

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5634111号  
(P5634111)

(45) 発行日 平成26年12月3日(2014.12.3)

(24) 登録日 平成26年10月24日(2014.10.24)

(51) Int. Cl. F I  
**HO4N 5/91 (2006.01)** HO4N 5/91 N  
**G11B 27/034 (2006.01)** G11B 27/034

請求項の数 4 (全 12 頁)

<p>(21) 出願番号 特願2010-104185 (P2010-104185)                  (22) 出願日 平成22年4月28日(2010.4.28)                  (65) 公開番号 特開2011-234226 (P2011-234226A)                  (43) 公開日 平成23年11月17日(2011.11.17)                  審査請求日 平成25年4月26日(2013.4.26)</p>	<p>(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 100090273                  弁理士 園分 孝悦                  (72) 発明者 菊地 徹                  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 梅本 章子</p>
---	---

最終頁に続く

(54) 【発明の名称】 映像編集装置、映像編集方法及びプログラム

(57) 【特許請求の範囲】

【請求項1】

複数の映像フレームからなる第一の映像データの一部に続き、前記第一の映像データとは異なる複数の映像フレームからなる第二の映像データの一部を連結させる映像編集装置であって、

前記第一の映像データを構成する複数の映像フレームと前記第二の映像データを構成する複数の映像フレームとから、同一人物が検出されるかを判定する判定手段と、

前記判定手段によって前記映像フレームから同一人物が検出されたと判定されない場合に、前記第一の映像データを構成する映像フレームから検出される前記映像フレームの被写体である人物の視線方向と、前記第二の映像データを構成する映像フレームが撮像された時点での、前記映像フレームの被写体と前記第二の映像データを撮像した撮像装置との位置関係に対応するカメラアングルとの関連性に基づいて、前記第一の映像データを構成する複数の映像フレームの何れかと前記第二の映像データを構成する複数の映像フレームの何れかとから、相関が高い映像フレームの組み合わせを選択する選択手段と、

前記選択手段により相関が高い映像フレームの組み合わせとして選択された前記第一の映像データにおける映像フレームと前記第二の映像データにおける映像フレームとを、連結させる連結手段と、  
 を有する映像編集装置。

【請求項2】

更に、ユーザによる操作入力の検出或いは予め規定された選択条件に基づき、3つ以上

の複数の映像データの中から、前記第一の映像データと前記第二の映像データとして用いる映像データを選択する選択手段を有する請求項1に記載の映像編集装置。

【請求項3】

複数の映像フレームからなる第一の映像データの一部に続き、前記第一の映像データとは異なる複数の映像フレームからなる第二の映像データの一部を連結させる映像編集装置が実行する映像編集方法であって、

前記第一の映像データを構成する複数の映像フレームと前記第二の映像データを構成する複数の映像フレームとから、同一人物が検出されるかを判定する判定ステップと、

前記判定ステップによって前記映像フレームから同一人物が検出された場合と判定されない場合に、前記第一の映像データを構成する映像フレームから検出される前記映像フレームの被写体である人物の視線方向と、前記第二の映像データを構成する映像フレームが撮像された時点での、前記映像フレームの被写体と前記第二の映像データを撮像した撮像装置との位置関係に対応するカメラアングルとの関連性に基づいて、前記第一の映像データを構成する複数の映像フレームの何れかと前記第二の映像データを構成する複数の映像フレームの何れかとから、相関が高い映像フレームの組み合わせを選択する選択ステップと、

前記選択ステップにより相関が高い映像フレームの組み合わせとして選択された前記第一の映像データにおける映像フレームと前記第二の映像データにおける映像フレームとを、連結させる連結ステップと、

を含む映像編集方法。

【請求項4】

コンピュータに読み込ませ実行させることで、前記コンピュータを請求項1又は2に記載の映像編集装置として機能させるコンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

複数の映像データを連結した編集映像データを生成するものに関する。

【背景技術】

【0002】

近年、デジタルビデオカメラ、デジタルカメラ、デジタル一眼レフカメラ、携帯電話、携帯情報端末、携帯音楽プレイヤー等、様々な機器に動画撮影機能が搭載されている。また、ハードディスクドライブやメモリカード等のデータ記憶媒体の大容量化も進み、個人が所有する映像データも膨大になりつつある。更に、インターネット上に映像データをアップロードし視聴し合うサービスや、テレビのネットワーク対応により、撮り貯めた映像を視聴する機会が増加している。

撮影した映像を視聴する場合には、撮影時間と同じだけの時間が必要となってしまう。そこでユーザは、チャプタメニューと呼ばれる頭出し機能やファイル単位のサムネイル表示等で見たい映像データを選択している。更に、映像データ内に不要区間が存在する場合にはユーザは、早送りをするとといった操作を行う。

また、このような操作は面倒で手間がかかるため、映像編集を行うユーザも多い。ユーザは、映像編集では、映像データの視聴したい区間（以下、カットと呼ぶ）を連結し、必要に応じてカットにグラフィック映像等の合成、カットの繋ぎ目にトランジションと呼ばれる映像効果、BGM等を付与し編集映像データを作成する。

しかしながら、映像編集には編集に関する専門的な知識や映像編集ツールの使用方法の知識が必要であり、一般的なユーザには難しい作業となる。

また、編集や編集ツールに関する知識のあるユーザであっても、膨大な映像データから必要な映像データを選択し、更に映像データのカットを選択するといった作業は面倒な作業である。

【0003】

そこで、簡単に映像編集を行うための技術として、テンプレートファイルに基づき映像編集を自動化する技術が開示されている。テンプレートファイルには、映像データが挿入

される素材枠、エフェクト、BGM、及びトランジションが、タイムコードに応じて予め定義されている。映像編集装置は、テンプレートファイルの素材枠それぞれのメタデータを解析すると共に、映像データのメタデータも解析し、メタデータが一致する映像データを、テンプレートファイルの素材枠に挿入する（特許文献1参照）。

また、映像フレーム間の動きベクトルを解析し、同じ向き、同じ大きさの動きベクトルが集まっている領域を追尾中の被写体であると判定することで、被写体を捉えているシーンを自動で抽出、編集する技術が開示されている。ここで、編集映像データは、複数のシーンから構成される。前記技術では、各シーンを編集順に並べるだけでなく、重要度に基づいて並び替えたり、同じようなシーンを繋ぎ合わせて纏めたりする（特許文献2参照）。

10

また、検索条件に対応する解析規則に従って検索結果として得られた映像データを解析し、表示パターンに関連し選択された表示規則と解析結果に基づいて、検索結果として表示する表示内容を決定する技術が開示されている（特許文献3参照）。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開2009-55152号公報

【特許文献2】特開2008-131617号公報

【特許文献3】特開2006-139382号公報

【発明の概要】

20

【発明が解決しようとする課題】

【0005】

しかしながら、特許文献1による手法では、メタデータが一致する映像データを、テンプレートファイルの素材枠に挿入しているため、挿入された映像データ間の繋がりがよくない場合がある。例えば、映像データ間で被写体の位置が突然ジャンプしてしまったり、被写体の移動方向が急に逆向きになったりする場合である。

また、特許文献2による手法では、動きベクトルに基づき同じようなシーンを繋ぎ合わせてはいるものの、映像データ間での被写体位置のジャンプは防ぐことができない。また、同じようなシーンであっても意味的な繋がりが無く不自然な映像になる場合もある。

また、特許文献3による手法では、表示パターンによって意味的な繋がりは考慮されているものの、被写体位置のジャンプは発生してしまう。また、映像データ単位で解析が行われるため、意味的な繋がりを考慮してカット同士を繋ぐことはできない。例えば、映像データ中で、視線を向けたカットの直後に視線の先の被写体のカットを繋ぐといった場合である。

30

繋がりが不適切な映像は、視聴者に混乱を与え内容が理解し難くなる。

【0006】

本発明はこのような問題点に鑑みなされたもので、編集映像におけるカット間の繋がりが自然になるようにすることを目的とする。

【課題を解決するための手段】

【0007】

40

本発明は、複数の映像フレームからなる第一の映像データの一部に続き、前記第一の映像データとは異なる複数の映像フレームからなる第二の映像データの一部を連結させる映像編集装置であって、前記第一の映像データを構成する複数の映像フレームと前記第二の映像データを構成する複数の映像フレームとから、同一人物が検出されるかを判定する判定手段と、前記判定手段によって前記映像フレームから同一人物が検出されたと判定されない場合に、前記第一の映像データを構成する映像フレームから検出される前記映像フレームの被写体である人物の視線方向と、前記第二の映像データを構成する映像フレームが撮像された時点での、前記映像フレームの被写体と前記第二の映像データを撮像した撮像装置との位置関係に対応するカメラアングルとの関連性に基づいて、前記第一の映像データを構成する複数の映像フレームの何れかと前記第二の映像データを構成する複数の映像

50

フレームの何れかから、相関が高い映像フレームの組み合わせを選択する選択手段と、前記選択手段により相関が高い映像フレームの組み合わせとして選択された前記第一の映像データにおける映像フレームと前記第二の映像データにおける映像フレームとを、連結させる連結手段と、を有する。

【発明の効果】

【0008】

本発明によれば、任意の複数の映像データに対して、編集された映像データを閲覧する者に、第一の映像データを構成する映像フレームの被写体である人物の視線の先に、第二の映像データを構成する映像フレームの被写体が存在することを連想させるような自然な編集を容易に行う事が可能となる。

10

【図面の簡単な説明】

【0009】

【図1】映像編集装置のハードウェア構成の一例を示す図である。

【図2】映像編集装置の機能構成の一例を示す図である。

【図3】テンプレート情報の一例を示す図である。

【図4】メタデータの一例を示す図である。

【図5】プレイリストの一例を示す図である。

【図6】カット点選択結果の一例を示す図である。

【図7】処理の流れを説明するための図である。

【図8】映像フレーム間相関評価の一例を示す図である。

20

【図9】映像編集処理の一例を示すフローチャートである。

【図10】カット点選択処理の一例を示すフローチャートである。

【図11】映像フレーム間相関評価処理の一例を示すフローチャートである。

【発明を実施するための形態】

【0010】

以下、本発明の実施形態について図面に基づいて説明する。

【0011】

<実施形態1>

図1は、映像編集装置（コンピュータ）のハードウェア構成の一例を示す図である。映像編集装置100は、記憶媒体に記憶された映像データを編集し編集映像データを生成する装置である。映像編集装置100は、バス101、ROM（Read Only Memory）102、キーボード103、記憶媒体104から構成される。更に、映像編集装置100は、CPU（Central Processing Unit）105、RAM（Random Access Memory）106、ディスプレイ107から構成される。

30

バス101は、CPU105が他のハードウェアモジュールとの間でデータを読み書きするための制御モジュールである。本実施形態では単一のバス101に全てのハードウェアモジュールが接続されているが、接続されるハードウェアモジュールの種別に応じてバスが複数あってもよい。

ROM102は、映像編集装置100を起動するためのプログラムが記憶されている。映像編集装置100に電源が投入されるとROM102のプログラムがCPU105により読み出し、実行され、各ハードウェアモジュールの初期設定やOS（Operating System）の起動等を行う。

40

キーボード103は一般的なキーボードであり、アルファベット、数字、Enter、Back Space、ESC、TAB等のユーザからの入力をCPU105に伝える。尚、キーボード103はタッチパネル、マウス、機器固有のボタン等、他の入力デバイスであってよい。

【0012】

記憶媒体104は、映像データや編集映像データ、後述する図3を用いて説明するテンプレート等の各種データを記憶する記憶媒体である。また、OSや本映像編集装置の機能

50

ブロックの処理プログラムも記憶する。より具体的には、記憶媒体104は、ハードディスクドライブ、SSD(Solid State Disk)、CD-ROM等である。尚、記憶媒体104は、映像編集装置100の筐体に収められていても、ネットワーク等で接続されたものであってもよい。

CPU105は、映像編集装置の起動プログラム、OS、機能ブロックの処理プログラムを実行する演算器である。

RAM106は、CPU105が各種プログラムを実行する際に、データの一時保存を行う記憶媒体である。本実施形態のカット点選択結果やプレイリストはRAM106に記憶される。

ディスプレイ107は、映像データの表示、編集映像データの表示、各種グラフィカルユーザインタフェースの表示を行う。具体的には液晶ディスプレイ、プラズマディスプレイ、CRT(Cathode Ray Tube)ディスプレイ等である。

CPU105がプログラムを実行することによって、後述する映像編集装置100の機能及び後述するフローチャートに係る処理が実現される。

#### 【0013】

図2は、映像編集装置の機能構成の一例を示す図である。映像編集装置100は、映像データ選択部201、カット繋ぎ評価部202、カット点設定部203、編集映像データ生成部204から構成される。

映像データ220は、編集対象となる複数の映像データであり、ビデオカメラのON/OFF等によって、予め撮影単位毎に個別のファイルに分割され、記憶媒体104に記憶される。

テンプレート情報210は、生成したい編集映像の特徴に応じて、映像データ220を選択するための条件が記載されたテンプレート情報であり、記憶媒体104に記憶される。

図3は、テンプレート情報210の一例を示す図である。テンプレート情報210は、カットID、選択条件の各情報から構成される。カットIDはカットを一意に識別する識別子であり、カットの出現順にシーケンシャルな番号が振られる。選択条件は、各カットに対応した映像データを選択するための条件である。尚、テンプレート情報の構成、構造は一例であり、これに限られるものではない。例えば、選択条件として、「入場シーン」等、ユーザ付与メタデータが条件として記載されてもよい。

メタデータ240は、映像データ220の内容に関する情報が記載されたメタデータテーブルであり、映像編集処理に先立ち予め作成され記憶媒体104に記憶される。

図4は、メタデータ240の一例を示す図である。メタデータ240は、映像データ名、顔サイズ、人数、ユーザ付与(ユーザ付与メタデータ)の各情報から構成される。映像データ名は、映像データを一意に識別可能な情報であり、ファイル名である。顔サイズは、映像データから検出した顔の大きさであり、「大」、「中」、「小」等に分類する。人数は、映像データから検出した顔の数であり「多」、「少」、「1」等に分類する。ユーザ付与は、ユーザが任意に設定可能な情報であり、映像データの内容を表すキーワード等が設定される。

尚、メタデータの種類、値等は一例であり、これに限られるものではない。また、テーブルではなく、データベースを用いてもよい。また、メタデータ240は、映像編集処理中に作成されてもよい。

編集映像データ230は、編集結果として生成される編集映像データであり、記憶媒体104に記憶される。

#### 【0014】

映像データ選択部201は、テンプレート情報210とメタデータ240とに基づき、第1のプレイリスト900aを作成する。図5(a)は、第1のプレイリスト900aの一例を示し、第1のプレイリスト900aは、カットID、映像データ名の各情報から構成される。カットIDは、カットを一意に識別する識別子であり、テンプレート情報210のカットIDと同じ値となる。映像データ名は、映像データを一意に識別可能な情報で

10

20

30

40

50

あり、テンプレート情報 210 に記載された選択条件とメタデータ 240 のメタデータとがマッチした映像データ名となる。

カット繋ぎ評価部 202 は、第 1 のプレイリスト 900 a と映像データ 220 とからカット点選択結果 800 を生成する。

カット点選択結果 800 は、第 1 のプレイリスト 900 a で隣接する 2 つの映像データの映像フレーム間で相関の高い組み合わせをカット点として記載した情報である。映像フレーム間の相関の高い組み合わせは、視覚的又は意味的な相関の高い組み合わせである。視覚的な相関が高い組み合わせとは、被写体の動作の連続性、被写体の移動方向の同一性、被写体の位置の同一性、視覚的な類似性の高い組み合わせである。意味的な相関が高い組み合わせとは、視線方向が一致する組み合わせである。カット繋ぎ評価部 202 は、映像データ 220 から 1 つ以上の映像フレームの画像の解析、或いは映像区間に予め付与されたメタデータを用いて評価を行う。図 6 は、カット点選択結果 800 の一例を示す図である。カット点選択結果 800 は、時間軸上で前方の映像データ名、後方の映像データ名、前方の映像データのカット点、後方の映像データのカット点から構成される。ここで、前方の映像データのカット点は、前方の映像データの区間のアウト点を示しており、後方の映像データのカット点は、後方の映像データの区間のイン点を示している。図 6 より、映像データ video\_\_1.mpg と video\_\_2.mpg とのカット点として video\_\_1.mpg の 48.00 秒と video\_\_2.mpg の 3.13 秒との映像フレームが選択されていることが分かる。

#### 【0015】

尚、本実施形態では、カット繋ぎ評価部 202 は、カット点選択結果 800 を、第 1 のプレイリスト 900 a に設定された映像データの中で隣接する 2 つの映像データに対して生成しているが、映像データ 220 の全ての組み合わせに対して生成してもよい。

カット点設定部 203 は、カット点選択結果 800 に基づき、第 1 のプレイリスト 900 a にカット点を追加した第 2 のプレイリスト 900 b を生成する。

図 5 (b) に第 2 のプレイリスト 900 b を示しており、第 2 のプレイリスト 900 b は、カット ID、映像データ名、イン点、アウト点の各情報から構成される。ここで、カット ID、映像データ名は、第 1 のプレイリスト 900 a と同じ値となる。イン点、アウト点は、第 1 のプレイリスト 900 a で隣接する 2 つの映像データ名とカット点選択結果 800 の前方、後方の映像データ名とが一致するカット点となる。

編集映像データ生成部 204 は、第 2 のプレイリスト 900 b に設定された映像データ、イン点、アウト点に基づき、映像データの区間を読み出し連結して編集映像データ 230 を生成する。

#### 【0016】

図 7 は、本実施形態の処理の流れを説明するための図である。

220 - 1、220 - 2、220 - 3、220 - 4 は、編集対象となる映像データである。図 7 (a) は、映像データ選択部 201 により映像データが選択された状態を示している。図 7 (b) は、カット繋ぎ評価部 202 によって隣接する映像データ、例えば 220 - 1 と 220 - 2 との間のカット繋ぎが評価され、カット点設定部 203 によりカット点が設定された状態を示している。映像データ 220 - 1 と 220 - 2 とのカット点として映像フレーム 300 - 1、300 - 2 が選択されている。図 7 (c) は、カット点設定部 203 により設定されたカット点に基づき編集映像データ 230 - 1 が生成された状態を示している。例えば、映像データ 220 - 1 と 220 - 2 とは映像フレーム 300 - 1 と 300 - 2 とで連結されている。

図 7 に示すように、映像編集装置 100 は、2 つの映像データを連結する際に、相関が高いと評価した映像フレーム、例えば 300 - 1 と 300 - 2 とを連結するよう制御するため、繋がりやすい編集映像を生成することができる。

#### 【0017】

図 8 は、映像フレーム間相関評価の一例を示す図である。

図 8 (a) は、同一被写体の動作の連続性に基づく相関の評価の一例である。図 8 (a)

10

20

30

40

50

)で300-13と300-14とは前方の映像データの映像フレーム、300-15、300-16、300-17は後方の映像データの映像フレームである。カット繋ぎ評価部202は、動作の連続性の評価では、まず、動きベクトルから動きのある領域を抽出し、動きのある領域の画像のマッチングによって同一被写体の同一動作であるか否かを判定する。続いて、カット繋ぎ評価部202は、同一被写体の同一動作である場合には、動き周辺のオブジェクトとの距離変化が映像フレーム間で連続的か否かで動きが連続か否かを評価する。図8(a)では、動きのある領域がカップの画像であり、カット繋ぎ評価部202は、顔又は口との距離変化から300-14と300-17との連続性が高いと評価する。また、カット繋ぎ評価部202は、300-13に対しては300-16が、連続性が高いと評価する。

10

図8(b)は、映像フレームの類似性に基づき相関を評価する一例である。カット繋ぎ評価部202は、映像フレームの類似性では、映像フレームの画像の色彩、明度、輝度の分布を比較する既知の手法を適用する。図8(b)において、映像フレーム300-19と300-21とは被写体の前を人や物が横切り、映像フレームが前景に覆われた状態を示している。カット繋ぎ評価部202は、このように類似性の高い映像フレームは相関が高いと評価する。

#### 【0018】

図8(c)は、同一被写体の位置の同一性に基づき相関を評価する一例である。この例では、カット繋ぎ評価部202は、同一人物であると認識された顔の検出位置を被写体位置として同一性を評価している。カット繋ぎ評価部202は、映像フレーム300-24と300-26との被写体位置の同一性が高いため、相関が高いと評価する。尚、カット繋ぎ評価部202は、顔検出及び人物の認識は既知の手法を用いて行う。

20

図8(d)は、同一被写体の移動方向の同一性に基づき相関を評価する一例である。カット繋ぎ評価部202は、同一人物の顔検出位置の変化に基づき、同一被写体の移動を検出し移動方向の同一性を評価する。カット繋ぎ評価部202は、映像フレーム300-30と300-32とは方向の同一性が高いため、相関が高いと評価する。

図8(e)は、被写体の視線の方向と被写体位置との同一性に基づき相関を評価する一例である。カット繋ぎ評価部202は、被写体の視線方向の検出では、顔検出範囲と目、口等の部位の位置及び瞳の位置検出に基づく手法等、既知の手法を適用する。カット繋ぎ評価部202は、被写体の方向の検出では、近年カメラに組み込まれつつある電子磁石や傾きセンサの情報を使用する手法やユーザ付与メタデータを使用する既知の手法を適用する。映像フレーム300-34で被写体向きが上で、300-36、300-37が、飛行機が飛ぶ映像であるため、カット繋ぎ評価部202は、300-34と300-36、300-37とが、相関が高いと評価する。

30

#### 【0019】

図9は、映像編集処理の一例を示すフローチャートである。

映像編集処理が開始されると、映像データ選択部201は、記憶媒体104に記憶されたテンプレート情報210から選択条件を読み出す(S1001)。続いて、映像データ選択部201は、選択条件とメタデータ240とに基づき映像データを選択し、第1のプレイリスト900aを作成する(S1002)。次に、カット繋ぎ評価部202は、第1のプレイリスト900aから隣接する2つの映像データを読み出す(S1003)。ここで、第1のプレイリスト900aの映像データが終了していない場合は(S1004、No)、カット繋ぎ評価部202は、映像フレーム間の相関を評価しカット点選択結果800を作成する(S1005)。尚、カット点の選択処理の詳細に関しては図10で説明する。

40

#### 【0020】

続いて、カット点設定部203は、相関の高い映像フレームが存在する場合は(S1006、Yes)、相関の高いフレームをカット点として追加した第2のプレイリストを作成し(S1008)、S1003からの処理を繰り返す。カット点設定部203は、相関の高い映像フレームが存在しない場合は(S1006、No)、S1003からの処理を

50

繰り返す。S1004において、プレイリストの映像データへの処理が終了した場合(S1004、Yes)、編集映像データ生成部204は、プレイリスト900bに基づいて編集映像データを生成し(S1011)、図9に示す処理を終了する。

映像編集装置100は、S1005とS1008とで、映像フレーム間の相関の評価結果に基づき相関の高い映像フレームをイン点、アウト点として選択するので、繋がりやすい編集映像を生成できる。

尚、本実施形態では、テンプレート情報210に記載された選択条件に基づき映像データを選択する例を示したが、映像データ選択部201は、ユーザの操作入力に基づき映像データが選択するようにしてもよい。また、S1005では第1のプレイリスト900aで隣接する2つの映像データに対してカット点の選択を行ったが、カット繋ぎ評価部202は、映像データ全ての組み合わせに付いて行ってもよい。また、S1008では相関の高いフレームをカット点として設定しているが、相関の高いフレームが複数の場合、カット点設定部203は、任意の1組を選択するか、相関の高さを定量化し、数値の最も高い(相関の最も高い)1組を選択してもよい。

#### 【0021】

図10は、カット点選択処理の一例を示すフローチャートである。カット点選択処理では、カット点選択結果800が作成される。

カット点選択処理が開始されると、カット点選択結果800が存在しない場合には、カット繋ぎ評価部202は、カット点選択結果800の初期化を行う(S1101)。次に、カット繋ぎ評価部202は、S1003で選択された映像データのうち前方の映像データの最後尾の映像フレームを読み出す(S1102)。次に、カット繋ぎ評価部202は、後方の映像データの先頭の映像フレームを読み出す(S1103)。後方の映像データの映像フレームが終了でない、即ち、後方の映像データの最後尾まで読み出し済みでない場合(S1104のNo)、カット繋ぎ評価部202は、S1102とS1103とで読み出した2つの映像フレーム間の相関を評価する(S1105)。ここで、相関の評価処理の詳細は、図11を用いて後述する。続いて、S1105で相関が高いと評価した場合、カット繋ぎ評価部202は、その映像データ名、カット点をカット点選択結果800に追加する(S1106)。次に、カット繋ぎ評価部202は、後方の映像データの映像フレームを1フレームだけインクリメントして読み出し(S1107)、S1104からの処理を繰り返す。

#### 【0022】

S1104で、後方の映像データの映像フレームが終了、即ち、後方の映像データの最後尾まで読み出し済みの場合(S1104のYes)、カット繋ぎ評価部202は、前方の映像データの映像フレームを1フレームだけデクリメントして読み出す(S1111)。前方の映像データの映像フレームが終了、即ち、前方の映像データの先頭まで読み出し済みの場合(S1112のYes)、カット繋ぎ評価部202は、図10に示す処理を終了する。終了でない場合(S1112のNo)、カット繋ぎ評価部202は、S1103からの処理を繰り返す。

以上の処理により、映像編集装置100は、前方の映像データの映像フレームと後方の映像データの映像フレームとの間で相関を評価し、相関が高い映像フレームをカット点として選択することができる。

尚、本実施形態では、全ての映像フレームの組み合わせに対して相関を評価する例を記載したが、カット繋ぎ評価部202は、予め規定された範囲の映像フレームの組み合わせに対して評価を行ってもよい。

#### 【0023】

図11は、映像フレーム間相関評価処理の一例を示すフローチャートである。

映像フレーム間相関評価処理は、S1105から呼び出される。処理が開始されると、カット繋ぎ評価部202は、前後の映像データの各々の映像フレームについて、1コマ前の映像フレームとの間の動きベクトルを算出し、動きベクトルの方向と大きさが同程度の領域を動き領域として選択する(S601)。続いて、カット繋ぎ評価部202は、前後

10

20

30

40

50

の映像データの映像フレーム間で、カラーヒストグラム、テクスチャ、形状がマッチする領域を同一動き領域として選択する（S602）。ここで、カラーヒストグラムは、色相、彩度、明度の各々を例えば16次元に縮退したものである。また、テクスチャは、ウェーブレット変換により画像を周波数分解し、ベクトル化したものである。また、形状は、領域の外接円からの距離を正規化しヒストグラム化したものである。S602によって同一動き領域が存在した場合（S603、Yes）、カット繋ぎ評価部202は、S611に進む。S611では、カット繋ぎ評価部202は、前後の映像データの映像フレーム毎に、エッジを抽出し、同一動き領域の近傍で単位面積辺りのエッジの数が多く、両方の映像フレーム間でエッジの形状の類似性が高い1つ以上の領域を基準領域として選択する。続いて、カット繋ぎ評価部202は、前後の映像データの映像フレーム間で、基準領域と同一動き領域の距離を基準領域の面積等で正規化した正規化距離の連続性が高ければ動きによる相関が高いと評価する（S612）。カット繋ぎ評価部202は、距離の連続性の高さを、前方の映像データの映像フレームの1コマ後の映像フレームと後方の映像データの映像フレームとの正規化距離の同一性の高さで判定する。また、カット繋ぎ評価部202は、前方の映像データの映像フレームと後方の映像データの1コマ前の映像フレームとの正規化距離で判定を行ってもよい。

10

## 【0024】

S612に続いて、或いはS603で同一動き領域がないと判定した場合（S603、No）、カット繋ぎ評価部202は、前後の映像データの映像フレーム間で、カラーヒストグラム、テクスチャの類似度を算出する（S604）。類似度が所定値以上の場合（S605、Yes）、カット繋ぎ評価部202は、類似性による相関が高いと評価する（S621）。

20

S621に続いて、或いはS605で類似度が所定以上でないとして判定した場合（S605、No）、カット繋ぎ評価部202は、前後の映像データの映像フレームの各々について顔認識による人物検出を行う（S606）。ここで、カット繋ぎ評価部202は、顔認識では一般的な手法を用いる。前後の映像データの映像フレーム間で、同一人物が存在する場合（S607、Yes）、カット繋ぎ評価部202は、前後の映像データの映像フレーム間で同一人物の検出位置の差が所定範囲内であれば位置による相関が高いと評価する（S631）。続いて、カット繋ぎ評価部202は、前後の映像データの映像フレーム間で同一人物の移動方向の差が所定範囲内であれば方向による相関が高いと評価する（S632）。ここで、カット繋ぎ評価部202は、移動方向は人物検出位置の時間変化から算出する。

30

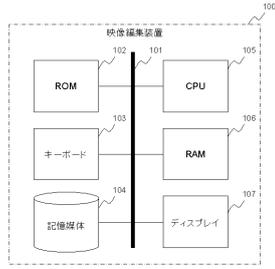
S632に続いて、或いはS607で同一人物が存在しないと判定した場合（S607、No）、カット繋ぎ評価部202は、前方の映像データの映像フレームで視線方向の検出を行う（S608）。ここで、カット繋ぎ評価部202は、視線方向を、黒目、まぶた等の検出結果を目の3Dモデルにマッピングすることにより算出する。続いて、上下方向の視線が検出された場合（S609、Yes）、カット繋ぎ評価部202は、後方映像データの映像フレームのカメラアングルが視線方向と同程度であれば視線による相関が高いと評価し（S641）、図11に示す処理を終了する。ここで、カメラアングルは3軸の加速度センサによって撮影時に付与されているものとする。また、S609で上下方向の視線がないと判定された場合（S609、No）、カット繋ぎ評価部202は、図11に示す処理を終了する。

40

以上の説明の通り、本実施形態によれば、第1のプレイリストで隣接する2つの映像データの映像フレーム間で相関の高い組み合わせをカット点として選択することができる。

S601からS607により、視覚的な相関が高い組み合わせ（被写体の動作の連続性、被写体の移動方向の同一性、被写体の位置の同一性、視覚的な類似性の高い組み合わせ）をカット点として選択することができる。そして、S608からS610により、意味的な相関が高い組み合わせ（視線方向が一致する組み合わせ）をカット点として選択することができる。

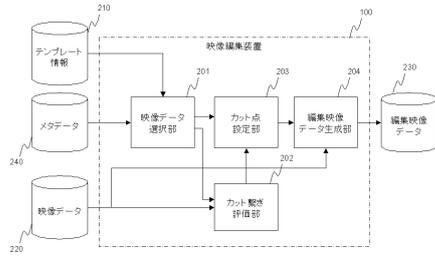
【図1】



【図3】

カットID	選択条件
1	顔サイズ=小 AND 人数=多
2	顔サイズ=中 AND 人数=少
3	顔サイズ=大 AND 人数=1
⋮	⋮
⋮	⋮

【図2】



【図4】

映像データ名	顔サイズ	人数	ユーザ付与
video_1.mpg	小	多	なし
video_2.mpg	中	少	なし
video_3.mpg	大	1	なし
video_4.mpg	なし	なし	飛行機 上空
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮

【図5】

900a

カットID	映像データ名
1	video_1.mpg
2	video_2.mpg
3	video_3.mpg
⋮	⋮
⋮	⋮

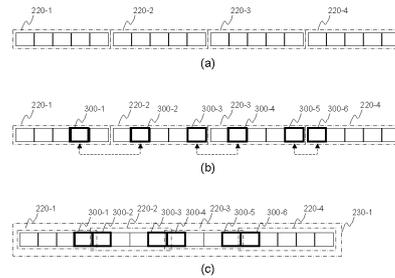
(a)

900b

カットID	映像データ名	イン点	アウト点
1	video_1.mpg	0:00:00.00	0:00:48.00
2	video_2.mpg	0:00:03.13	0:00:24.30
3	video_3.mpg	0:00:02.06	0:00:38.13
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮

(b)

【図7】

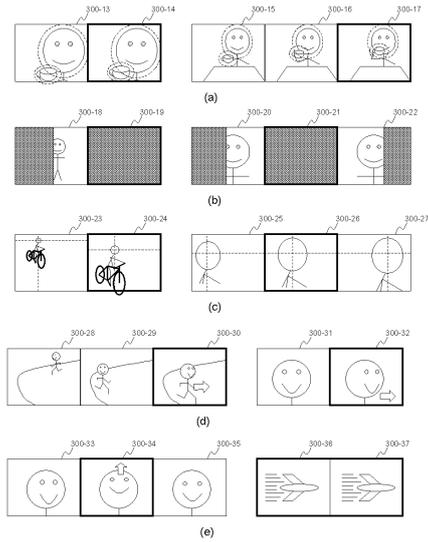


【図6】

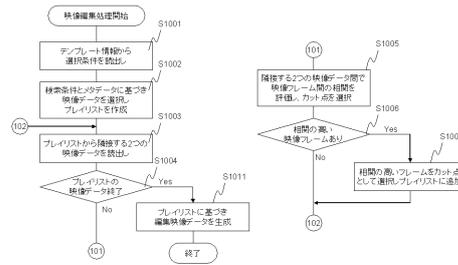
800

別映像データ名	該映像データ名	別映像データカット点	該映像データカット点
video_1.mpg	video_2.mpg	0:00:48.00	0:00:03.13
video_1.mpg	video_2.mpg	0:00:47.86	0:00:04.03
video_2.mpg	video_3.mpg	0:00:24.30	0:00:02.06
video_3.mpg	video_4.mpg	0:00:38.13	0:00:05.26
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮

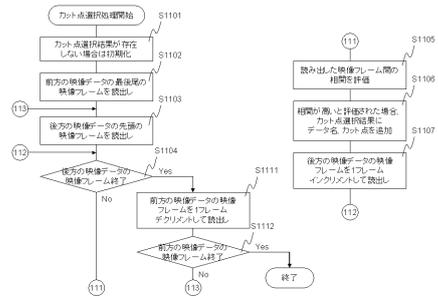
【図 8】



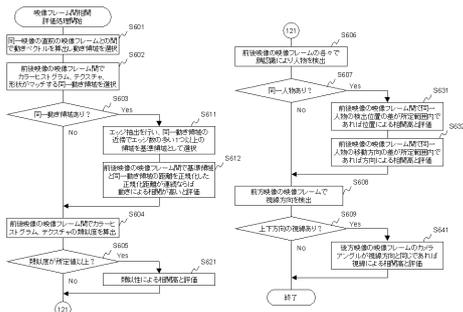
【図 9】



【図 10】



【図 11】



## フロントページの続き

- (56)参考文献 特開2007-184869(JP,A)  
特開2000-187737(JP,A)  
特開平09-149318(JP,A)  
特開2007-266838(JP,A)  
特開2007-082091(JP,A)  
特開2003-078868(JP,A)  
特開2003-069947(JP,A)  
鳶田聡, 知っておきたいキーワード(第32回)映像文法, 映像情報メディア学会誌, (社)映像情報メディア学会, 2008年 9月 1日, 第62巻, 第9号, p.1403-1405

## (58)調査した分野(Int.Cl., DB名)

H04N	5/76	-	5/956
G11B	20/10	-	20/12
G11B	27/00	-	27/34