

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6588801号
(P6588801)

(45) 発行日 令和1年10月9日(2019.10.9)

(24) 登録日 令和1年9月20日(2019.9.20)

(51) Int.Cl.

F 1

| | | |
|-------------|-----------|-------------|
| HO4N 19/587 | (2014.01) | HO4N 19/587 |
| HO4N 19/31 | (2014.01) | HO4N 19/31 |
| HO4N 19/40 | (2014.01) | HO4N 19/40 |

請求項の数 12 (全 14 頁)

(21) 出願番号

特願2015-214800 (P2015-214800)

(22) 出願日

平成27年10月30日 (2015.10.30)

(65) 公開番号

特開2017-85495 (P2017-85495A)

(43) 公開日

平成29年5月18日 (2017.5.18)

審査請求日

平成30年10月17日 (2018.10.17)

(73) 特許権者 000001007

キヤノン株式会社

東京都大田区下丸子3丁目30番2号

(74) 代理人 100076428

弁理士 大塚 康徳

(74) 代理人 100115071

弁理士 大塚 康弘

(74) 代理人 100112508

弁理士 高柳 司郎

(74) 代理人 100116894

弁理士 木村 秀二

(74) 代理人 100130409

弁理士 下山 治

(74) 代理人 100134175

弁理士 永川 行光

最終頁に続く

(54) 【発明の名称】 画像処理装置、画像処理方法、及び、プログラム

(57) 【特許請求の範囲】

【請求項 1】

時間的な階層構造を有するピクチャで構成される第1の動画と第2の動画とを結合して第3の動画を生成する画像処理装置であって、

前記第1の動画と前記第2の動画の時間的な階層構造を比較する比較手段と、

前記比較により前記時間的な階層構造が不一致と判定された場合に、前記第2の動画が有する時間的な階層構造を前記第1の動画が有する時間的な階層構造と一致させるように前記第2の動画を変換する変換手段と、

前記第1の動画と、前記変換手段が変換した前記第2の動画とを結合して前記第3の動画を生成する結合手段と

を備え、

前記変換手段は、前記第2の動画の変換を、該第2の動画の復号を伴わずに該第2の動画が有するピクチャ数を増加、または、減少させることにより行い、

前記時間的な階層構造における各階層は、フレームレートと関連づけられ、

前記時間的な階層構造が不一致の場合とは、前記時間的な階層構造に含まれる階層の数が異なる場合と、前記時間的な階層構造に含まれる各階層と関連づけられたフレームレートが異なる場合との少なくともいずれかの場合であり、

前記第1の動画と、前記第2の動画とが有する前記時間的な階層の最下位の階層のフレームレートが異なっている場合に、前記変換手段は、前記第2の動画の最下位の階層のフレームレートを変更して、前記第1の動画の最下位の階層のフレームレートと一致させた

後、前記第2の動画の変換を行うことを特徴とする画像処理装置。

【請求項2】

前記変換手段は、前記ピクチャの数を増加させる場合に、前記第2の動画が有するピクチャを参照する、該ピクチャのコピーピクチャを挿入することにより前記第2の動画の変換を行うことを特徴とする請求項1に記載の画像処理装置。

【請求項3】

前記コピーピクチャは、前記時間的な階層構造において、前記第2の動画が有するピクチャが属する階層の上位の階層のピクチャとして挿入されることを特徴とする請求項2に記載の画像処理装置。

【請求項4】

前記変換手段は、前記ピクチャの数を減少させる場合に、前記第2の動画が有するピクチャのうち、前記第1の動画が有しない階層に属するピクチャを破棄することにより前記第2の動画の変換を行うことを特徴とする請求項1又は2に記載の画像処理装置。

【請求項5】

前記変換手段は、前記第1の動画が有しない階層は、前記第2の動画の時間的な階層構造における最上位の階層を含むことを特徴とする請求項4に記載の画像処理装置。

【請求項6】

前記時間的な階層構造において、上下に隣接する2つの階層において、上位の階層は下位の階層の2倍のフレームレートを有することを特徴とする請求項1に記載の画像処理装置。

10

【請求項7】

前記第1の動画の最下位の階層のフレームレートが、前記第2の動画の最下位の階層のフレームレートよりも高い場合、

前記変換手段は、前記第2の動画が有する階層のうち、前記最下位の階層から、前記第1の動画の最下位の階層に対応する階層までに属するピクチャを統合することにより、前記第2の動画の最下位の階層のフレームレートを前記第1の動画の最下位の階層のフレームレートと一致させることを特徴とする請求項1に記載の画像処理装置。

20

【請求項8】

前記第1の動画の最下位の階層のフレームレートが、前記第2の動画の最下位の階層のフレームレートよりも低い場合に、

30

前記変換手段は、前記第2の動画を復号したのち、前記第1の動画の最下位の階層に対応する階層を有する動画として再符号化することにより、前記第2の動画の最下位の階層のフレームレートを前記第1の動画の最下位の階層のフレームレートと一致させることを特徴とする請求項1又は7に記載の画像処理装置。

【請求項9】

動画を復号する復号手段を更に備え、

前記復号手段は、前記第3の動画を復号することを特徴とする請求項1乃至8のいずれか1項に記載の画像処理装置。

【請求項10】

被写体を撮影して画像を生成する撮像手段と、

40

前記画像を時間的な階層構造を有するピクチャとして符号化して動画を生成する符号化手段と、

請求項1乃至9のいずれか1項に記載の画像処理装置の各手段とを備えることを特徴とする撮像装置。

【請求項11】

コンピュータを請求項1乃至9のいずれか1項に記載の画像処理装置の各手段として機能させることを特徴とするプログラム。

【請求項12】

時間的な階層構造を有するピクチャで構成される第1の動画と第2の動画とを結合して第3の動画を生成する画像処理方法であって、

50

前記第1の動画と前記第2の動画の時間的な階層構造を比較する比較工程と、
前記比較により前記時間的な階層構造が不一致と判定された場合に、前記第2の動画が
有する時間的な階層構造を該第1の動画が有する時間的な階層構造と一致させるように前
記第2の動画を変換する変換工程と、

前記第1の動画と、前記変換工程で変換された前記第2の動画とを結合して前記第3の
動画を生成する結合工程と
を備え、

前記変換工程では、前記第2の動画の変換が、該第2の動画の復号を伴わずに該第2の
動画が有するピクチャ数を増加、又は、減少させることにより行われ、

前記時間的な階層構造における各階層は、フレームレートと関連づけられ、

10

前記時間的な階層構造が不一致の場合とは、前記時間的な階層構造に含まれる階層の数
が異なる場合と、前記時間的な階層構造に含まれる各階層と関連づけられたフレームレー
トが異なる場合との少なくともいずれかの場合であり、

前記第1の動画と、前記第2の動画とが有する前記時間的な階層の最下位の階層のフレ
ームレートが異なっている場合に、前記変換工程では、前記第2の動画の最下位の階層の
フレームレートを変更して、前記第1の動画の最下位の階層のフレームレートと一致させ
た後、前記第2の動画の変換が行われることを特徴とする画像処理方法。

【発明の詳細な説明】

【技術分野】

【0001】

20

本発明は、画像処理装置、画像処理方法、及び、プログラムに関する。

【背景技術】

【0002】

近年、デジタルビデオカメラや動画撮影機能付きデジタルカメラの普及により、動画データを扱う機会が増加している。これに伴い、様々な動画編集アプリケーションが提供されている。こうした動画編集の1つに、複数の動画を組み合わせて1つの動画を生成する編集がある。例えば、特許文献1では、異なるフレームレートを持つ複数の動画を結合編集する際に、フレームレートの互換性が取れるように、一方のフレームレートを変換する方法を開示する。

【先行技術文献】

30

【特許文献】

【0003】

【特許文献1】特開2008-278492号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

特許文献1が開示する方法では、異なるフレームレートの動画を結合する場合、変換元の符号化データを一旦復号してから変換先のフレームレートに従って再符号化していた。このように対象となる符号化データの全体を復号し、再符号化する処理を編集の度毎に行うと、処理時間が膨大となり処理の効率化を図ることが困難になってしまう。

40

【0005】

そこで本発明は、異なるフレームレートの動画を結合編集する場合に効率的な処理を可能にすることを目的とする。

【課題を解決するための手段】

【0006】

上記課題を解決するための本発明は、時間的な階層構造を有するピクチャで構成される第1の動画と第2の動画とを結合して第3の動画を生成する画像処理装置であって、

前記第1の動画と前記第2の動画の時間的な階層構造を比較する比較手段と、

前記比較により前記時間的な階層構造が不一致と判定された場合に、前記第2の動画が
有する時間的な階層構造を前記第1の動画が有する時間的な階層構造と一致させるように

50

前記第2の動画を変換する変換手段と、

前記第1の動画と、前記変換手段が変換した前記第2の動画とを結合して前記第3の動画を生成する結合手段とを備え、

前記変換手段は、前記第2の動画の変換を、該第2の動画の復号を伴わずに該第2の動画が有するピクチャ数を増加、または、減少させることにより行い、

前記時間的な階層構造における各階層は、フレームレートと関連づけられ、

前記時間的な階層構造が不一致の場合とは、前記時間的な階層構造に含まれる階層の数が異なる場合と、前記時間的な階層構造に含まれる各階層と関連づけられたフレームレートが異なる場合との少なくともいずれかの場合であり、

前記第1の動画と、前記第2の動画とが有する前記時間的な階層の最下位の階層のフレームレートが異なっている場合に、前記変換手段は、前記第2の動画の最下位の階層のフレームレートを変更して、前記第1の動画の最下位の階層のフレームレートと一致させた後、前記第2の動画の変換を行うことを特徴とする。10

【発明の効果】

【0007】

本発明によれば、異なるフレームレートの動画を結合編集する場合に効率的な処理が可能となる。

【図面の簡単な説明】

【0008】

【図1】発明の実施形態に対応する動画符号化装置の構成例を示すブロック図。20

【図2】発明の第1の実施形態に対応する処理の一例を示すフローチャート。

【図3】発明の第1の実施形態に対応する、編集対象動画のフレームレートを変更する処理の一例を説明する図。

【図4】発明の第1の実施形態に対応する、編集対象動画のフレームレートを変更する処理の他の例を説明する図。

【図5】発明の第2の実施形態に対応する処理の一例を示すフローチャート。

【図6】発明の第2の実施形態に対応する、編集対象動画の最下位の階層を基準動画と一致させるための処理の一例を説明する図。

【図7】発明の第2の実施形態に対応する、再符号化により動画の時間的階層構造を再構成する例を説明する図。30

【図8】発明の第2の実施形態に対応する、編集対象動画のフレームレートを変更する処理の他の例を説明する図。

【発明を実施するための形態】

【0009】

以下、図面を参照して、本発明の実施形態を詳細に説明する。

【0010】

【実施形態1】

図1に本発明の第1の実施形態に対応する動画符号化装置の構成例を示す。図1に示す動画符号化装置は、入力される映像信号に対して符号化処理を施して記録する他、記録した複数の符号化データの結合編集が可能となっている。当該動画符号化装置は、例えばデジタルカメラ、デジタルビデオカメラとして実現することができる。また、それ以外に、例えばパーソナルコンピュータ、携帯電話、スマートフォン、PDA、タブレット端末などの任意の情報処理端末或いは撮像装置として実現することもできる。なお、図1は、デジタルカメラ等として機能する場合を考慮して撮像部102を含む構成を示した。しかし、発明の実施形態として、フレームレートや時間的階層構造の異なる動画を結合編集するための画像編集装置、或いは、画像処理装置として、撮像部102を有しない構成で実現してもよい。40

【0011】

図1を参照しながら、本実施形態の動画符号化装置100によるH.265規格に基づいて符号化した複数の動画の結合編集処理について説明する。なお、本実施形態ではH.50

265 規格に基づいた符号化方式を例に説明するが、H.264等の他の時間スケーラブル符号化方式にも適用可能である。

【0012】

図1に示す操作部101は、ユーザの指令を受け付け、指令に応じた制御信号を生成し、CPU103に供給する。例えば、操作部101は、ユーザ指令を受け付ける入力デバイスとして、ボタンやタッチパネル等からなる。撮像部102は、CMOSイメージセンサなどの撮像素子とA/D変換器を含んで構成される。撮像素子は、操作部101からの記録指示に従って被写体を撮影して得られた光学像を光電変換してアナログ画像信号を出力し、A/D変換器が当該画像信号をA/D変換して画像データを生成する。画像データは、メモリ104に出力され、記憶される。CPU103は、操作部101から入力される制御信号に基づき、本実施形態に対応する動画符号化装置100を構成する各処理部を制御する。メモリ104は、撮像部102から出力された画像データや、符号化部105または、復号化部106による処理後の動画データなどを一時的に記憶する。また、メモリ104には、CPU103のワークメモリとしても機能し、さらに後述する各種の処理を実行させるためのプログラムやデータも格納している。10

【0013】

符号化部105は、メモリ104に記憶された画像データを、時間スケーラブル符号化方式に従って符号化する。所定画素ブロック単位に画面内又は画面間予測符号化を行い、直交変換、量子化を行い、量子化されたデータをエントロピー符号化して符号化データを生成する。ここで、編集性の高い符号化データを生成するアプローチの1つとして、本実施形態では時間スケーラブル符号化により符号化データを生成する。時間スケーラブル符号化では、予めフレームレートに時間的な階層構造に基づくスケーラビリティを持たせることで、1つの符号化データから複数のフレームレートの動画像が生成できる。20

【0014】

例えば、120fpsのフレームレートを有する符号化データの場合、当該符号化データを構成するピクチャが属する時間的階層（時間レイヤ、或いは、レイヤ）として、30fps、60fps、120fpsのそれぞれのフレームレートに対応する階層を有することができる。また、当該階層構造では、上下に隣接し合う階層のフレームレートにつき、上位の階層のフレームレートが下位の階層のフレームレートの2倍となるように各階層の位置関係が決定される。より具体的に、例えば120fpsを構成する動画の1秒に相当する120枚のピクチャのうち、30枚が最下位の階層である30fpsのフレームレートと関連づけられた第1の階層に属することとなる。また、第1の階層の1つ上位の階層である60fpsのフレームレートと関連づけられた第2の階層には30枚のピクチャが属することとなる。これにより、第1の階層と第2の階層とを合わせて60枚のピクチャとなり、60fpsのフレームレートを達成することができる。更に、第2の階層の1つ上位の階層である120fpsのフレームレートと関連づけられた最上位の階層である第3の階層には残りの60枚のピクチャが属することとなる。これにより、第1から第3の階層までのピクチャを統合して120枚のピクチャとなり120fpsのフレームレートを達成することができる。30

【0015】

ただし、時間スケーラブル符号化では、各ピクチャの画像参照構造が階層的になるよう⁴⁰に符号化を行い、ピクチャ間の画像参照関係は、必ず自身のレイヤよりも上位の時間レイヤの画像を参照しないように制限した上で、各ピクチャに対して時間レイヤ情報を作成する。よって、上記例では、60fpsの第2の階層に属するピクチャは第1の階層に属するピクチャしか参照できない。また、120fpsの第3の階層に属するピクチャは、第1又は第2の階層に属するピクチャしか参照できない。符号化部105は、生成した符号化データをメモリ104に書き込む。

【0016】

復号化部106は、メディア108に記録された符号化データを読み込み、エントロピー復号化する。そして、逆量子化、逆直交変換を行い、復号した画像データをメモリ1050

4に書き出す。復号した画像データは、CPU103の指示に基づき、表示部107等に供給し、表示することができる。表示部107は、編集の進捗状況や、復号化部106で復号した画像データを表示するディスプレイである。メディア108は、ハードディスクドライブ、または不揮発性メモリ等で構成される記憶媒体であって、符号化データが記録される。内部バス109は、本実施形態に対応する動画符号化装置100を構成する各処理部を接続するインターネット・バスである。

【0017】

本実施形態の動画符号化装置100は、以上のように構成される。また、本実施形態における動画符号化装置100は、複数の動画の結合編集が可能となっており、以下の様に動作する。

10

【0018】

まず、ユーザが操作部101を介して編集開始の指示が入力すると、編集対象の候補となる動画の一覧画面が表示部107に表示される。ユーザは操作部101を操作して、表示された候補の中から編集対象とする動画を選択することができる。編集対象動画が選択されると、CPU103は、メディア108に記録した動画の中から、ユーザが指定する編集対象となる動画のフレームレート情報と時間レイヤ情報を取得する。ここで、ユーザは操作部101を介して、編集対象として選択した動画の中からフレームレート互換性基準となる動画（以下、「基準動画」ともいう）を更に選択する。フレームレート互換性基準となる動画とは、編集対象動画の間でフレームレートが異なる場合に行うフレームレート変換において基準となる動画である。その後CPU103は、編集対象動画のデータをメモリ104に書き出し、結合編集処理を開始する。結合編集処理が終了すると、CPU103は、結合編集した動画データをメモリ104に書き出し、メディア108に記録する。

20

【0019】

なお、結合編集処理では、編集対象となる動画間でフレームレートおよび時間レイヤ構造が異なる場合、編集対象となる動画の間でフレームレートの互換性を取る必要がある。そこで、本実施形態では、フレームレート互換性基準となる動画を除いた編集対象となる動画の時間レイヤ構造を変換してフレームレートの変換を行い、結合編集処理を実行する。

【0020】

30

以下、本実施形態に対応する結合編集処理の詳細について図2に示すフローチャートを参照して説明する。以下では、CPU103が行う複数の動画の結合編集において、動画間のフレームレートが異なる場合の時間レイヤ構造の変換方法について特に詳細に説明する。図2のフローチャートに対応する処理は、例えば、結合編集処理部として機能するCPU103が対応するプログラム（メモリ104に格納）を実行することにより実現できる。

【0021】

まず、S201において、編集対象としてユーザが選択した各動画のフレームレートと時間レイヤ構造の情報を、該動画を保存するメディア108から取得する。フレームレート(fps)は、ファイルフォーマット中に動画情報として埋め込まれている場合、ファイルフォーマットを参照することによって取得することが可能である。また、符号化データのヘッダ部のHRD parameters syntaxに含まれる、num_units_in_tickと、time_scaleから算出し、取得することも可能である。一方、時間レイヤ情報を取得する方法は、符号化データのヘッダ部のnal_unit_header Syntaxに含まれるnuh_temporal_ID_plus1から判別する方法や、VPS(Video Parameter Set) Syntaxに含まれるvps_MAX_sub_layers_minus1から取得する方法や、SPS(Sequence Parameter Set) Syntaxに含まれるsps_MAX_sub_layers_minus1から取得することが可能である。ここでは、編集対象となる動画のフレームレートと時間レイヤ構造との関係は以下の通りである。なお、Temporal ID(時間ID)は時間レイヤ番号を表す指示子である。

40

i) フレームレート：30fps

50

Temporal ID0 = 30 f p s

ii) フレームレート: 60 f p s

Temporal ID0 = 30 f p s

Temporal ID1 = 60 f p s

iii) フレームレート: 120 f p s

Temporal ID0 = 30 f p s

Temporal ID1 = 60 f p s

Temporal ID2 = 120 f p s

【0022】

次に S202 では、S201 で編集対象として選択した動画の中から、フレームレートの互換性を持たせるうえで基準となる基準動画をユーザが選択する。以下に説明する例では、上記 3 つの動画の中から 60 f p s の動画が基準動画として選択されたものとする。なお、フレームレート互換性基準（以下、単に「基準」ともいう。）となる動画の選択は、操作部 101 を介したユーザ指示に従うだけでなく、例えば編集対象となる動画の中でフレームレートの最も高い動画や、フレームレートの最も低い動画を基準として自動的に選択してもよい。また、選択された編集対象画像のフレームレートが不一致の場合にのみ実行してもよい。なお、選択方法自体が発明の本質的特徴ではないので、上述以外の方法であってもよい。

【0023】

S203 では、フレームレート変換対象となる動画（以下、「変換対象動画」ともいう。）のフレームレート（以下、「FRc」という。）と、基準動画のフレームレート（以下、「FRr」という。）とで、フレームレートが等しいかどうかを判定する。FRc と FRr とが等しい場合、S208 に処理を進め、異なる場合は、S204 に処理を進める。

【0024】

S204 では、FRc と FRr とで、フレームレートの大小比較を行う。FRc の方が小さい場合、S205 に処理を進め、大きい場合には、S206 に処理を進める。ここでは、基準動画は 60 f p s と設定している。従って、30 f p s の動画は、S205 に処理を進め、また、120 f p s の動画は、S206 に処理を進める。

【0025】

S205 では、変換対象動画のフレームレート FRc が、基準動画のフレームレート FRr より小さい場合のフレームレート変換処理及び時間レイヤ構造変換処理を行う。具体的には、フレーム補間によってフレームレートを増加させる処理で、先頭ピクチャからピクチャ毎に直前に表示されるピクチャを繰り返して表示するピクチャデータ（以下、「コピーピクチャデータ」という。）を生成し挿入する。即ち、同一内容のピクチャが 2 回ずつ表示されるように編集することで、30 f p s の動画が 60 f p s の動画にアップコンバートされる。なお、コピーピクチャデータは、フレームレート不足を補うための補間データで、そのデータ構成は P ピクチャもしくは B ピクチャの形式で、直前のピクチャを参照し、そのまま保持する符号化データである。これらのコピーピクチャデータは、時間レイヤ指示子である Temporal ID を、コピーピクチャ参照元の Temporal ID + 1 の上位時間レイヤとして扱う。ここで S205 処理は、30 f p s の動画を 60 f p s の動画に変換する処理であって、S205 処理によって変換した時間レイヤ構造を図 3 に示す。

【0026】

図 3 (a) は 30 f p s を有する変換対象動画に対してコピーピクチャデータを挿入して、フレームレートを 60 f p s に向上させる例を示している。ここでは、30 f p s のフレームレートを構成する各ピクチャのコピーピクチャが P ピクチャとして生成され、それぞれ生成元となったピクチャを参照している。なお、図 3 では簡単のために P ピクチャのみで生成しているが、B ピクチャとして生成してもよい。生成されたコピーピクチャの Temporal ID は、元のピクチャの Temporal ID である 0 よりも 1 つ上のレイヤを示す値として 1 となる。図 3 (b) は、図 3 (a) のような変換処理を行う際の基準動画の一例を示す。

10

20

30

40

50

しており、ここでは 60fps の動画が示されている。

【0027】

次に S206 では、変換対象動画のフレームレート F_{Rc} が、基準動画のフレームレート F_{Rr} より大きい場合のフレームレート変換及び時間レイヤ構造変換処理を行う。具体的には、ピクチャ破棄によってフレームレートを減少させる処理で、先頭ピクチャからピクチャ毎に基準動画のフレームレートより高い上位時間レイヤのピクチャを破棄する。ここでは、120fps の動画を 60fps の動画に変換する処理であって、S206 処理によって変換した時間レイヤ構造を図4 に示す。

【0028】

図4(a)は120fps を有する変換対象動画に対してピクチャを破棄してフレームレートを 60fps に減少させる例を示している。ここでは、基準動画のフレームレートが 60fps であるので、120fps (Temporal ID=2) の時間レイヤに属するピクチャが破棄されている。図4(b)は、図4(a)のような変換処理を行う際の基準動画の一例を示しており、ここでは 60fps の動画が示されている。

10

【0029】

S207 では、S205 もしくは S206 でフレームレート変換対象となる動画が時間レイヤ構造を再構成したことによって変更したフレームレート情報と時間レイヤ情報を更新する。フレームレートに関しては、符号化データのヘッダ部のHRD parameters syntax に含まれる、num_units_in_tick と、time_scale 等の情報を更新する。また時間レイヤ情報に関しては、符号化データのヘッダ部のnal_unit_header Syntax に含まれるnuh_temporal_ID_plus1、VPS(Video Parameter Set) Syntax に含まれるvps_MAX_sub_layers_minus1、SPS(Sequence Parameter Set) Syntax に含まれるsps_MAX_sub_layers_minus1 等の情報を更新する。

20

【0030】

続く S208 では、変換対象動画のフレームレート変換が全て終わったかどうかを判定する。もし、変換対象動画の全てについて処理が終わっている場合は、S210 に処理を進め、終わっていない場合は、S209 へ処理を進める。ここで変換対象動画とは、基準動画を除いた全ての編集対象動画を指す。ここでは基準動画以外の編集対象動画が 2 つなので、S203 から S207 までの処理を 2 回実行することとなる。続く S209 では、未処理の変換対象動画を選択するし、S203 に戻って処理を繰り返す。一方、未編集の変換対象動画が存在せず、全てについて変換処理が終了している場合は S210 において、編集対象となる動画を結合して 1 つの動画ファイルを生成する。そして、時間レイヤ構造を再構成したことによって変更したフレームレート情報をファイルフォーマット中に記録する。

30

【0031】

以上のように、動画符号化装置 100 は、時間スケーラブルに符号化した複数の動画間でフレームレートが異なる場合、動画の復号を行わず、該動画に含まれるピクチャ数を増加、又は、減少させることにより時間レイヤ構造変換を行ってフレームレートの互換性を取る。まず、メディア 108 に保存されている編集対象となる動画のフレームレートと時間レイヤ構造を取得し、編集対象となる動画の中から、フレームレートの互換性を持たせる基準となる動画を選択する。そして、編集対象となる動画は、フレームレートの互換性基準となる動画と同一の時間レイヤ構造となるように時間レイヤ構造を再構成することで、フレームレートの互換性を持った動画が生成できる。また、フレームレート変換処理には、ピクチャ間の画像参照関係を変更することができないため、再符号化処理が不要な結合編集が可能となる。

40

【0032】

[実施形態 2]

次に発明の第 2 の実施形態について説明する。本実施形態に対応する動画符号化装置の構成は、図 1 の動画符号化装置と同様であるため、その構成に関する説明は省略する。上述の第 1 の実施形態では、編集対象動画の最下位時間レイヤのフレームレートが等しい場

50

合を説明した。これに対し、本実施形態では、編集対象動画の最下位時間レイヤのフレームレートが異なる場合の時間レイヤ構造変換方法について説明する。本実施形態ではフレームレート変換対象となる動画の最下位時間レイヤが示すフレームレートが、フレームレート互換性基準となる動画の最下位時間レイヤが示すフレームレートよりも大きい場合は、再符号化によって時間レイヤ構造を変換する。その一方で、フレームレート変換対象となる動画の最下位時間レイヤが示すフレームレートが、フレームレート互換性基準となる動画の最下位時間レイヤが示すフレームレートよりも小さい場合は、Temporal IDの書き換えと、ヘッダ情報の書き換えのみで時間レイヤ構造の変換処理を行う。

【0033】

以下、本実施形態に対応する変換処理の詳細を図5に示すフローチャートを参照して説明する。図5のフローチャートに対応する処理は、例えば、結合編集処理部として機能するCPU103が対応するプログラム（メモリ104に格納）を実行することにより実現できる。なお、本実施形態は、第1の実施形態の処理を包含しつつ、最下位時間レイヤのフレームレートを統一する処理が含まれる。そこで、主として本実施形態特有の処理について詳細に説明し、第1の実施形態と同様の処理に関しては説明を省略する。

【0034】

まず、S501では、編集対象となる動画を選択し、各動画のフレームレートと時間レイヤ構造を取得する。当該処理は第1の実施形態のS201に相当する。ここでは、編集対象となる動画のフレームレートと時間レイヤ構造との関係は以下に示す通りである。

i) フレームレート：60fps

Temporal ID0 = 30fps

Temporal ID1 = 60fps

ii) フレームレート：120fps

Temporal ID0 = 60fps

Temporal ID1 = 120fps

iii) フレームレート：240fps

Temporal ID0 = 120fps

Temporal ID1 = 240fps

【0035】

S502では、S501で編集対象として選択した動画の中から基準動画を選択する。当該処理は第1の実施形態のS202に相当する。ここでは、フレームレートが120fpsの動画を基準動画として選択するものとする。続くS503では、変換対象動画の再下位時間レイヤのフレームレート（以下、「FRc_low」という）と、基準動画の最下位時間レイヤのフレームレート（以下、「FRr_low」という）が等しいかどうかを判定する。FRc_lowとFRr_lowとが等しい場合はS508に処理を進め、異なる場合にはS504に処理を進める。ここでは、FRc_lowはいずれもFRr_lowと異なるためS504に処理を進める。

【0036】

S504では、変換対象動画のFRc_lowと基準動画のFRr_lowとで大小比較を行う。FRr_lowよりFRc_lowの方が小さい場合、S505に処理を進め、大きい場合にはS506に処理を進める。上記例では基準動画は120fpsなのにに対し、60fpsの変換対象動画のFRc_lowは30fpsであるため、S505に進む。一方、240fpsの変換対象動画のFRc_lowは120fpsであるため、S506に進む。

【0037】

続くS505では、Temporal IDの書き換えにより、変換対象動画のFRc_lowを基準動画のFRr_lowに合わせる。これは、後の結合編集時に時間レイヤ指示子であるTemporal IDが異なることで、動画の各時間レイヤが示すフレームレートの整合が取れないことを防ぐための処理である。上記の例で、変換対象となるのは60fpsの動画である。即ち、Temporal ID = 1の時間レイヤに含まれる各ピクチャをTemporal ID = 0に書き換え、Temporal ID = 0とTemporal ID = 1の時間レイヤを統合することで、最下位時間レイヤが示すフレ

10

20

30

40

50

ームレートを 60 f p s に変換する。S 5 0 5 の処理によって再構成した時間レイヤ構造を図 6 に示す。図 6 (a) は、書き換え前にTemporal ID=1の時間レイヤに含まれていたPピクチャが、Temporal ID = 0の時間レイヤに書き換えることを示している。この書き換えの結果、図 6 (b) に示すようにTemporal ID = 0が示す時間レイヤのフレームレートが 60 f p s となる。

【0038】

また、時間レイヤ構造の変更に伴い、符号化データのヘッダ部のnal_unit_header syntaxに含まれるnal_unit_type(N A L ユニットの種類を表す識別子)とVUI parameters syntaxに含まれるhrd_parameters(仮想参照デコーダで使用するパラメータ群)を書き換える。ここでは、時間レイヤ統合前にTemporal ID = 1としていたピクチャは、ピクチャの画像参照関係を表すnal_unit_type情報をTRAIL_N(nal_unit_type=2)からTRAIL_R(nal_unit_type=1)に書き換える。そして、新たに生成したTemporal ID = 0に相当するH R D 情報を、時間レイヤ構造変換前のTemporal ID = 1に相当するH R D 情報に置き換える。10

【0039】

S 5 0 6 では、再符号化により、変換対象動画のFRc_lowを基準動画のFRr_lowと一致させる。これは、S 5 0 5 で説明したように、後の結合編集時にTemporal IDが異なることで、編集対象となる動画間の各時間レイヤが示すフレームレートの整合が取れないことを防ぐための処理である。ここでは、240 f p s の変換対象動画のFRc_lowを、120 f p s の基準動画のFRc_lowと統一する。具体的に、120 f p s と 240 f p s の2つの時間レイヤ構造を、再符号化によって、60 f p s 、120 f p s 、240 f p s の3つの時間レイヤ構造に変換する。図 7 は、その結果として得られる再構成後の変換対象動画の時間レイヤ構造を示す。図 7において、Temporal ID = 0の時間レイヤに含まれるピクチャは 60 f p s を有し、120 f p s 、240 f p s のフレームレートは、それぞれTemporal IDが繰り上がって、Temporal ID = 1、Temporal ID = 2となる。これにより、フレームレートが 240 f p s の変換対象動画のFRc_lowが基準動画のFRr_lowと同じ 60 f p s となった。20

【0040】

次に S 5 0 7 では、S 5 0 5 もしくは S 5 0 6 で変換対象動画の時間レイヤ構造を再構成したことによって変更したフレームレート情報と時間レイヤ情報を更新する。当該処理は、第 1 の実施形態の S 2 0 7 における処理に相当する。各ヘッダ部に格納される情報を更新した後、図 2 の S 2 0 3 に処理を進める。これ以降の処理は、図 2 に示した S 2 0 3 以降の処理と同様である。S 2 0 3 では、変換対象動画のフレームレートFRcと、基準動画のフレームレートFRrとで、フレームレートが等しいか比較を行う。上記の例では、変換対象動画はいずれも基準動画とフレームレートが異なるため S 2 0 4 に処理を進める。30

【0041】

S 2 0 4 では、FRcとFRrとの大小比較を行い、変換対象動画のフレームレートFRcの方が小さい場合は、S 2 0 5 に処理を進め、大きい場合には、S 2 0 6 に処理を進める。ここでは、基準動画のフレームレートFRrを 120 f p s と設定しているので、60 f p s の変換対象動画は、S 2 0 5 で処理され、240 f p s の変換対象動画は S 2 0 6 で処理される。40

【0042】

S 2 0 5 ではコピーピクチャによるピクチャ補間でフレームレートを増加させる処理を行う。ここでは、60 f p s の変換対象動画を 120 f p s に変換する処理を行い、図 8 (a) に示すような時間レイヤ構造を生成することになる。図 8 (a) では、60 f p s のフレームレートを構成する各ピクチャのコピーピクチャが P ピクチャとして生成され、それぞれ生成元となったピクチャを参照している。生成されたコピーピクチャのTemporal IDは、元のピクチャのTemporal IDである 0 よりも1つ上の時間レイヤを示す値として 1 となる。

【0043】

次に S 2 0 6 では、ピクチャ破棄によってフレームレートを減少させる処理を行う。50

こでは、複数の時間レイヤ構造を持つ 240 f p s の変換対象動画を 120 f p s の動画に変換する処理を行い、図 8 (b) に示すような時間レイヤ構造を生成することになる。図 8 (b) では 3 つの時間レイヤのうち、 240 f p s の時間レイヤに属するピクチャを破棄して、 60 f p s の Temporal ID = 0 の時間レイヤと、 120 f p s の Temporal ID = 1 の時間レイヤのみを残している。これ以降、第 1 の実施形態で記載したのと同様に S 2 0 7 から S 2 1 0 の処理を行う。なお、本実施形態ではフレームレート変換対象となる動画の数が 2 つであるため、 S 5 0 3 から S 5 0 7 及び S 2 0 3 から S 2 0 7 までの処理を 2 回実行する。

【 0 0 4 4 】

以上によれば、編集対象動画の間で最下位時間レイヤのフレームレートが異なる場合であっても、変換対象動画の FRc_low を基準動画の FRr_low と同一になるように時間レイヤ構造を再構成することで、結合編集時の Temporal ID のミスマッチを防止することができる。これにより、複数の動画を結合編集した場合にフレームレートの互換性を持った動画が生成できる。また、本実施形態では FRc_low が FRr_low より大きい場合にのみ再符号化を行い、 FRc_low が FRr_low より小さい場合は Temporal ID とヘッダ情報の書き換えのみで時間レイヤ構造を変換するので、より少ない再符号化でフレームレートの互換性を持たせることができる。

【 0 0 4 5 】

(その他の実施例)

本発明は、上述の実施形態の 1 以上の機能を実現するプログラムを、ネットワーク又は記憶媒体を介してシステム又は装置に供給し、そのシステム又は装置のコンピュータにおける 1 つ以上のプロセッサーがプログラムを読み出し実行する処理でも実現可能である。また、 1 以上の機能を実現する回路（例えば、 A S I C ）によっても実現可能である。

【 符号の説明 】

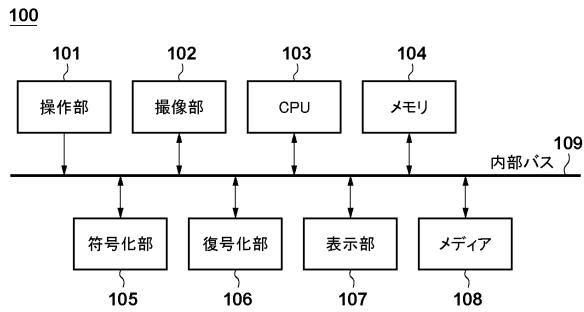
【 0 0 4 6 】

1 0 0 : 動画符号化装置、 1 0 1 : 操作部、 1 0 2 : 撮像部、 1 0 3 : C P U 、 1 0 4 : メモリ、 1 0 5 : 符号化部、 1 0 6 : 復号化部、 1 0 7 : 表示部、 1 0 8 : メディア

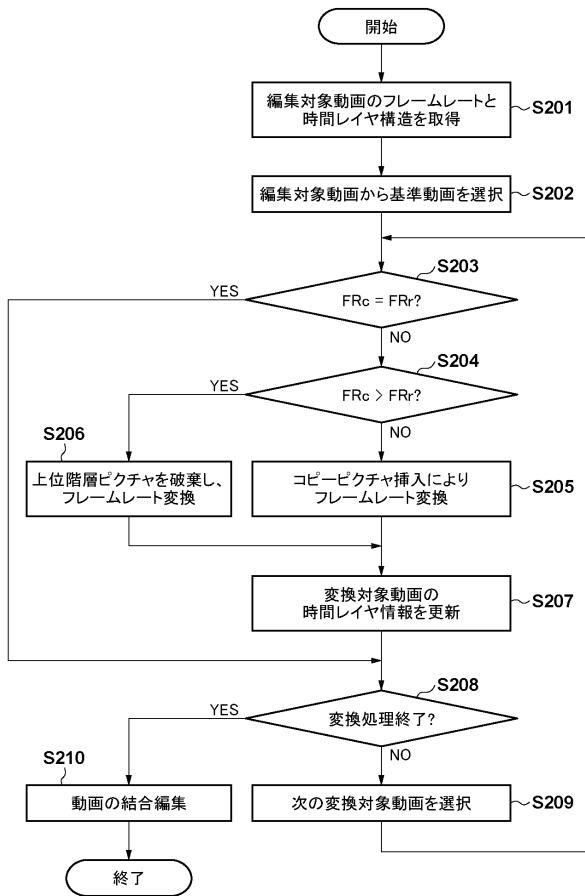
10

20

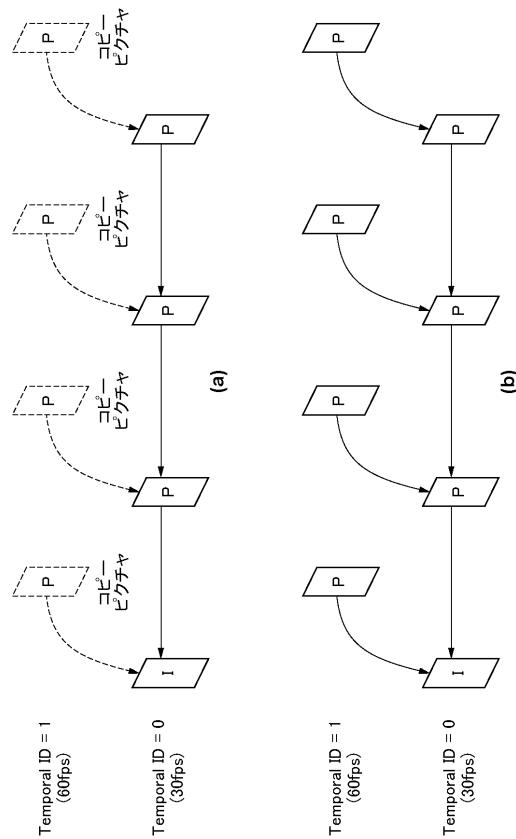
【図1】



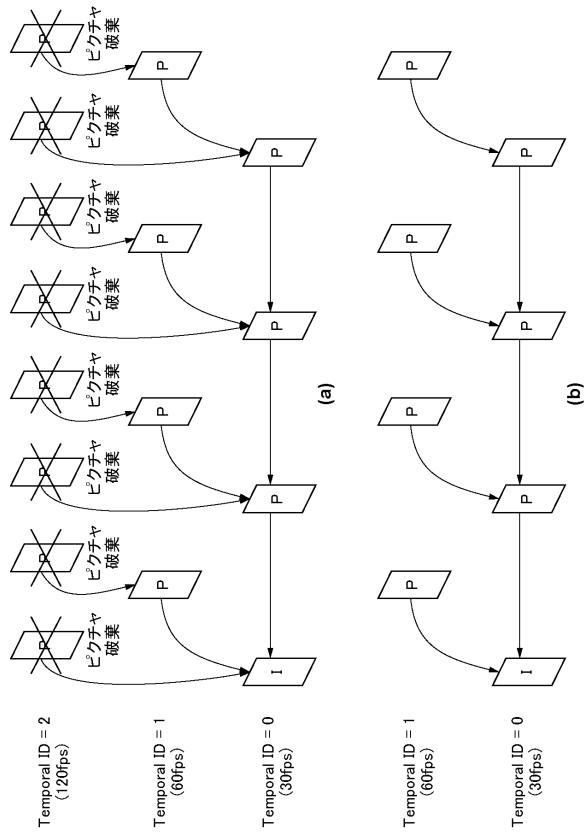
【図2】



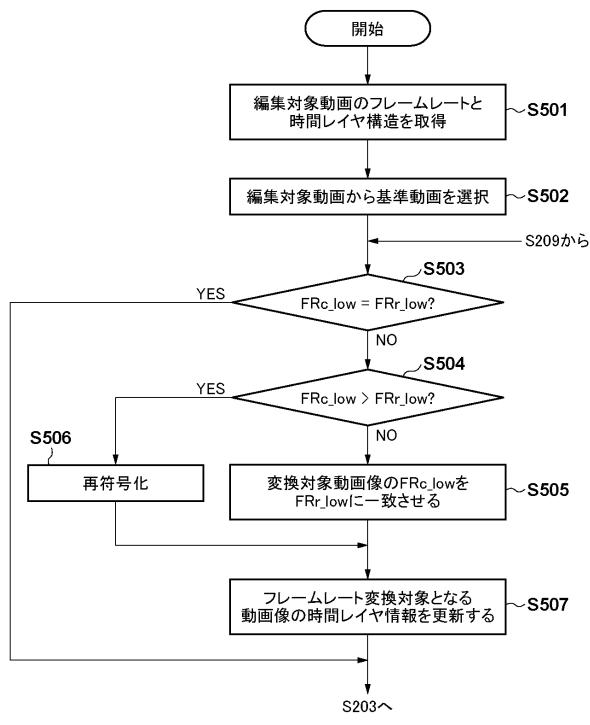
【図3】



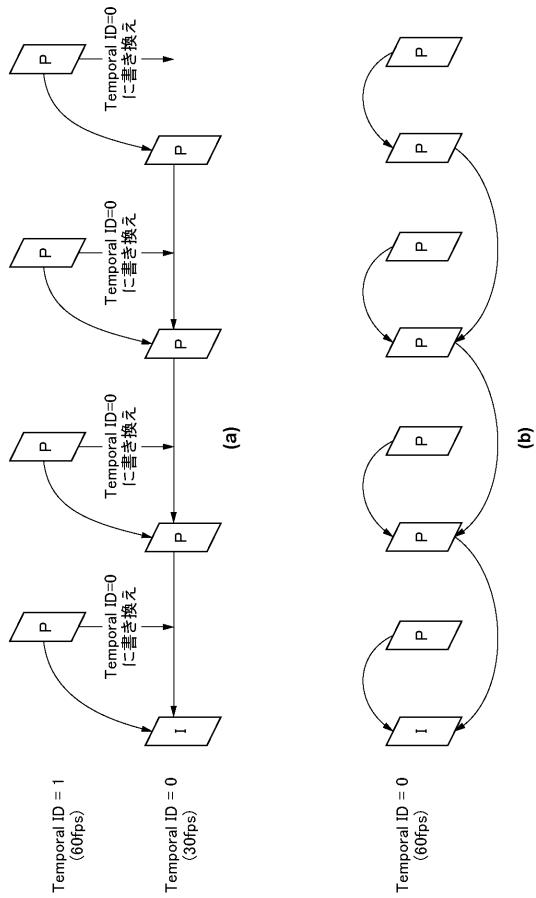
【図4】



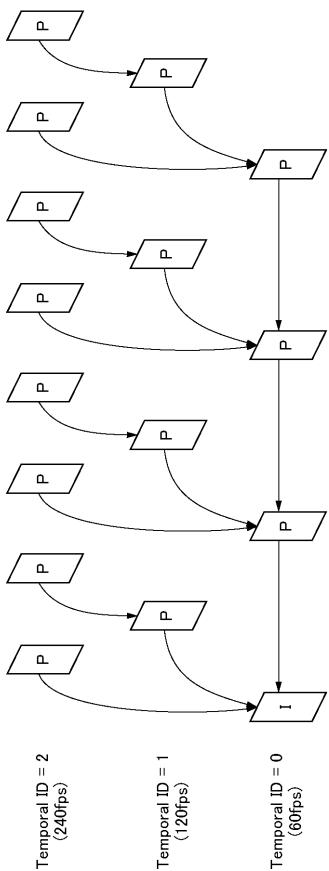
【図5】



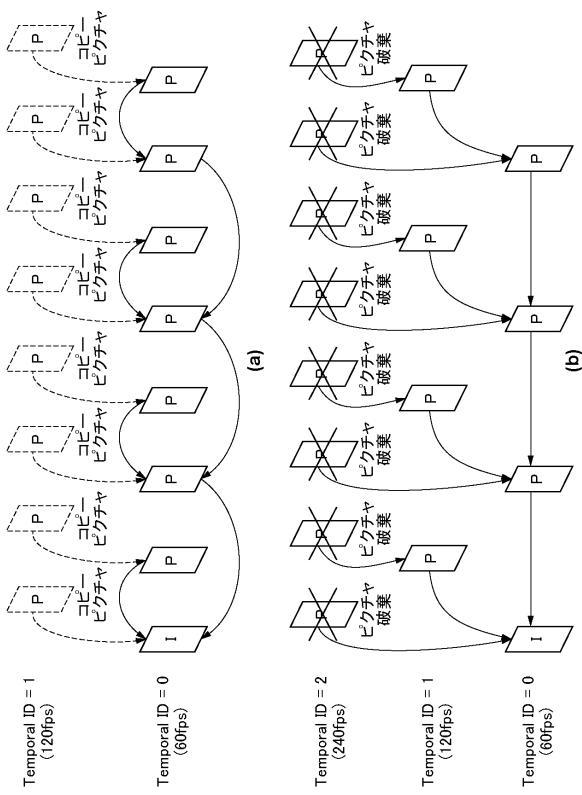
【図6】



【図7】



【図8】



フロントページの続き

(72)発明者 鈴木 遼太

東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

審査官 岩井 健二

(56)参考文献 特開2014-220599(JP,A)

特開2015-080035(JP,A)

国際公開第2014/050597(WO,A1)

国際公開第2008/060262(WO,A1)

国際公開第2007/076486(WO,A2)

米国特許出願公開第2005/0195897(US,A1)

(58)調査した分野(Int.Cl., DB名)

H04N 19/00 - 19/98