

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6330667号
(P6330667)

(45) 発行日 平成30年5月30日(2018.5.30)

(24) 登録日 平成30年5月11日(2018.5.11)

(51) Int.Cl.	F 1
HO4N 19/70	(2014.01) HO4N 19/70
HO4N 19/31	(2014.01) HO4N 19/31
HO4N 21/434	(2011.01) HO4N 21/434
HO4N 21/435	(2011.01) HO4N 21/435

請求項の数 5 (全 46 頁)

(21) 出願番号	特願2015-1993 (P2015-1993)	(73) 特許権者	000002185 ソニー株式会社 東京都港区港南1丁目7番1号
(22) 出願日	平成27年1月8日(2015.1.8)	(74) 代理人	100093241 弁理士 宮田 正昭
(62) 分割の表示	特願2014-7306 (P2014-7306) の分割 原出願日 平成26年1月17日(2014.1.17)	(74) 代理人	100101801 弁理士 山田 英治
(65) 公開番号	特開2015-97418 (P2015-97418A)	(74) 代理人	100095496 弁理士 佐々木 榮二
(43) 公開日	平成27年5月21日(2015.5.21)	(74) 代理人	100086531 弁理士 澤田 俊夫
審査請求日	平成29年1月12日(2017.1.12)	(74) 代理人	110000763 特許業務法人大同特許事務所
(31) 優先権主張番号	特願2013-167168 (P2013-167168)		
(32) 優先日	平成25年8月9日(2013.8.9)		
(33) 優先権主張国	日本国(JP)		

最終頁に続く

(54) 【発明の名称】送信装置、送信方法、受信装置、受信方法、符号化装置および符号化方法

(57) 【特許請求の範囲】

【請求項 1】

動画像データを構成する各ピクチャの画像データが階層符号化されて生成された、低階層側のピクチャの符号化画像データを持つ第1のビデオストリームと高階層側のピクチャの符号化画像データを持つ第2のビデオストリームを含むと共に、上記第1のビデオストリームおよび上記第2のビデオストリームにそれぞれ対応して、各ビデオストリームに含まれる各ピクチャの符号化画像データに対応する階層の最大および最小の値を示す階層範囲情報を含み、さらに、上記第1のビデオストリームに対応して該第1のビデオストリームのレベル指定値が挿入された第1のデスクリプタと、上記第2のビデオストリームに対応して上記第1のビデオストリームおよび上記第2のビデオストリームを合わせたビデオストリームのレベル指定値が挿入された第2のデスクリプタを含む多重化ストリームを受信する受信部と、

上記階層範囲情報を、上記第1のデスクリプタおよび第2のデスクリプタに挿入されているビデオストリームのレベル指定値に基づいて、上記受信された多重化ストリームに含まれる上記第1のビデオストリーム、あるいは上記第1のビデオストリームおよび上記第2のビデオストリームの双方から、最下位の階層から所望の階層までのピクチャの符号化画像データを取り出し、復号化処理を行って得られた各ピクチャの画像データのフレームレートを表示能力に合わせる処理部を備える

受信装置。

【請求項 2】

10

20

上記処理部は、デコード能力に応じて選択されたビデオストリームが持つ所定階層組のピクチャの符号化画像データを取り出し、復号化処理を行う

請求項 1 に記載の受信装置。

【請求項 3】

上記符号化画像データは NAL ユニット構造を有し、上記第 1 のビデオストリームの SPS の NAL ユニットに、上記第 1 のビデオストリームのレベル指定値が挿入されている
請求項 1 に記載の受信装置。

【請求項 4】

上記高階層側のピクチャの符号化画像データのデコードタイミングが上記低階層側のピクチャの符号化画像データのデコードタイミングの中間タイミングとなるように、各ピクチャの符号化画像データにタイムスタンプが付与されている
10

請求項 1 に記載の受信装置。

【請求項 5】

受信部が、動画像データを構成する各ピクチャの画像データが階層符号化されて生成された、低階層側のピクチャの符号化画像データを持つ第 1 のビデオストリームと高階層側のピクチャの符号化画像データを持つ第 2 のビデオストリームを含むと共に、上記第 1 のビデオストリームおよび上記第 2 のビデオストリームにそれぞれ対応して、各ビデオストリームに含まれる各ピクチャの符号化画像データに対応する階層の最大および最小の値を示す階層範囲情報を含み、さらに、上記第 1 のビデオストリームに対応して該第 1 のビデオストリームのレベル指定値が挿入された第 1 のデスクリプタと、上記第 2 のビデオストリームに対応して上記第 1 のビデオストリームおよび上記第 2 のビデオストリームを合わせたビデオストリームのレベル指定値が挿入された第 2 のデスクリプタを含む多重化ストリームを受信する受信ステップと、
20

処理部が、上記階層範囲情報と、上記第 1 のデスクリプタおよび第 2 のデスクリプタに挿入されているビデオストリームのレベル指定値に基づいて、上記受信された多重化ストリームに含まれる上記第 1 のビデオストリーム、あるいは上記第 1 のビデオストリームおよび上記第 2 のビデオストリームの双方から、最下位の階層から所望の階層までのピクチャの符号化画像データを取り出し、復号化処理を行って得られた各ピクチャの画像データのフレームレートを表示能力に合わせる処理ステップを有する

受信方法。

30

【発明の詳細な説明】

【技術分野】

【0001】

本技術は、送信装置、送信方法、受信装置、受信方法、符号化装置および符号化方法に関する。詳しくは、本技術は、動画像データを構成する各ピクチャの画像データを階層符号化して送信する送信装置等に関する。

【背景技術】

【0002】

圧縮動画を、放送、ネット等でサービスする際、受信機のデコード能力によって再生可能なフレーム周波数の上限が制限される。従って、サービス側は普及している受信機の再生能力を考慮して、低フレーム周波数のサービスのみに制限したり、高低複数のフレーム周波数のサービスを同時提供したりする必要がある。
40

【0003】

受信機は、高フレーム周波数のサービスに対応するには、高コストとなり、普及の阻害要因となる。初期に低フレーム周波数のサービス専用の安価な受信機のみ普及していて、将来サービス側が高フレーム周波数のサービスを開始する場合、新たな受信機が無いと全く視聴不可能であり、サービスの普及の阻害要因となる。

【0004】

例えば、HEVC (High Efficiency Video Coding) において、動画像データを構成する各ピクチャの画像データを階層符号化することによる時間方向スケーラビリティが提案
50

されている（非特許文献1参照）。受信側では、N A L（Network Abstraction Layer）ユニットのヘッダに挿入されているテンポラルID（temporal_id）に基づき、各ピクチャの階層を識別でき、デコード能力に対応した階層までの選択的なデコードが可能となる。

【先行技術文献】

【非特許文献】

【0005】

【非特許文献1】Gary J. Sullivan, Jens-Rainer Ohm, Woo-Jin Han, Thomas Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard" IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 22, NO. 12, pp. 1649-1668, DECEMBER 2012

10

【発明の概要】

【発明が解決しようとする課題】

【0006】

本技術の目的は、受信側においてデコード能力に応じた良好なデコード処理を可能とすることがある。

【課題を解決するための手段】

【0007】

本技術の概念は、

動画像データを構成する各ピクチャの画像データを複数の階層に分類し、該分類された各階層の画像データを符号化し、該符号化された各階層のピクチャの画像データを持つビデオストリームを生成する画像符号化部と、

20

上記生成されたビデオストリームを含む所定フォーマットのコンテナを送信する送信部とを備え

上記画像符号化部は、

高階層ほどピクチャ毎の符号化画像データのデコード時間間隔が小さくなるように設定されたデコードタイミング情報を上記各階層のピクチャの符号化画像データに付加する送信装置にある。

【0008】

本技術において、画像符号化部により、動画像データを構成する各ピクチャの画像データが符号化されてビデオストリーム（符号化ストリーム）が生成される。この場合、動画像データを構成する各ピクチャの画像データが複数の階層に分類されて符号化され、各階層のピクチャの画像データを持つビデオストリームが生成される。この場合、各階層のピクチャの符号化画像データに、高階層ほどピクチャ毎の符号化画像データのデコード時間間隔が小さくなるように設定されたデコードタイミング情報、例えばデコードタイムスタンプが付加される。

30

【0009】

送信部により、上述のビデオストリームを含む所定フォーマットのコンテナが送信される。例えば、コンテナは、デジタル放送規格で採用されているトランスポートストリーム（M P E G - 2 T S）であってもよい。また、例えば、コンテナは、インターネットの配信などで用いられるM P 4、あるいはそれ以外のフォーマットのコンテナであってもよい。

40

【0010】

例えば、画像符号化部は、各階層のピクチャの符号化画像データを持つ単一のビデオストリームを生成し、複数の階層を2以上の所定数の階層組に分割し、各階層組のピクチャの符号化画像データに、所属階層組を識別するための識別情報を付加する、ようにされてもよい。この場合、例えば、識別情報は、ビットストリームのレベル指定値であり、高階層側の階層組ほど高い値とされる、ようにされてもよい。

【0011】

また、例えば、画像符号化部は、複数の階層を2以上の所定数の階層組に分割し、各階

50

層組のピクチャの符号化画像データをそれぞれ持つ所定数のビデオストリームを生成する、ようにされてもよい。この場合、例えば、画像符号化部は、各階層組のピクチャの符号化画像データに、所属階層組を識別するための識別情報を付加する、ようにされてもよい。そして、この場合、例えば、識別情報は、ビットストリームのレベル指定値であり、高階層側の階層組ほど高い値とされる、ようにされてもよい。

【0012】

このように本技術においては、高階層ほどピクチャ毎の符号化画像データのデコード時間間隔が小さくなるように設定されたデコードタイミング情報が各階層のピクチャの符号化画像データに付加されるものである。そのため、受信側においてデコード性能に応じた良好なデコード処理が可能となる。例えば、デコード能力が低い場合であっても、バッファ破たんを招くことなく、低階層のピクチャの符号化画像データを選択的にデコードすることが可能となる。

10

【0013】

なお、本技術において、例えば、画像符号化部は、各階層のピクチャの符号化画像データを持つ単一のビデオストリームを生成するか、あるいは複数の階層を2以上の所定数の階層組に分割し、各階層組のピクチャの符号化画像データをそれぞれ持つ所定数のビデオストリームを生成し、コンテナのレイヤに、このコンテナに含まれるビデオストリームの構成情報を挿入する情報挿入部をさらに備える、ようにされてもよい。この場合、例えば、受信側では、コンテナに含まれるビデオストリームの構成情報に基づいて、ビデオストリームの構成を容易に把握でき、適切なデコード処理を行うことが可能となる。

20

【0014】

また、本技術において、例えば、送信部は、複数の階層を2以上の所定数の階層組に分割し、低階層側の階層組のピクチャの符号化画像データをコンテナするパケットの優先度ほど高く設定する、ようにされてもよい。この場合、例えば、受信側では、このパケットの優先度に基づいて、自身のデコード能力に応じた階層組のピクチャの符号化画像データのみをバッファに取り込むことが可能となる。

【0015】

また、本技術の他の概念は、

動画像データを構成する各ピクチャの画像データが複数の階層に分類されて符号化されることで得られた各階層のピクチャの符号化画像データを持つビデオストリームを含む所定フォーマットのコンテナを受信する受信部を備え、

30

上記各階層のピクチャの符号化画像データには、高階層ほどピクチャ毎の符号化画像データのデコード時間間隔が小さくなるように設定されたデコードタイミング情報が付加されており、

上記受信されたコンテナに含まれる上記ビデオストリームから選択された所定階層以下の階層のピクチャの符号化画像データを、上記デコードタイミング情報が示すデコードタイミングでデコードして、上記所定階層以下の階層のピクチャの画像データを得る処理部をさらに備える

受信装置にある。

【0016】

40

本技術において、受信部により、所定フォーマットのコンテナが受信される。このコンテナには、動画像データを構成する各ピクチャの画像データが複数の階層に分類されて符号化されることで得られた各階層のピクチャの画像データを持つビデオストリームが含まれている。各階層のピクチャの符号化画像データには、高階層ほどピクチャ毎の符号化画像データのデコード時間間隔が小さくなるように設定されたデコードタイミング情報、例えばデコードタイムスタンプが付加されている。

【0017】

処理部により、受信コンテナに含まれるビデオストリームから選択された所定階層以下の階層のピクチャの符号化画像データがデコードされて各ピクチャの画像データが得られる。この場合、各ピクチャの符号化画像データのデコードは、

50

遅くとも、それぞれ、付加されているデコードタイミング情報が示すデコードタイミングで行われる。

【0018】

例えば、受信されたコンテナには、各階層のピクチャの符号化画像データを持つ単一のビデオストリームが含まれており、複数の階層は2以上の所定数の階層組に分割され、低階層側の階層組のピクチャの符号化画像データをコンテナするパケットの優先度ほど高く設定されており、処理部は、デコード能力に応じて選択された優先度のパケットでコンテナされた所定階層組のピクチャの符号化画像データをバッファに取り込んでデコードする、ようにされてもよい。

【0019】

また、例えば、受信されたコンテナには、複数の階層が分割されて得られた2以上の所定数の階層組のピクチャの画像データをそれぞれ持つ所定数のビデオストリームが含まれており、処理部は、デコード能力に応じて選択されたビデオストリームが持つ所定階層組のピクチャの符号化画像データをバッファに取り込んでデコードする、ようにされてもよい。

【0020】

このように本技術においては、各階層のピクチャの符号化画像データには高階層ほどピクチャ毎の符号化画像データのデコード時間間隔が小さくなるように設定されたデコードタイムスタンプが付加されており、選択された所定階層以下の階層のピクチャの符号化画像データがそれに付加されているデコードタイミング情報が示すデコードタイミングで行われるものである。そのため、デコード性能に応じた良好なデコード処理が可能となる。例えば、デコード能力が低い場合であっても、バッファ破たんを招くことなく、低階層のピクチャの符号化画像データを選択的にデコードすることが可能となる。

【0021】

なお、本技術において、例えば、処理部で得られる各ピクチャの画像データのフレームレートを表示能力に合わせるポスト処理部をさらに備える、ようにされてもよい。この場合、デコード能力が低い場合であっても、高表示能力にあったフレームレートの画像データを得ることが可能となる。

【0022】

また、本技術の他の概念は、

動画像データを構成する各ピクチャの画像データを複数の階層に分類し、該分類された各階層のピクチャの画像データを符号化し、該符号化された各階層のピクチャの画像データを持つビデオストリームを生成する画像符号化部を備え、

上記画像符号化部は、上記複数の階層を2以上の所定数の階層組に分割し、各階層組に対応したサブストリームのそれぞれにビットストリームのレベル指定値を挿入し、

上記各階層組に対応したサブストリームのそれぞれに挿入されるビットストリームのレベル指定値は、自己の階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値とされる

符号化装置にある。

【0023】

本技術において、画像符号化部により、動画像データを構成する各ピクチャの画像データが複数の階層に分類され、この分類された各階層のピクチャの画像データが符号化され、この符号化された各階層のピクチャの画像データを持つビデオストリームが生成される。

【0024】

この場合、複数の階層は2以上の所定数の階層組に分割され、各階層組に対応したサブストリームのそれぞれにビットストリームのレベル指定値が挿入される。そして、この場合、各階層組に対応したサブストリームのそれぞれに挿入されるビットストリームのレベル指定値は、自己の階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値とされる。

10

20

30

40

50

【0025】

例えば、画像符号化部は、各階層組に対応したサブストリームのそれぞれを含む所定数のビデオストリームを生成する、ようにされてもよい。また、例えば、画像符号化部は、各階層組に対応したサブストリームの全てを含む単一のビデオストリームを生成する、ようにされてもよい。

【0026】

このように本技術においては、各階層組に対応したサブストリームのそれぞれにビットストリームのレベル指定値が挿入され、その値は自己の階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値とされるものである。そのため、ビデオストリームの受信側では、各サブストリームのデコードが可能か否かの判断を、挿入されているビットストリームのレベル指定値に基づいて容易に判断することが可能となる。

10

【0027】

また、本技術の他の概念は、

動画像データを構成する各ピクチャの画像データを複数の階層に分類し、該分類された各階層のピクチャの画像データを符号化し、該符号化された各階層のピクチャの画像データを持つビデオストリームを生成する画像符号化部を備え、

上記画像符号化部は、上記複数の階層を2以上の所定数の階層組に分割し、各階層組に対応したサブストリームのそれぞれにビットストリームのレベル指定値を挿入し、

上記各階層組に対応したサブストリームのそれぞれに挿入されるビットストリームのレベル指定値は、当該階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値とされ、

20

上記生成されたビデオストリームを含む所定フォーマットのコンテナを送信する送信部と、

上記コンテナのレイヤに、各階層組のサブストリームに挿入されるビットストリームのレベル指定値が当該階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値であることを示すフラグ情報を挿入する情報挿入部をさらに備える

送信装置にある。

【0028】

本技術において、画像符号化部により、動画像データを構成する各ピクチャの画像データが複数の階層に分類され、この分類された各階層のピクチャの画像データが符号化され、この符号化された各階層のピクチャの画像データを持つビデオストリームが生成される。

30

【0029】

この場合、複数の階層は2以上の所定数の階層組に分割され、各階層組に対応したサブストリームのそれぞれにビットストリームのレベル指定値が挿入される。そして、この場合、各階層組に対応したサブストリームのそれぞれに挿入されるビットストリームのレベル指定値は、当該階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値とされる。

【0030】

送信部により、生成されたビデオストリームを含む所定フォーマットのコンテナが送信される。情報挿入部により、コンテナのレイヤに、各階層組のサブストリームに挿入されるビットストリームのレベル指定値が当該階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値であることを示すフラグ情報が挿入される。

40

【0031】

このように本技術においては、受信側では、コンテナのレイヤに挿入されるフラグ情報により、各階層組のサブストリームに挿入されるビットストリームのレベル指定値が、当該階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値であることがわかる。そのため、受信側では、所定の階層組以下の各サブストリームに含まれる全ての階層のピクチャを含むレベル値を階層毎のレベル指定値を用いるなどして確認する処理が不要となり、デコード処理の効率化を図ることが可能となる。

50

【発明の効果】

【0032】

本技術によれば、デコード能力に応じた良好なデコード処理が可能となる。なお、ここに記載された効果は必ずしも限定されるものではなく、本開示中に記載されたいずれかの効果であってもよい。

【図面の簡単な説明】

【0033】

【図1】実施の形態としての送受信システムの構成例を示すブロック図である。

【図2】送信装置の構成例を示すブロック図である。

【図3】エンコーダで行われる階層符号化の一例を示す図である。 10

【図4】NALユニットヘッダの構造例(Syntax)およびその構造例における主要なパラメータの内容(Semantics)を示す図である。

【図5】HEVCによる各ピクチャの符号化画像データの構成を説明するための図である。

【図6】階層符号化の際のエンコード、デコード、表示順序と遅延の一例を示す図である。

【図7】階層符号化の符号化ストリームと、指定階層における表示期待(表示順)を示す図である。

【図8】エンコーダ入力順と、指定階層におけるデコーダ出力の表示順を示す図である。

【図9】階層符号化の際のピクチャのエンコードタイミング(デコード時にはデコードタイミングとなる)の一例を示す図である。 20

【図10】エンコーダの単一のビデオストリーム(符号化ストリーム)の出力例を示す図である。

【図11】エンコーダのベースストリーム(B-stream)と拡張ストリーム(E-stream)の2つのビデオストリーム(符号化ストリーム)の出力例を示す図である。

【図12】エンコーダの構成例を示すブロック図である。

【図13】エンコーダの処理フローの一例を示す図である。

【図14】HEVCデスクリプタ(HEVC_descriptor)の構造例(Syntax)を示す図である。

【図15】HEVCデスクリプタの構造例における主要な情報の内容(Semantics)を示す図である。 30

【図16】スケーラビリティ・エクステンション・デスクリプタ(scability_extension_descriptor)の構造例(Syntax)を示す図である。

【図17】スケーラビリティ・エクステンション・デスクリプタの構造例における主要な情報の内容(Semantics)を示す図である。

【図18】TSパケットの構造例(Syntax)を示す図である。

【図19】VPSに含まれるビットレートのレベル指定値(general_level_idc)と、TSパケットヘッダの「transport_priority」の設定値との関係を示す図である。

【図20】マルチプレクサの構成例を示すブロック図である。

【図21】マルチプレクサの処理フローの一例を示す図である。 40

【図22】単一ストリームによる配信を行う場合のトランスポートストリームTSの構成例を示す図である。

【図23】単一ストリームによる配信を行う場合のトランスポートストリームTSの具体的な構成例を示す図である。

【図24】複数ストリーム(2ストリーム)による配信を行う場合のトランスポートストリームTSの構成例を示す図である。

【図25】2ストリームによる配信を行う場合のトランスポートストリームTSの具体的な構成例を示す図である。

【図26】2ストリームによる配信を行う場合のトランスポートストリームTSの他の具体的な構成例を示す図である。 50

【図27】受信装置の構成例を示すブロック図である。

【図28】デマルチプレクサの構成例を示すブロック図である。

【図29】トランSPORTストリームTSに単一のビデオストリーム(符号化ストリーム)が含まれている場合を示す図である。

【図30】トランSPORTストリームTSにベースストリームと拡張ストリームの2つのビデオストリーム(符号化ストリーム)が含まれている場合を示す図である。

【図31】デマルチプレクサの処理フロー(1フレーム)の一例を示す図である。

【図32】デマルチプレクサの処理フロー(2フレーム)の一例を示す図である。

【図33】デコーダの構成例を示すブロック図である。

【図34】受信装置におけるデコーダ処理能力を考慮したビデオストリーム毎のデコード処理手順の一例を示すフローチャートである。 10

【図35】ポスト処理部の構成例を示す図である。

【図36】デコーダ、ポスト処理部の処理フローの一例を示す図である。

【発明を実施するための形態】

【0034】

以下、発明を実施するための形態(以下、「実施の形態」とする)について説明する。なお、説明は以下の順序で行う。

1. 実施の形態

2. 変形例

【0035】

<1. 実施の形態>

[送受信システム]

図1は、実施の形態としての送受信システム10の構成例を示している。この送受信システム10は、送信装置100と、受信装置200とを有する構成となっている。

【0036】

送信装置100は、コンテナとしてのトランSPORTストリームTSを放送波に載せて送信する。このトランSPORTストリームTSには、動画像データを構成する各ピクチャの画像データが複数の階層に分類され、各階層のピクチャの画像データの符号化データを持つビデオストリームが含まれる。この場合、例えば、H.264/AVC、H.265/HVECなどの符号化が施され、被参照ピクチャが自己階層および/または自己階層よりも低い階層に所属するように符号化される。 30

【0037】

各階層のピクチャの符号化画像データに、ピクチャ毎に、所属階層を識別するための階層識別情報が付加される。この実施の形態においては、各ピクチャのNALユニット(nal_unit)のヘッダ部分に、階層識別情報(temporal_id)を意味する“nuh_temporal_id_plus1”が配置される。このように階層識別情報が付加されることで、受信側では、所定階層以下の階層の符号化画像データを選択的に取り出してデコード処理を行うことができる。

【0038】

トランSPORTストリームTSには、各階層のピクチャの符号化画像データを持つ単一のビデオストリーム、あるいは複数の階層が2以上の所定数の階層組に分割され、各階層組のピクチャの符号化画像データをそれぞれ持つ所定数のビデオストリームが含まれる。また、このトランSPORTストリームTSには、階層符号化の階層情報と、ビデオストリームの構成情報が挿入される。この情報は、トランSPORTレイヤに挿入される。この情報により、受信側では、階層構成やストリーム構成を容易に把握でき、適切なデコード処理を行うことが可能となる。 40

【0039】

また、上述したように複数の階層が所定数の階層組に分割され、低階層側の階層組のピクチャの符号化画像データをコンテナするTSパケット(トランSPORTストリームパケット)の優先度が高く設定される。この優先度により、受信側では、自身のデコード能力

10

20

30

40

50

に応じた階層組のピクチャの符号化画像データのみをバッファに取り込んで処理することが可能となる。

【0040】

また、上述したように複数の階層が所定数の階層組に分割され、各階層組のピクチャの符号化画像データに、所属階層組を識別するための識別情報が付加される。この識別情報として、例えば、ビットストリームのレベル指定値 (level_idc) が利用され、高階層側の階層組ほど高い値とされる。

【0041】

受信装置 200 は、送信装置 100 から放送波に載せて送られてくる上述のトランスポートストリーム TS を受信する。受信装置 200 は、このトランSPORTストリーム TS に含まれるビデオストリームから自身のデコード能力に応じて選択的に所定階層以下の階層の符号化画像データを取り出してデコードし、各ピクチャの画像データを取得して、画像再生を行う。
10

【0042】

例えば、上述したように、トランSPORTストリーム TS に、複数の階層のピクチャの符号化画像データを持つ単一のビデオストリームが含まれている場合がある。その場合、デコード能力に応じて選択された優先度の TS パケットでコンテナされた所定階層組のピクチャの符号化画像データがバッファに取り込まれてデコードされる。

【0043】

また、例えば、上述したように、トランSPORTストリーム TS に、複数の階層が分割されて得られた 2 以上の所定数の階層組のピクチャの符号化画像データをそれぞれ持つ所定数のビデオストリームが含まれている場合がある。その場合、デコード能力に応じて選択されたビデオストリームが持つ所定階層組のピクチャの符号化画像データがバッファに取り込まれてデコードされる。
20

【0044】

また、受信装置 200 は、上述のようにデコードして得られた各ピクチャの画像データのフレームレートを表示能力に合わせるポスト処理を行う。このポスト処理により、例えば、デコード能力が低い場合であっても、高表示能力にあったフレームレートの画像データを得ることが可能となる。

【0045】

「送信装置の構成」

図 2 は、送信装置 100 の構成例を示している。この送信装置 100 は、CPU (Central Processing Unit) 101 と、エンコーダ 102 と、圧縮データバッファ (cpb : coded picture buffer) 103 と、マルチプレクサ 104 と、送信部 105 を有している。CPU 101 は、制御部であり、送信装置 100 の各部の動作を制御する。
30

【0046】

エンコーダ 102 は、非圧縮の動画像データを入力して、階層符号化を行う。エンコーダ 102 は、この動画像データを構成する各ピクチャの画像データを複数の階層に分類する。そして、エンコーダ 102 は、この分類された各階層のピクチャの画像データを符号化し、各階層のピクチャの符号化画像データを持つビデオストリームを生成する。エンコーダ 102 は、例えば、H.264 / AVC、H.265 / HEVC などの符号化を行う。この際、エンコーダ 102 は、参照するピクチャ (被参照ピクチャ) が、自己階層および / または自己階層よりも下位の階層に所属するように、符号化する。
40

【0047】

図 3 は、エンコーダ 102 で行われる階層符号化の一例を示している。この例は、0 から 4 までの 5 階層に分類され、各階層のピクチャの画像データに対して符号化が施された例である。

【0048】

縦軸は階層を示している。階層 0 から 4 のピクチャの符号化画像データを構成する NAL ユニット (nal_unit) のヘッダ部分に配置される temporal_id (階層識別情報) として

、それぞれ、0から4が設定される。一方、横軸は表示順(P O C : picture order of composition)を示し、左側は表示時刻が前で、右側は表示時刻が後になる。

【0049】

図4(a)は、NALユニットヘッダの構造例(Syntax)を示し、図4(b)は、その構造例における主要なパラメータの内容(Semantics)を示している。「Forbidden_zero_bit」の1ビットフィールドは、0が必須である。「Nal_unit_type」の6ビットフィールドは、NALユニットタイプを示す。「Nuh_layer_id」の6ビットフィールドは、0を前提とする。「Nuh_temporal_id_plus1」の3ビットフィールドは、temporal_idを示し、1を加えた値(1~7)をとる。

【0050】

図3に戻って、矩形枠のそれぞれがピクチャを示し、数字は、符号化されているピクチャの順、つまりエンコード順(受信側ではデコード順)を示している。例えば、「2」から「17」の16個のピクチャによりサブ・ピクチャグループ(Sub group of pictures)が構成されており、「2」はそのサブ・ピクチャグループの先頭のピクチャとなる。「1」は前のサブ・ピクチャグループのピクチャである。このサブ・ピクチャグループがいくつか集まってGOP(Group Of Pictures)となる。

【0051】

GOPの先頭ピクチャの符号化画像データは、図5に示すように、AUD、VPS、SPS、PPS、PSEI、SLICE、SSEI、EOSのNALユニットにより構成される。一方、GOPの先頭ピクチャ以外のピクチャは、AUD、PPS、PSEI、SLICE、SSEI、EOSのNALユニットにより構成される。VPSはSPSと共に、シーケンス(GOP)に一度、PPSは毎ピクチャで伝送可能とされている。

【0052】

図3に戻って、実線矢印は、符号化におけるピクチャの参照関係を示している。例えば、「2」のピクチャは、Pピクチャであり、「1」のピクチャを参照して符号化される。また、「3」のピクチャは、Bピクチャであり、「1」、「2」のピクチャを参照して符号化される。同様に、その他のピクチャは、表示順で近くのピクチャを参照して符号化される。なお、階層4のピクチャは、他のピクチャからの参照がない。

【0053】

エンコーダ102は、各階層のピクチャの符号化画像データを持つビデオストリームを生成する。例えば、エンコーダ102は、複数の階層を2以上の所定数の階層組に分割し、各階層組に対応したサブストリームのそれぞれを含む所定数のビデオストリームを生成するか、または、各階層組に対応したサブストリームの全てを含む単一のビデオストリームを生成する。

【0054】

例えば、図3の階層符号化の例において、階層0から3が低階層の階層組とされ、階層4が高階層の階層組とされて2つの階層組に分割されるとき、2つのサブストリームが存在する。すなわち、階層0から3のピクチャの符号化画像データを持つサブストリームと、階層4のピクチャの符号化画像データを持つサブストリームである。この場合、エンコーダ102は、2つのサブストリームを含む単一のビデオストリーム、または2つのサブビデオストリームをそれぞれ含む2つのビデオストリームを生成する。

【0055】

エンコーダ102は、生成するビデオストリームの数によらず、上述したように、複数の階層を2以上の所定数の階層組に分割し、各階層組のピクチャの符号化画像データに、所属階層組を識別するための識別情報を付加する。この場合、例えば、識別情報として、SPS(Sequence Parameter Set)と、EPS(Enhanced Sequence Parameter Set)に含まれるビットストリームのレベル指定値である「general_level_idc」が利用される。

【0056】

SPSは、従来周知のNALユニットであり、最下位の階層組のサブストリーム、つまり

10

20

30

40

50

リベースサブストリームに、シーケンス(G O P)毎に含まれる。一方、 E S P S は、新規定義する N A L ユニットであり、最下位より上位の階層組のサブストリーム、つまりエンハンスサブストリームに、シーケンス(G O P)毎に含まれる。 S P S , E S P S に含まれる「 general_level_idc 」の値は、高階層側の階層組ほど高い値とされる。

【 0 0 5 7 】

なお、サブレイヤ(sublayer)毎に「 sub_layer_level_idc 」を S P S , E S P S で送ることができるので、階層組を識別する識別情報として、この「 sub_layer_level_idc 」を用いることも可能である。以上は S P S だけでなく V P S においても供給される。

【 0 0 5 8 】

この場合、各階層組のサブストリームの S P S , E S P S に挿入される「 general_level_idc 」の値は、自己の階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値とされる。例えば、図 3 の階層符号化の例において、階層 0 から 3 の階層組のサブストリームの S P S に挿入される「 general_level_idc 」の値は、階層 0 から 3 のピクチャのみを含むレベル値とされる。例えば、そのフレームレートが 6 0 P であるときは、“ 1 e v e l 5 . 1 ”とされる。また、例えば、図 3 の階層符号化の例において、階層 4 の階層組のサブストリームの E S P S に挿入される「 general_level_idc 」の値は、階層 0 から 4 の全てのピクチャを含むレベル値とされる。例えば、そのフレームレートが 1 2 0 P であるときは、“ 1 e v e l 5 . 2 ”とされる。

【 0 0 5 9 】

図 6 は、階層符号化の際のエンコード、デコード、表示順序と遅延の一例を示している。この例は、上述の図 3 の階層符号化例に対応している。この例は、全階層(全レイヤ)を、フル時間解像度で階層符号化する場合を示している。図 6 (a)はエンコーダ入力を示す。図 6 (b)に示すように、16ピクチャ分の遅延をもって、各ピクチャがエンコード順にエンコードされて、符号化ストリームが得られる。また、図 6 (b)はデコーダ入力も示し、各ピクチャがデコード順にデコードされる。そして、図 6 (c)に示すように、4ピクチャの遅延をもって、各ピクチャがデコードの画像データが表示順に得られる。

【 0 0 6 0 】

図 7 (a)は、上述の図 6 (b)に示す符号化ストリームと同様の符号化ストリームを、階層 0 から 2 、階層 3 、階層 4 の 3 段階に分けて示している。ここで、「 T i d 」は、 temporal_id を示している。図 7 (b)は、階層 0 から 2 、つまり T i d = 0 ~ 2 の部分階層の各ピクチャを選択的にデコードする場合の表示期待(表示順)を示している。また、図 7 (c)は、階層 0 から 3 、つまり T i d = 0 ~ 3 の部分階層の各ピクチャを選択的にデコードする場合の表示期待(表示順)を示している。さらに、図 7 (d)は、階層 0 から 4 、つまり T i d = 0 ~ 4 の全階層の各ピクチャを選択的にデコードする場合の表示期待(表示順)を示している。

【 0 0 6 1 】

図 7 (a)の符号化ストリームをデコード能力別にデコード処理するには、時間解像度がフルレートのデコード能力が必要となる。しかし、 T i d = 0 ~ 2 のデコードを行う場合、符号化されたフルの時間解像度に対して、 1 / 4 のデコード能力をもつデコーダが処理可能とすべきである。また、 T i d = 0 ~ 3 のデコードを行う場合、符号化されたフルの時間解像度に対して、 1 / 2 のデコード能力をもつデコーダが処理可能とすべきである。

【 0 0 6 2 】

しかし、階層符号化において参照される低階層に属するピクチャが連続し、それらが時間解像度でフルなタイミングで符号化されると、部分デコードするデコーダの能力が追いつかないことになる。図 7 (a)の A の期間がそれに該当する。 T i d = 0 ~ 2 、あるいは T i d = 0 ~ 3 の部分的な階層をデコードするデコーダは、表示の例で示すような、時間軸が 1 / 4 あるいは 1 / 2 の能力でデコード・表示を行うため、 A の期間符号化された時間解像度がフルで連続するピクチャのデコードはできない。その間、 c p b にエンコーダの想定外のバッファ占有量になる。

10

20

30

40

50

【0063】

T_a は $T_{id} = 0 \sim 2$ をデコードするデコーダにおけるピクチャ毎のデコード処理に要する時間を示す。 T_b は $T_{id} = 0 \sim 3$ をデコードするデコーダにおけるピクチャ毎のデコード処理に要する時間を示す。 T_c は $T_{id} = 0 \sim 4$ (全階層) をデコードするデコーダにおけるピクチャ毎のデコード処理に要する時間を示す。これらの各時間の関係は、 $T_a > T_b > T_c$ となる。

【0064】

そこで、この実施の形態においては、階層符号化の低階層に属するピクチャは、ピクチャ毎のデコード間隔を大きく取り、高階層に行くにつれ、デコード間隔が小さくなるようにバッファ制御を行う。その際、階層数に対してのミニマムデコード能力 (Target minimum decoder capability) を定義する。例えば、図3の階層符号化の例において、ミニマムデコード能力が階層2までデコードできる能力とすると、5階層のうちの1/4の時間解像度で階層0から2のピクチャをデコードできるように、エンコード時の間隔をとり、後述のマルチプレクサ104で多重化する際に、そのデコード時刻の差をデコードタイムスタンプ (DTS : decoding Time stamp) の値に反映させる。

10

【0065】

図3の階層符号化例のように、階層数が0～4の5つである場合、階層0から2に属するピクチャの間隔はフル解像度の4倍の時間間隔とし、階層3に属するピクチャの間隔は、フル解像度の2倍の時間間隔とし、階層4に属するピクチャの間隔は、フル解像度の時間間隔とする。

20

【0066】

一方で、エンコーダ102は、階層間でピクチャのエンコード (=デコード) のタイミングが重ならないようにする。すなわち、エンコーダ102は、上述の方法で各ピクチャのエンコードを行う際、低階層のピクチャと高階層のピクチャとでエンコードタイミングが重なる場合には、より多くのピクチャから参照される低階層のピクチャのエンコードを優先し、高階層のピクチャは、それに準じたタイミングとする。ただし、最高階層に属するピクチャは非参照のBピクチャなので、デコードしてそのまま表示 (つまりdpb (decoded picture buffer) に貯めない) とするタイミングとなるよう制御することが可能となる。

30

【0067】

図8(a)はエンコーダ入力順を示している(図6(a)と同じ)。また、図8(b)～(d)は、表示順(システムレイヤとしてはPTSに相当)を示している(図7(b)～(d)と同じ)。

【0068】

図9は、階層符号化の際のピクチャのエンコードタイミング(デコード時にはデコードタイミングとなる)の一例を示している。この例は、上述の図3の階層符号化例に対応している。そして、この例は、ミニマムデコード能力を、階層2までデコードできる能力としたものである。実線による下線が付された部分は、1つのSGP (Sub Group of Picture) に属するピクチャ(「2」～「17」の16個のピクチャ)を示している。また、実線矩形枠で示すピクチャは現在のSGPに属しているが、破線矩形枠で示すピクチャは現在のSGPには属してはおらず、現在のSGPに属するピクチャによる予測には影響を与えない。

40

【0069】

この場合、階層0から2に属するピクチャ、つまりピクチャ「2」、「3」、「4」、「11」・・・の間隔は、フル解像度の4倍の時間間隔である T_a とされる。また、階層3に属するピクチャ、つまり「5」、「8」、「12」・・・の間隔は、基本的には、フル解像度の2倍の時間間隔である T_b とされる。

【0070】

しかし、「8」のピクチャのタイミングは、「11」のピクチャのタイミングとの重なりを避けるために、エンコードのタイミングが、次の時間間隔位置とされる。以下、同様

50

に、「12」、「15」のピクチャのタイミングも、階層0から2に属するピクチャとの重なりを避けるように調整される。この結果、階層3に属するピクチャのタイミングは、階層0から2に属するピクチャのタイミングの中間とされる。

【0071】

また、階層4に属するピクチャ、つまり「6」、「7」、「9」…の間隔は、基本的には、フル解像度の時間間隔であるTcとされる。しかし、階層0から3に属する各ピクチャのタイミングとの重なりを避けるように調整される結果、階層4に属するピクチャのタイミングは、階層0から3に属するピクチャのタイミングの中間とされる。

【0072】

図示のように、1SGP期間で、1SGP分のピクチャ（「2」～「17」の16個のピクチャ）のエンコード処理が行われる。これは、上述したように低階層に属するピクチャのエンコード間隔を大きく取る場合であっても実時間処理が可能であることを示している。10

【0073】

図10は、エンコーダ102の出力例を示している。この例は、エンコーダ102が单一のビデオストリーム（符号化ストリーム）を出力する例である。この例は、図3の階層符号化例に対応し、図9に示すタイミングで各ピクチャがエンコードされた場合の例である。

【0074】

ビデオストリームは、階層0から4に属する各ピクチャの符号化画像データがエンコード順（符号化順）に配列されたものになっている。なお、このビデオストリームを受信側でデコードするに当たっては、現在のSGP（太線枠のピクチャ）に属する被参照ピクチャ（階層0から3のピクチャ）はデコード後に非圧縮データバッファ（dpb：decoded picture buffer）に滞留し、他のピクチャからの参照に備える。20

【0075】

図11は、エンコーダ102の出力例を示している。この例は、エンコーダ102がベースストリーム（B_str）と拡張ストリーム（E_str）の2つのビデオストリーム（符号化ストリーム）を出力する例である。この例は、図3の階層符号化例に対応し、図9に示すタイミングで各ピクチャがエンコードされた場合の例である。

【0076】

ベースストリーム（B-stream）は、階層0から3に属する各ピクチャの符号化画像データがエンコード順（符号化順）に配列されたものになっている。また、拡張ストリーム（E-stream）は、階層4に属する各ピクチャの符号化画像データがエンコード順（符号化順）に配列されたものになっている。なお、このビデオストリームを受信側でデコードするに当たっては、現在のSGP（太線枠のピクチャ）に属する被参照ピクチャ（階層0から3のピクチャ）はデコード後に非圧縮画像データバッファ（dpb：decoded picture buffer）に滞留し、他のピクチャからの参照に備える。30

【0077】

図12は、エンコーダ102の構成例を示している。このエンコーダ102は、テンポラルID発生部121と、バッファ遅延制御部122と、HRD（Hypothetical Reference Decoder）設定部123と、パラメータセット/SEIエンコード部124と、スライスエンコード部125と、NALパケット化部126を有している。40

【0078】

テンポラルID発生部121には、CPU101から、階層数（Number of layers）の情報が供給される。テンポラルID発生部121は、この階層数の情報に基づいて、階層数に応じたtemporal_idを発生する。例えば、図3の階層符号例においては、temporal_id = 0～4が発生される。

【0079】

バッファ遅延制御部122には、CPU101から、ミニマムデコード能力（minimum_target_decoder_level_idc）の情報が供給されると共に、テンポラルID発生部121で50

発生される temporal_id が供給される。バッファ遅延制御部 122 は、階層毎に、各ピクチャの「cpb_removal_delay」、「dpb_output_delay」を計算する。

【0080】

この場合、階層数に対しての、それをデコードするターゲットデコーダのミニマムデコード能力が指定されることで、被参照の低階層ピクチャのエンコードタイミング、そしてデコード即表示の高階層ピクチャのエンコードタイミングが決定される(図9参照)。このエンコードタイミングは、受信側で圧縮データバッファ(c p b : coded picture buffer)から読み出されるデコードタイミングと同じ意味を示す。

【0081】

「cpb_removal_delay」は、ピクチャが属する階層を反映して決められる。例えば、階層数を N とし、temporal_id(T i d) が 0 ~ N - 1 の範囲の値をとるものとする。また、ミニマムデコード能力は、temporal_id = K の階層のピクチャをデコードできる能力とする。バッファ遅延制御部 122 は、各階層におけるピクチャのエンコード間隔 D を、以下の式(1)で求め、「cpb_removal_delay」、「dpb_output_delay」に反映させる。

【0082】

$$\begin{aligned}
 D &= 2^{**}(N-1 - K) && (Tid \leq K) \\
 D &= 2^{**}(N-1 - Tid) && (K < Tid < N - 1) \\
 D &= \text{入力シーケンス間隔} && (Tid = N ? 1) \\
 &&& \dots (1)
 \end{aligned}$$

【0083】

なお、階層間で、時間的にエンコードタイミングが重なってしまう場合は、低階層側が優先的にエンコードされ、高階層側は上記の式で割り当てられる、次のタイムスロット(timeslot)でエンコードされる。

【0084】

H R D (Hypothetical Reference Decoder) 設定部 123 には、バッファ遅延制御部 122 で計算された各階層のピクチャの「cpb_removal_delay」、「dpb_output_delay」が供給されると共に、C P U 1 0 1 からストリーム数(Number of streams)の情報が供給される。H R D 設定部 123 は、これらの情報に基づいて H R D 設定を行う。

【0085】

パラメータセット/ S E I エンコード部 124 には、H R D 設定情報と共に、temporal_id が供給される。パラメータセット/ S E I エンコード部 124 は、符号化するストリーム数に応じて、各階層のピクチャの V P S 、 S P S (E S P S) 、 P P S などのパラメータセットと S E I を生成する。

【0086】

例えば、「cpb_removal_delay」と「dpb_output_delay」を含むピクチャ・タイミング・S E I (Picture timing SEI) が生成される。また、例えば、「initial_cpb_removal_time」を含むバッファリング・ピリオド・S E I (Buffereing Perifod SEI) が生成される。バッファリング・ピリオド・S E I は、G O P の先頭のピクチャ(アクセスユニット)に対応して生成される。

【0087】

「initial_cpb_removal_time」は、圧縮データバッファ(c p b)から G O P (Group Of Picture)の先頭のピクチャの符号化画像データをデコードする際に取り出す時刻(初期時刻)を示す。「cpb_removal_delay」は、各ピクチャの符号化画像データを圧縮データバッファ(c p b)から取り出す時間であり、「initial_cpb_removal_time」と合わせて時刻が決まる。また、「dpb_output_delay」は、デコードして非圧縮データバッファ(d p b)に入つてから取り出す時間を示す。

【0088】

スライスエンコード部 125 は、各階層のピクチャの画像データをエンコードしてスライスデータ(slice segment header, slice segment data)を得る。スライスデコード部 125 は、フレームバッファにより、時間方向の予測の状態を表す情報として、「Predic

10

20

30

40

50

tion Unit」の予測先ピクチャのインデックスを示す「ref_idx_10_active(ref_idx_11_active)」を、「slice segment header」に挿入する。これにより、デコード時には、temporal_idで示される階層レベルと共に、被参照ピクチャが決定される。また、スライスデコード部125は、現在のスライス(slice)のインデックスを、「short_term_ref_pic_set_idx」、あるいは「it_idx_sps」として、「slice segment header」に挿入する。

【0089】

NALパケット化部126は、パラメータセット/SEIエンコード部124で生成されたパラメータセットおよびSEIと、スライスエンコード部125で生成されるスライスデータに基づき、各階層のピクチャの符号化画像データを生成し、ストリーム数に応じた数のビデオストリーム(符号化ストリーム)を出力する。

10

【0090】

その際、ピクチャごとに、その階層を示すtemporal_idがNALユニットヘッダに付される(図4参照)。また、temporal_idで示される階層に属するピクチャは、サブレイヤ(sub_layer)として括られ、サブレイヤごとのビットレートのレベル指定値「Level_idc」が「sublayer_level_idc」とされて、VPSやSPS(ESPS)に挿入される。

【0091】

図13は、エンコーダ102の処理フローを示す。エンコーダ102は、ステップST1において、処理を開始し、その後に、ステップST2の処理に移る。このステップST2において、エンコーダ102は、階層符号化における階層数Nを設定する。次に、エンコーダ102は、ステップST3において、各階層のピクチャのtemporal_idを0~(N-1)とする。

20

【0092】

次に、エンコーダ102は、ステップST4において、対象デコーダのうち、最小能力のデコーダがデコードできる階層レベルKを、0~N-1の範囲内に設定する。そして、エンコーダ102は、ステップST5において、バッファ遅延制御部122で、各階層におけるピクチャエンコード間隔Dを、上述の数式(1)で求める。

【0093】

次に、エンコーダ102は、ステップST6において、階層間でピクチャのエンコードタイミングが時間的に重なるか否かを判断する。エンコードタイミングが重なるとき、エンコーダ102は、ステップST7において、低階層のピクチャを優先して符号化し、高階層のピクチャは、次のエンコード間隔Dのタイミングでエンコードする。その後、エンコーダ102は、ステップST8の処理に移る。

30

【0094】

ステップST6でエンコードタイミングが重ならないとき、エンコーダ102は、直ちに、ステップST8の処理に移る。このステップST8において、エンコーダ102は、ステップST5で求めた各階層のピクチャのエンコード間隔Dを「cpb_removal_delay」、「dpb_output_delay」に反映し、HRD設定、パラメータセット/SEIのエンコード、スライスエンコードを行い、NALユニットとして多重化ブロックへ転送する。その後、エンコーダ102は、ステップST9において、処理を終了する。

40

【0095】

図2に戻って、圧縮データバッファ(cpb)103は、エンコーダ102で生成された、各階層のピクチャの符号化データを含むビデオストリームを、一時的に蓄積する。マルチプレクサ104は、圧縮データバッファ103に蓄積されているビデオストリームを読み出し、PESパケット化し、さらにトランスポートパケット化して多重し、多重化ストリームとしてのトランスポートストリームTSを得る。

【0096】

このトランスポートストリームTSには、各階層のピクチャの符号化画像データを持つ単一のビデオストリーム、あるいは複数の階層が2以上の所定数の階層組に分割され、各階層組のピクチャの符号化画像データをそれぞれ持つ所定数のビデオストリームが含まれる。マルチプレクサ104は、トランスポートストリームTSに、階層情報、ストリーム

50

構成情報を挿入する。

【0097】

トランスポートストリーム TS には、PSI (Program Specific Information) として、PMT (Program Map Table) が含まれている。この PMT には、各ビデオストリームに関連した情報を持つビデオエレメンタリ・ループ (video ES1 loop) が存在する。このビデオエレメンタリ・ループには、各ビデオストリームに対応して、ストリームタイプ、パケット識別子 (PID) 等の情報が配置されると共に、そのビデオストリームに関連する情報を記述するデスクリプタも配置される。

【0098】

マルチプレクサ 104 は、このデスクリプタの一つとして、HEVC デスクリプタ (HEVC_descriptor) を挿入し、さらに、新たに定義するスケーラビリティ・エクステンション・デスクリプタ (scalability_extension_descriptor) を挿入する。

【0099】

図 14 は、HEVC デスクリプタ (HEVC_descriptor) の構造例 (Syntax) を示している。また、図 15 は、その構造例における主要な情報の内容 (Semantics) を示している。

【0100】

「descriptor_tag」の 8 ビットフィールドは、デスクリプタタイプを示し、ここでは、HEVC デスクリプタであることを示す。「descriptor_length」の 8 ビットフィールドは、デスクリプタの長さ (サイズ) を示し、デスクリプタの長さとして、以降のバイト数を示す。

【0101】

「level_idc」の 8 ビットフィールドは、ビットレートのレベル指定値を示す。また、「temporal_layer_subset_flag = 1」であるとき、「temporal_id_min」の 5 ビットフィールドと、「temporal_id_max」の 5 ビットフィールドが存在する。「temporal_id_min」は、対応するビデオストリームに含まれる階層符号化データの最も低い階層の temporal_id の値を示す。「temporal_id_max」は、対応するビデオストリームが持つ階層符号化データの最も高い階層の temporal_id の値を示す。

【0102】

「level_constrained_flag」の 1 ビットフィールドは、新たに定義するものであり、該当サブストリーム (substream) に SPS あるいは EPPS が存在し、その要素の “general_level_idc” は、そのサブストリームが含む temporal_id (階層識別情報) 以下のピクチャ (Picture) を含むレベル (Level) 値をもつことを示す。“1” は、該当サブストリームに SPS あるいは EPPS が存在し、その要素の “general_level_idc” は、そのサブストリームが含む temporal_id 以下のピクチャを含むレベル値をもつ、ことを示す。“0” は、対象となるサービスを構成するサブストリーム群の中には SPS が 1 つ存在し、その “general_level_idc” は、当該サブストリームのみならず、同一サービスの下の他のサブストリームも含むレベル値を示す。

【0103】

「scalability_id」の 3 ビットフィールドは、新たに定義するものであり、複数のビデオストリームがスケーラブルなサービスを供給する際、個々のストリームに付されるスケーラビリティを示す ID である。“0” はベースストリームを示し、“1” ~ “7” はベースストリームからのスケーラビリティの度合いによって増加する ID である。

【0104】

図 16 は、スケーラビリティ・エクステンション・デスクリプタ (scalability_extension_descriptor) の構造例 (Syntax) を示している。また、図 17 は、その構造例における主要な情報の内容 (Semantics) を示している。

【0105】

「scalability_extension_descriptor_tag」の 8 ビットフィールドは、デスクリプタタイプを示し、ここでは、スケーラビリティ・エクステンション・デスクリプタであること

10

20

30

40

50

を示す。「scalability_extension_descriptor_length」の8ビットフィールドは、デスクリプタの長さ（サイズ）を示し、デスクリプタの長さとして、以降のバイト数を示す。「extension_stream_existing_flag」の1ビットフィールドは、別ストリームによる拡張サービスがあることを示すフラグである。“1”は拡張ストリームがあることを示し、“0”は拡張ストリームがないことを示す。

【0106】

「extension_type」の3ビットフィールドは、拡張のタイプを示す。“001”は、拡張が、時間方向スケーラブルであることを示す。“010”は、拡張が、空間方向スケーラブルであることを示す。“011”は、拡張が、ビットレートスケーラブルであることを示す。

10

【0107】

「number_of_streams」の4ビットフィールドは、配信サービスに関与するストリームの総数を示す。「scalability_id」の3ビットフィールドは、複数のビデオストリームがスケーラブルなサービスを供給する際、個々のストリームに付されるスケーラビリティを示すIDである。“0”はベースストリームを示し、“1”～“7”はベースストリームからのスケーラビリティの度合いによって増加するIDである。「minimum_target_decoder_level_idc」の8ビットフィールドは、該当ストリームが対象とするデコーダの能力を示す。この情報は、受信機において、デコーダがストリームをデコードする前に符号化ピクチャの想定デコードタイミングがデコーダのpictureデコード処理能力の範囲を超えていないかどうかの判断に利用する。

20

【0108】

上述したように、この実施の形態において、SPS, ESPSに含まれるビットレートのレベル指定値(general_level_idc)などは、複数の階層を2以上の所定数の階層組に分割した際の所属階層組の識別情報として利用される。各階層組のレベル指定値の値は、この階層組のピクチャと、この階層組より低階層側の全ての階層組のピクチャとからなるフレームレートに対応した値とされる。

【0109】

マルチプレクサ104は、低階層側の階層組のピクチャの符号化画像データをコンテナするTSパケットの優先度ほど高く設定する。マルチプレクサ104は、例えば、複数の階層を低階層組と高階層組に二分する場合、TSパケットヘッダの「transport_priority」の1ビットフィールドを利用する。

30

【0110】

図18は、TSパケットの構造例(Syntax)を示している。「transport_priority」の1ビットフィールドは、ベースレイヤ、つまり低階層側の階層組のピクチャの符号化画像データをコンテナするTSパケットの場合は“1”に設定され、ノンベースレイヤ、つまり高階層側の階層組のピクチャの符号化画像データをコンテナするTSパケットの場合は“0”に設定される。

【0111】

図19は、SPS, ESPSのNALユニットに含まれるビットレートのレベル指定値(general_level_idc)と、TSパケットヘッダの「transport_priority」の設定値との関係を示している。受信側では、これらの情報の一方あるいは双方を用いて、低階層側の階層組のピクチャの符号化画像データと、高階層側の階層組のピクチャの符号化画像データとを、分別することが可能となる。

40

【0112】

図20は、マルチプレクサ104の構成例を示している。TSプライオリティ発生部141と、セクションコーディング部142と、PESパケット化部143-1～143-Nと、スイッチ部144と、トランスポートパケット化部145を有している。

【0113】

PESパケット化部143-1～143-Nは、それぞれ、圧縮データバッファ103に蓄積されているビデオストリーム1～Nを読み込み、PESパケットを生成する。この際、

50

P E S パケット化部 1 4 3 -1 ~ 1 4 3 -N は、ビデオストリーム 1 ~ N の H R D 情報を元に D T S (Decoding Time Stamp) 、 P T S (Presentation Time Stamp) のタイムスタンプを P E S ヘッダに付与する、この場合、各ピクチャの「cpu_removal_delay」、「dpb_output_delay」が参照され、 S T C (System Time Clock) 時刻に同期した精度で、各々 D T S 、 P T S に変換され、 P E S ヘッダの所定位置に配置される。

【 0 1 1 4 】

スイッチ部 1 4 4 は、 P E S パケット化部 1 4 3 -1 ~ 1 4 3 -N で生成された P E S パケットを、パケット識別子 (P I D) に基づいて選択的に取り出し、トランスポートパケット化部 1 4 5 に送る。トランスポートパケット化部 1 4 5 は、 P E S パケットをペイロードに含む T S パケットを生成し、トランスポートストリームを得る。

10

【 0 1 1 5 】

T S プライオリティ発生部 1 4 1 には、 C P U 1 0 1 から、階層数 (Number of layers) とストリーム数 (Number of streams) の情報が供給される。T S プライオリティ発生部 1 4 1 は、階層数で示される複数の階層を 2 以上の所定数の階層組に分割した場合における、各階層組の優先度を発生する。例えば、2 分割される場合には、T S パケットヘッダの「transport_priority」の 1 ビットフィールドに挿入すべき値が発生される (図 1 9 参照)。

【 0 1 1 6 】

T S プライオリティ発生部 1 4 1 は、各階層組の優先度の情報を、トランスポートパケット化部 1 4 5 に送る。トランスポートパケット化部 1 4 5 は、この情報に基づいて、各 T S パケットの優先度を設定する。この場合、上述したように、低階層側の階層組のピクチャの符号化画像データをコンテナする T S パケットの優先度ほど高く設定する。

20

【 0 1 1 7 】

セクションコーディング部 1 4 2 には、 C P U 1 0 1 から、階層数 (Number of layers) と、ストリーム数 (Number of streams) と、最小ターゲットデコーダ・レベル (Minimum_target_decoder_level_idc) の情報が供給される。セクションコーディング部 1 4 2 は、この情報に基づいて、トランスポートストリーム T S に挿入すべき各種のセクションデータ、例えば、上述した H E V C デスクリプタ (HEVC_descriptor) 、スケーラビリティ・エクステンション・デスクリプタ (scalability_extension_descriptor) などを生成する。

30

【 0 1 1 8 】

セクションコーディング部 1 4 2 は、各種セクションデータを、トランスポートパケット化部 1 4 5 に送る。トランスポートパケット化部 1 4 5 は、このセクションデータを含む T S パケットを生成し、トランスポートストリーム T S に挿入する。

【 0 1 1 9 】

図 2 1 は、マルチプレクサ 1 0 4 の処理フローを示す。この例は、複数の階層を低階層組と高階層組の 2 つに分割する例である。マルチプレクサ 1 0 4 は、ステップ S T 1 1 において、処理を開始し、その後に、ステップ S T 1 2 の処理に移る。このステップ S T 1 2 において、マルチプレクサ 1 0 4 は、ビデオストリーム (ビデオエレメンタリストリーム) の各ピクチャの temporal_id_ 、構成する符号化ストリーム数を設定する。

40

【 0 1 2 0 】

次に、マルチプレクサ 1 0 4 は、ステップ S T 1 3 において、低階層組のピクチャ、あるいは低階層組のピクチャを含むビデオストリームを多重化する際の「 transport_priority 」を “ 1 ” に設定する。また、マルチプレクサ 1 0 4 は、ステップ S T 1 4 において、 H R D 情報 (cpu_removal_delay、dpb_output_delay) を参照して、 D T S 、 P T S を決め、 P E S ヘッダに挿入する。

【 0 1 2 1 】

次に、マルチプレクサ 1 0 4 は、ステップ S T 1 5 において、シングルストリーム (单一ビデオストリーム) か否かを判断する。シングルストリームであるとき、マルチプレクサ 1 0 4 は、ステップ S T 1 6 において、1 つの P I D (パケット識別子) で多重化処理

50

を進めることとし、その後に、ステップST17の処理に移る。一方、シングルストリームでないとき、マルチプレクサ104は、ステップST18において、複数のパケットPID（パケット識別子）で多重化処理を進めることとし、その後に、ステップST17の処理に移る。

【0122】

このステップST17において、マルチプレクサ104は、HEVCデスクリプタ、スケーラビリティ・エクステンション・デスクリプタなどをコーディングする。そして、マルチプレクサ104は、ステップST19において、ビデオストリームをPESペイロードに挿入してPESパケット化し、その後、ステップST20において、トランスポートパケット化し、トランスマルチストリームTSを得る。その後、マルチプレクサ104は、ステップST21において、処理を終了する。
10

【0123】

図22は、単一ビデオストリームによる配信を行う場合のトランスポートストリームTSの構成例を示している。このトランスポートストリームTSには、単一ビデオストリームが含まれている。すなわち、この構成例では、複数の階層のピクチャの例えはHEVCによる符号化画像データを持つビデオストリームのPESパケット「video PES1」が存在すると共に、オーディオストリームのPESパケット「audio PES1」が存在する

【0124】

この単一のビデオストリームには、階層符号化の複数の階層が2以上の所定数の階層組に分割されて得られた所定数のサブストリームが含まれる。ここで、最下位の階層組のサブストリーム（ベースサブストリーム）にはSPSが含まれ、最下位より上位の階層組のサブストリーム（エンハンスサブストリーム）にはESPSが含まれる。そして、SPS, ESPSの要素の“general_level_idc”的値は、自己の階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値とされる。
20

【0125】

各ピクチャの符号化画像データには、VPS、SPS、ESPS、SEIなどのNALユニットが存在する。上述したように、各ピクチャのNALユニットのヘッダには、そのピクチャの階層を示すtemporal_idが挿入されている。また、例えは、SPS, ESPSにはビットレートのレベル指定値(general_level_idc)が含まれている。また、例えは、ピクチャ・タイミング・SEI(Picture timing SEI)には、「cpb_removal_delay」と「dpb_output_delay」が含まれている。
30

【0126】

なお、各ピクチャの符号化画像データをコンテナするTSパケットのヘッダに「transport_priority」の1ビットの優先度を示すフィールドが存在する。この「transport_priority」により、コンテナする符号化画像データが、低階層組のピクチャのものか、あるいは高階層組のピクチャのものかが識別可能である。

【0127】

また、トランスポートストリームTSには、PSI(Program Specific Information)として、PMT(Program Map Table)が含まれている。このPSIは、トランスポートストリームに含まれる各エレメンタリストリームがどのプログラムに属しているかを記した情報である。
40

【0128】

PMTには、プログラム全体に関連する情報を記述するプログラム・ループ(Program loop)が存在する。また、PMTには、各エレメンタリストリームに関連した情報を持つエレメンタリ・ループが存在する。この構成例では、ビデオエレメンタリ・ループ(video ES1 loop)が存在すると共に、オーディオエレメンタリ・ループ(audio ES1 loop)が存在する。

【0129】

ビデオエレメンタリ・ループには、ビデオストリーム(video PES1)に対応して、ストリームタイプ、パケット識別子(PID)等の情報が配置されると共に、そのビデオストリ
50

ームに関連する情報を記述するデスクリプタも配置される。このデスクリプタの一つとして、上述した H E V C デスクリプタ (HEVC_descriptor)、スケーラビリティ・エクステンション・デスクリプタ (scalability_extension_descriptor) が挿入される。

【 0 1 3 0 】

図 2 3 は、図 3 の階層符号化の例において、階層 0 ~ 3 のピクチャでベースサブストリーム (B stream) が生成され、階層 4 のピクチャでエンハンスサブストリーム (E stream) が生成される場合を示している。この場合、ベースサブストリームに含まれる各ピクチャは 6 0 P を構成し、エンハンスサブストリーム (E stream) に含まれる各ピクチャは、ベースサブストリームに含まれる各ピクチャに追加されて P E S 全体で 1 2 0 P を構成する。

10

【 0 1 3 1 】

ベースサブストリームのピクチャは、「 A U D 」、「 V P S 」、「 S P S 」、「 P P S 」、「 P S E I 」、「 S L I C E 」、「 S S E I 」、「 E O S 」などの N A L ユニットにより構成される。「 V P S 」、「 S P S 」は、例えば、 G O P の先頭ピクチャに挿入される。 S P S の要素の “ general_level_idc ” の値は、“ level5.1 ” とされる。なお、「 E O S 」はなくてもよい。

【 0 1 3 2 】

一方、エンハンスサブストリームのピクチャは、「 A U D 」、「 E S P S 」、「 P P S 」、「 P S E I 」、「 S L I C E 」、「 S S E I 」、「 E O S 」などの N A L ユニットにより構成される。なお、「 E S P S 」は、例えば、 G O P の先頭ピクチャに挿入される。 E S P S の要素の “ general_level_idc ” の値は、“ level5.2 ” とされる。なお、「 P S E I 」、「 S S E I 」、「 E O S 」はなくてもよい。

20

【 0 1 3 3 】

「 video ES1 loop 」には、ビデオストリーム (video PES1) に対応して、ストリームタイプ、パケット識別子 (PID) 等の情報が配置されると共に、そのビデオストリームに関連する情報を記述するデスクリプタも配置される。このストリームタイプは、ベースストリームを示す “ 0 x 2 4 ” とされる。また、デスクリプタの一つとして、上述した H E V C デスクリプタが挿入される。

【 0 1 3 4 】

「 level_constrained_flag 」の 1 ビットフィールドは、“ 1 ” とされる。これにより、「該当サブストリームに S P S あるいは E S P S が存在し、その要素の “ general_level_idc ” は、そのサブストリームが含む temporal_id 以下のピクチャを含むレベル値をもつ」ことが示される。また、「 level_idc 」の値は、ビデオストリーム (video PES1) の全体のレベル値を示す “ level5.2 ” とされる。また、「 temporal_id_min 」は 0 とされ、「 temporal_id_max 」は 4 とされ、ビデオストリーム (video PES1) に階層 0 ~ 4 のピクチャが含まれていることが示される。

30

【 0 1 3 5 】

このような単一ビデオストリームによる配信が行われる場合、受信側では、「 level_constrained_flag 」、 S P S 、 E S P S の要素の “ general_level_idc ” などに基づいて、各サブストリームが自身のデコーダ処理能力の範囲内にあるか否かが判断され、範囲内にあるサブストリームのデコードが行われる。

40

【 0 1 3 6 】

図 2 4 は、複数ストリーム、ここでは 2 ストリームによる配信を行う場合のトランスポートストリーム T S の構成例を示している。このトランスポートストリーム T S には、2 つのビデオストリームが含まれている。すなわち、この構成例では、複数の階層が低階層組と高階層組の 2 つの階層組に分割され、2 つの階層組のピクチャの例えば H E V C による符号化画像データを持つビデオストリームの P E S パケット「 video PES1 」、「 video PES2 」が存在すると共に、オーディオストリームの P E S パケット「 audio PES1 」が存在する。

【 0 1 3 7 】

50

この 2 つのビデオストリームには、階層符号化の複数の階層が 2 つの階層組に分割されて得られた 2 つのサブストリームのそれぞれが含まれる。ここで、下位側の階層組のサブストリーム（ベースサブストリーム）には S P S が含まれ、上位側の階層組のサブストリーム（エンハンスサブストリーム）には E S P S が含まれる。そして、S P S, E S P S の要素の “general_level_idc” の値は、自己の階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値とされる。

【 0 1 3 8 】

各ピクチャの符号化画像データには、S P S、E S P S などの N A L ユニットが存在する。上述したように、各ピクチャの N A L ユニットのヘッダには、そのピクチャの階層を示す temporal_id が挿入されている。また、例えば、S P S、E S P S にはビットレートのレベル指定値 (general_level_idc) が含まれている。また、例えば、ピクチャ・タイミング・S E I (Picture timing SEI) には、「cpb_removal_delay」と「dpb_output_delay」が含まれている。10

【 0 1 3 9 】

また、各ピクチャの符号化画像データをコンテナする T S パケットのヘッダに「transport_priority」の 1 ビットの優先度を示すフィールドが存在する。この「transport_priority」により、コンテナする符号化画像データが、低階層組のピクチャのものか、あるいは高階層組のピクチャのものかが識別可能である。

【 0 1 4 0 】

また、トランスポートストリーム T S には、P S I (Program Specific Information) として、P M T (Program Map Table) が含まれている。この P S I は、トランスポートストリームに含まれる各エレメンタリストリームがどのプログラムに属しているかを記した情報である。20

【 0 1 4 1 】

P M T には、プログラム全体に関連する情報を記述するプログラム・ループ (Program loop) が存在する。また、P M T には、各エレメンタリストリームに関連した情報を持つエレメンタリ・ループが存在する。この構成例では、2 つのビデオエレメンタリ・ループ (video ES1 loop, video ES2 loop) が存在すると共に、オーディオエレメンタリ・ループ (audio ES1 loop) が存在する。

【 0 1 4 2 】

各ビデオエレメンタリ・ループには、ビデオストリーム (video PES1, video PES2) に対応して、ストリームタイプ、パケット識別子 (PID) 等の情報が配置されると共に、そのビデオストリームに関連する情報を記述するデスクリプタも配置される。このデスクリプタの一つとして、上述した H E V C デスクリプタ (HEVC_descriptor)、スケーラビリティ・エクステンション・デスクリプタ (scalability_extension_descriptor) が挿入される。30

【 0 1 4 3 】

図 25 は、図 3 の階層符号化の例において、階層 0 ~ 3 のピクチャでベースサブストリーム (B stream) が生成され、階層 4 のピクチャでエンハンスサブストリーム (E stream) が生成される場合を示している。この場合、ベースサブストリームに含まれる各ピクチャは 6 0 P を構成し、エンハンスサブストリーム (E stream) に含まれる各ピクチャは、ベースサブストリームに含まれる各ピクチャに追加されて P E S 全体で 1 2 0 P を構成する。40

【 0 1 4 4 】

ベースサブストリームのピクチャは、「A U D」、「V P S」、「S P S」、「P P S」、「P S E I」、「S L I C E」、「S S E I」、「E O S」などの N A L ユニットにより構成される。なお、「V P S」、「S P S」は、例えば、G O P の先頭ピクチャに挿入される。S P S の要素の “general_level_idc” の値は、“level5.1” とされる。なお、「E O S」はなくてもよい。

【 0 1 4 5 】

10

20

30

40

50

一方、エンハンスサブストリームのピクチャは、「AUD」、「ESPS」、「PPS」、「PSEI」、「SLICE」、「SSEI」、「EOS」などのNALユニットにより構成される。なお、「ESPS」は、例えば、GOPの先頭ピクチャに挿入される。ESPSの要素の“general_level_idc”的値は、“level5.2”とされる。なお、「PSEI」、「SSEI」、「EOS」はなくてもよい。

【0146】

「video ES1 loop」には、ビデオストリーム(video PES1)に対応して、ストリームタイプ、パケット識別子(PID)等の情報が配置されると共に、そのビデオストリームに関連する情報を記述するデスクリプタも配置される。このストリームタイプは、ベースストリームを示す“0x24”とされる。また、デスクリプタの一つとして、上述したHEVCデスクリプタが挿入される。

【0147】

「level_constrained_flag」の1ビットフィールドは、“1”とされる。これにより、「該当サブストリームにSPSあるいはESPSが存在し、その要素の“general_level_idc”は、そのサブストリームが含むtemporal_id以下のピクチャを含むレベル値をもつ」ことが示される。また、「level_idc」の値は、ベースサブストリーム(B stream)のレベル値を示す“level5.1”とされる。また、「temporal_id_min」は0とされ、「temporal_id_max」は3とされ、ベースサブストリーム(B stream)に階層0～3のピクチャが含まれていることが示される。

【0148】

「video ES2 loop」には、ビデオストリーム(video PES2)に対応して、ストリームタイプ、パケット識別子(PID)等の情報が配置されると共に、そのビデオストリームに関連する情報を記述するデスクリプタも配置される。このストリームタイプは、エンハンスストリームを示す“0x25”とされる。また、デスクリプタの一つとして、上述したHEVCデスクリプタが挿入される。

【0149】

「level_constrained_flag」の1ビットフィールドは、“1”とされる。これにより、「該当サブストリームにSPSあるいはESPSが存在し、その要素の“general_level_idc”は、そのサブストリームが含むtemporal_id以下のピクチャを含むレベル値をもつ」ことが示される。また、「level_idc」の値は、ベースサブストリーム(B stream)およびエンハンスストリーム(E stream)のレベル値を示す“level5.2”とされる。また、「temporal_id_min」は4とされ、「temporal_id_max」は4とされ、エンハンスストリーム(E stream)に階層4のピクチャが含まれていることが示される。

【0150】

このような複数ビデオストリームによる配信が行われる場合、受信側では、「level_constrained_flag」、「SPS」、「ESPS」の要素の“general_level_idc”などに基づいて、各サブストリームが自身のデコーダ処理能力の範囲内にあるか否かが判断され、範囲内にあるサブストリームのデコードが行われる。

【0151】

図26は、図3の階層符号化の例において、階層0～3のピクチャでベースサブストリーム(B stream)が生成され、階層4のピクチャでエンハンスサブストリーム(E stream)が生成される場合におけるトランスポートストリームTSの他の構成例を示している。この場合、ベースサブストリームに含まれる各ピクチャは60Pを構成し、エンハンスサブストリーム(E stream)に含まれる各ピクチャは、ベースサブストリームに含まれる各ピクチャに追加されてPES全体で120Pを構成する。

【0152】

ベースサブストリームのピクチャは、「AUD」、「VPS」、「SPS」、「PPS」、「PSEI」、「SLICE」、「SSEI」、「EOS」などのNALユニットにより構成される。なお、「VPS」、「SPS」は、例えば、GOPの先頭ピクチャに挿入される。SPSの要素の“general_level_idc”的値は、“level5.2”とされる。この

10

20

30

40

50

場合、S P S の要素の “sub_layer_level_present_flag” は “1” とされ、“sublayer_level_idc[3]” で、ベースサブストリームのレベル値 “level5.1” が示される。なお、「EOS」はなくてもよい。

【0153】

エンハンスサブストリームのピクチャは、「AUD」、「PPS」、「SLICE」などのNALユニットにより構成される。しかし、図25におけるような「ESPS」のNALユニットは存在しない。

【0154】

「video ES1 loop」には、ビデオストリーム(video PES1)に対応して、ストリームタイプ、パケット識別子(PID)等の情報が配置されると共に、そのビデオストリームに関連する情報を記述するデスクリプタも配置される。このストリームタイプは、ベースストリームを示す“0x24”とされる。また、デスクリプタの一つとして、上述したHEVCデスクリプタが挿入される。

10

【0155】

図25におけるような「level_constrained_flag」は存在しない。「level_idc」の値は、ベースサブストリーム(B stream)のレベル値を示す“level5.1”とされる。また、「temporal_id_min」は0とされ、「temporal_id_max」は3とされ、ベースサブストリーム(B stream)に階層0～3のピクチャが含まれていることが示される。

【0156】

「video ES2 loop」には、ビデオストリーム(video PES2)に対応して、ストリームタイプ、パケット識別子(PID)等の情報が配置されると共に、そのビデオストリームに関連する情報を記述するデスクリプタも配置される。このストリームタイプは、エンハンスストリームを示す“0x25”とされる。また、デスクリプタの一つとして、上述したHEVCデスクリプタが挿入される。

20

【0157】

図25におけるような「level_constrained_flag」は存在しない。「level_idc」の値は、ベースサブストリーム(B stream)およびエンハンスストリーム(E stream)のレベル値を示す“level5.2”とされる。また、「temporal_id_min」は4とされ、「temporal_id_max」は4とされ、エンハンスストリーム(E stream)に階層4のピクチャが含まれていることが示される。

30

【0158】

このような複数ビデオストリームによる配信が行われる場合、受信側では、S P S の要素の “general_level_idc”、“sublayer_level_idc” などに基づいて、各サブストリームが自身のデコーダ処理能力の範囲内にあるか否かが判断され、範囲内にあるサブストリームのデコードが行われる。

【0159】

図2に戻って、送信部105は、トランスポートストリームTSを、例えば、QPSK/O FDM等の放送に適した変調方式で変調し、RF変調信号を送信アンテナから送信する。

【0160】

図2に示す送信装置100の動作を簡単に説明する。エンコーダ102には、非圧縮の動画像データが入力される。エンコーダ102では、この動画像データに対して、階層符号化が行われる。すなわち、エンコーダ102では、この動画像データを構成する各ピクチャの画像データが複数の階層に分類されて符号化され、各階層のピクチャの符号化画像データを持つビデオストリームが生成される。この際、参照するピクチャが、自己階層および/または自己階層よりも下位の階層に所属するように、符号化される。

40

【0161】

エンコーダ102では、各階層のピクチャの符号化画像データを持つビデオストリームが生成される。例えば、エンコーダ102では、複数の階層が2以上の所定数の階層組に分割され、各階層組に対応したサブストリームのそれぞれを含む所定数のビデオストリーム

50

ムが生成されるか、または、各階層組に対応したサブストリームの全てを含む単一のビデオストリームが生成される。

【0162】

エンコーダ102では、各階層組のピクチャの符号化画像データに、所属階層組を識別するための識別情報が付加される。この場合、例えば、識別情報として、S P S, E S P Sの要素である「general_level_idc」が利用される。S P Sは、最下位の階層組のサブストリーム（ベースサブストリーム）に、シーケンス（G O P）毎に含まれる。一方、E S P Sは、最下位より上位の階層組のサブストリーム（エンハンスサブストリーム）に、シーケンス（G O P）毎に含まれる。S P S, E S P Sに含まれる「general_level_idc」の値は、高階層側の階層組ほど高い値とされる。例えば、各階層組のサブストリームのS P S, E S P Sに挿入される「general_level_idc」の値は、自己の階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値とされる。

【0163】

エンコーダ102で生成された、各階層のピクチャの符号化データを含むビデオストリームは、圧縮データバッファ（c p b）103に供給され、一時的に蓄積される。マルチプレクサ104では、圧縮データバッファ103に蓄積されているビデオストリームが読み出され、P E Sパケット化され、さらにトランスポートパケット化されて多重され、多重化ストリームとしてのトランスポートストリームT Sが得られる。

【0164】

このトランスポートストリームT Sには、各階層のピクチャの符号化画像データを持つ単一のビデオストリーム、あるいは2以上の所定数のビデオストリームが含まれる。マルチプレクサ104では、トランスポートストリームT Sに、階層情報、ストリーム構成情報が挿入される。すなわち、マルチプレクサ104では、各ビデオストリームに対応したビデオエレメンタリ・ループに、H E V Cデスクリプタ（HEVC_descriptor）、スケーラビリティ・エクステンション・デスクリプタ（scalability_extension_descriptor）が挿入される。

【0165】

また、マルチプレクサ104では、低階層側の階層組のピクチャの符号化画像データをコンテナするT Sパケットの優先度ほど高く設定される。マルチプレクサ104では、例えば、複数の階層を低階層組と高階層組に二分される場合、T Sパケットヘッダの「transport_priority」の1ビットフィールドが利用されて優先度が設定される。

【0166】

マルチプレクサ104で生成されるトランスポートストリームT Sは、送信部105に送られる。送信部105では、このトランスポートストリームT Sが、例えば、Q P S K / O F D M等の放送に適した変調方式で変調され、R F変調信号が送信アンテナから送信される。

【0167】

「受信装置の構成」

図27は、受信装置200の構成例を示している。この受信装置200は、C P U（Central Processing Unit）201と、受信部202と、デマルチプレクサ203と、圧縮データバッファ（c p b : coded picture buffer）204を有している。また、この受信装置200は、デコーダ205と、非圧縮データバッファ（d p b : decoded picture buffer）206と、ポスト処理部207を有している。C P U201は、制御部を構成し、受信装置200の各部の動作を制御する。

【0168】

受信部202は、受信アンテナで受信されたR F変調信号を復調し、トランスポートストリームT Sを取得する。デマルチプレクサ203は、トランスポートストリームT Sから、デコード能力（Decoder temporal layer capability）に応じた階層組のピクチャの符号化画像データを選択的に取り出し、圧縮データバッファ（c p b : coded picture buffer）204に送る。

10

20

30

40

50

【0169】

図28は、デマルチプレクサ203の構成例を示している。このデマルチプレクサ203は、PCR抽出部231と、タイムスタンプ抽出部232と、セクション抽出部233と、TSプライオリティ抽出部234と、PESペイロード抽出部235と、ピクチャ選択部236を有している。

【0170】

PCR抽出部231は、PCR(Program Clock Reference)が含まれるTSパケットからPCRを抽出し、CPU201に送る。タイムスタンプ抽出部232は、ピクチャ毎にPESヘッダに挿入されているタイムスタンプ(DTS、PTS)を抽出し、CPU201に送る。セクション抽出部233は、トランSPORTストリームTSからセクションデータを抽出し、CPU201に送る。このセクションデータには、上述したHEVCデスクリプタ(HEVC_descriptor)、スケーラビリティ・エクステンション・デスクリプタ(scability_extension_descriptor)などが含まれている。

10

【0171】

TSプライオリティ抽出部234は、各TSパケットに設定されている優先度情報を抽出する。この優先度は、上述したように、複数の階層を2以上の所定数の階層組に分割した場合における各階層組の優先度であり、低階層側の階層組ほど高く設定されている。例えば、低階層組と高階層組に2分されている場合、TSパケットヘッダの「transport_priority」の1ビットフィールドの値が抽出される。この値は、低階層組では“1”とされ、高階層組では“0”に設定されている。

20

【0172】

PESペイロード抽出部235は、トランSPORTストリームTSからPESペイロード、つまり、各階層のピクチャの符号化画像データを抽出する。ピクチャ選択部236は、PESペイロード抽出部235で取り出される各階層のピクチャの符号化画像データから、デコード能力(Decoder temporal layer capability)に応じた階層組のピクチャの符号化画像データを選択的に取り出し、圧縮データバッファ(cpb : coded picture buffer)204に送る。この場合、ピクチャ選択部236は、セクション抽出部233で得られる階層情報、ストリーム構成情報、TSプライオリティ抽出部234で抽出される優先度情報を参照する。

30

【0173】

例えば、トランSPORTストリームTSに含まれるビデオストリーム(符号化ストリーム)のフレームレートが120fpsである場合を考える。例えば、複数の階層が低階層側の階層組と高階層側の階層組とに2分割され、各階層組のピクチャのフレームレートがそれぞれ60fpsであるとする。例えば、上述の図3に示す階層符号化例では、階層0から3は低階層側の階層組とされ、階層4は高階層側の階層組とされる。

【0174】

TSパケットのヘッダに含まれる「transport_priority」の1ビットフィールドは、ベースレイヤ、つまり低階層側の階層組のピクチャの符号化画像データをコンテナするTSパケットの場合には“1”に設定され、ノンベースレイヤ、つまり高階層側の階層組のピクチャの符号化画像データをコンテナするTSパケットの場合には“0”に設定されている。

40

【0175】

この場合、トランSPORTストリームTSに、各階層のピクチャの符号化データを持つ単一のビデオストリーム(符号化ストリーム)(図10参照)が含まれている場合がある。また、この場合、トランSPORTストリームTSに、低階層側の階層組のピクチャの符号化画像データを持つベースストリーム(B-stream)と、高階層側の階層組のピクチャの符号化画像データを持つエンハンスストリーム(E-stream)の2つのビデオストリーム(符号化ストリーム)が含まれている場合(図11参照)がある。

【0176】

ピクチャ選択部236は、例えば、デコード能力が、120P(120fps)に対応

50

している場合、全階層のピクチャの符号化画像データを取り出し、圧縮データバッファ(`cpb`)204に送る。一方、ピクチャ選択部236は、例えば、デコード能力が、120Pに対応していないが60P(`60fps`)に対応している場合、低階層側の階層組のピクチャの符号化画像データのみを取り出し、圧縮データバッファ(`cpb`)204に送る。

【0177】

図29は、トランスポートストリームTSに単一のビデオストリーム(符号化ストリーム)が含まれている場合を示している。ここで、「High」は高階層側の階層組のピクチャを示し、「Low」は低階層側の階層組のピクチャを示す。また、「P」は「`transport_priority`」を示している。

10

【0178】

デコード能力が、120Pに対応している場合、ピクチャ選択部236は、全階層のピクチャの符号化画像データを取り出し、圧縮データバッファ(`cpb`)204に送り、領域1(`cpb_1`)に蓄積する。一方、デコード能力が、120Pに対応していないが60Pに対応している場合、「`transport_priority`」に基づくフィルタリングを行って、P=1である低階層側の階層組のピクチャだけを取り出し、圧縮データバッファ(`cpb`)204に送り、領域1(`cpb_1`)に蓄積する。

【0179】

図30は、トランスポートストリームTSにベースストリームと拡張ストリームの2つのビデオストリーム(符号化ストリーム)が含まれている場合を示している。ここで、「High」は高階層側の階層組のピクチャを示し、「Low」は低階層側の階層組のピクチャを示す。また、「P」は「`transport_priority`」を示している。また、ベースストリームのパケット識別子(PID)はPID1であり、拡張ストリームのパケット識別子(PID)はPID2であるとする。

20

【0180】

デコード能力が、120Pに対応している場合、ピクチャ選択部236は、全階層のピクチャの符号化画像データを取り出し、圧縮データバッファ(`cpb`)204に送る。そして、低階層側の階層組のピクチャの符号化画像データは領域1(`cpb_1`)に蓄積し、低階層側の階層組のピクチャの符号化画像データは領域2(`cpb_2`)に蓄積する。

【0181】

30

一方、デコード能力が、120Pに対応していないが120Pに対応している場合、パケット識別子(PID)に基づくフィルタリングを行って、PID1である低階層側の階層組のピクチャだけを取り出し、圧縮データバッファ(`cpb`)204に送り、領域1(`cpb_1`)に蓄積する。なお、この場合も、「`transport_priority`」に基づくフィルタリングを行ってもよい。

【0182】

図31は、デマルチプレクサ203の処理フローの一例を示している。この処理フローは、トランスポートストリームTSに単一のビデオストリーム(符号化ストリーム)が含まれている場合を示している。

【0183】

40

デマルチプレクサ203は、ステップST31において、処理を開始し、その後に、ステップST32の処理に移る。このステップST32において、CPU201から、デコード能力(`Decoder temporal layer capability`)が設定される。次に、デマルチプレクサ203は、ステップST33において、全階層をデコードする能力があるか否かを判断する。

【0184】

全階層をデコードする能力があるとき、デマルチプレクサ203は、ステップST34において、該当PIDフィルタを通過する全TSパケットをデマルチプレクスし、セクションパーサリング(`Section parsing`)を行う。その後、デマルチプレクサ203は、ステップST35の処理に移る。

50

【0185】

ステップST33で全階層をデコードする能力がないとき、デマルチプレクサ203は、ステップST36において、「transport_priority」が“1”的TSパケットをデマルチプレクスし、セクションパーサリング(Section parsing)を行う。その後、デマルチプレクサ203は、ステップST35の処理に移る。

【0186】

ステップST35において、デマルチプレクサ203は、対象となるPIDのセクションの中で、HEVCデスクリプタ(HEVC_descriptor)、スケーラビリティ・エクステンション・デスクリプタ(scalability_extension_descriptor)を読み、拡張ストリームの有無、スケーラブルタイプ、ストリームの数とID、temporal_idの最大、最小値、最小ターゲットデコーダ・レベルを得る。

10

【0187】

次に、デマルチプレクサ203は、ステップST37で、PIDの対象となる符号化ストリームを圧縮データバッファ(cpb)204へ転送すると共に、DTS、PTSを、CPU201に通知する。デマルチプレクサ203は、ステップST37の処理の後、ステップST38において、処理を終了する。

20

【0188】

図32は、デマルチプレクサ203の処理フローの一例を示している。この処理フローは、トランスポートストリームTSにベースストリームと拡張ストリームの2つのビデオストリーム(符号化ストリーム)が含まれている場合を示している。

20

【0189】

デマルチプレクサ203は、ステップST41において、処理を開始し、その後に、ステップST42の処理に移る。このステップST42において、CPU201から、デコード能力(Decoder temporal layer capability)が設定される。次に、デマルチプレクサ203は、ステップST43において、全階層をデコードする能力があるか否かを判断する。

【0190】

全階層をデコードする能力があるとき、デマルチプレクサ203は、ステップST44において、該当PIDフィルタを通過する全TSパケットをデマルチプレクスし、セクションパーサリング(Section parsing)を行う。その後、デマルチプレクサ203は、ステップST45の処理に移る。

30

【0191】

ステップST43で全階層をデコードする能力がないとき、デマルチプレクサ203は、ステップST46において、PID=PID1のTSパケットをデマルチプレクスし、セクションパーサリング(Section parsing)を行う。その後、デマルチプレクサ203は、ステップST45の処理に移る。

【0192】

ステップST45において、デマルチプレクサ203は、対象となるPIDのセクションの中で、HEVCデスクリプタ(HEVC_descriptor)、スケーラビリティ・エクステンション・デスクリプタ(scalability_extension_descriptor)を読み、拡張ストリームの有無、スケーラブルタイプ、ストリームの数とID、temporal_idの最大、最小値、最小ターゲットデコーダ・レベルを得る。

40

【0193】

次に、デマルチプレクサ203は、ステップST47で、PIDの対象となる符号化ストリームを圧縮データバッファ(cpb)204へ転送すると共に、DTS、PTSを、CPU201に通知する。デマルチプレクサ203は、ステップST47の処理の後、ステップST48において、処理を終了する。

【0194】

図27に戻って、圧縮データバッファ(cpb)204は、デマルチプレクサ203で取り出されるビデオストリーム(符号化ストリーム)を、一時的に蓄積する。デコーダ20

50

5は、圧縮データバッファ204に蓄積されているビデオストリームから、デコードすべき階層として指定された階層のピクチャの符号化画像データを取り出す。そして、デコーダ205は、取り出された各ピクチャの符号化画像データを、それぞれ、そのピクチャのデコードタイミングでデコードし、非圧縮データバッファ(dpb)206に送る。

【0195】

ここで、デコーダ205には、CPU201からデコードすべき階層がtemporal_idで指定される。この指定階層は、デマルチブレクサ203で取り出されるビデオストリーム(符号化ストリーム)に含まれる全階層、あるいは低階層側の一部の階層とされ、CPU201により自動的に、あるいはユーザ操作に応じて設定される。また、デコーダ205には、CPU201から、DTS(Decoding Time stamp)に基づいて、デコードタイミングが与えられる。なお、デコーダ205は、各ピクチャの符号化画像データをデコードする際に、必要に応じて、非圧縮データバッファ206から被参照ピクチャの画像データを読み出して利用する。

【0196】

図33は、デコーダ205の構成例を示している。このデコーダ205は、テンポラルID解析部251と、対象階層選択部252と、ストリーム統合部253と、デコード部254を有している。テンポラルID解析部251は、圧縮データバッファ204に蓄積されているビデオストリーム(符号化ストリーム)を読み出し、各ピクチャの符号化画像データのNALユニットヘッダに挿入されているtemporal_idを解析する。

【0197】

対象階層選択部252は、圧縮データバッファ204から読み出されたビデオストリームから、テンポラルID解析部251の解析結果に基づいて、デコードすべき階層として指定された階層のピクチャの符号化画像データを取り出す。この場合、対象階層選択部252からは、圧縮データバッファ204から読み出されたビデオストリームの数および指定階層に応じて、単一または複数のビデオストリーム(符号化ストリーム)が出力される。

【0198】

ストリーム統合部253は、対象階層選択部252から出力される所定数のビデオストリーム(符号化ストリーム)を一つに統合する。デコード部254は、ストリーム統合部253で統合されたビデオストリーム(符号化ストリーム)が持つ各ピクチャの符号化画像データを、順次デコードタイミングでデコードし、非圧縮データバッファ(dpb)206に送る。

【0199】

この場合、デコード部254はデマルチブレクサ203から得られるLevel_constraint_d_flagにより、SPS, ESSPSの解析を行って、「general_level_idc」、「sublayer_level_idc」などを把握し、ストリームあるいはサブストリームが自身のデコーダ処理能力範囲内でデコードし得るものかどうかを確認する。また、この場合、デコード部254は、SEIの解析を行って、例えば、「initial_cpb_removal_time」、「cpb_removal_delay」を把握し、CPU201からのデコードタイミングが適切か確認する。

【0200】

また、デコード部254は、スライス(Slice)のデコードを行う際に、スライスヘッダ(Slice header)から、時間方向の予測先を表す情報として、「ref_idx_10_active(ref_idx_11_active)」を取得し、時間方向の予測を行う。なお、デコード後のピクチャは、スライスヘッダ(slice header)から得られる「short_term_ref_pic_set_idx」、あるいは「it_idx_sps」が指標とされて、他のピクチャによる被参照として処理される。

【0201】

図34のフローチャートは、受信装置200におけるデコーダ処理能力を考慮したビデオストリーム毎のデコード処理手順の一例を示している。受信装置200は、ステップST61で処理を開始し、ステップST62において、HEVCデスクリプタ(HEVC_descriptor)を読む。

10

20

30

40

50

【0202】

次に、受信装置200は、ステップST63において、HEVCデスクリプタに“level_constrained_flag”が存在するか判断する。存在するとき、受信装置200は、ステップST64において、“level_constrained_flag”が“1”であるか判断する。“1”であるとき、受信装置200は、ステップST65の処理に進む。

【0203】

このステップST65において、受信装置200は、該当するPIDのPESパケット(PES packet)のタイムスタンプを参照し、ペイロード(payload)部分のビデオストリームのSPSあるいはEPSを読む。そして、受信装置200は、ステップST66において、SPSあるいはEPSの要素である「general_level_idc」を読む。 10

【0204】

次に、受信装置200は、ステップST67において、「general_level_idc」がデコーダ処理能力範囲内か判断する。デコーダ処理能力範囲内であるとき、受信装置200は、ステップST68において、該当するストリームあるいはサブストリームをデコードする。その後に、受信装置200は、ステップST69において、処理を終了する。一方、ステップST67でデコーダ処理能力範囲内でないとき、受信装置200は、直ちにステップST69に進み、処理を終了する。

【0205】

また、ステップST63で“level_constrained_flag”が存在しないとき、あるいはステップST64で“level_constrained_flag”が“0”であるとき、受信装置200は、ステップST70の処理に移る。このステップST70において、受信装置200は、該当するPIDのPESパケット(PES packet)のタイムスタンプを参照し、ペイロード(payload)部分のビデオストリームのSPSを読む。一方、該当するビデオストリームにSPSが存在しない場合は、temporal_layerが下位のピクチャ(Picture)を含むサブストリームのSPSを参照する。 20

【0206】

次に、受信装置200は、ステップST71において、SPSの要素である「general_level_idc」を読む。そして、受信装置200は、ステップST72において、「general_level_idc」がデコーダ処理能力範囲内か判断する。デコーダ処理能力範囲内であるとき、受信装置200は、ステップST73の処理に移る。 30

【0207】

一方、デコーダ処理能力範囲内でないとき、受信装置200は、ステップST74において、SPSの要素の「Sublayer_level_idc」をチェックする。そして、受信装置200は、ステップST75において、「Sublayer_level_idc」がデコーダ処理能力範囲内である「Sublayer」が存在するか判断する。存在しないとき、受信装置200は、直ちにステップST69に進み、処理を終了する。一方、存在するとき、受信装置200は、ステップST73の処理に移る。

【0208】

このステップST73において、受信装置200は、ストリームの全体、あるいは該当するサブレイヤ(Sublayer)部分を、temporal_id値を参照してデコードする。その後、受信装置200は、ステップST69において、処理を終了する。 40

【0209】

図27に戻って、非圧縮データバッファ(dpб)206は、デコード205でデコードされた各ピクチャの画像データを、一時的に蓄積する。ポスト処理部207は、非圧縮データバッファ(dpб)206から表示タイミングで順次読み出された各ピクチャの画像データに対して、そのフレームレートを、表示能力に合わせる処理を行う。この場合、CPU201から、PTS(Presentation Time stamp)に基づいて、表示タイミングが与えられる。

【0210】

例えば、デコード後の各ピクチャの画像データのフレームレートが120fpsであつ 50

て、表示能力が 120 f p s であるとき、ポスト処理部 207 は、デコード後の各ピクチャの画像データをそのままディスプレイに送る。また、例えば、デコード後の各ピクチャの画像データのフレームレートが 120 f p s であって、表示能力が 60 f p s であるとき、ポスト処理部 207 は、デコード後の各ピクチャの画像データに対して時間方向解像度が 1/2 倍となるようにサブサンプル処理を施し、60 f p s の画像データとしてディスプレイに送る。

【0211】

また、例えば、デコード後の各ピクチャの画像データのフレームレートが 60 f p s であって、表示能力が 120 f p s であるとき、ポスト処理部 207 は、デコード後の各ピクチャの画像データに対して時間方向解像度が 2 倍となるように補間処理を施し、120 f p s の画像データとしてディスプレイに送る。また、例えば、デコード後の各ピクチャの画像データのフレームレートが 60 f p s であって、表示能力が 60 f p s であるとき、ポスト処理部 207 は、デコード後の各ピクチャの画像データをそのままディスプレイに送る。

【0212】

図 35 は、ポスト処理部 207 の構成例を示している。この例は、上述したようにデコード後の各ピクチャの画像データのフレームレートが 120 f p s あるいは 60 f p s であって、表示能力が 120 f p s あるいは 60 f p s である場合に対処可能とした例である。

【0213】

ポスト処理部 207 は、補間部 271 と、サブサンプル部 272 と、スイッチ部 273 を有している。非圧縮データバッファ 206 からのデコード後の各ピクチャの画像データは、直接スイッチ部 273 に入力され、あるいは補間部 271 で 2 倍のフレームレートとされた後にスイッチ部 273 に入力され、あるいはサブサンプル部 272 で 1/2 倍のフレームレートとされた後にスイッチ部 273 に入力される。

【0214】

スイッチ部 273 には、CPU 201 から、選択情報が供給される。この選択情報は、CPU 201 が、表示能力を参照して自動的に、あるいは、ユーザ操作に応じて発生する。スイッチ部 273 は、選択情報に基づいて、入力のいずれかを選択的に出力とする。これにより、非圧縮データバッファ (d p b) 206 から表示タイミングで順次読み出された各ピクチャの画像データのフレームレートは、表示能力に合ったものとされる。

【0215】

図 36 は、デコーダ 205、ポスト処理部 207 の処理フローの一例を示している。デコーダ 205、ポスト処理部 207 は、ステップ ST 51 において、処理を開始し、その後に、ステップ ST 52 の処理に移る。このステップ ST 52 において、デコーダ 205 は、圧縮データバッファ (c p b) 204 に蓄積されているデコード対象のビデオストリームを読み出し、temporal_idに基づいて、CPU 201 からデコード対象として指定される階層のピクチャを選択する。

【0216】

次に、デコーダ 205 は、ステップ ST 53 において、選択された各ピクチャの符号化画像データをデコードタイミングで順次デコードし、デコード後の各ピクチャの画像データを非圧縮データバッファ (d p b) 206 に転送して、一時的に蓄積する。次に、ポスト処理部 207 は、ステップ ST 54 において、非圧縮データバッファ (d p b) 206 から、表示タイミングで各ピクチャの画像データを読み出す。

【0217】

次に、ポスト処理部 207 は、読み出された各ピクチャの画像データのフレームレートが表示能力にあっているか否かを判断する。フレームレートが表示能力に合っていないとき、ポスト処理部 207 は、ステップ ST 56 において、フレームレートを表示能力に合わせて、ディスプレイに送り、その後、ステップ ST 57 において、処理を終了する。一方、フレームレートが表示能力に合っているとき、ポスト処理部 207 は、ステップ ST

10

20

30

40

50

58において、フレームレートそのままでディスプレイに送り、その後、ステップS T 57において、処理を終了する。

【0218】

図27に示す受信装置200の動作を簡単に説明する。受信部202では、受信アンテナで受信されたRF変調信号が復調され、トランSPORTストリームTSが取得される。このトランSPORTストリームTSは、デマルチプレクサ203に送られる。デマルチプレクサ203では、トランSPORTストリームTSから、デコード能力(Decoder temporal layer capability)に応じた階層組のピクチャの符号化画像データが選択的に取り出され、圧縮データバッファ(cpb)204に送られ、一時的に蓄積される。

【0219】

デコーダ205では、圧縮データバッファ204に蓄積されているビデオストリームから、デコードすべき階層として指定された階層のピクチャの符号化画像データが取り出される。そして、デコーダ205では、取り出された各ピクチャの符号化画像データが、それぞれ、そのピクチャのデコードタイミングでデコードされ、非圧縮データバッファ(dpb)206に送られ、一時的に蓄積される。この場合、各ピクチャの符号化画像データがデコードされる際に、必要に応じて、非圧縮データバッファ206から被参照ピクチャの画像データが読み出されて利用される。

【0220】

非圧縮データバッファ(dpb)206から表示タイミングで順次読み出された各ピクチャの画像データは、ポスト処理部207に送られる。ポスト処理部207では、各ピクチャの画像データに対して、そのフレームレートを、表示能力に合わせるための補間あるいはサブサンプルが行われる。このポスト処理部207で処理された各ピクチャの画像データは、ディスプレイに供給され、その各ピクチャの画像データによる動画像の表示が行われる。

【0221】

以上説明したように、図1に示す送受信システム10においては、送信側において、階層毎にエンコード間隔が算出され、高階層ほどピクチャ毎の符号化画像データのデコード時間間隔が小さくなるように設定されたデコードタイムスタンプが各階層のピクチャの符号化画像データに付加されるものである。そのため、例えば、受信側においてデコード能力に応じた良好なデコード処理が可能となる。例えば、デコード能力が低い場合であっても、圧縮データバッファ204のバッファ破たんを招くことなく、低階層のピクチャの符号化画像データを選択的にデコードすることが可能となる。

【0222】

また、図1に示す送受信システム10においては、送信側において、トランSPORTストリームTSのレイヤに、スケーラビリティ・エクステンション・デスクリプタ(scalability_extension_descriptor)等が挿入されるものである。そのため、例えば、受信側では、階層符号化における階層情報、トランSPORTストリームTSに含まれるビデオストリームの構成情報などを容易に把握でき、適切なデコード処理を行うことが可能となる。

【0223】

また、図1に示す送受信システム10においては、送信部において、複数の階層を2以上の所定数の階層組に分割し、低階層側の階層組のピクチャの符号化画像データをコンテナするTSパケットの優先度ほど高く設定されるものである。例えば、2分割の場合、「transport_priority」の1ビットフィールドは、ベースレイヤ、つまり低階層側の階層組のピクチャの符号化画像データをコンテナするTSパケットの場合は“1”に設定され、ノンベースレイヤ、つまり高階層側の階層組のピクチャの符号化画像データをコンテナするTSパケットの場合は“0”に設定される。そのため、例えば、受信側では、このTSパケットの優先度に基づいて、自身のデコード能力に応じた階層組のピクチャの符号化画像データのみを圧縮データバッファ(cpb)204に取り込むことが可能となり、バッファ破たんを回避することが容易となる。

【0224】

10

20

30

40

50

また、図1に示す送受信システム10においては、送信側において、各階層組に対応したサブストリームのそれぞれにビットストリームのレベル指定値が挿入され、その値は自己の階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値とされるものである。そのため、ビデオストリームの受信側では、各サブストリームのデコードが可能か否かの判断を、挿入されているビットストリームのレベル指定値に基づいて容易に判断することが可能となる。

【0225】

また、図1に示す送受信システム10においては、送信側において、トランスポートストリームTSのレイヤ（コンテナのレイヤ）に、各階層組のサブストリームに挿入されるビットストリームのレベル指定値が、当該階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値であることを示すフラグ情報（level_constrained_flag）が挿入されるものである。そのため、受信側では、このフラグ情報により、各階層組のサブストリームに挿入されるビットストリームのレベル指定値が当該階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値であることがわかり、sublayer_level_idcを用いた確認処理が不要となり、デコード処理の効率化を図ることが可能となる

【0226】

<2. 変形例>

なお、上述実施の形態においては、送信装置100と受信装置200からなる送受信システム10を示したが、本技術を適用し得る送受信システムの構成は、これに限定されるものではない。例えば、受信装置200の部分が、例えば、（HDMI（High-Definition Multimedia Interface）などのデジタルインターフェースで接続されたセットトップボックスおよびモニタの構成などであってもよい。

【0227】

また、上述実施の形態においては、コンテナがトランスポートストリーム（MPEG-2 TS）である例を示した。しかし、本技術は、インターネット等のネットワークを利用して受信端末に配信される構成のシステムにも同様に適用できる。インターネットの配信では、MP4やそれ以外のフォーマットのコンテナで配信されることが多い。つまり、コンテナとしては、デジタル放送規格で採用されているトランスポートストリーム（MPEG-2 TS）、インターネット配信で使用されているMP4などの種々のフォーマットのコンテナが該当する。

【0228】

また、本技術は、以下のような構成を取ることもできる。

（1）動画像データを構成する各ピクチャの画像データを複数の階層に分類し、該分類された各階層のピクチャの画像データを符号化し、該符号化された各階層のピクチャの画像データを持つビデオストリームを生成する画像符号化部と、

上記生成されたビデオストリームを含む所定フォーマットのコンテナを送信する送信部とを備え、

上記画像符号化部は、

高階層ほどピクチャ毎の符号化画像データのデコード時間間隔が小さくなるように設定されたデコードタイミング情報を上記各階層のピクチャの符号化画像データに付加する送信装置。

（2）上記画像符号化部は、

上記各階層のピクチャの符号化画像データを持つ単一のビデオストリームを生成し、

上記複数の階層を2以上の所定数の階層組に分割し、各階層組のピクチャの符号化画像データに、所属階層組を識別するための識別情報を付加する

前記（1）に記載の送信装置。

（3）上記識別情報は、ビットストリームのレベル指定値であり、高階層側の階層組ほど高い値とされる

前記（2）に記載の送信装置。

（4）上記画像符号化部は、

10

20

30

40

50

上記複数の階層を 2 以上の所定数の階層組に分割し、
上記各階層組のピクチャの符号化画像データをそれぞれ持つ上記所定数のビデオストリームを生成する

前記 (1) に記載の送信装置。

(5) 上記画像符号化部は、

上記各階層組のピクチャの符号化画像データに、所属階層組を識別するための識別情報を付加する

前記 (4) に記載の送信装置。

(6) 上記識別情報は、ビットストリームのレベル指定値であり、高階層側の階層組ほど高い値とされる

10

前記 (5) に記載の送信装置。

(7) 上記画像符号化部は、

上記各階層のピクチャの符号化画像データを持つ単一のビデオストリームを生成するか、あるいは上記複数の階層を 2 以上の所定数の階層組に分割し、各階層組のピクチャの符号化画像データをそれぞれ持つ上記所定数のビデオストリームを生成し、

上記コンテナのレイヤに、該コンテナに含まれるビデオストリームの構成情報を挿入する情報挿入部をさらに備える

前記 (1) から (6) のいずれかに記載の送信装置。

(8) 上記送信部は、

上記複数の階層を 2 以上の所定数の階層組に分割し、低階層側の階層組のピクチャの符号化画像データをコンテナするパケットの優先度ほど高く設定する

20

前記 (1) から (8) のいずれかに記載の送信装置。

(9) 動画像データを構成する各ピクチャの画像データを複数の階層に分類し、該分類された各階層のピクチャの画像データを符号化し、該符号化された各階層のピクチャの画像データを持つビデオストリームを生成する画像符号化ステップと、

送信部により上記生成されたビデオストリームを含む所定フォーマットのコンテナを送信する送信ステップとを有し、

上記画像符号化ステップでは、

高階層ほどピクチャ毎の符号化画像データのデコード時間間隔が小さくなるように設定されたデコードタイミング情報を上記各階層のピクチャの符号化画像データに付加する

30

送信方法。

(10) 動画像データを構成する各ピクチャの画像データが複数の階層に分類されて符号化されることで得られた各階層のピクチャの符号化画像データを持つビデオストリームを含む所定フォーマットのコンテナを受信する受信部を備え、

上記各階層のピクチャの符号化画像データには、高階層ほどピクチャ毎の符号化画像データのデコード時間間隔が小さくなるように設定されたデコードタイミング情報が付加されており、

上記受信されたコンテナに含まれる上記ビデオストリームから選択された所定階層以下の階層のピクチャの符号化画像データを、上記デコードタイミング情報が示すデコードタイミングでデコードして、上記所定階層以下の階層のピクチャの画像データを得る処理部をさらに備える

40

受信装置。

(11) 上記受信されたコンテナには、上記各階層のピクチャの符号化画像データを持つ単一のビデオストリームが含まれており、

上記複数の階層は 2 以上の所定数の階層組に分割され、低階層側の階層組のピクチャの符号化画像データをコンテナするパケットの優先度ほど高く設定されており、

上記処理部は、デコード能力に応じて選択された優先度のパケットでコンテナされた所定階層組のピクチャの符号化画像データをバッファに取り込んでデコードする

前記 (10) に記載の受信装置。

(12) 上記受信されたコンテナには、上記複数の階層が分割されて得られた 2 以上の

50

所定数の階層組のピクチャの符号化画像データをそれぞれ持つ上記所定数のビデオストリームが含まれてあり、

上記処理部は、デコード能力に応じて選択されたビデオストリームが持つ所定階層組のピクチャの符号化画像データをバッファに取り込んでデコードする

前記(10)に記載の受信装置。

(13) 上記処理部で得られる各ピクチャの画像データのフレームレートを表示能力に合わせるポスト処理部をさらに備える

前記(10)から(12)のいずれかに記載の受信装置。

(14) 受信部により動画像データを構成する各ピクチャの画像データが複数の階層に分類されて符号化されることで得られた各階層のピクチャの符号化画像データを持つビデオストリームを含む所定フォーマットのコンテナを受信する受信ステップを有し、

上記各階層のピクチャの符号化画像データには、高階層ほどピクチャ毎の符号化画像データのデコード時間間隔が小さくなるように設定されたデコードタイミング情報が付加されており、

上記受信されたコンテナに含まれる上記ビデオストリームから選択された所定階層以下の階層のピクチャの符号化画像データを、上記デコードタイミング情報が示すデコードタイミングでデコードして、上記所定階層以下の階層のピクチャの画像データを得る処理ステップをさらに有する

受信方法。

(15) 動画像データを構成する各ピクチャの画像データを複数の階層に分類し、該分類された各階層のピクチャの画像データを符号化し、該符号化された各階層のピクチャの画像データを持つビデオストリームを生成する画像符号化部と、

上記生成されたビデオストリームを含む所定フォーマットのコンテナを送信する送信部とを備え、

上記画像符号化部は、

上記各階層のピクチャの符号化画像データを持つ単一のビデオストリームを生成するか、あるいは上記複数の階層を2以上の所定数の階層組に分割し、各階層組のピクチャの符号化画像データをそれぞれ持つ上記所定数のビデオストリームを生成し、

上記コンテナのレイヤに、該コンテナに含まれるビデオストリームの構成情報を挿入する情報挿入部をさらに備える

送信装置。

(16) 動画像データを構成する各ピクチャの画像データを複数の階層に分類し、該分類された各階層のピクチャの画像データを符号化し、該符号化された各階層のピクチャの画像データを持つビデオストリームを生成する画像符号化部と、

上記生成されたビデオストリームを含む所定フォーマットのコンテナを送信する送信部とを備え、

上記送信部は、

上記複数の階層を2以上の所定数の階層組に分割し、低階層側の階層組のピクチャの符号化画像データをコンテナするパケットの優先度ほど高く設定する

送信装置。

(17) 動画像データを構成する各ピクチャの画像データを複数の階層に分類し、該分類された各階層のピクチャの画像データを符号化し、該符号化された各階層のピクチャの画像データを持つビデオストリームを生成する画像符号化部を備え、

上記画像符号化部は、上記複数の階層を2以上の所定数の階層組に分割し、各階層組に対応したサブストリームのそれぞれにビットストリームのレベル指定値を挿入し、

上記各階層組に対応したサブストリームのそれぞれに挿入されるビットストリームのレベル指定値は、自己の階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値とされる

符号化装置。

(18) 上記画像符号化部は、

10

20

30

40

50

上記各階層組に対応したサブストリームのそれぞれを含む上記所定数のビデオストリームを生成する、または上記各階層組に対応したサブストリームの全てを含む単一のビデオストリームを生成する

前記(17)に記載の符号化装置。

(19) 画像符号化部により、動画像データを構成する各ピクチャの画像データを複数の階層に分類し、該分類された各階層のピクチャの画像データを符号化し、該符号化された各階層のピクチャの画像データを持つビデオストリームを生成する画像符号化ステップを有し、

上記画像符号化ステップでは、上記複数の階層を 2 以上の所定数の階層組に分割し、各階層組に対応したサブストリームのそれぞれにビットストリームのレベル指定値を挿入し

上記各階層組に対応したサブストリームのそれぞれに挿入されるビットストリームのレベル指定値は、自己の階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値とされる

符号化方法。

(20) 動画像データを構成する各ピクチャの画像データを複数の階層に分類し、該分類された各階層のピクチャの画像データを符号化し、該符号化された各階層のピクチャの画像データを持つビデオストリームを生成する画像符号化部を備え、

上記画像符号化部は、上記複数の階層を2以上の所定数の階層組に分割し、各階層組に対応したサブストリームのそれぞれにビットストリームのレベル指定値を挿入し、

上記各階層組に対応したサブストリームのそれぞれに挿入されるピットストリームのレベル指定値は、当該階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値とされ、

上記生成されたビデオストリームを含む所定フォーマットのコンテナを送信する送信部と、

上記コンテナのレイヤに、各階層組のサブストリームに挿入されるピットストリームのレベル指定値が当該階層組以下の階層組に含まれる全ての階層のピクチャを含むレベル値であることを示すフラグ情報を挿入する情報挿入部をさらに備える

送信装置。

[0 2 2 9]

本技術の主な特徴は、階層毎にエンコード間隔を算出し、高階層ほどピクチャ毎の符号化画像データのデコード時間間隔が小さくなるように設定されたデコードタイムスタンプが各階層のピクチャの符号化画像データに付加することで、受信側においてデコード性能に応じた良好なデコード処理を可能としたことである（図9参照）。また、本技術の主な特徴は、複数の階層を2以上の所定数の階層組に分割し、低階層側の階層組のピクチャの符号化画像データをコンテナするパケットの優先度ほど高く設定することで、受信側において優先度をもとに自身のデコード能力に応じた階層組のピクチャの符号化画像データのみをバッファに取り込み、バッファ破綻を回避可能にしたことである（図19参照）。

【符号の説明】

【 0 2 3 0 】

1.0: : : 送受信システム

100 · · · 送信装置

101 · · · CPU

102 ● ● ● テンコーダ

1.0.3. . . . 壓縮データバッファ (cph)

104 : : : ヌル毛ブレタ共

105 · · · 送信部

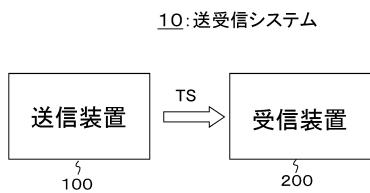
1.3.1.1. テン森ラル T.D. 発生部

133: バッコヌ延制御部

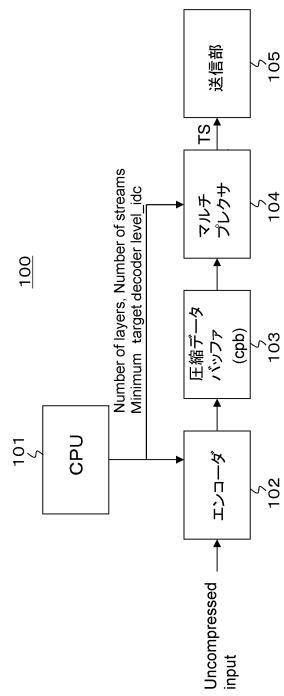
1 2 3 : : : H B D 設 実 部

1 2 4	・	・	・	パラメータセット / S E I エンコード部	
1 2 5	・	・	・	スライスエンコード部	
1 2 6	・	・	・	N A L パケット化部	
1 4 1	・	・	・	T S プライオリティ発生部	
1 4 2	・	・	・	セクションコーディング部	
1 4 3	-1	～	1 4 3	-N	・
1 4 4	・	・	・	PE S パケット化部	
2 0 0	・	・	・	トランスポートパケット化部	
2 0 0	・	・	・	受信装置	
2 0 1	・	・	・	C P U	10
2 0 2	・	・	・	受信部	
2 0 3	・	・	・	デマルチブレクサ	
2 0 4	・	・	・	圧縮データバッファ (c p b)	
2 0 5	・	・	・	デコーダ	
2 0 6	・	・	・	非圧縮データバッファ (d p b)	
2 0 7	・	・	・	ポスト処理部	
2 3 1	・	・	・	P C R 抽出部	
2 3 2	・	・	・	タイムスタンプ抽出部	
2 3 3	・	・	・	セクション抽出部	
2 3 4	・	・	・	T S プライオリティ抽出部	
2 3 5	・	・	・	PE S ペイロード抽出部	20
2 3 6	・	・	・	ピクチャ選択部	
2 5 1	・	・	・	テンポラル I D 解析部	
2 5 2	・	・	・	対象階層選択部	
2 5 3	・	・	・	ストリーム統合部	
2 5 4	・	・	・	デコード部	
2 7 1	・	・	・	補間部	
2 7 2	・	・	・	サブサンプル部	
2 7 3	・	・	・	スイッチ部	

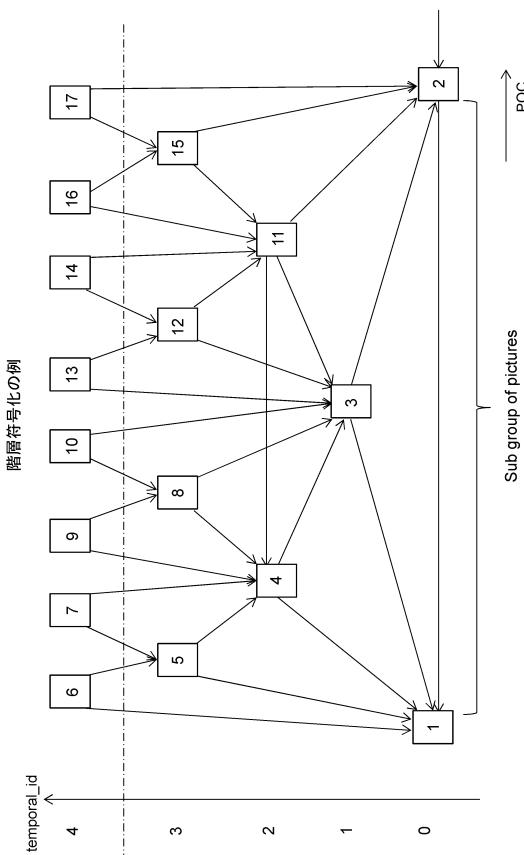
【図1】



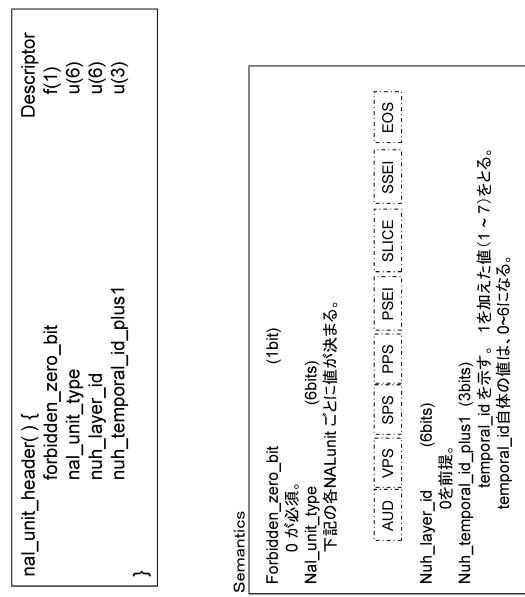
【図2】



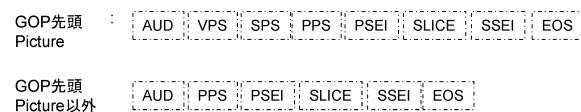
【図3】



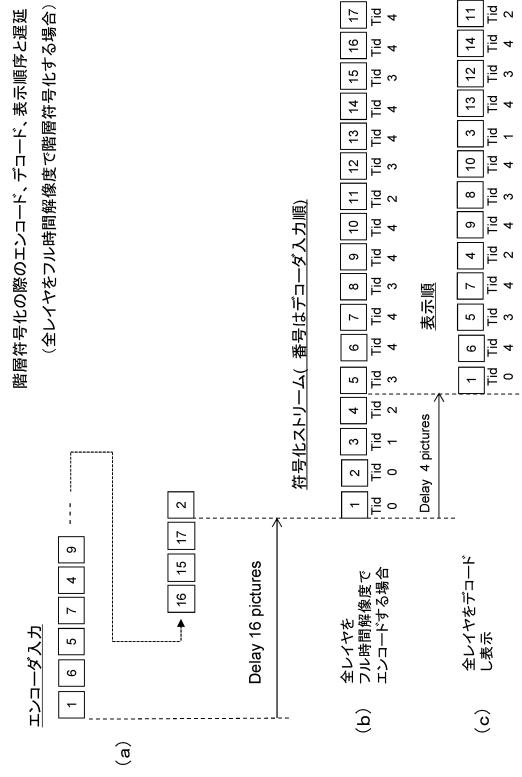
【図4】



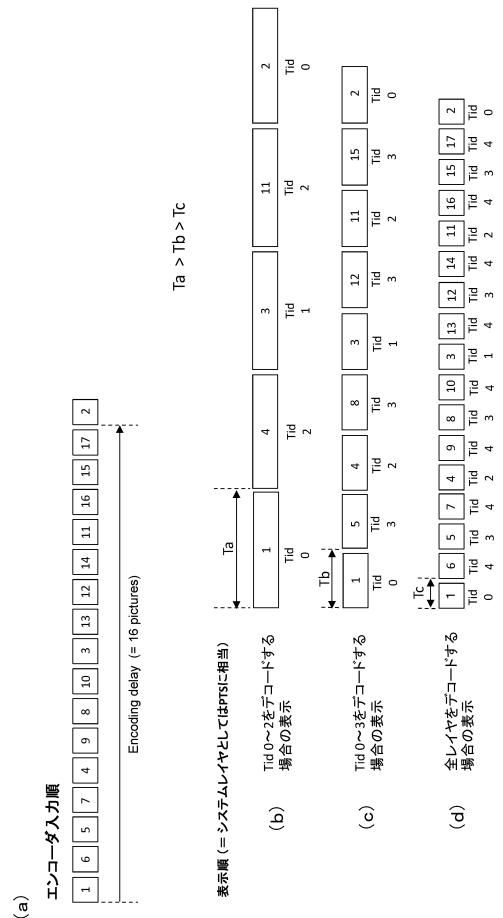
【図5】



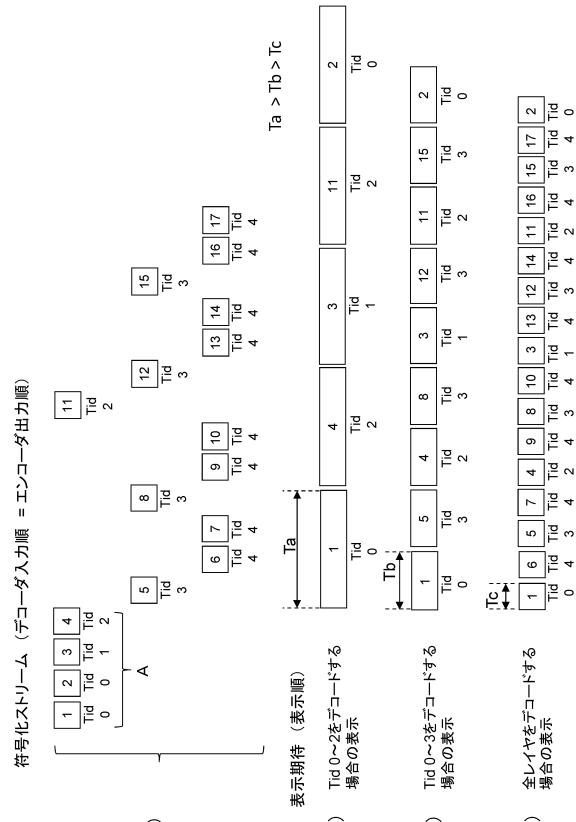
【図 6】



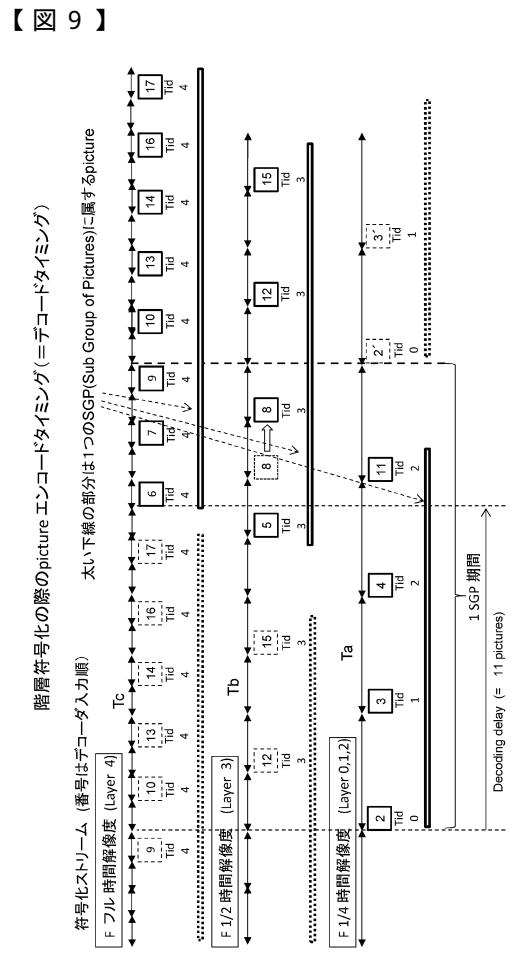
【図 8】



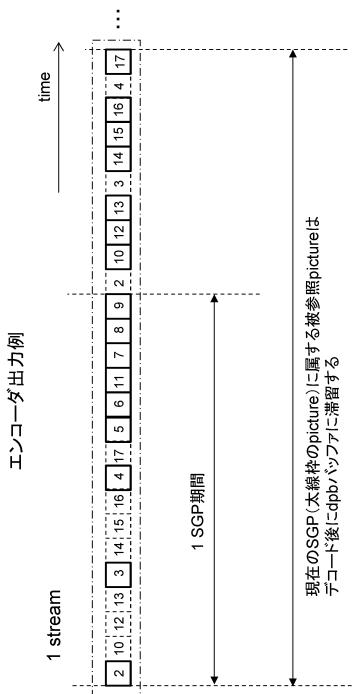
【図 7】



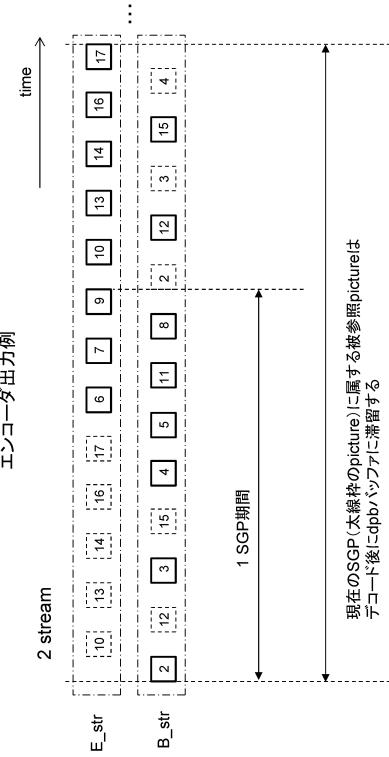
【図 9】



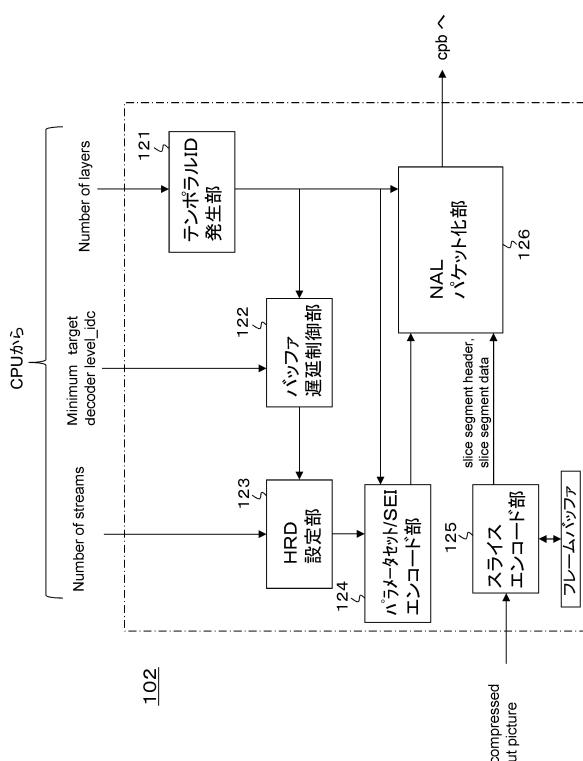
【図10】



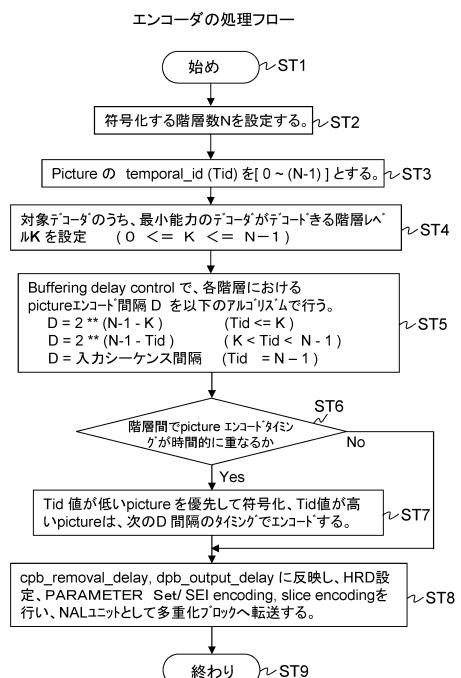
【図11】



【図12】



【図13】



【図 1 4】

```

HEVC_descriptor() {
    descriptor_tag
    descriptor_length
    profile_space
    tier_flag
    profile_idc
    reserved_zero_16bits
    level_idc
    profile_compatibility_indication
    temporal_layer_subset_flag
    HEVC_still_present_flag
    HEVC_24hr_picture_present_flag
    frame_packing_arrangement_SEI_present_flag
    level_constrained_flag
    scalability_id
    if ( temporal_layer_subset_flag == '1' ) {
        temporal_id_min
        reserved
        temporal_id_max
        reserved
    }
}

```

【図 1 5】

Semantics	
level_constrained_flag (1bit)	該当substreamにSPSあるいはPPSが存在し、その要素のGeneral_level_idcは、そのsubstreamが含むTemporal_id 以下のPictureを含むLevel値をもつことを示す。
1	該当substreamにSPSあるいはPPSが存在し、その要素のGeneral_level_idcは、そのsubstreamが含むTemporal_id 以下のPictureを含むLevel値をもつ。
0	対象となるサービスを構成するサービス・ストリーム群の中にSPSがつづけられ、そのgeneral_level_idcは、当該サービス・ストリームのみならず、同一サービスの下の他のサービス・ストリームも含むレベル値を示す。
scalability_id (3bits)	複数のビデオストリームが一つのスケーラブルなサービスを供給する際、個々のストリームに付されるスケーラビリティを示す ID
0	Base stream
1~7	Base streamからのscalabilityの度合いによって増加するID

【図 1 6】

scalability_extension_descriptor syntax

Syntax	No. of Bits	Format
scalability_extension_descriptor() {		
scalability_extension_descriptor_tag	8	uimsbf
scalability_extension_descriptor_length	8	uimsbf
extension_stream_existing_flag	1	bslbf
extension_type	3	bslbf
number_of_streams	4	bslbf
scalability_id	3	bslbf
minimum_target_decoder_level_idc	8	bslbf
reserved	5	0x1f
}		

【図 1 7】

scalability_extension_descriptor semantics	
extension_stream_existing_flag (1bit)	別ストリームによる拡張サービスがあることを示すフラグ
1	拡張ストリームがある
0	拡張ストリームがない
extension_type (3bits)	拡張ストリームのタイプを示す
001	拡張は時間方向スケーラブル
010	拡張は空間方向スケーラブル
011	拡張はビットレースケーラブル
others	reserved
number_of_streams (4bits)	配信サービスに関するストリームの総数
0000	reserved
0001	1 stream
⋮	
1111	15 streams
scalability_id (3bits)	複数のビデオストリームが一つのスケーラブルなサービスを供給する際、個々のストリームに付されるスケーラビリティを示す ID
0	Base stream
1~7	Base streamからのscalabilityの度合いによって増加するID
minimum_target_decoder_level_idc (8bits)	該当するストリームが対象とする最小デコード能力のlevel値を示す
general_level_idcの定義と一致。	

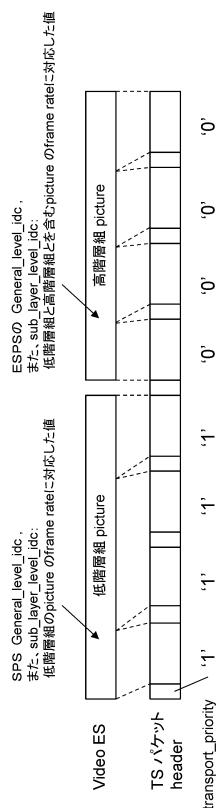
【図18】

```

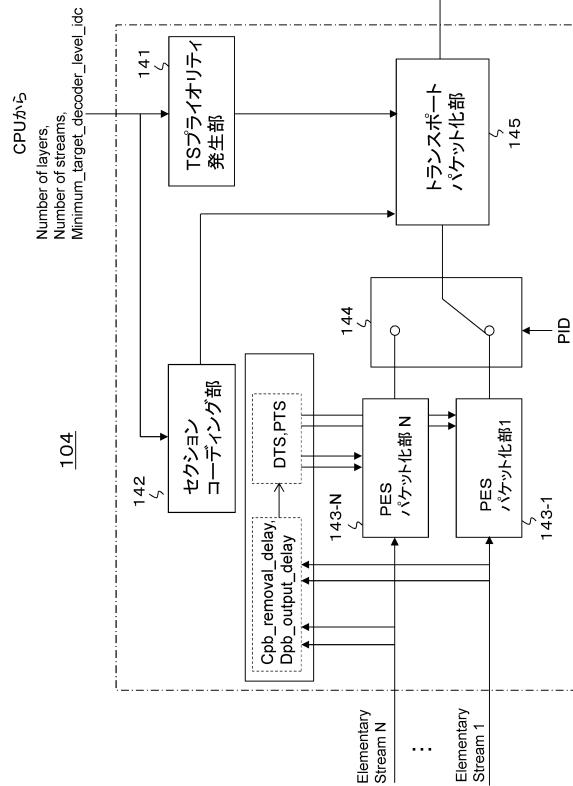
Syntax
transport_packet(){
    sync_byte
    transport_error_indicator
    payload_unit_start_indicator
    transport_priority
    PID
    transport_scrambling_control
    adaptation_field_control
    continuity_counter
    if(adaptation_field_control=='10' ||
adaptation_field_control=='11'){
        adaptation_field()
    }
    if(adaptation_field_control=='01' ||
adaptation_field_control=='11'){
        for (i=0;i<N;i++){
            data_byte
        }
    }
}

```

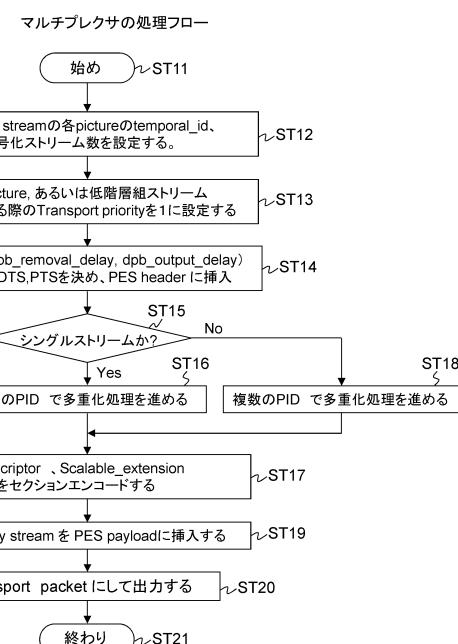
【図19】



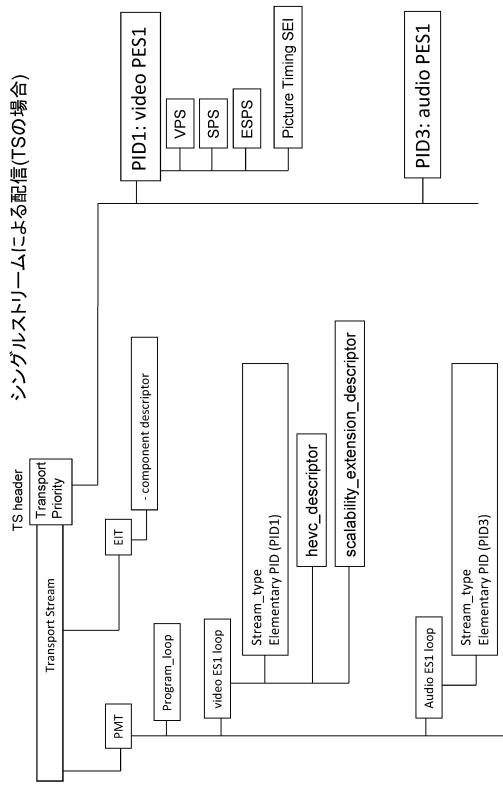
【図20】



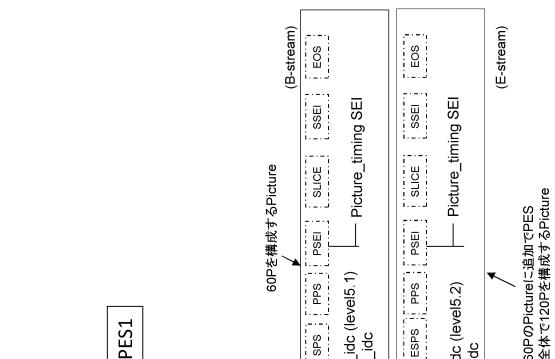
【図21】



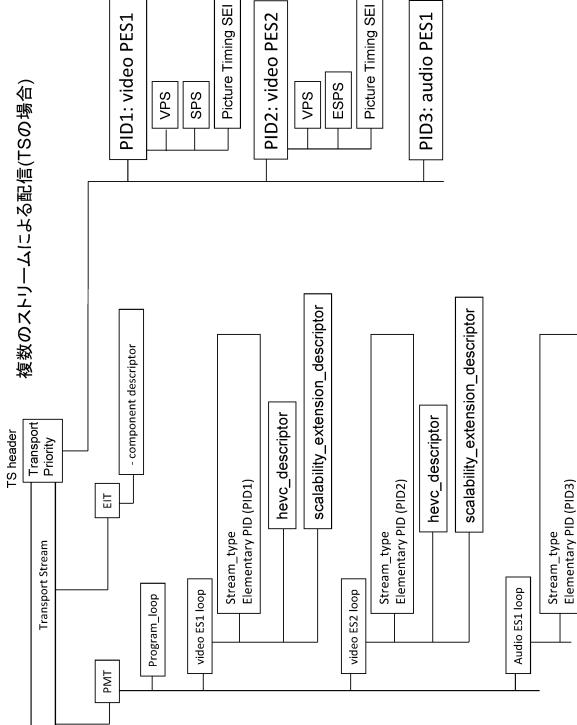
【図 2 2】



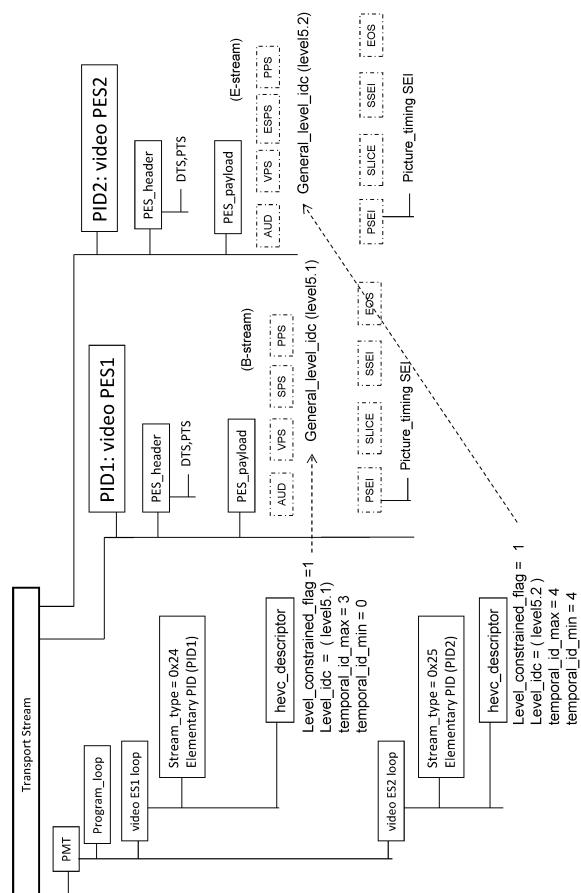
【図 2 3】



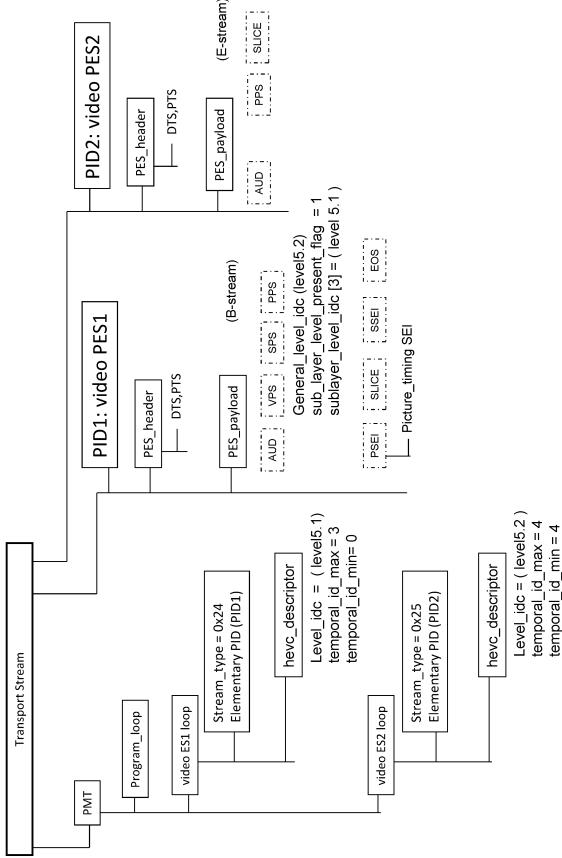
【図 2 4】



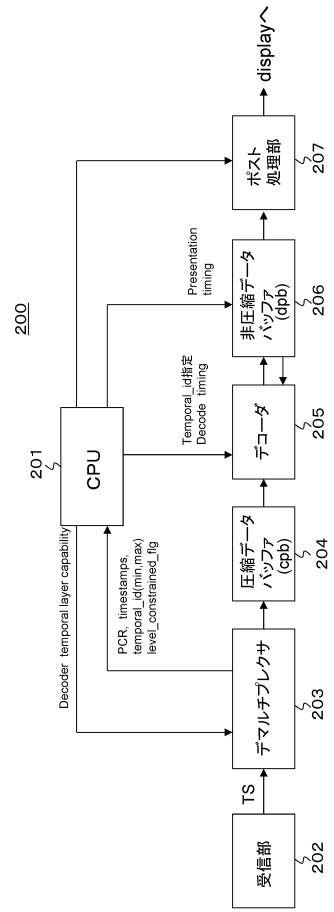
【図 2 5】



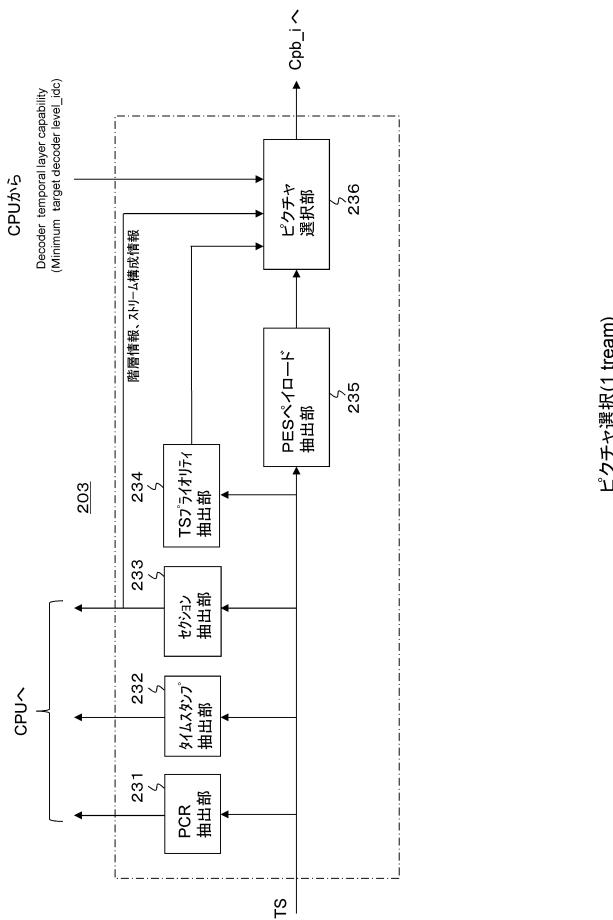
【図26】



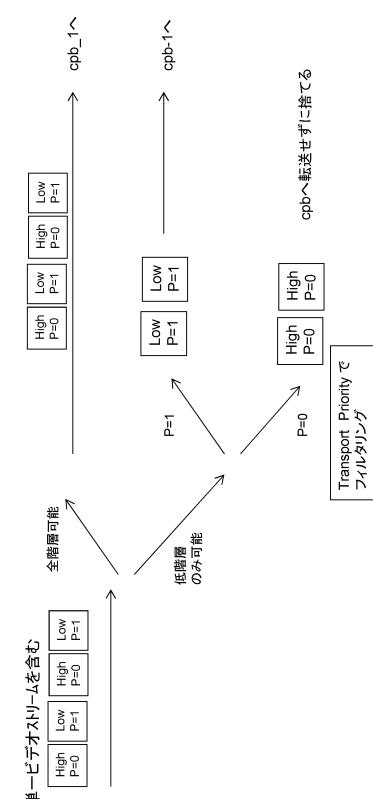
【 図 27 】



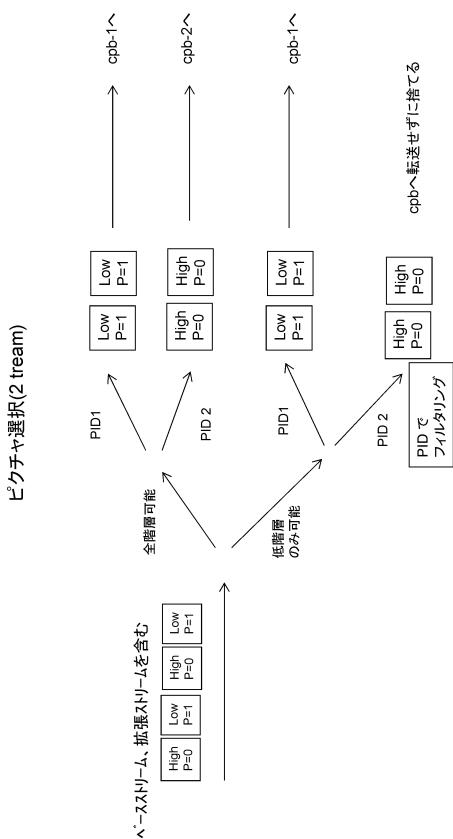
【図28】



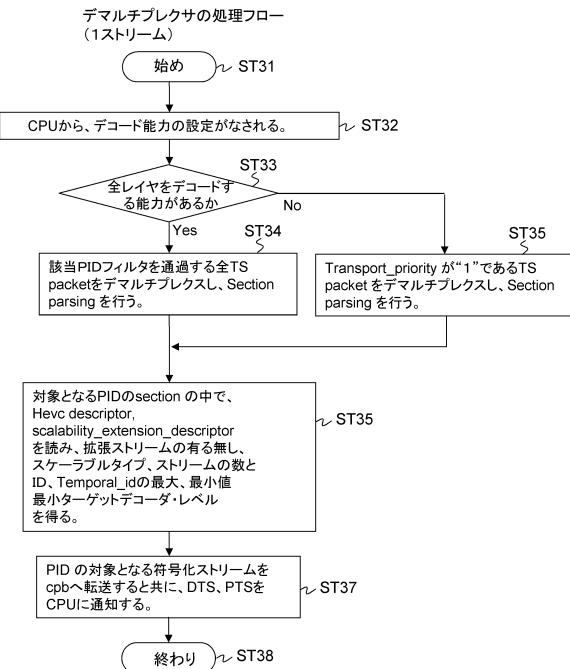
【図29】



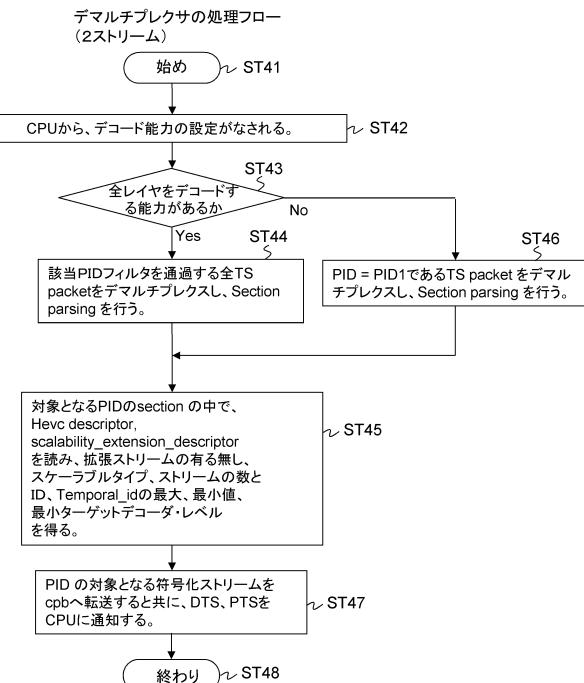
【図30】



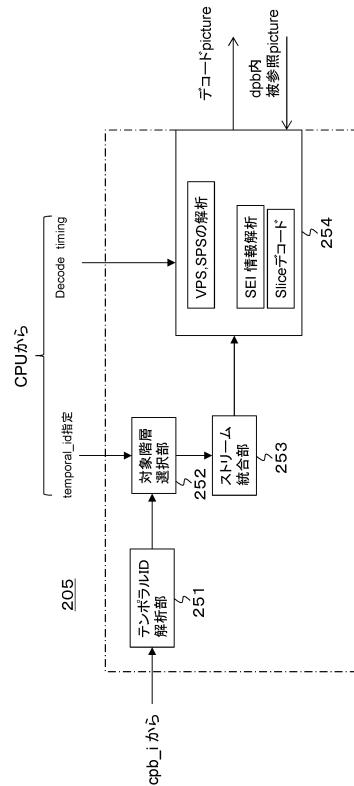
【図31】



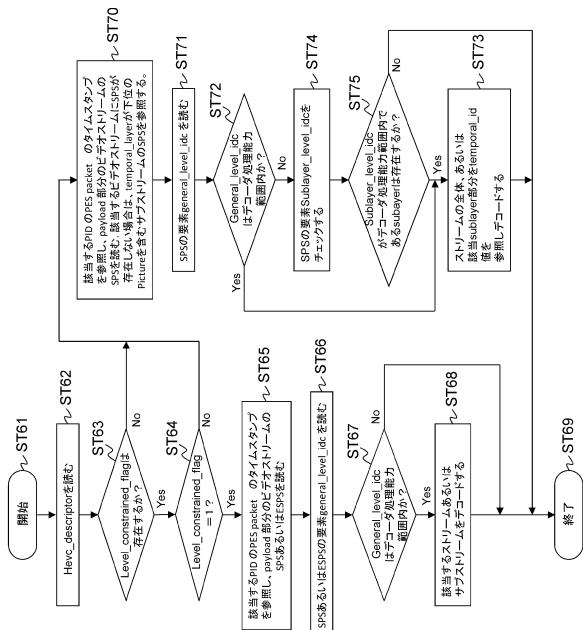
【図32】



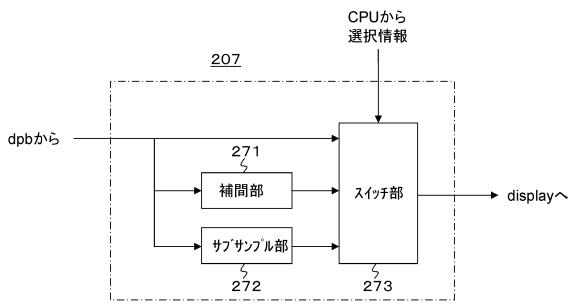
【図33】



【図34】

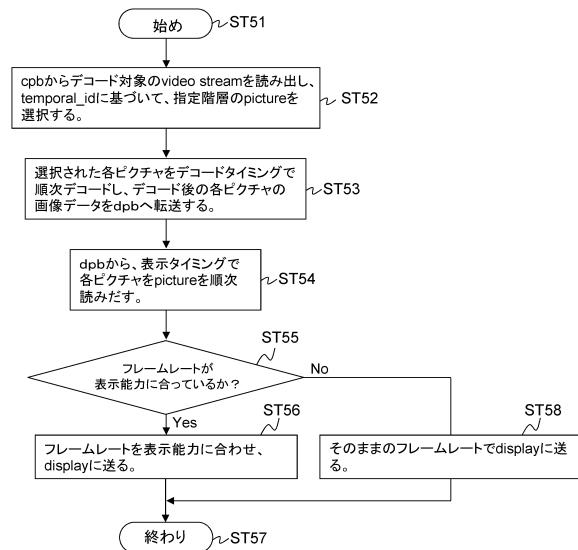


【図35】



【図36】

デコーダ、ポスト処理部の処理フロー



フロントページの続き

(72)発明者 塚越 郁夫
東京都港区港南1丁目7番1号 ソニー株式会社内

審査官 牛丸 太希

(56)参考文献 特開2002-010251(JP, A)
国際公開第03/075524(WO, A1)
特開2005-192013(JP, A)
国際公開第2012/023281(WO, A1)
国際公開第2010/032636(WO, A1)
特表2006-515132(JP, A)
特開2004-088244(JP, A)
再公表特許第11/052045(JP, A1)
特開2011-035747(JP, A)
米国特許出願公開第2006/0256851(US, A1)
Thomas Rusert, et al., High level syntax for scalability support in HEVC, JCTVC-F491, ITU-T, 2011年 7月16日, pp1~9
ISO/IEC JTC1/SC29, Transport of High Efficiency Video Coding (HEVC) Streams over ITU-T Rec H.222.0, ISO/IEC 13818-1:201X/PDAM3(E) Rec. ITU-T H.222.0/Amd. 3, ISO/IEC JTC1/SC29, 2012年 5月24日, pp1~9, <http://itscj.ipsj.or.jp/sc29/open/29n12806t.doc>

(58)調査した分野(Int.Cl., DB名)

H04N 19/00-19/98
H04N 21/00-21/858