によって決定され得るピクセル差分に関して、コーディングされるべきビデオブロックの P U に厳密に一致することが判明しているブロックである。いくつかの例では、ビデオエンコーダ 20 は、 D P B 64 に記憶された参照ピクチャのサブ整数ピクセル位置の値を計算し得る。たとえば、ビデオエンコーダ 20 は、参照ピクチャの 4 分の 1 ピクセル位置、 8 分の 1 ピクセル位置、または他の分数ピクセル位置の値を補間し得る。したがって、動き推定ユニット 42 は、フルピクセル位置と分数ピクセル位置とに対して動き探索を実行し、分数ピクセル精度で動きベクトルを出力し得る。30

【 0 0 9 3 】

[0097] 動き推定ユニット 42 は、 P U の位置を参照ピクチャの予測ブロックの位置と比較することによって、インターフォーマンスされたスライスにおけるビデオブロックの P U のための動きベクトルを計算する。参照ピクチャは、第 1 の参照ピクチャリスト (リスト 0) または第 2 の参照ピクチャリスト (リスト 1) から選択されてよく、それらの参照ピクチャリストの各々は、 D P B 64 に記憶された 1 つまたは複数の参照ピクチャを識別する。動き推定ユニット 42 は、計算された動きベクトルをエントロピー符号化ユニット 56 と動き補償ユニット 44 とに送る。40

【 0 0 9 4 】

[0098] 動き補償ユニット 44 によって実行される動き補償は、動き推定によって決定された動きベクトルに基づいて予測ブロックをフェッチまたは生成すること、場合によってはサブピクセル精度への補間を実行することを伴い得る。現在のビデオブロックの P U のための動きベクトルを受信すると、動き補償ユニット 44 は、動きベクトルが参照ピクチャリストの 1 つにおいて指す予測ブロックの位置を特定し得る。ビデオエンコーダ 20 は、コーディングされている現在のビデオブロックのピクセル値から予測ブロックのピクセル値を減算し、ピクセル差分値を形成することによって残差ビデオブロックを形成する。ピクセル差分値は、ブロックの残差データを形成し、ルーマとクロマの両方の差分成分を含み得る。加算器 50 は、この減算演算を実行する 1 つまたは複数のコンポーネントを表す。動き補償ユニット 44 はまた、ビデオスライスのビデオブロックを復号する際にビデオデコーダ 30 が使用するための、ビデオブロックおよびビデオスライスと関連付けられるシンタックス要素を生成し得る。50

【0095】

[0099] イントラ予測処理ユニット46は、上で説明されたように、動き推定ユニット42および動き補償ユニット44によって実行されるインター予測の代替として、現在のブロックをイントラ予測し得る。特に、イントラ予測処理ユニット46は、現在のブロックを符号化するために使用すべきイントラ予測モードを決定し得る。いくつかの例では、イントラ予測処理ユニット46は、たとえば、別個の符号化パスの間に、様々なイントラ予測モードを使用して現在のブロックを符号化することができ、イントラ予測処理ユニット46（または、いくつかの例では、モード選択ユニット40）は、試験されたモードから使用するのに適切なイントラ予測モードを選択することができる。たとえば、イントラ予測処理ユニット46は、様々な試験されたイントラ予測モードに対するレートひずみ分析を使用してレートひずみ値を計算し、試験されたモードの中で最良のレートひずみ特性を有するイントラ予測モードを選択し得る。レートひずみ分析は、一般に、符号化されたブロックと、符号化されたブロックを生成するために符号化された元の符号化されていないブロックとの間のひずみ（または誤差）の量、ならびに符号化されたブロックを生成するために使用されたビットレート（すなわち、ビットの数）を決定する。イントラ予測処理ユニット46は、符号化された様々なブロックのひずみおよびレートから比を計算し、どのイントラ予測モードがブロックの最良のレートひずみ値を示すかを決定し得る。

【0096】

[0100] いずれの場合も、ブロックのためのイントラ予測モードを選択した後に、イントラ予測処理ユニット46は、ブロックのための選択されたイントラ予測モードを示す情報をエントロピー符号化ユニット56に与え得る。エントロピー符号化ユニット56は、本開示の技法に従って、選択されたイントラ予測モードを示す情報を符号化し得る。ビデオエンコーダ20は、複数のイントラ予測モードインデックステーブルおよび複数の修正されたイントラ予測モードインデックステーブル（コードワードマッピングテーブルとも呼ばれる）と、様々なブロックの符号化コンテキストの定義と、コンテキストの各々について使用すべき、最確イントラ予測モード、イントラ予測モードインデックステーブル、および修正されたイントラ予測モードインデックステーブルの指示とを含み得る構成データを、送信されるビットストリームに含め得る。

【0097】

[0101] 予測処理ユニット41が、インター予測またはイントラ予測のいずれかを介して、現在のビデオブロックのための予測ブロックを生成した後に、ビデオエンコーダ20は、現在のビデオブロックから予測ブロックを減算することによって残差ビデオブロックを形成する。残差ブロック中の残差ビデオデータは、1つまたは複数のT U中に含まれ、変換処理ユニット52に適用され得る。変換処理ユニット52は、離散コサイン変換（DCT）または概念的に同様の変換などの変換を使用して、残差ビデオデータを残差変換係数に変換する。変換処理ユニット52は、残差ビデオデータをピクセル領域からの周波数領域などの変換領域に変換し得る。

【0098】

[0102] 変換処理ユニット52は、得られた変換係数を量子化ユニット54に送り得る。量子化ユニット54は、ビットレートをさらに低減するために変換係数を量子化する。量子化処理は、係数の一部またはすべてと関連付けられるビット深度を低減し得る。量子化の程度は、量子化パラメータを調整することによって変更され得る。いくつかの例では、量子化ユニット54は次いで、量子化された変換係数を含む行列の走査を実行し得る。代替的に、エントロピー符号化ユニット56が、走査を実行し得る。

【0099】

[0103] 量子化の後、エントロピー符号化ユニット56は、量子化された変換係数をエントロピー符号化する。たとえば、エントロピー符号化ユニット56は、コンテキスト適応型可変長コーディング（C A V L C）、コンテキスト適応型バイナリ算術コーディング（C A B A C）、シンタックスペースコンテキスト適応型バイナリ算術コーディング（S B A C）、確率間隔区分エントロピー（P I P E）コーディング、または別のエントロピー

符号化方法もしくは技法を実行し得る。エントロピー符号化ユニット 56 によるエントロピー符号化の後に、符号化されたビットストリームは、ビデオデコーダ 30 に送信され、またはビデオデコーダ 30 による後の送信または取り出しのためにアーカイブされ得る。エントロピー符号化ユニット 56 はまた、コーディングされている現在のビデオスライスのための動きベクトルと他のシンタックス要素とをエントロピー符号化することができる。

【 0 1 0 0 】

[0104]逆量子化ユニット 58 および逆変換処理ユニット 60 は、それぞれ逆量子化および逆変換を適用して、参照ピクチャの参照ブロックとして後で使用するためにピクセル領域において残差ブロックを再構築する。動き補償ユニット 44 は、残差ブロックを参照ピクチャリストの 1 つの中の参照ピクチャの 1 つの予測ブロックに加算することによって参照ブロックを計算することができる。動き補償ユニット 44 はまた、再構築された残差ブロックに 1 つまたは複数の補間フィルタを適用して、動き推定において使用するためのサブ整数ピクセル値を計算し得る。加算器 62 は、再構築された残差ブロックを動き補償ユニット 44 によって生成された動き補償予測ブロックに加算して、D P B 64 に記憶するための参照ブロックを生成する。参照ブロックは、後続のビデオフレームまたはピクチャ中のブロックをインター予測するために、動き推定ユニット 42 および動き補償ユニット 44 によって参照ブロックとして使用され得る。

【 0 1 0 1 】

[0105]本開示の態様によれば、ビデオエンコーダ 20 は、マルチレイヤビデオデータの 2 つ以上のレイヤに補足強化情報 (S E I) メッセージが適用されることを決定し、マルチレイヤビデオデータの 2 つ以上のレイヤと関連付けられる複数のインスタンスを有するシンタックス要素を S E I メッセージが参照することを決定するように構成され得る。そのような事例では、ビデオエンコーダ 20 は、複数のインスタンスのすべてを同じ値に設定することができる。

【 0 1 0 2 】

[0106]本開示の態様によれば、ビデオエンコーダ 20 はまた、アクセスユニットの第 1 のピクチャに対する第 1 の V C L N A L ユニットを生成するように構成され得る。第 1 の V C L N A L ユニットは、第 1 のスライスタイプを含む。ビデオエンコーダ 20 は、アクセスユニットの第 2 のピクチャに対する第 2 の V C L N A L ユニットを生成することができる。第 2 の V C L N A L ユニットは、第 2 のスライスタイプを含む。ビデオエンコーダ 20 は、第 1 のスライスタイプおよび第 2 のスライスタイプに基づいて、A U D N A L ユニットを生成することができる。

【 0 1 0 3 】

[0107]図 3 は、本開示で説明される技法を実装し得る例示的なビデオデコーダ 30 を示すブロック図である。図 3 は、説明のために与えられており、本開示で広く例示され説明される技法に対する限定ではない。説明のために、本開示は、H E V C コーディングの状況においてビデオデコーダ 30 を説明する。しかしながら、本開示の技法は、他のコーディング規格またはコーディング方法に適用可能であり得る。

【 0 1 0 4 】

[0108]図 3 の例では、ビデオデコーダ 30 は、エントロピー復号ユニット 80 と、予測処理ユニット 81 と、逆量子化ユニット 86 と、逆変換処理ユニット 88 と、加算器 90 と、フィルタユニット 91 と、参照ピクチャメモリ 92 とを含む。予測処理ユニット 81 は、動き補償ユニット 82 と、イントラ予測処理ユニット 84 とを含む。ビデオデコーダ 30 は、いくつかの例では、図 2 からのビデオエンコーダ 20 に関して説明された符号化パスとは全般に逆の復号パスを実行し得る。

【 0 1 0 5 】

[0109]復号処理の間、ビデオデコーダ 30 は、符号化されたビデオスライスのビデオブロックと、関連するシンタックス要素とを表す、符号化されたビデオビットストリームをビデオエンコーダ 20 から受信する。ビデオデコーダ 30 は、ネットワークエンティティ

10

20

30

40

50

78から符号化されたビデオビットストリームを受信し得る。ネットワークエンティティ78は、たとえば、上で説明された技法の1つまたは複数を実装するように構成されたサーバ、MANE、ビデオエディタ/スプライサ、または他のそのようなデバイスであり得る。ネットワークエンティティ78は、ビデオエンコーダ20のようなビデオエンコーダを含んでもよく、または含まなくてもよい。本開示で説明される技法のいくつかは、ネットワークエンティティ78が符号化されたビデオビットストリームをビデオデコーダ30に送信するよりも前に、ネットワークエンティティ78によって実施され得る。いくつかのビデオ復号システムでは、ネットワークエンティティ78およびビデオデコーダ30は別個のデバイスの一部であり得るが、他の事例では、ネットワークエンティティ78に関して説明される機能は、ビデオデコーダ30を備える同じデバイスによって実行され得る。

10

【0106】

[0110]復号処理の間、ビデオデコーダ30は、符号化されたビデオスライスのビデオブロックと、関連するシンタックス要素とを表す、符号化されたビデオビットストリームをビデオエンコーダ20から受信する。ビデオデコーダ30は、受信された符号化されたビデオビットストリームをビデオデータメモリ79に記憶する。ビデオデータメモリ79は、ビデオデコーダ30のコンポーネントによって復号されるべき符号化されたビデオビットストリームのようなビデオデータを記憶し得る。ビデオデータメモリ79に記憶されたビデオデータは、たとえば、リンク16を介して、記憶デバイス26から、または、カメラのようなローカルビデオソースから、または、物理データ記憶媒体にアクセスすることによって取得され得る。ビデオデータメモリ79は、符号化されたビデオビットストリームからの符号化されたビデオデータを記憶するコーディングピクチャバッファ(CPB)を形成し得る。DPB94は、たとえば、イントラコーディングモードまたはインターフォーマンスモードでビデオデコーダ30によってビデオデータを復号する際に使用するための参照ビデオデータを記憶する参照ピクチャメモリであり得る。ビデオデータメモリ79およびDPB94は、DRAM、SDRAM、MRAM、RRAM、または他のタイプのメモリデバイスのような、様々なメモリデバイスのいずれかによって形成され得る。ビデオデータメモリ79およびDPB94は、同じメモリデバイスまたは別個のメモリデバイスによって与えられ得る。様々な例では、ビデオデータメモリ79は、ビデオデコーダ30の他のコンポーネントとともにオンチップであるか、またはそれらのコンポーネントに対してオフチップであり得る。

20

【0107】

[0111]ビデオデコーダ30のエントロピー復号ユニット80は、量子化された係数と、動きベクトルと、他のシンタックス要素とを生成するために、ビデオデータメモリ79に記憶されているビデオデータを復号する。エントロピー復号ユニット80は、動きベクトルと他のシンタックス要素とを予測処理ユニット81に転送する。ビデオデコーダ30は、ビデオスライスレベルおよび/またはビデオブロックレベルでシンタックス要素を受信することができる。

30

【0108】

[0112]ビデオスライスがイントラコーディングされた(I)スライスとしてコーディングされるとき、予測処理ユニット81のイントラ予測処理ユニット84は、シグナリングされたイントラ予測モードと、現在のフレームまたはピクチャの前に復号されたブロックからのデータとに基づいて、現在のビデオスライスのビデオブロックのための予測データを生成し得る。ビデオフレームがインターフォーマンスモードでイントラコーディングされる(すなわち、BまたはP)スライスとしてコーディングされるとき、予測処理ユニット81の動き補償ユニット82は、エントロピー復号ユニット80から受信された動きベクトルおよび他のシンタックス要素に基づいて、現在のビデオスライスのビデオブロックのための予測ブロックを生成する。予測ブロックは、参照ピクチャリストのうちの1つの中の、参照ピクチャの1つから生成され得る。ビデオデコーダ30は、参照ピクチャメモリ92に記憶された参照ピクチャに基づいて、デフォルトの構築技法を使用して、参照フレームリスト、すなわち、リスト

40

50

ト 0 とリスト 1 とを構築することができる。

【 0 1 0 9 】

[0113]動き補償ユニット 8 2 は、動きベクトルと他のシンタックス要素とを構文解析することによって現在のビデオスライスのビデオブロックのための予測情報を決定し、復号されている現在のビデオブロックのための予測ブロックを生成するために予測情報を使用する。たとえば、動き補償ユニット 8 2 は、ビデオスライスのビデオブロックをコーディングするために使用される予測モード(たとえば、イントラまたはインター予測)と、インター予測スライスタイプ(たとえば、BスライスまたはPスライス)と、スライスの参照ピクチャリストの1つまたは複数のための構築情報と、スライスの各々のインター符号化されたビデオブロックのための動きベクトルと、スライスの各々のインターコーディングされたビデオブロックのためのインター予測ステータスと、現在のビデオスライス中のビデオブロックを復号するための他の情報とを決定するために、受信されたシンタックス要素のいくつかを使用する。

【 0 1 1 0 】

[0114]動き補償ユニット 8 2 はまた、補間フィルタに基づいて補間を実行し得る。動き補償ユニット 8 2 は、ビデオブロックの符号化の間にビデオエンコーダ 2 0 によって使用されるような補間フィルタを使用して、参照ブロックのサブ整数ピクセルの補間された値を計算し得る。この場合、動き補償ユニット 8 2 は、受信されたシンタックス要素からビデオエンコーダ 2 0 によって使用される補間フィルタを決定し、その補間フィルタを使用して予測ブロックを生成し得る。

【 0 1 1 1 】

[0115]逆量子化ユニット 8 6 は、ピットストリーム中で与えられ、エントロピー復号ユニット 8 0 によって復号された量子化された変換係数を逆量子化(inverse quantize)、すなわち、逆量子化(de-quantize)する。逆量子化処理は、量子化の程度を決定し、同様に、適用されるべき逆量子化の程度を決定するための、ビデオスライス中のビデオブロックごとにビデオエンコーダ 2 0 によって計算される量子化パラメータの使用を含み得る。逆変換ユニット 8 8 は、ピクセル領域において残差ブロックを生成するために、逆変換、たとえば、逆DCT、逆整数変換、または概念的に同様の逆変換処理を変換係数に適用する。

【 0 1 1 2 】

[0116]動き補償ユニット 8 2 が、動きベクトルおよび他のシンタックス要素に基づいて現在のビデオブロックのための予測ブロックを生成した後、ビデオデコーダ 3 0 は、逆変換処理ユニット 8 8 からの残差ブロックを動き補償ユニット 8 2 によって生成された対応する予測ブロックと加算することによって、復号されたビデオブロックを形成する。加算器 9 0 は、この加算演算を実行する1つまたは複数のコンポーネントを表す。所望される場合、ピクセル移行を平滑化するために、または別様にビデオ品質を向上させるために、(コーディンググループの中とコーディンググループの後のいずれかの)ループフィルタも使用され得る。

【 0 1 1 3 】

[0117]フィルタユニット 9 1 は、デブロッキングフィルタ、適応ループフィルタ(ALF)、およびサンプル適応オフセット(SAO)フィルタのような、1つまたは複数のループフィルタを表すことが意図されている。図3では、フィルタユニット 9 1 はループ内フィルタであるとして示されているが、他の構成では、フィルタユニット 9 1 はループ後フィルタとして実装され得る。所与のフレームまたはピクチャ中の復号されたビデオブロックは、次いで、後続の動き補償のために使用される参照ピクチャを記憶する参照ピクチャメモリ 9 2 に記憶される。参照ピクチャメモリ 9 2 はまた、図1のディスプレイデバイス 3 1 のようなディスプレイデバイス上で後の表示のために、復号されたビデオを記憶する。

【 0 1 1 4 】

[0118]本開示の態様によれば、ビデオデコーダ 3 0 は、非INBLとINBLとを含

む複数のレイヤを備えるコーディングされたビデオデータを受信するように構成され得る。ビデオデコーダ30は、第1の表現フォーマットパラメータを含むコーディングされたビデオデータと関連付けられるVPSを受信し、第2の表現フォーマットパラメータを含むINBLと関連付けられるSPSを受信することができる。ビデオデコーダ30は、第1の表現フォーマットパラメータに基づいて非INBLを復号し、第2の表現フォーマットパラメータに基づいてINBLを復号することができる。

【0115】

[0119]図4は、ネットワーク120の一部を形成するデバイスの例示的なセットを示すブロック図である。この例では、ネットワーク120は、ルーティングデバイス124A、124B(ルーティングデバイス124)と、トランスコーディングデバイス126とを含む。ルーティングデバイス124およびトランスコーディングデバイス126は、ネットワーク120の一部を形成し得る少数のデバイスを表すことが意図される。スイッチ、ハブ、ゲートウェイ、ファイアウォール、ブリッジ、および他のそのようなデバイスのような、他のネットワークデバイスも、ネットワーク120に含まれ得る。その上、サーバデバイス122とクライアントデバイス128との間のネットワーク経路に沿って追加のネットワークデバイスが設けられ得る。いくつかの例では、サーバデバイス122はソースデバイス12(図1)に対応し得るが、クライアントデバイス128は宛先デバイス14(図1)に対応し得る。

【0116】

[0120]一般に、ルーティングデバイス124は、ネットワーク120を通じてネットワークデータを交換するための1つまたは複数のルーティングプロトコルを実装する。いくつかの例では、ルーティングデバイス124は、プロキシまたはキャッシュ動作を実行するように構成され得る。したがって、いくつかの例では、ルーティングデバイス124はプロキシデバイスと呼ばれることがある。一般に、ルーティングデバイス124は、ネットワーク120を通じたルートを発見するためにルーティングプロトコルを実行する。そのようなルーティングプロトコルを実行することによって、ルーティングデバイス124Bは、それ自体からサーバデバイス122へのルーティングデバイス124Aを介したネットワークルートを発見し得る。ルーティングデバイス124の1つまたは複数は、本開示の1つまたは複数の態様を使用するMANEを備え得る。

【0117】

[0121]たとえば、MANEは、非INBLとINBLとを含む複数のレイヤを含むコーディングされたビデオデータを受信するように構成され得る。MANEは、第1の表現フォーマットパラメータを含むコーディングされたビデオデータと関連付けられるVPSを受信することができる。MANEは、第2の表現フォーマットパラメータを含むINBLと関連付けられるSPSを受信することができる。MANEは、第1の表現フォーマットパラメータに基づいて非INBLを処理し、第2の表現フォーマットパラメータに基づいてINBLを処理することができる。第2の表現フォーマットパラメータコンプライアンスに基づいてINBLを処理することの一部として、MANEは、第2の表現フォーマットパラメータに基づいて、INBLをベースレイヤとなるように書き換えることができる。書き換え処理の一部として、MANEは、INBLを、たとえばHEVCバージョン1に適合するベースレイヤへと、トランスコーディング(または変換)する。MANEは、INBLを入力として受信し、HEVCバージョン1のデコーダに出力を送信する。

【0118】

[0122]本開示の技法は、ルーティングデバイス124およびトランスコーディングデバイス126のようなネットワークデバイスによって実施され得るが、クライアントデバイス128によっても実施され得る。このように、ルーティングデバイス124、トランスコーディングデバイス126、およびクライアントデバイス128は、本開示の技法を実行するように構成されるデバイスの例を表す。その上、図1のデバイス、ならびに図2に示されるビデオエンコーダ20および図3に示されるビデオデコーダ30も、本開示の技法を実行するように構成され得る例示的なデバイスである。

10

20

30

40

50

【0119】

[0123]図5は、本開示の技法による、マルチレイヤビデオデータを処理する方法を示すフロー チャートである。図5の技法は、一般的なビデオ処理デバイスに関して説明される。ビデオ処理デバイスは、たとえば、後処理エンティティ27、ネットワークエンティティ28、ビデオデコーダ30、ルーティングデバイス124、またはトランスコーディングデバイス126のような、ビデオ処理デバイスに対応し得る。ビデオ処理デバイスは、非INBLとINBLとを含むコーディングされたビデオデータを受信する(140)。ビデオ処理デバイスは、コーディングされたビデオデータのために、第1の表現フォーマットパラメータを含むVPSを受信する(142)。ビデオ処理デバイスは、第2の表現フォーマットパラメータを含むINBLと関連付けられるSPSを受信する(144)。ビデオ処理デバイスは、第1の表現フォーマットパラメータに基づいて非INBLを処理し(146)、第2の表現フォーマットパラメータに基づいてINBLを処理する(148)。第1の表現フォーマットパラメータのうちの表現フォーマットパラメータの少なくとも1つの値は、第2の表現フォーマットパラメータのうちの表現フォーマットパラメータの少なくとも1つの値と異なり得る。

【0120】

[0124]ビデオ処理デバイスがMANEである場合、ビデオ処理デバイスは、第2の表現フォーマットパラメータに基づいてINBLをベースレイヤとなるように書き換えることによって、第2の表現フォーマットパラメータに基づいてINBLを処理することができる。INBLは、0よりも大きなレイヤ識別情報(たとえば、nuh_1ayr_id)を有することがあり、INBLを書き換えることの一部として、ビデオ処理デバイスは、INBLのレイヤ識別情報を0に設定することができるので、0に等しいレイヤ識別情報を伴うベースレイヤを作成する。INBLから書き換えられたベースレイヤは、マルチレイヤビデオをサポートしないHEVCバージョン1のデコーダによって復号可能であり得る。

【0121】

[0125]ビデオ処理デバイスがビデオデコーダである場合、ビデオ処理デバイスは、INBLを復号することによって、第2の表現フォーマットパラメータに基づいてINBLをプロセシングすることができる。INBLを復号した後で、ビデオデコーダは、INBLおよび非INBLに基づいて、復号されたビデオを出力または表示することができる。

【0122】

[0126]図6は、本開示の技法による、マルチレイヤビデオデータを符号化する方法を示すフロー チャートである。図6の技法は、ビデオエンコーダ20に関して説明される。ビデオエンコーダ20は、マルチレイヤビデオデータの2つ以上のレイヤにSEIメッセージが適用されることを決定する(150)。ビデオエンコーダ20は、マルチレイヤビデオデータの2つ以上のレイヤと関連付けられる複数のインスタンスを有するシンタックス要素をSEIメッセージが参照することを決定する(152)。マルチレイヤビデオデータの2つ以上のレイヤと関連付けられる複数のインスタンスを有するシンタックス要素をSEIメッセージが参照することに応答して、ビデオエンコーダ20は、複数のインスタンスのすべてを同じ値に設定する(154)。シンタックス要素の複数のインスタンスは、マルチレイヤビデオデータの第1のレイヤと関連付けられる第1のシーケンスパラメータセット中のシンタックス要素の第1のインスタンスと、マルチレイヤビデオデータの第2のレイヤと関連付けられる第2のシーケンスパラメータセット中のシンタックス要素の第2のインスタンスとを含み得る。シンタックス要素は、field_seq_flagシンタックス要素、general_progressive_source_flagシンタックス要素、またはgeneral_interlaced_source_flagシンタックス要素のいずれであってもよい。

【0123】

[0127]図7は、本開示の技法による、マルチレイヤビデオデータを符号化する方法を示すフロー チャートである。図7の技法は、ビデオエンコーダ20に関して説明される。ビ

デオエンコーダ20は、アクセスユニットの第1のピクチャに対する第1のVCL_NALユニットを生成する(160)。第1のVCL_NALユニットは、第1のスライスタイルを含む。ビデオエンコーダ20は、アクセスユニットの第2のピクチャに対する第2のVCL_NALユニットを生成する(162)。第2のVCL_NALユニットは、第2のスライスタイルを含む。ビデオエンコーダ20は、第1のスライスタイルおよび第2のスライスタイルに基づいて、AUD_NALユニットを生成する(164)。

【0124】

[0128]ビデオエンコーダ20は、第1のスライスタイルと第2のスライスタイルの両方をアクセスユニットが含むことを示すピクチャタイプシンタックス要素をAUD_NALユニットにおいて示すことによって、AUD_NALユニットを生成することができる。第1のAUD_NALユニットは、アクセスユニットと先行するアクセスユニットとの境界を特定するために使用され得る。AUD_NALユニットは、加えて、または代替的に、アクセスユニット中のすべてのピクチャに対するピクチャタイプを示すために使用され得る。第1のスライスタイルおよび第2のスライスタイルは、Iスライス、Pスライス、およびBスライスからなる群から選択されてよく、または、第1のスライスタイルおよび第2のスライスタイルは、IスライスおよびPスライスからなる群から選択されてよい。

10

【0125】

[0129]図8は、本開示の技法による、マルチレイヤビデオデータを符号化する方法を示すフローチャートである。図8の技法は、ビデオデコーダ30に関して説明される。ビデオデコーダ30は、第1のアクセスユニットに対する第1のAUD_NALユニットを受信する(170)。第1のAUD_NALユニットのためのレイヤ識別子は0に等しい。ビデオデコーダ30は、第1のアクセスユニットに対する第2のAUD_NALユニットを受信する(172)。第2のAUD_NALユニットのためのレイヤ識別子は0よりも大きい。ビデオデコーダ30は、第1のアクセスユニットを復号する(174)。

20

【0126】

[0130]第1のアクセスユニット中の第2のAUD_NALユニットは、復号順序において、第1のAUD_NALユニットと、第1のアクセスユニット中の少なくとも1つのビデオコーディングレイヤ(VCL)NALユニットとに後続し得る。第1のAUD_NALユニットは、復号順序において、第1のアクセスユニット中で最初のNALユニットであり得る。第1のAUD_NALユニットは、第1のアクセスユニットと先行するアクセスユニットとの境界を特定することができる。いくつかの例では、第1のAUD_NALユニット以外の、第1のアクセスユニット中の他のAUD_NALユニットは、0に等しいレイヤ識別子を有しないことがある。

30

【0127】

[0131]1つまたは複数の例では、説明される機能は、ハードウェア、ソフトウェア、ファームウェア、またはそれらの任意の組合せで実装され得る。ソフトウェアで実装される場合、機能は、1つまたは複数の命令またはコードとしてコンピュータ可読媒体上に記憶されるか、またはコンピュータ可読媒体を介して送信され、ハードウェアベースの処理ユニットによって実行され得る。コンピュータ可読媒体は、データ記憶媒体などの有形媒体に対応するコンピュータ可読記憶媒体、または、たとえば、通信プロトコルに従って、ある場所から別の場所へのコンピュータプログラムの転送を可能にする任意の媒体を含む通信媒体を含み得る。このようにして、コンピュータ可読媒体は、一般に、(1)非一時的である有形のコンピュータ可読記憶媒体または(2)信号または搬送波などの通信媒体に相当し得る。データ記憶媒体は、本開示で説明される技法の実装のために命令、コードおよび/またはデータ構造を取り出すために、1つもしくは複数のコンピュータまたは1つもしくは複数のプロセッサによってアクセスされ得る任意の利用可能な媒体であり得る。コンピュータプログラム製品は、コンピュータ可読媒体を含み得る。

40

【0128】

[0132]限定ではなく例として、そのようなコンピュータ可読記憶媒体は、RAM、ROM、EEPROM(登録商標)、CD-ROMもしくは他の光ディスク記憶装置、磁気デ

50

イスク記憶装置もしくは他の磁気記憶デバイス、フラッシュメモリ、または、命令またはデータ構造の形式で所望のプログラムコードを記憶するために使用されコンピュータによってアクセスされ得る任意の他の媒体を備え得る。また、いかなる接続もコンピュータ可読媒体と適切に呼ばれる。たとえば、命令が、ウェブサイト、サーバ、または他のリモートソースから、同軸ケーブル、光ファイバーケーブル、ツイストペア、デジタル加入者回線（D S L）、または赤外線、無線、およびマイクロ波などのワイヤレス技術を使用して送信される場合、同軸ケーブル、光ファイバーケーブル、ツイストペア、D S L、または赤外線、無線、およびマイクロ波などのワイヤレス技術は、媒体の定義に含まれる。ただし、コンピュータ可読記憶媒体およびデータ記憶媒体は、接続、搬送波、信号、または他の一時的媒体を含まないが、代わりに非一時的有形記憶媒体を対象とすることを理解されたい。本明細書で使用されるディスク（disk）およびディスク（disc）は、コンパクトディスク（disc）（C D）、レーザーディスク（登録商標）（disc）、光ディスク（disc）、デジタル多用途ディスク（disc）（D V D）、フロッピー（登録商標）ディスク（disk）およびB l u - r a yディスク（disc）を含み、ディスク（disk）は、通常、データを磁気的に再生し、ディスク（disc）は、データをレーザーで光学的に再生する。上記の組合せもコンピュータ可読媒体の範囲に含まれるべきである。

【0129】

[0133]命令は、1つまたは複数のデジタル信号プロセッサ（D S P）、汎用マイクロプロセッサ、特定用途向け集積回路（A S I C）、フィールドプログラマブル論理アレイ（F P G A）、または他の等価な集積回路もしくはディスクリート論理回路のような、1つまたは複数のプロセッサによって実行され得る。したがって、本明細書で使用される「プロセッサ」という用語は、上述の構造または本明細書で説明された技法の実装に好適な任意の他の構造のいずれかを指し得る。加えて、いくつかの態様では、本明細書で説明される機能は、符号化および復号のために構成された専用のハードウェアおよび／もしくはソフトウェアモジュール内で与えられ、または複合コーデックに組み込まれ得る。また、本技法は、1つまたは複数の回路または論理要素において完全に実装され得る。

【0130】

[0134]本開示の技法は、ワイヤレスハンドセット、集積回路（I C）またはI Cのセット（たとえば、チップセット）を含む、多種多様なデバイスまたは装置において実装され得る。本開示では、開示される技法を実行するように構成されたデバイスの機能的態様を強調するために、様々なコンポーネント、モジュール、またはユニットが説明されたが、それらのコンポーネント、モジュール、またはユニットを、必ずしも異なるハードウェアユニットによる実現を必要とするとは限らない。むしろ、上で説明されたように、様々なユニットが、好適なソフトウェアおよび／またはファームウェアとともに、上で説明された1つまたは複数のプロセッサを含めて、コーデックハードウェアユニットにおいて組み合わされてよく、または相互動作するハードウェアユニットの集合によって与えられてよい。

【0131】

[0135]様々な例が説明されてきた。これらおよび他の例は以下の特許請求の範囲内に入る。

以下に、本願出願の当初の特許請求の範囲に記載された発明を付記する。

[C 1]

マルチレイヤビデオデータを符号化する方法であって、

アクセスユニットの第1のピクチャのための第1のビデオコーディングレイヤ（V C L）ネットワーク抽象化レイヤ（N A L）ユニットを生成することと、前記第1のV C L N A Lユニットは第1のスライスタイルを備える、

前記アクセスユニットの第2のピクチャのための第2のV C L N A Lユニットを生成することと、前記第2のV C L N A Lユニットは第2のスライスタイルを備える、

前記第1のスライスタイルおよび第2のスライスタイルに基づいて、アクセスユニットデリミタ（A U D）N A Lユニットを生成することとを備える、方法。

10

20

30

40

50

[C 2]

前記 AUD_NAL ユニットを生成することが、前記第 1 のスライスタイプと前記第 2 のスライスタイプの両方をアクセスユニットが含むことを示すピクチャタイプシンタックス要素を前記 AUD_NAL ユニットにおいて示すことを備える、 C 1 に記載の方法。

[C 3]

前記 AUD_NAL ユニットが、前記アクセスユニット中のすべてのピクチャに対するピクチャタイプを示す、 C 1 に記載の方法。

[C 4]

前記第 1 のスライスタイプおよび前記第 2 のスライスタイプが、 I スライス、 P スライス、および B スライスからなる群から選択される、 C 1 に記載の方法。

10

[C 5]

前記第 1 のスライスタイプおよび前記第 2 のスライスタイプが、 I スライスおよび P スライスからなる群から選択される、 C 1 に記載の方法。

[C 6]

前記第 1 の AUD_NAL ユニットが、前記アクセスユニットと先行するアクセスユニットとの境界を特定する、 C 1 に記載の方法。

[C 7]

ワイヤレス通信デバイス上で実行可能であり、前記ワイヤレス通信デバイスが、前記マルチレイヤビデオデータを記憶するように構成されるメモリと、前記メモリに記憶されている前記マルチレイヤビデオデータを処理するための命令を実行するように構成される 1 つまたは複数のプロセッサと、前記マルチレイヤビデオデータを送信するように構成される送信機とを備える、 C 1 に記載の方法。

20

[C 8]

前記ワイヤレス通信デバイスが携帯電話であり、前記送信機が、セルラー通信規格に従って変調されたデータとして前記マルチレイヤビデオデータを送信するように構成される、 C 7 に記載の方法。

[C 9]

ビデオデータを処理するためのデバイスであって、マルチレイヤビデオデータのビットストリームの少なくとも一部分を記憶するように構成されるメモリと、 1 つまたは複数のプロセッサとを備え、前記 1 つまたは複数のプロセッサが、アクセスユニットの第 1 のピクチャのための第 1 のビデオコーディングレイヤ (VCL) ネットワーク抽象化レイヤ (NAL) ユニットを生成することと、前記第 1 の VCL_NAL ユニットは第 1 のスライスタイプを備える、前記アクセスユニットの第 2 のピクチャのための第 2 の VCL_NAL ユニットを生成することと、前記第 2 の VCL_NAL ユニットは第 2 のスライスタイプを備える、前記第 1 のスライスタイプおよび第 2 のスライスタイプに基づいて、アクセスユニットデリミタ (AUD_NAL) ユニットを生成することとを行うように構成される、デバイス。

30

[C 10]

前記 AUD_NAL ユニットを生成するために、前記 1 つまたは複数のプロセッサが、前記第 1 のスライスタイプと前記第 2 のスライスタイプの両方をアクセスユニットが含むことを示すピクチャタイプシンタックス要素を前記 AUD_NAL ユニットにおいて生成する、 C 9 に記載のデバイス。

40

[C 11]

前記 AUD_NAL ユニットが、前記アクセスユニット中のすべてのピクチャに対するピクチャタイプを示す、 C 9 に記載のデバイス。

[C 12]

前記第 1 のスライスタイプおよび前記第 2 のスライスタイプが、 I スライス、 P スライ

50

ス、およびBスライスからなる群から選択される、C 9に記載のデバイス。

[C 13]

前記第1のスライスタイプおよび前記第2のスライスタイプが、IスライスおよびPスライスからなる群から選択される、C 9に記載のデバイス。

[C 14]

前記第1のAUD_NALユニットが、前記アクセスユニットと先行するアクセスユニットとの境界を特定する、C 9に記載のデバイス。

[C 15]

ワイヤレス通信デバイスであり、

前記マルチレイヤビデオデータを送信するように構成される送信機をさらに備える、C 9に記載のデバイス。

10

[C 16]

前記ワイヤレス通信デバイスが携帯電話を備え、前記送信機が、セルラー通信規格に従って変調されたデータとして前記マルチレイヤビデオデータを送信するように構成される、C 15に記載のデバイス。

[C 17]

マルチレイヤビデオデータを符号化するための装置であって、

アクセスユニットの第1のピクチャのための第1のビデオコーディングレイヤ(VCL)ネットワーク抽象化レイヤ(NAL)ユニットを生成するための手段と、前記第1のVCL_NALユニットは第1のスライスタイプを備える、

20

前記アクセスユニットの第2のピクチャのための第2のVCL_NALユニットを生成するための手段と、前記第2のVCL_NALユニットは第2のスライスタイプを備える

、前記第1のスライスタイプおよび第2のスライスタイプに基づいて、アクセスユニットデリミタ(AUD_NAL)ユニットを生成するための手段とを備える、装置。

[C 18]

前記AUD_NALユニットを生成するための前記手段が、前記第1のスライスタイプと前記第2のスライスタイプの両方をアクセスユニットが含むことを示すピクチャタイプシンタックス要素を前記AUD_NALユニットにおいて示すための手段を備える、C 17に記載の装置。

30

[C 19]

前記AUD_NALユニットが、前記アクセスユニット中のすべてのピクチャに対するピクチャタイプを示す、C 17に記載の装置。

[C 20]

前記第1のスライスタイプおよび前記第2のスライスタイプが、Iスライス、Pスライス、およびBスライスからなる群から選択される、C 17に記載の装置。

[C 21]

前記第1のスライスタイプおよび前記第2のスライスタイプが、IスライスおよびPスライスからなる群から選択される、C 17に記載の装置。

[C 22]

前記第1のAUD_NALユニットが、前記アクセスユニットと先行するアクセスユニットとの境界を特定する、C 17に記載の装置。

40

[C 23]

1つまたは複数のプロセッサによって実行されると、前記1つまたは複数のプロセッサに、

アクセスユニットの第1のピクチャのための第1のビデオコーディングレイヤ(VCL)ネットワーク抽象化レイヤ(NAL)ユニットを生成することと、前記第1のVCL_NALユニットは第1のスライスタイプを備える、

前記アクセスユニットの第2のピクチャのための第2のVCL_NALユニットを生成することと、前記第2のVCL_NALユニットは第2のスライスタイプを備える、

50

前記第1のスライスタイルおよび第2のスライスタイルに基づいて、アクセスユニットデリミタ（AUD）NALユニットを生成することを行わせる命令を記憶した、コンピュータ可読記憶媒体。

[C 2 4]

前記 AUD NAL ユニットを生成するために、前記 1 つまたは複数のプロセッサが、前記第 1 のスライスタイルと前記第 2 のスライスタイルの両方をアクセスユニットが含むことを示すピクチャタイプシンタックス要素を前記 AUD NAL ユニットに含める、C 2 3 に記載のコンピュータ可読記憶媒体。

[C 2 5]

前記 A U D N A L ユニットが、前記アクセスユニット中のすべてのピクチャに対するピクチャタイプを示す、C 2 3 に記載のコンピュータ可読記憶媒体。

[C 2 6]

前記第1のスライスタイルおよび前記第2のスライスタイルが、Iスライス、Pスライス、およびBスライスからなる群から選択される、C23に記載のコンピュータ可読記憶媒体。

[C 2 7]

前記第1のスライスタイルおよび前記第2のスライスタイルが、IスライスおよびPスライスからなる群から選択される、C23に記載のコンピュータ可読記憶媒体。

[C 2 8]

前記第1のAUD/NALユニットが、前記アクセスユニットと先行するアクセスユニットとの境界を特定する、C23に記載のコンピュータ可読記憶媒体。

10

20

【 义 1 】

1

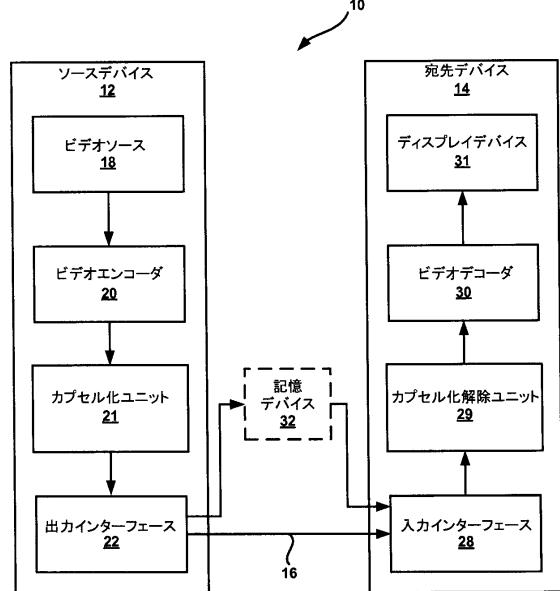


FIG. 1

【义2】

図 2

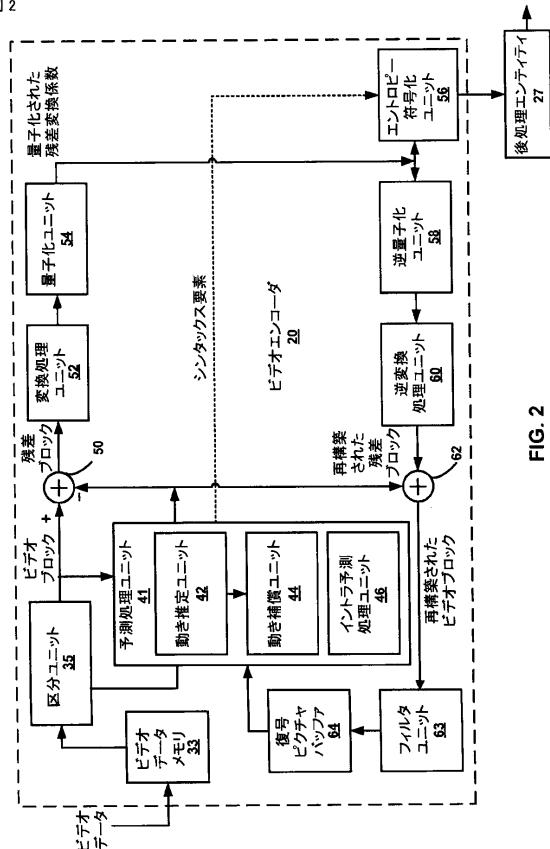
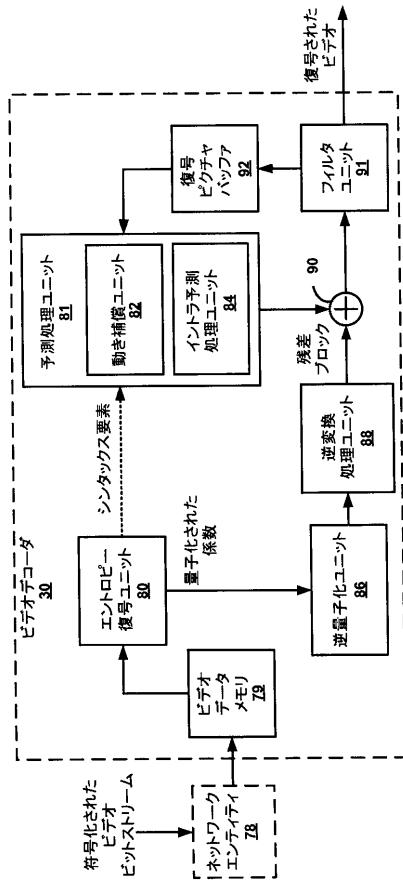


FIG. 2

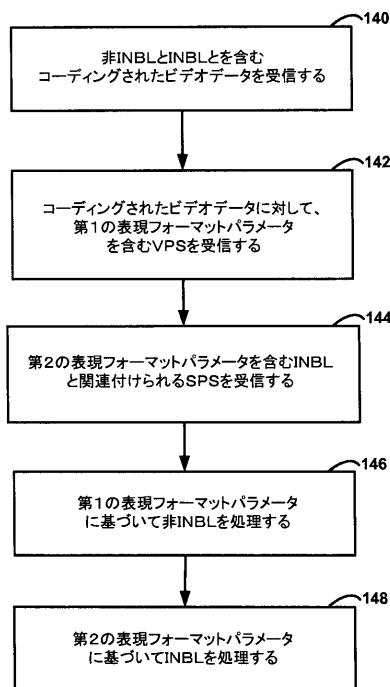
【 図 3 】

3



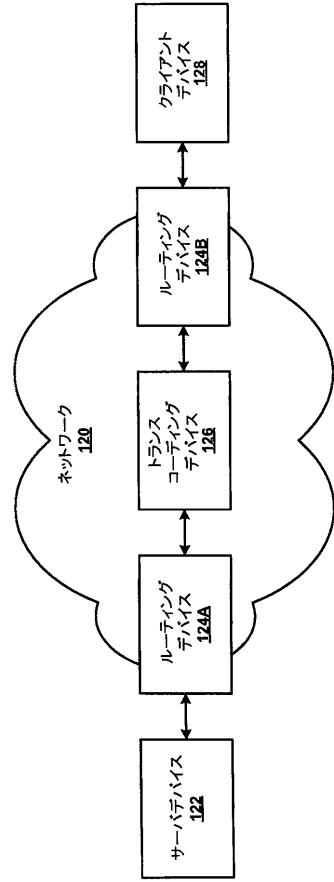
【図5】

図 5



【 図 4 】

四 4



【 四 6 】

图 6

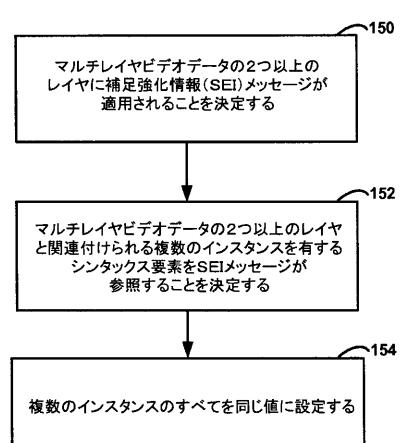
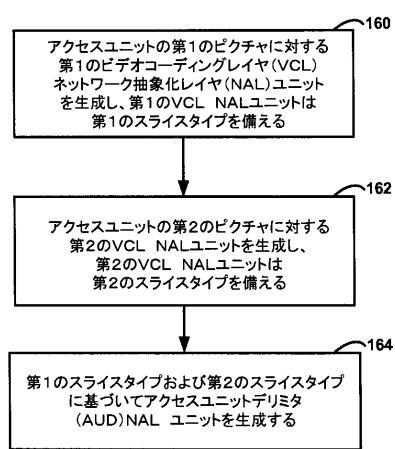


FIG. 5

FIG. 6

【図7】

図7



【図8】

図8

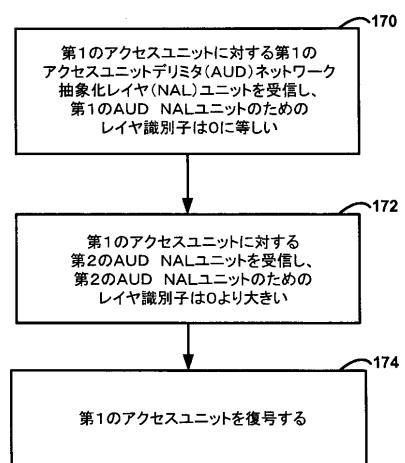


FIG. 7

FIG. 8

フロントページの続き

前置審査

(72)発明者 ワン、イエ - クイ

アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775

(72)発明者 ヘンドリー、フヌ

アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775

(72)発明者 ラマスブラモニアン、アダルシュ・クリシュナン

アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775

審査官 岩井 健二

(56)参考文献 特開2010-245766 (JP, A)

国際公開第2015/053286 (WO, A1)

国際公開第2014/055537 (WO, A1)

High efficiency video coding, Recommendation ITU-T H.265, ITU-T, 2013年 4月, H.265 (04/2013), pp.1-12,30,35-36,39-41,53-57,71,76-80

Jianle Chen et al., Preliminary version of High efficiency video coding (HEVC) scalable extension Draft 6, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 17th Meeting: Valencia, ES, 2014年 4月, JCTVC-C-Q1008_v1, pp.45-47,61,63-67,89,91-93

Hendry, Ye-Kui Wang and Adarsh K. Ramasubramonian, MV-HEVC/SHVC HLS: More miscellanous cleanups, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 18th Meeting: Sapporo, JP, 2014年 7月, JCTVC-R0279v2, pp.1-5

(58)調査した分野(Int.Cl., DB名)

H04N 19/00 - 19/98