

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6694829号
(P6694829)

(45) 発行日 令和2年5月20日(2020.5.20)

(24) 登録日 令和2年4月22日(2020.4.22)

(51) Int.Cl.

F 1

H04N 5/14 (2006.01)

H04N 5/14 O 4 O

G06T 7/246 (2017.01)

G06T 7/246

H04N 5/76 (2006.01)

H04N 5/76

請求項の数 10 (全 29 頁)

(21) 出願番号 特願2016-572515 (P2016-572515)
 (86) (22) 出願日 平成27年6月10日 (2015.6.10)
 (65) 公表番号 特表2017-528016 (P2017-528016A)
 (43) 公表日 平成29年9月21日 (2017.9.21)
 (86) 國際出願番号 PCT/US2015/034992
 (87) 國際公開番号 WO2015/191650
 (87) 國際公開日 平成27年12月17日 (2015.12.17)
 審査請求日 平成30年5月16日 (2018.5.16)
 (31) 優先権主張番号 14/303,466
 (32) 優先日 平成26年6月12日 (2014.6.12)
 (33) 優先権主張国・地域又は機関
米国(US)

(73) 特許権者 314015767
マイクロソフト テクノロジー ライセンシング、エルエルシー
アメリカ合衆国 ワシントン州 98052 レッドモンド ワン マイクロソフト ウェイ
(74) 代理人 100107766
弁理士 伊東 忠重
(74) 代理人 100070150
弁理士 伊東 忠彦
(74) 代理人 100091214
弁理士 大貫 進介

最終頁に続く

(54) 【発明の名称】ルールに基づくビデオ重要度解析

(57) 【特許請求の範囲】

【請求項 1】

ビデオファイルをデコードして、ビデオフレーム及び該ビデオフレームに関連付けられたオーディオデータを取得することと、

前記ビデオフレーム及び前記ビデオフレームに関連付けられた前記オーディオデータの2つ以上の高レベルの特徴と、前記ビデオフレームの低レベルの特徴とを、局所的ルール及び大局的ルールのうち少なくとも1つに基づいて解析して、前記ビデオフレームの2つ以上の特徴スコアを取得することであって、前記2つ以上の高レベルの特徴は、前記ビデオフレーム内の人間の顔の動き又は位置に関連する少なくとも1つの特徴と、前記オーディオデータが、雑音を含むか、発話を含むか又は音楽を含むかと、を含み、前記低レベルの特徴は、露出品質、彩度品質、色相多様性、ビデオフレームの揺れ、平均明度、色エンタロピー、及び隣接するビデオフレーム間におけるヒストグラム差のうち少なくとも1つを含む、ことと、

前記ビデオフレームの前記2つ以上の特徴スコアを組み合わせて、前記ビデオフレームのフレーム重要度スコアを導出することと、

前記ビデオファイルを、前記ビデオフレームの前記フレーム重要度スコアと前記ビデオファイルのさらなるビデオフレームの他フレーム重要度スコアとに少なくとも基づいて、異なるセクション重要度値のビデオセクションに区分すること、

を含む動作をプロセッサに実行させるプログラム。

【請求項 2】

10

前記動作は、前記セクション重要度値に従って前記ビデオセクションをランク付けすること、をさらに含む、請求項1に記載のプログラム。

【請求項3】

前記動作は、

前記ビデオファイル内のすべてのビデオフレームの前記フレーム重要度スコアに少なくとも基づいて前記ビデオファイルのビデオ重要度値を算出することと、

前記ビデオファイルの前記ビデオ重要度値とさらなるビデオファイルの他ビデオ重要度値とに少なくとも基づいて、前記さらなるビデオファイルに対して前記ビデオファイルをランク付けすることと、

をさらに含む、請求項1に記載のプログラム。

10

【請求項4】

前記解析することは、

前記ビデオフレームに対して顔検出を実行して1つ以上の人間の顔を検出することと、顔追跡を実行して前記1つ以上の人間の顔を追跡することと、

前記ビデオフレーム内の少なくとも1つの人間の顔の顔サイズ、顔位置、又は1つ以上の顔面特徴に少なくとも基づいて、前記ビデオフレームに否定的スコアバイアス又は肯定的スコアバイアスを割り当てることであって、前記1つ以上の顔面特徴は、前記少なくとも1つの人間の顔内の笑み及び開いた目のうち少なくとも1つの、欠如又は存在を含むことと、

前記1つ以上の人間の顔のうち少なくとも1つの人間の顔に対して、顔認識及び顔グループ化のうち少なくとも1つを実行することと、

顔認識データ及び顔グループ化データのうち少なくとも1つに少なくとも基づいて1以上の主要人物を識別することと、

前記ビデオフレーム内における主要人物の存在に対して、前記ビデオフレームにさらなる肯定的スコアバイアスを割り当てることと、

前記否定的スコアバイアスと前記肯定的スコアバイアスとについて考慮する前記ビデオフレームの顔重要度スコアを計算することと、

を含む、請求項1に記載のプログラム。

【請求項5】

ビデオファイルをデコードして、ビデオフレーム及び該ビデオフレームに関連付けられたオーディオデータを取得することと、

前記ビデオフレーム及び前記ビデオフレームに関連付けられた前記オーディオデータの2つ以上の高レベルの特徴と、前記ビデオフレームの低レベルの特徴とを、局所的ルール及び大局的ルールのうち少なくとも1つに基づいて解析して、前記ビデオフレームの2つ以上の特徴スコアを取得することであって、前記2つ以上の高レベルの特徴は、前記ビデオフレーム内の人間の顔の動き又は位置に関連する少なくとも1つの特徴と、前記オーディオデータが、雑音を含むか、発話を含むか又は音楽を含むかと、を含み、前記低レベルの特徴は、露出品質、彩度品質、色相多様性、ビデオフレームの揺れ、平均明度、色エンタロピー、及び隣接するビデオフレーム間におけるヒストグラム差のうち少なくとも1つを含む、ことと、

30

前記ビデオフレームの前記2つ以上の特徴スコアを組み合わせて、前記ビデオフレームのフレーム重要度スコアを導出することと、

前記ビデオファイルのビデオ重要度値を、前記ビデオフレームの前記フレーム重要度スコアと前記ビデオファイルのさらなるビデオフレームの他フレーム重要度スコアとに少なくとも基づいて算出することと、

前記ビデオファイルを、前記ビデオファイル内の複数のビデオフレームに関連付けられた、前記ビデオファイル内の複数のビデオフレームの前記フレーム重要度スコアに少なくとも基づいて、異なるセクション重要度のビデオセクションに区分することと、

前記ビデオファイルの前記ビデオ重要度値とさらなるビデオファイルの他ビデオ重要度値とに少なくとも基づいて、前記さらなるビデオファイルに対して前記ビデオファイルを

40

50

ランク付けすることと、
を含む、コンピュータにより実施される方法。

【請求項 6】

前記ビデオファイルを区分することは、
前記ビデオファイルの前記ビデオフレーム内の特徴点を、前記ビデオフレームの複数のダウンサンプリングされた画像間における共通点を検出することによって、検出することと、
前記ビデオフレーム内の前記特徴点を、前記ビデオファイルの前のビデオフレーム内の特徴点のさらなるセットに対してマッチさせることと、

前記ビデオフレーム間における前記特徴点の動きを説明するホモグラフィ変換マトリクスを算出することと、

前記ホモグラフィ変換マトリクスから前記ビデオフレームのモーションデータを計算することであって、前記モーションデータは、前記特徴点の動きのズーミング及び垂直並進成分の大きさ及び方向を含む、ことと、

前記ビデオフレームの前記モーションデータを、前記ビデオファイル内の1つ以上の残りのフレームのモーションデータと組み合わせることと、

前記ビデオファイル内の前記ビデオフレームの前記モーションデータの移動平均を算出することと、

ゼロ交差ゾーンを確認して、前記移動平均に少なくとも基づいて前記モーションデータの局所的ピーク及び谷情報を見つけることと、

前記ビデオファイルのビデオセクションを、前記モーションデータの前記局所的ピーク及び前記谷情報に少なくとも基づいて決定することと、

を含む、請求項5に記載のコンピュータにより実施される方法。

【請求項 7】

前記解析することは、
ビデオフレームに対して顔検出を実行して1つ以上の人間の顔を検出することと、
顔追跡を実行して前記1つ以上の人間の顔を追跡することと、

前記ビデオフレーム内の少なくとも1つの人間の顔の顔サイズ、顔位置、又は1つ以上の顔面特徴に少なくとも基づいて、前記ビデオフレームに否定的スコアバイアス又は肯定的スコアバイアスを割り当てることであって、前記1つ以上の顔面特徴は、前記少なくとも1つの人間の顔内の笑み及び開いた目のうち少なくとも1つの、欠如又は存在を含む、ことと、

前記1つ以上の人間の顔のうち少なくとも1つの人間の顔に対して、顔認識及び顔グループ化のうち少なくとも1つを実行することと、

顔認識データ及び顔グループ化データのうち少なくとも1つに少なくとも基づいて1以上の主要人物を識別することと、

前記ビデオフレーム内における主要人物の存在に対して、前記ビデオフレームに肯定的スコアバイアスを割り当てることと、

前記否定的スコアバイアスと前記肯定的スコアバイアスとについて考慮する前記ビデオフレームの顔重要度スコアを計算することと、

を含む、請求項5に記載のコンピュータにより実施される方法。

【請求項 8】

システムであって、
1つ以上のプロセッサと、
複数の動作を実行するように前記1つ以上のプロセッサにより実行可能である複数のコンピュータ実行可能コンポーネントを含むメモリと、

を含み、前記複数の動作は、

ビデオファイルをデコードして、ビデオフレーム及び該ビデオフレームに関連付けられたオーディオデータを取得することと、

前記ビデオフレーム及び前記ビデオフレームに関連付けられた前記オーディオデータ

10

20

30

40

50

の 2 つ以上の高レベルの特徴と、前記ビデオフレームの低レベルの特徴とを、局所的ルール及び大局的ルールのうち少なくとも 1 つに基づいて解析して、前記ビデオフレームの 2 つ以上の特徴スコアを取得することであって、前記 2 つ以上の高レベルの特徴は、前記ビデオフレーム内の人間の顔の動き又は位置に関連する少なくとも 1 つの特徴と、前記オーディオデータが、雑音を含むか、発話を含むか又は音楽を含むかと、を含み、前記低レベルの特徴は、露出品質、彩度品質、色相多様性、ビデオフレームの揺れ、平均明度、色エントロピー、及び隣接するビデオフレーム間におけるヒストグラム差のうち少なくとも 1 つを含む、ことと、

前記ビデオフレームの前記 2 つ以上の特徴スコアを組み合わせて、前記ビデオフレームのフレーム重要度スコアを導出することと、

10

前記ビデオファイルを、前記ビデオファイル内の複数のビデオフレームに関連付けられた、前記ビデオファイル内の複数のビデオフレームの前記フレーム重要度スコアに少なうとも基づいて、異なるセクション重要度のビデオセクションに区分することと、

前記ビデオフレームと前記ビデオファイル内のさらなるビデオフレームとの前記フレーム重要度スコアに少なくとも基づいて算出された、前記ビデオセクションのセクション重要度値に従って、前記ビデオセクションをランク付けすることと、

を含む、システム。

【請求項 9】

前記複数の動作は、

前記ビデオファイル内の前記複数のビデオフレームの前記フレーム重要度スコアに少なうとも基づいて、前記ビデオファイルのビデオ重要度値を算出することと、

20

前記ビデオファイルの前記ビデオ重要度値とさらなるビデオファイルの他ビデオ重要度値とに少なくとも基づいて、前記さらなるビデオファイルに対して前記ビデオファイルをランク付けすることと、

をさらに含む、請求項 8 に記載のシステム。

【請求項 10】

請求項 1 乃至 4 いずれか一項に記載のプログラムを記憶している記憶媒体。

【発明の詳細な説明】

【背景技術】

【0001】

30

消費者は、自身のスマートフォン及びパーソナルビデオレコーダーを用いて、ビデオをしばしば捕捉する。しかしながら、これら消費者のうち小さいパーセンテージのみが、自身のビデオを編集し、他の人々と共有する。さらに、ビデオが、多くの計画なしに思いつきで一般に撮影され、ごくわずかの興味深い瞬間を含み得るとき、消費者は、こうしたビデオの編集を退屈なプロセスであると思うことがある。時間とともに、より多くのビデオが消費者によって生成されるとき、この消費者は、ビデオの内容を思い出すのに困難を感じることもある。大抵のビデオ再生装置が、ビデオのサムネイル画像表現を提供し得るが、こうしたサムネイル画像表現は、ビデオの内容に対する十分な手がかりを提供しないことがある。

【発明の概要】

【0002】

40

本明細書に説明されるのは、ビデオファイルのルールに基づく解析を実行し、ビデオファイルのセクションをその重要度に基づいてランク付けする手法である。該手法は、さらに、ビデオファイルの集合のルールに基づく解析を実行し、複数のビデオファイルをその重要度に基づいてランク付けすることを含み得る。ビデオファイル又はビデオセクションの重要度は、該ビデオファイル又はビデオセクションが視聴者において発生させることを予期される関心の量に相関し得る。様々な実施形態において、ルールに基づく解析は、ビデオファイル又はビデオセクションの重要度を、該ビデオファイル又はビデオセクション内のビデオフレームの主観的な重要度及び / 又は客観的な技術的品質に基づいて評価することができる。

50

【 0 0 0 3 】

本発明の概要は、以下で詳細な説明においてさらに説明される概念のうち選択されたものを簡素化された形式で紹介するよう提供される。本発明の概要は、請求される対象事項の重要な特徴又は必須の特徴を識別するものではなく、請求される対象事項の範囲を限定するために使用されるものでもない。

【 図面の簡単な説明 】**【 0 0 0 4 】**

詳細な説明は、添付図面を参照して説明される。図面において、参照番号のうち最も左の（複数の）桁は、参照番号が最初に出現する図を識別する。異なる図面における同じ参照番号の使用は、類似の又は同一のアイテムを示す。

10

【図1】ルールに基づくビデオ解析エンジンを使用して、ビデオファイルの集合内のビデオファイルのビデオセクション又はビデオファイルをその重要度に基づいてランク付けする一例示的なスキームを例示するブロック図である。

【図2】ビデオファイルの集合内のビデオファイルのビデオセクション又はビデオファイルをその重要度に基づいてランク付けするルールに基づくビデオ解析エンジンの例示的なコンポーネントを示す一例示的な図である。

【図3】複数のビデオフレーム内の例示的な特徴点を整列させるホモグラフ変換の使用を示す一例示的な図である。

【図4】ルールに基づくビデオ解析を使用してビデオファイル内の特徴を解析して、ビデオファイルのビデオセクションをその重要度に基づいてランク付けする一例示的な処理を例示するフロー図である。

20

【図5】ルールに基づくビデオ解析を使用してビデオファイルの特徴を解析してビデオファイルを重要度に基づいてランク付けする一例示的な処理を例示するフロー図である。

【図6】ビデオフレームの顔重要度スコアを計算する一例示的な処理を例示するフロー図である。

【図7】特徴点の動きを解析することによってビデオ内的重要ビデオセクションを決定する一例示的な処理を例示するフロー図である。

【 発明を実施するための形態 】**【 0 0 0 5 】**

本明細書に説明されるのは、ビデオファイルのルールに基づく解析を実行してビデオファイルのセクションをその重要度（importance）に基づいてランク付けする手法である。上記手法は、さらに、ビデオファイルの集合のルールに基づく解析を実行してビデオファイルをその重要度に基づいてランク付けすることを含み得る。ビデオファイル又はビデオセクションの重要度は、該ビデオファイル又はビデオセクションが視聴者において発生させることを予期される関心の量に相関し得る。様々な実施形態において、ルールに基づく解析は、ビデオファイル又はビデオセクションの重要度を、該ビデオファイル又はビデオセクション内のビデオフレームの主観的な重要度及び／又は客観的な技術的品質に基づいて評価することができる。主観的な重要度の一例は、視聴者が、該視聴者に知られている人間の顔を表すビデオフレームを、上記視聴者が知らない人間の顔を表すビデオフレームよりもより重要と考えることであり得る。客観的な技術的品質の一例は、ビデオフレームの露出品質であり得る。

30

【 0 0 0 6 】

ルールに基づく解析は、フレームベースで複数の低レベルの特徴及び高レベルの特徴について、ビデオファイルのオーディオコンテンツ及びビデオコンテンツを解析することを含む。例えば、低レベルの特徴には、露出品質(exposure quality)、彩度品質(saturation quality)、ビデオフレームの揺れ(shakiness)、平均明度(average brightness)、色エントロピー、及び／又は隣接するビデオフレーム間におけるヒストグラム差などの特徴が含まれ得る。高レベルの特徴には、ビデオフレーム内で検出される人間の顔の数量、位置、及び／又は顔面特徴などの特徴が含まれ得る。解析には、局所的ルール及び大局的ルールの適用が含まれ得る。局所的ルールは、ビデオフレームの特徴解析結果の生

40

50

成の間に適用され得、大局的ルールは、ビデオファイル全体の特徴解析結果の生成の間に適用され得る。

【0007】

上記ルールは、特徴解析からの結果を提供し、組み合わせて、重要度スコアを生成することができる。重要度スコアは、ビデオファイルのフレーム、ビデオのセクション、及び／又はビデオファイルのその全体について生成されることができる。これら重要度スコアは、ビデオファイルのセクション及び／又はビデオファイルをランク付けするのに使用されることができる。重要度スコアは、ビデオファイルの視聴、編集、及び共有を容易にするのに使用されることができる。例えば、消費者が、ウェブサイト上で共有するために、最も高い重要度スコアを有するビデオファイルのセットを選択することができる。別の例において、アプリケーションが、最も高い重要度スコアを有するビデオファイルのセクションを一緒にまとめて(stitch)、ハイライトビデオファイルを作成することができる。10

【0008】

いくつかの例において、ビデオファイルは、検出されたカメラモーションの量に基づいて、様々な重要性の度合いを有するビデオセクションに区分されることができる。カメラモーションは、ビデオファイルの中のビデオフレーム内で検出される特徴点の動きを介して検出されることができる。いくつかの例において、ビデオセクションの重要度は、該ビデオセクションが視聴者から引き出すことを予期される関心の量に相関し得る。

【0009】

本明細書に説明される手法の使用は、ユーザがビデオファイルを該ユーザに対するその重要度に基づいてランク付けすることを可能にすることができる。こうしたランク付けに基づいて、ユーザは、どのビデオファイルを保持すべきかとどのビデオファイルを削除すべきかとを決めることができる。いくつかの例において、ユーザは、さらに、ビデオファイルのランク付けを使用して、特定のビデオファイルをオンライン共有ウェブサイト上に投稿すべきかどうかを決定することができる。本明細書に説明される手法は、さらに、ビデオファイルの重要度セクションを表現するサムネイル画像表現を提示することができ、したがって、ユーザは、一見しただけで、ビデオファイルのうち興味深い部分を言うことができる。こうした情報は、ビデオファイルを編集することにおいてユーザを支援して、ビデオファイルのコンテンツ品質を向上させ、あるいは個別セクションを強調することができる。様々な実施形態に従ってビデオファイルのルールに基づく解析を実行する手法の例が、以下で図1～図7を参照して説明される。2030

【0010】

例示的なスキーム

図1は、ルールに基づくビデオ解析エンジンを使用してビデオファイルの集合の中のビデオファイルのビデオセクション又はビデオファイルを重要度に基づいてランク付けする一例示的なスキーム100を例示するプロック図である。例示的なスキーム100は、ビデオ解析エンジン102を含み得る。ビデオ解析エンジン102は、1つ以上のコンピューティング装置104上で実行することができる。1つ以上のコンピューティング装置104には、汎用目的コンピュータ、例えば、デスクトップコンピュータ、タブレットコンピュータ、ラップトップコンピュータ、サーバなどが含まれ得る。しかしながら、他の実施形態において、コンピューティング装置104には、スマートフォン、ゲームコンソール、又は任意の他の電子装置が含まれてもよい。複数のコンピューティング装置104が、様々なプロセッサ、例えば、中央プロセッサユニット(CPU)、グラフィックスプロセッサユニット(GPU)、他のタイプのプロセッサ、又は上記のうち任意のものの組み合わせなどを含み得る。40

【0011】

ビデオ解析エンジン102は、ビデオ集合106のルールに基づく解析を実行することができる。ビデオ集合106は、複数のビデオファイル、例えば、ビデオファイル108(1)～108(N)などを含み得る。ルールに基づく解析は、フレームベースで複数の低レベルの特徴110及び複数の高レベルの特徴112について、ビデオファイル10850

(1)～108(N)のオーディオコンテンツ及びビデオコンテンツを解析することを含み得る。例えば、複数の低レベルの特徴110には、露出品質、彩度品質、及びビデオフレームの揺れなどの特徴が含まれ得る。複数の高レベルの特徴112には、ビデオフレーム内で検出される人間の顔の数量、位置、及び顔の特徴などの特徴が含まれ得る。

【0012】

ルールに基づく解析を実行することによって、ビデオ解析エンジン102は、ビデオ集合106のうち、ビデオ108(1)などの一ビデオファイルの複数セクションの重要度スコアと、ビデオファイル108(1)～108(N)などの複数ビデオファイルの重要度スコアとを生成することができる。このことに応じて、ビデオ解析エンジン102は、ビデオのセクションをその重要度スコアに従ってランク付けすることができる。例えば、ビデオファイル108(1)は、ビデオセクション114及びビデオセクション116を含み得、これらセクションは、その重要度スコア、例えば、重要度スコア118及び120に従ってそれぞれランク付けされる。ビデオセクションがいったんランク付けされると、ビデオ解析エンジン102は、ビデオセクションのサムネイル画像表現を表示することができ、サムネイル画像表現の選択が、対応するビデオセクションをメディアプレーヤ122に再生させることができる。

10

【0013】

ビデオ解析エンジン102は、さらに、ビデオ集合106の中のビデオファイルをその重要度スコアに従ってランク付けすることができる。例えば、ビデオ集合106のビデオファイル108(1)～108(N)は、その重要度スコア124(1)～124(N)に従ってランク付けされることがある。ビデオファイルがいったんランク付けされると、ビデオ解析エンジン102は、ビデオファイルのサムネイル画像表現を表示することができ、サムネイル画像表現の選択が、対応するビデオファイル又は対応するビデオファイルのセクションをメディアプレーヤ122に再生させることができる。

20

【0014】

例示的なコンポーネント

図2は、ビデオファイルの集合の中のビデオファイルのセクション又はビデオファイルをその重要度に基づいてランク付けするルールに基づくビデオ解析エンジン102の例示的なコンポーネントを示す一例示的な図である。ビデオ解析エンジン102は、1つ以上のコンピューティング装置104によって実装されることができる。コンピューティング装置104は、1つ以上のプロセッサ202、インターフェース204、及びメモリ206を含み得る。プロセッサ202の各々は、シングルコアプロセッサ又はマルチコアプロセッサであり得る。インターフェース204には、ユーザインターフェース及びネットワークインターフェースが含まれ得る。ユーザインターフェースは、データ出力装置（例えば、ビジュアルディスプレイ、オーディオスピーカ）と、1つ以上のデータ入力装置とを含み得る。データ入力装置には、これらに限られないが、キーパッド、キーボード、マウス装置、ジェスチャを受け入れるタッチスクリーン、マイクロフォン、音声又は発話認識装置、及び任意の他の適切な装置又は他の電子／ソフトウェア選択方法のうち、1つ又は複数の組み合わせが含まれ得る。

30

【0015】

ネットワークインターフェースには、コンピューティング装置104がネットワークを介してデータを送信し及び受信することを可能にする有線及び／又は無線の通信インターフェースコンポーネントが含まれ得る。様々な実施形態において、無線インターフェースコンポーネントには、これらに限られないが、セルラー、Wi-Fi、Ultra-wideband(UWB)、パーソナルエリアネットワーク（例えば、Bluetooth（登録商標））、衛星送信、及び／又は同様のものが含まれ得る。有線インターフェースコンポーネントには、直接I/Oインターフェース、例えば、イーサネット（登録商標）インターフェース、シリアルインターフェース、ユニバーサルシリアルバス(USB)インターフェース、及び／又は同様のものなどが含まれ得る。このようなものとして、コンピューティング装置104はネットワーク能力を有し得る。例えば、コンピューティング

40

50

装置 104 は、1つ以上のネットワーク、例えば、インターネット、モバイルネットワーク、ワイドエリアネットワーク、ローカルエリアネットワークなどを介して、他の電子装置（例えば、ラップトップコンピュータ、デスクトップコンピュータ、モバイルフォン、サーバ等）との間でデータを交換することができる。

【0016】

メモリ 206 は、コンピュータ読取可能媒体、例えば、コンピュータ記憶媒体などを用いて実装され得る。コンピュータ読取可能媒体には、少なくとも 2 タイプのコンピュータ読取可能媒体、すなわち、コンピュータ記憶媒体及び通信媒体が含まれる。コンピュータ記憶媒体には、コンピュータ読取可能命令、データ構造、プログラムモジュール、又は他のデータなどの情報の記憶のために任意の方法又はテクノロジーにおいて実装される揮発性及び不揮発性の、取外し可能及び取り外し不能の媒体が含まれる。コンピュータ記憶媒体には、これらに限られないが、RAM、ROM、EEPROM、フラッシュメモリ、若しくは他のメモリテクノロジー、CD-ROM、デジタル多用途ディスク（DVD）、若しくは他の光学記憶装置、磁気カセット、磁気テープ、磁気ディスク記憶装置、若しくは他の磁気記憶装置、又は、コンピューティング装置によるアクセスのための情報を記憶することに使用され得る任意の他の非伝送媒体が含まれる。対照的に、通信媒体は、コンピュータ読取可能命令、データ構造、プログラムモジュール、又は他のデータを、搬送波などの変調されたデータ信号、又は他の伝送メカニズムにおいて具現化し得る。本明細書において定義されるとき、コンピュータ記憶媒体は通信媒体を含まない。

【0017】

コンピューティング装置 104 のメモリ 206 は、オペレーティングシステム 208 と、ビデオ解析エンジン 102 を実装するモジュールとを記憶し得る。オペレーティングシステム 208 は、コンピューティング装置 104 が様々な入力（例えば、ユーザ制御、ネットワークインターフェース、及び / 又はメモリ装置）を介してデータを受信し、該データをプロセッサ 202 を用いて処理して出力を生成することを可能にするコンポーネントを含み得る。オペレーティングシステム 208 は、出力を提示する（例えば、電子ディスプレイ上に画像を表示する、メモリにデータを記憶する、データを別の電子装置に送信する等）1つ以上のコンポーネントをさらに含み得る。オペレーティングシステム 208 は、ユーザがインターフェース 204 を用いてビデオ解析エンジン 102 のモジュールと相互作用することを可能にすることができる。さらに、オペレーティングシステム 208 は、オペレーティングシステムに一般に関連付けられる様々な他の機能を実行する他のコンポーネントを含んでもよい。

【0018】

上記モジュールには、ビデオデコーダモジュール 210、低レベル解析モジュール 212、高レベル解析モジュール 214、モーション解析モジュール 216、重要度算出モジュール 218、ビデオ区分モジュール 220、ビデオランク付けモジュール 222、及びユーザインターフェースモジュール 224 が含まれ得る。モジュールの各々が、特定のタスクを実行し又は特定の抽象データタイプを実装するルーチン、プログラム、命令、オブジェクト、及び / 又はデータ構造を含み得る。さらに、データストア 226 がメモリ 206 内に存在し得る。低レベル解析モジュール 212 と高レベル解析モジュール 214 の各々は、局所的ルール又は大局的ルールを適用して、ビデオファイル内の特徴データ、すなわち、ビデオデータ及び / 又はオーディオデータの重要度を解析することができる。局所的ルールは、単一のビデオフレーム内の特徴データに基づく、該単一のビデオフレームの重要度の割り当てに影響し得る。逆に、大局的ルールは、ビデオファイルのうち複数のビデオフレーム内の特徴データに基づく、該ビデオファイル全体の重要度の割り当てに影響し、あるいは、ビデオファイル内のビデオフレームのセットにわたり共有される特徴に基づく、該ビデオフレームのセットの各々についての重要度の割り当てに影響し得る。

【0019】

ビデオデコーダモジュール 210 は、ビデオファイル、例えば、ビデオファイル 108 (1)などをデコードして、ビデオフレーム及び / 又は該ビデオフレームの各々に関連付

10

20

30

40

50

けられたオーディオデータを取得することができる。ビデオデコーダモジュール210は、様々なコーデックを使用して、ビデオファイル、例えば、H.264、MPEG-2、MPEG-4等をデコードすることができる。

【0020】

低レベル解析モジュール212は、各々のデコードされたビデオフレームを低レベルの特徴について解析して、特徴スコアを作り出すことができる。様々な実施形態において、低レベルの特徴には、露出品質、彩度品質、色相多様性(hue variety)、揺れ、平均明度、色エントロピー、及び／又は隣接するビデオフレーム間におけるヒストグラム差が含まれ得る。低レベル解析モジュール212は、ビデオフレームの露出、彩度、及び色相を示すヒストグラムを導出するアルゴリズムを使用することができる。露出品質の解析において、低レベル解析モジュール212は、ビデオフレームの露出バランスの露出ヒストグラムを解析することができる。低レベル解析モジュール212は、局所的ルールに従って露出バランスに基づいてビデオフレームに露出格付けスコアを割り当てることができ、これにおいて、より高い露出バランスがより高い露出格付けスコアを結果としてもたらし得る。逆に、ビデオフレームのより低い露出バランスは、より低い露出格付けスコアを結果としてもたらし得る。

10

【0021】

彩度品質の解析において、低レベル解析モジュール212は、ビデオフレームの彩度ヒストグラム、例えば、HSV色空間の彩度ヒストグラムなどを解析することができる。この解析に基づいて、低レベル解析モジュール212は、局所的ルールに従って彩度ヒストグラムの中の値の中間セット内の彩度の量を反映する彩度スコアを計算することができる。このようなものとして、上記中間レンジ内におけるより大きい彩度は、ビデオフレームのより高い彩度スコアを結果としてもたらす。逆に、上記中間レンジ内におけるより小さい彩度は、ビデオフレームのより低い彩度スコアを結果としてもたらす。

20

【0022】

色相多様性の解析において、低レベル解析モジュール212は、ビデオフレームの色相ヒストグラムのバランスを査定することができる。低レベル解析モジュール212は、さらに、局所的ルールに基づいて色相スコアを割り当てることができる。このことに応じて、低レベル解析モジュール212は、ビデオフレームの色相がうまくバランスをとられている、すなわち、より高い多様性の色を示すとき、より高い色相スコアを割り当てることができる。逆に、低レベル解析モジュールは、ビデオフレームの色相があまりバランスをとられていない、すなわち、より低い多様性の色を示すとき、より低い色相スコアを割り当てることができる。

30

【0023】

揺れの解析において、低レベル解析モジュール212は、モーション解析モジュール216を使用してフレーム間における特徴点の動きを追跡し、その動きをモデル化する変換を生成することができる。特徴点は、複数のビデオフレームにわたり確実に場所を特定され(located)得る、ビデオフレーム内の関心点である。特徴点は、該特徴点が2次元(2D)構造を含み、且つx方向及びy方向において局所化され(localized)得るという点で、区別可能である。低レベル解析モジュール212は、上記変換を使用して、特徴点モーションの大きさ及び方向に関連した局所的及び大局的トレンドを解析することができる。局所的トレンド及び大局的トレンドは、ビデオファイル内の揺れを、局所的なフレームごとの変化の属性として定量化することができる。ビデオファイルの揺れは、以下でモーション類別解析に関して説明されるとおり、モーション解析モジュール216によって決定されることがある。このことに応じて、低レベル解析モジュール212は大局的ルールを適用して、ビデオファイル内の揺れの量に対して反比例する揺れスコアをビデオファイルに割り当てることができ、したがって、より大きい揺れはより低い揺れスコアを結果としてもたらし、逆もまた同様である。

40

【0024】

平均明度の解析において、低レベル解析モジュール212は、ビデオフレーム内のすべ

50

ての画素の輝度成分の平均を算出することができる。様々な実施形態において、低レベル解析モジュール212は、ビデオフレームのグレースケール画像表現における画素値を平均することができる。例えば、画素値は0～255に及んでもよく、これにおいて、0は色の黒に対応し、255は色の白に対応する。いくつかの実施形態において、低レベル解析モジュール212は、毎n番目の画素から画素値を読み出すようにさらに最適化されることがある。換言すると、低レベル解析モジュール212は、解析を実行する間、x方向及び/又はy方向において所定数の画素をスキップすることができる。ビデオフレーム内の画素の画素値の平均に基づいて、低レベル解析モジュール212は、ビデオフレームの明度を決定することができる。このことに応じて、低レベル解析モジュール212は局所的ルールを適用して、ビデオフレームの平均画素値が明度の所定中間レンジに入るとき¹⁰、ビデオフレームの平均画素値に比例する明度スコアを割り当てることができる。しかしながら、低レベル解析モジュール212は、平均画素値が外側に入る、すなわち、明度の所定中間レンジよりもより高いか又はより低いとき、明度の所定中間レンジに入る平均画素値に割り当てられるいかなる明度スコアよりもより低い明度スコアを割り当てることができる。平均画素値が減少するとき、明度の所定中間レンジの下限よりもより小さい間、上記明度スコアは小さくなり得る。さらに、平均画素値が増加するとき、明度の所定中間レンジの上限よりもより大きい間、上記明度スコアは小さくなり得る。

【0025】

色エントロピーの解析において、低レベル解析モジュール212は、ビデオフレーム内の色エントロピーの量を決定することができる。色エントロピーの量は、ビデオフレーム内の色間ににおける差の指標である。低レベル解析モジュール212によって割り当てられる色エントロピー値は、ビデオフレーム内の色エントロピーの実際量に依存して、“0”から“1”におよび得る。例えば、低レベル解析モジュール212は、ビデオフレームの画素が最大量の差を有するとき、ビデオフレームに“1”的色エントロピー値を割り当てることができる。低レベル解析モジュール212は、ビデオフレームの画素が最小量の差を有するとき、ビデオフレームに“0”的色エントロピー値を割り当てることができる。様々な実施形態において、低レベル解析モジュール212は、色空間（例えば、RGB色空間又はHSV色空間）の色ドメインヒストグラムを算出することによって色エントロピー値を決定することができる。²⁰

【0026】

上記実施形態において、低レベル解析モジュール212は、複数の色次元を捕捉するヒストグラムを最初作成することができる。例えば、RGB色空間において、R、G、及びBの各々は256のとり得る値を有し得、この場合、ヒストグラムは $256 \times 256 \times 256$ バケットを有し得る。少なくとも1つの実施形態において、バケットは、バケットサイズ及び/又は処理スピードを最適化することに対してさらに定量化されることがある。例えば、サイズが $256 \times 256 \times 256$ であってもよく、したがって、複数の色値が同じバケット内に入ることになる。ゆえに、1つの例において、ヒストグラムは、コードにおいて下記のフォーマットで表わされることができる：int Histogram[256 * 256 * 256]。このことは、ヒストグラム配列がRGB色空間におけるすべてのとり得る色のための要素を有することを意味する。このことに応じて、低レベル解析モジュール212が画素を読み出すとき、低レベル解析モジュール212は、下記として値を設定することができる：³⁰

【0027】

【数1】

```

int IndexInHistogramForColor =
pixelColor.red + (256 * pixelColor.green) + (256*256*
pixelColor.blue);

Histogram[IndexInHistogramForColor] =
Histogram[IndexInHistogramForColor]+1; // この色を有する1つのさら
                                             
なる画素が観察されるとき、そのカウントをインクリメントする

```

10

上記ステップがビデオフレーム内の各画素についていったん実行されると、低レベル解析モジュール212はヒストグラムを正規化することができる。換言すると、低レベル解析モジュール212は、各値をヒストグラムのサイズを用いて分割する(divide)ことができ、したがってヒストグラム内の値は0と1との間になり、値は最大で1まで増大する。結果として、最も高い値を有する要素が、ビデオフレーム内で最も頻繁に発生する。

20

【0028】

ヒストグラム内の画素のエントロピーは、すべての(Histogram[n] * log(Histogram[n]))の総和として、下記のとおり定式化されることができる：

【0029】

【数2】

$$H(X) = \sum_i P(x_i) I(x_i) = - \sum_i P(x_i) \log_b P(x_i)$$

30

低レベル解析モジュール212は、ヒストグラムからのエントロピーを最大のとり得るエントロピーで除算することによって、相対色エントロピー（すなわち、ビデオフレームの、他のビデオフレームに対する色エントロピー）を決定することができる。最大のとり得るエントロピーは、すべてのHistogram[n]が同じ値を有するとき、すなわち、(1.0 /

とり得る値の数)に等しいとして定義されることができる。ビデオフレームの相対色エントロピー値がいったん取得されると、低レベル解析モジュール212は局所的ルールを適用して、ビデオフレームの相対色エントロピーからビデオフレームのエントロピースコアを取得することができる。様々な実施形態において、ビデオフレームのエントロピースコアは、ビデオフレームの相対色エントロピー値に対して正比例してもよく、例えば、より高い相対エントロピー値がより高いエントロピースコアを結果としてもたらし、逆もまた同様である。

40

【0030】

ヒストグラム差の解析において、低レベル解析モジュール212は、2つの隣接するビデオフレーム間におけるヒストグラム差を決定することができる。様々な実施形態において、低レベル解析モジュール212は、各ビデオフレームを複数のセル（例えば、10×10セル）に分割することができる。ビデオフレームt及び隣接ビデオフレームt+1の各セルについて、低レベル解析モジュール212は色ヒストグラム（例えば、RGBヒストグラム）を算出することができる。その後、低レベル解析モジュール212は、ビデオフレームt内の各セルについて、そのヒストグラムと隣接ビデオフレームt+1内で対応する位置を有するセルのヒストグラムとにおける差を計算することができる。2つの隣接

50

ビデオフレーム内のセルペアのヒストグラム間における差がさらに標準化されて（例えば、2乗され、正規化され、かつ／あるいは平均される等）、上記2つの隣接フレームの最終的なヒストグラム差値を取得してもよく、これにおいて、その値は“0”と“1”との間におよび得る。2つの隣接ビデオフレームのヒストグラム差値がいったん取得されると、低レベル解析モジュール212は局所的ルールを適用して、ヒストグラム差値からビデオフレームtのヒストグラム差スコアを取得することができる。様々な実施形態において、ビデオフレームのヒストグラム差スコアはヒストグラム差に正比例してもよく、例えば、より高いヒストグラム差値がより高いヒストグラム差スコアを結果としてもたらし、逆もまた同様である。

【0031】

10

少なくともいくつかの実施形態において、低レベル解析モジュール212は、解析のいくつかを最適化してビデオファイルの処理をスピードアップすることができる。例えば、低レベル解析モジュール212は、ビデオファイル内のすべてのビデオフレームというよりもビデオファイル内のビデオフレームのサブセットについて、解析（例えば、露出格付け解析、色相多様性解析等）を実行することができる。低レベル解析モジュール212は、さらに、元のフレームの縮小された（scaled down）バージョンに対して解析のうち1つ以上を実行していくらかの効率を得ることができる。例えば、ビデオフレームの色相多様性及び彩度解析が、ビデオフレームのダウンサンプリングされたバージョンに対して実行されることができる。別の例において、露出品質解析が、ビデオフレームのモノクロの且つダウンサンプリングされたバージョンに対して実行されてもよい。低レベル解析モジュール212はさらに、複数の解析を並列に又は実質的に並列に実行してもよい。例えば、彩度品質解析及び色相多様性解析が並列に実行されてもよい。

【0032】

20

高レベル解析モジュール214は、各々のデコードされたビデオフレームを高レベルの特徴について解析することができる。少なくとも1つの実施形態において、高レベル特徴解析には、顔検出、顔追跡、顔認識、顕著性（saliency）解析、オーディオパワー解析、オーディオ分類解析、発話解析、及びモーション解析が含まれ得る。

【0033】

30

顔検出において、高レベル解析モジュール214は、デコードされたビデオフレームを解析して、人間の顔がビデオフレーム内に提示されているかどうかを検出することができる。検出された顔は、ビデオフレームを捕捉したカメラに對面しているか、又はカメラに對して横向きであり得る。上記検出に基づいて、高レベル解析モジュール214は、検出された顔のリストを、ビデオフレーム内におけるその位置、各顔によりカバーされるビデオフレームのエリア、及び検出における信頼度を示す各顔の検出信頼度スコアと共に生成することができる。

【0034】

様々な実施形態において、高レベル解析モジュール214は局所的ルールを適用して、検出された顔のサイズに基づいて、ビデオフレームのサイズのパーセンテージとして、ビデオフレームの顔重要度スコアを算出することができる。2つのビデオフレーム上で検出される同じサイズを有する顔は、同じ顔重要度スコアを割り当てられることができる。しかしながら、ビデオフレームt₂上の顔がビデオフレームt₁上の顔よりもより大きい場合、ビデオフレームt₂の顔重要度スコアはより高くなることになり、なぜならば、より大きい顔はより小さい顔よりもより重要と考えられるからである。他の実施形態において、高レベル解析モジュール214は、検出された顔のサイズが最小サイズ閾値と最大サイズ閾値との間である場合、顔重要度スコアを算出するように構成されてもよい。逆に、サイズが最小サイズ閾値よりもより小さいか又は最大サイズ閾値よりもより大きい顔は、高レベル解析モジュール214による顔重要度スコア算出に対して妥当でないと考えられてもよく、あるいは、否定的（negative）スコアバイアスが上記の発生に対して対応ビデオフレームに割り当てられてもよい。

【0035】

40

50

別法として、又は同時的に、ビデオフレームの顔重要度を算出する局所的ルールは、ビデオフレーム内に表される各顔の顔面特徴を考慮に入れてもよい。例えば、顔面特徴には、顔が笑っているか否か、又は目が開いているか否か等が含まれ得る。ゆえに、笑っている顔について対応ビデオフレームに対して肯定的(positive)スコアバイアスが割り当てられることができ、一方、顔が笑っていないときに否定的スコアバイアスが割り当てられることがある。同様にして、開いた目を有する顔について対応ビデオフレームに対して肯定的スコアバイアスが割り当てられることができ、目が閉じられているときに否定的スコアバイアスが割り当てられることがある。

【0036】

顔認識において、高レベル解析モジュール214は顔認識アルゴリズムを使用して、ビデオフレーム内で検出される各人の顔を、人の既知のアイデンティティ(identity)にマッチさせることができる。いくつかの実施形態において、高レベル解析モジュール214は、既知の顔の知識データベースを使用して、人間の顔を既知の人間にマッチさせることができる。別法として、又は同時に、高レベル解析モジュール214は、ユーザインターフェースモジュール24を使用してユーザインターフェース制御を提供することができる。上記ユーザインターフェース制御は、ユーザが、1つ以上の認識された顔の各々に、アイデンティティ、該顔の重要度格付け、及び/又は該顔を有する人間の視聴者に対する関係でタグ付けすることを可能にする。少なくとも1つの実施形態において、顔に関して視聴者により提供される情報は、知識データベースに追加されることができる。

【0037】

顔追跡において、高レベル解析モジュール214は、複数のビデオフレームにわたり人間の顔を追跡することができる。こうして、高レベル解析モジュール214は、ビデオファイル内に存在する顔のセットを確認し、さらに、各顔がビデオファイル内に出現する頻度を追跡することができる。さらに、顔グループ化において、高レベル解析モジュール214は、追跡される顔をグループ化して、異なるビデオフレーム上で検出される顔が同じ人間に属するかどうかを決定することができる。様々な実施形態において、高レベル解析モジュール214は、ビデオフレーム内の検出された顔の各々について、顔面特徴のセットを取得することができる。高レベル解析モジュール214は、検出された顔の顔面特徴のセットを比較して、検出された顔を顔面特徴類似度に従ってグループへとグループ化することができる。高レベル解析モジュール214は、顔の各グループの重要度を、各グループ内の顔の数に従って評価することができる。各グループ内の顔の数は、ビデオファイル内における顔の行き渡り(prevalence)に正比例する。さらに、より多くの行き渡りは、その顔を有する人間のより高い重要度を示し得、逆もまた同様である。このことに応じて、顔のグループのためのグループ重要度スコアが、グループ内の顔の顔重要度スコアの合計であり得る。このようなものとして、最高レベル解析モジュール214は、複数のビデオフレーム内に表されるとおりにグループ内の顔の顔重要度スコアを合計して、グループのためのグループ重要度スコアを算出することができる。

【0038】

例えば、最高グループ重要度スコアを有するグループ内に顔が出現する人は、主要な人物と考えられることができる。こうした情報に基づいて、高レベル解析モジュール214は大局的ルールをビデオフレームに適用することができ、したがって、主要人物を示すビデオフレームの顔重要度スコアがさらに肯定的にバイアスをかけられ、すなわち、高められることがある。さらに、複数のビデオフレーム内で捕捉されたとしてグループ内に含まれる顔の位置が、ビデオフレームに割り当てられる顔重要度スコアを低めてもよい。例えば、特定の人の顔を示すビデオフレームの重要度スコアが、該顔の、ビデオフレームの中心に対する、軸(例えば、x軸又はy軸)に沿った距離に従ってバイアスをかけられてもよく、したがって、中心に対してより近い顔がビデオフレームのより高い重要度を結果としてもたらし、逆もまた同様である。

【0039】

フレーム顕著性(saliency)解析において、高レベル解析モジュール214は、ビデオ

10

20

30

40

50

フレームの顕著な部分を検出することができる。例えば、ビデオフレームの顕著な部分は、動いているオブジェクトを捕捉することができる。顕著性解析に基づいて、高レベル解析モジュール 214 は局所的ルールを適用して、ビデオフレーム内のあらゆる画素の顕著性スコアを表示するヒートマップを生成することができる。ヒートマップは、マトリクス内に並べられたデータのグラフィカル表現であり、これにおいて、マトリクス内の個々の値は色を用いて表現される。高レベル解析モジュール 214 は、ビデオフレーム内の画素の顕著性スコアに基づくビデオフレームのフレーム顕著性スコアをさらに生成することができる。例えば、ビデオフレームのビデオフレーム顕著性スコアは、画素顕著性スコアの平均であってもよい。

【0040】

10

オーディオパワー解析において、高レベル解析モジュール 214 は、継続時間においてビデオフレームに対応するオーディオデータを査定し（例えば、1秒の 1 / 30 又は 1 / 60）、オーディオパワーの二乗平均平方根（root mean square）（RMS）値を算出することができる。オーディオパワーのより高い RMS 値は、対応するビデオフレームのより高い重要度を示し得、逆もまた同様である。ゆえに、高レベル解析モジュール 214 は、局所的ルールに従って、オーディオパワー重要度スコアを対応するビデオフレームに割り当てることができる。

【0041】

オーディオ分類解析において、高レベル解析モジュール 214 は、機械学習分類器を使用して、継続時間においてビデオフレームに対応するオーディオデータが異なるタイプのオーディオデータ（例えば、雑音、発話、又は音楽）を含むかどうかを決定することができる。異なるタイプのオーディオデータは、対応するビデオフレームの異なる重要度を反映し得る。局所的ルールに基づいて、高レベル解析モジュール 214 は、オーディオデータのタイプに基づいて、オーディオ分類重要度スコアを対応するビデオフレームに割り当てることができる。例えば、発話の存在が、高レベル解析モジュール 214 に、高いオーディオ分類重要度を対応ビデオフレーム割り当てさせてもよい。対照的に、音楽の存在が、対応ビデオフレームに対する中位のオーディオ分類スコアの割り当てを引き起こしてもよい。さらに、雑音の存在が、高レベル解析モジュール 214 に、低いオーディオ分類スコアを対応ビデオフレームに割り当てさせてもよい。

20

【0042】

30

モーション類別解析において、高レベル解析モジュール 214 は、モーション解析モジュール 216 を使用してビデオフレーム間における特徴点の動きを追跡し、その動きをモデル化する変換を生成することができる。高レベル解析モジュール 214 は、上記変換を使用して、動きの大きさ及び方向に関連した局所的トレンド及び全局的トレンドを解析することができる。ひいては、高レベル解析モジュール 214 は、局所的トレンド及び全局的トレンドを使用して、ビデオフレーム内で捕捉された揺れについて考慮し（account for）、シーンに対するカメラの意図的動作、例えば、ズーミング、パンニング等を決定することができる。

【0043】

様々な実施形態において、モーション解析モジュール 216 は、2つの隣接するフレームについて特徴点の場所を特定すること（locating）によって、モーション類別解析を開始することができる。特徴点は、画像の2次元（2D）又は3D変換においてでさえ識別可能なままである、該画像内の点であり得る。特徴点を検出するために、モーション解析モジュール 216 は、画像をダウンサンプリングし、より小さい次元のダウンサンプリングされた画像のピラミッドを作成することができる。それから、ダウンサンプリングされた画像はモーション解析モジュール 216 によって比較されて、ダウンサンプリングされた画像間における共通点（すなわち、特徴点）を決定することができる。様々な実施形態において、モーション解析モジュール 216 は、共通点を検出するためのいくつかの検出アルゴリズムのうち1つ以上、例えば、ラプラス（Laplace）検出アルゴリズム、ハリス（Harris）検出アルゴリズム、ヘシアン（Hessian）検出アルゴリズム、ヘシアンラプラ

40

50

ス検出アルゴリズム、ヘシアンアフィン (HessianAffine) 検出アルゴリズム、エッジ焦点 (EdgeFoci) 検出アルゴリズム等を使用することができる。

【0044】

2つの隣接するフレームについて特徴点がいったん識別されると、モーション解析モジュール216は、最大数の特徴点がマッチするように2つの隣接フレームを整列させる (aligns) 変換を決定することができる。変換は、ロバストパラメータ推定の一実装であるジオメトリックマッチングを用いて実行されることができる。変換は、マッチした特徴点から算出されるホモグラフィ (homography) 変換マトリクスを提供することができる。様々な実施形態において、モーション解析モジュール216は、ランダムサンプリング及びコンセンサス (Random Sampling and Consensus) (RANSAC) アルゴリズムを使用して、初期パラメータ推定と統計的インライア (statistical inliers) のリストとを取得することができ、これにおいて、初期パラメータ推定はさらに精緻化される。ホモグラフィ変換により整列させられるビデオフレーム内の様々なカメラモーションが、図3に例示される。
10

【0045】

図3は、複数のビデオフレーム内で例示的な特徴点を整列させるホモグラフ (homograph) 変換の使用を示す一例示的な図である。図示されるとおり、ビデオフレーム302～312の各々は、円形で表される識別された特徴点のグループをそれぞれ含み得る。例えば、ビデオフレーム302内の特徴点のグループは、円形314(1)～315(5)によって表される。特徴点のグループ内の各特徴点は、複数のビデオフレームを捕捉したカメラの動作にもかかわらず、この複数のビデオフレームにわたり互いに対する相対位置を保有し得る。例えば、変換されたビデオフレーム316は、ビデオフレーム302に対しての、カメラの追跡動作の結果である後続ビデオフレームであり得る。変換されたビデオフレーム318は、ビデオフレーム304に対しての、カメラのズーム動作の結果である後続ビデオフレームであり得る。変換されたビデオフレーム320は、ビデオフレーム306に対しての、ズーム／ドリー動作の結果である後続ビデオフレームであり得る。変換されたビデオフレーム312は、ビデオフレーム308に対しての、カメラのロール動作の結果である後続ビデオフレームであり得る。変換されたビデオフレーム314は、ビデオフレーム310に対しての、カメラの垂直パン／ピッチ／チルトの結果である後続ビデオフレームであり得る。変換されたビデオフレーム316は、ビデオフレーム312に対しての、カメラの水平パンニングの結果である後続ビデオフレームであり得る。
20
30

【0046】

しかしながら、ビデオフレーム302～312から変換されたビデオフレーム326～336を作り出したカメラによる動作にかかわらず、モーション解析モジュール216はホモグラフ変換を使用して、ビデオフレームとその対応する変換されたビデオフレームとの中の特徴点を整列させることができる。

【0047】

図2に戻ると、RANSACアルゴリズムは、特徴点マッチの最小サブセットから変形 (transformation) マトリクスパラメータを直接計算することができる。例えば、2つのフレーム間で対応する2つの特徴点から、相似変換 (similarity transform) (例えば、並進、回転、又はスケール変更 (scale)) が計算されることがある。候補ジオメトリック変形がいったん取得されると、RANSACアルゴリズムは、データセット内のすべての他の特徴点マッチに対して変形をテストすることと、十分な正確さで空間的に投影する (project) 特徴点であるインライアの数の総数を生成することによって、変形を検証することができる。換言すると、RANSACアルゴリズムは、点マッチの最小セットを最初ランダムに選び、このセットから変形パラメータを計算し、それから、インライアマッチの数を数えることによって上記パラメータを検証することができる。その後、RANSACアルゴリズムは、最良の変形を記録する。RANSACアルゴリズムは、変形パラメータの良好なセットを見つける確率がデータミスマッチレートを所与として所定の確率閾値に達するまで、上記手順を複数回繰り返すことができる。
40
50

【 0 0 4 8 】

いくつかの実施形態において、RANSACアルゴリズムは、ベイズ統計学のより良好な使用を行うように修正されることがある。変形パラメータの検証の間にインライアを数えるというよりも、RANSACアルゴリズムは、すべての特徴点マッチから、各ランダム変形について、ログ確率スコアを計算することができる。このスコアは2つの部分、すなわち、(1)パラメータと、共通に予期される値からパラメータがどれほど離れているかとに依存する、事前確率スコア、及び(2)特徴点マッチの再投影距離(re-projection distance)のロバスト関数(robust function)に基づく確率スコア、を含むことができる。上記スコアは、正しい場所に投影する特徴点の方を好むが、アウトライア(outliers)が共存することを許容する。

10

【 0 0 4 9 】

ホモグラフ変換から、モーション解析モジュール216は、ズーミング及び垂直並進成分(zooming and vertical translation components)の大きさ及び方向を抽出すると同時に、他の種類のモーションを無視することができる。上記大きさ及び方向は、2つの隣接するフレーム内に記録されるとおりの、カメラの意図的動作である。換言すると、上記大きさ及び方向は、第1のフレームから第2の隣接フレームへの実際の動きの変化の一次導関数(first order derivative)である。モーション解析モジュール216は、ビデオフレームの動きベクトルのモーションデータ移動平均を算出することによって、ビデオフレームを記録したカメラの揺れ動き(shakiness movement)を決定し、モーションデータ移動平均から上記カメラの意図的動きを減算することができる。モーションデータ移動平均の算出は、局所的分散(local variance)を抑制し、意図的動きを表す長期トレンドを維持する。換言すると、第1のフレームから第2のフレームへの意図的動きと全体的動き変化における差は、2つのフレームを記録したカメラの揺れ動きである。

20

【 0 0 5 0 】

フレームのセット内に記録されるズーミング及び垂直並進値の大きさは、これらフレームの重要度に関する手がかりを提供することができる。例えば、より高い値が、ビデオフレーム内の1つ以上のオブジェクトに関してのカメラモーションにおける加速を示し得る。さらに、より高い加速度を有するビデオフレームの領域がより重要であると仮定されてもよく、なぜならば、カメラがモーションにおける素早い変化を行って何らかのアクションを捕捉している可能性があるからである。このことに応じて、高レベル解析モジュール214は、モーション加速の量に基づいてモーション重要度スコアを各フレームに割り当てることができる。

30

【 0 0 5 1 】

モーション解析モジュール216は、意図的動きデータを解析して、ズーム及び垂直並進モーションの双方についてゼロ交差のまわりに置かれる(pivoted)複数の極大及び極小を決定することができる。いくつかの実施形態において、モーション解析モジュール216は、極大及び極小の場所を使用して、データをビデオセクションに区分することができる。別法として、又は同時に、モーション解析モジュール216は、極大及び極小の場所を使用して、ビデオファイルを、ペアにされた変化の方向を有するビデオセクションに区分することができ、上記のペアにされた方向は、例えば、ズームアウトとペアにされたズームイン、パンダウンとペアにされたパンアップなどである。

40

【 0 0 5 2 】

ビデオファイルの区分されたビデオセクションは、ビデオファイルのシーンにおける変化と一貫性があり得る。このことに応じて、ビデオファイル内のセクション境界がガイドラインとして使用されて、ビデオファイルを異なる重要度のビデオセクションに分割することができる。上記境界は、重要/非重要セクションの開始及び終了を、カメラの動作におけるシフト又はシーン内の活動の性質における変化が存在する時点に整列させることができる。さらに、モーション解析モジュール216は、セクションのズーム及びパンモーションの大きさを組み合わせ、平均することができる。ビデオセクションのズーム及びパンモーションの大きさの平均によって表現される加速の量がモーション解析モジュール2

50

16によって使用されて、フレームに関して上記で説明されたのと同じ仕方でビデオセクションにモーション重要度スコアを割り当てることができる。

【0053】

少なくともいくつかの実施形態において、高レベル解析モジュール214は、解析のいくつかを最適化して、ビデオファイルの処理をスピードアップすることができる。例えば、高レベル解析モジュール214は、各ビデオフレーム内の1つ以上の顔について、該ビデオフレームの、モノクロの及びダウンサンプリングされたバージョンを用いて、顔検出、顔追跡、及び／又は顔認識を実行することができる。高レベル解析モジュール214は、さらに、複数の解析を並列に又は実質的に並列に実行してもよい。例えば、顔追跡及び顔認識が並列に実行されることがある。

10

【0054】

重要度算出モジュール218は、ビデオファイルのビデオフレームについて生成される様々な特徴スコアを正規化し、ビデオファイルのビデオ重要度値を算出することができる。例えば、重要度算出モジュール218は、各ビデオフレームの正規化された特徴スコアのセット（例えば、顔重要度スコア、モーション重要度スコア、露出格付けスコア、彩度スコア等）を平均して、各ビデオフレームのフレーム重要度スコアを取得することができる。ビデオフレーム重要度スコアがさらに平均されて、ビデオファイルのビデオ重要度値を導出してもよい。いくつかの実施形態において、ビデオファイルのビデオ重要度値の算出は、ビデオフレームに関連付けられた1つ以上の特徴スコアにバイアスをかけることをさらに含むことができる。例えば、重要度算出モジュール218は、ビデオフレーム内における顔の存在がビデオフレームの色相多様性スコアよりもより高い度合で該フレームのフレーム重要度スコアに影響するように、肯定的バイアスを適用するよう構成されることがある。

20

【0055】

別の例において、重要度算出モジュール218は、ビデオファイルのビデオ重要度値を下記のとおり生成することができる：

【0056】

【数3】

30

$$frame_score = w_1 * Face\ Importance + w_2 * F_2 + w_3 * F_3 + \dots w_n * F_n \quad (1)$$

$$video_score = \frac{\sum(frame_scores)}{number\ of\ frames} \quad (2)$$

これにおいて、 w_i は重みであり、 F_i は特徴である。重みは、特徴の重要度を指図することができる。例えば、視聴者が明るいビデオをより好み、 F_2 がこのプロパティに相關する特徴である場合、重要度算出モジュール218は、 w_2 に対して、他の特徴の重みよりもより高い値を割り当てるよう構成されることがある。このバイアスは、他のシナリオにおいて適用されてもよい。さらなる例において、視聴者が特定の人を示すビデオファイルを優先的に選択したい場合、この視聴者は、重要度算出モジュール218を構成して、上記特定の人の顔を示すフレームにバイアスをかけて、こうしたビデオフレームのビデオフレーム重要度スコアを上げることができる。様々な実施形態において、重要度算出モジュール218は、ビデオファイルの特徴値 { F_1, F_2, \dots, F_n } をデータストア226に記憶することができる。ビデオファイルの特徴値の記憶は、異なる特徴がバイアスをかけられることになるシナリオにおいて、ビデオファイルについての将来の二重の解析を除外することができる。

40

50

【 0 0 5 7 】

少なくとも 1 つの他実施形態において、重要度算出モジュール 218 は、ビデオフレーム内に示される特徴に対して否定的バイアスを適用するように構成されることがある。例えば、揺れの量に比例する否定的バイアスが重要度算出モジュール 218 によって実施されて、揺れに比例してビデオフレームのビデオフレーム重要度スコアを低めることができる。

【 0 0 5 8 】

ビデオ区分モジュール 220 は、重要度に基づいてビデオファイルを複数のビデオセクションに区分することができる。いくつかの実施形態において、ビデオ区分モジュール 220 は、ビデオファイルの継続時間よりもより短い継続時間 t を有するビデオセクションを見つけることができる。こうした実施形態において、ビデオ区分モジュール 220 はウィンドウマス (window-mass) を算出することができ、ウィンドウマスは、ビデオファイルのうち($t * \text{フレームレート}$)のビデオフレームを有するウィンドウ内のビデオフレームにおけるフレーム重要度スコアの総和である。上記ウィンドウマスは、シフトする仕方で、ビデオファイルのすべてのビデオフレームについて、連続的に算出されることができる。このことに応じて、ビデオ区分モジュール 220 は、最も高いウィンドウマスを有するビデオフレームを、 t 秒の長さの重要ビデオセクションのうちの中心として選択することができる。他の実施形態において、ビデオ区分モジュール 220 は、モーション解析モジュール 216 に依存して、モーションデータに基づいてビデオファイルをビデオセクションに区分してもよい。ビデオ区分モジュール 220 がビデオファイルをビデオセクションにいったん区分すると、重要度算出モジュール 218 は、各ビデオセクションのセクション重要度値を、ビデオファイル全体に関してと同様の仕方で生成することができる。換言すると、重要度算出モジュール 218 は、ビデオセクション内のビデオフレームの正規化されたフレーム重要度スコアに基づいてセクション重要度値を生成することができる。いくつかの例において、重要度算出モジュール 218 は、ビデオセクションのセクション重要度値の生成の間、1 つ以上の特徴スコアに対するバイアスかけをさらに適用することができる。

【 0 0 5 9 】

様々な実施形態において、低レベル解析モジュール 212 と高レベル解析モジュール 214 と重要度算出モジュール 218 との各々は、ビデオセクション及び / 又はビデオファイルについて取得されたスコア、値、及び他の情報を、関連メタデータとしてデータストア 226 に記憶することができる。上記メタデータは、ビデオファイルに関連付けられる他のメタデータ、例えば、日付、場所、オンライン共有の数等と組み合わせられることがある。

【 0 0 6 0 】

ビデオランク付けモジュール 222 は、ビデオファイルのビデオセクションを、そのセクション重要度値に基づいてランク付けすることができる。別法として、又は同時的に、ビデオランク付けモジュール 222 は、ビデオファイルをそのビデオ重要度値に従ってランク付けすることができる。ランク付けは、最も重要なものから最も重要でないものへであってもよく、あるいは、逆もまた同様である。各々のランク付けされたビデオファイルについて、ビデオランク付けモジュール 222 は、ランク付けされたビデオファイル内のビデオセクションに関するメタデータをさらに記憶することができる。上記メタデータには、各ビデオセクションのランク付け、各ビデオセクションの開始及び終了時間、各ビデオセクションの継続時間、及び各ビデオセクションのセクション重要度値が含まれ得る。いくつかの実施形態において、ビデオランク付けモジュール 222 は、ビデオファイル又はビデオセクションについて、さらなる値をさらに算出することができる。こうした値には、重要度密度が含まれてもよく、重要度密度は、重要度スコア閾値を超える重要度スコアを有するビデオファイル又はビデオセクション内のビデオフレームのパーセンテージを反映することができる。さらに、上記値には品質密度が含まれてもよく、品質密度は、対応する閾値を超える否定的又は肯定的特徴を有するビデオファイル又はビデオセクション

10

20

30

40

50

内のフレームのパーセンテージを反映することができる。上記否定的又は肯定的特徴には、揺れ、過剰露出、過少露出等が含まれ得る。ビデオランク付けモジュール 222 は、個々のビデオファイル及び／又はビデオセクションのランク付けを生成するのに使用される複数のタイプのスコアと他の計算された値とを、関連メタデータとしてデータストア 226 に記憶することができる。

【0061】

ビデオファイル又はビデオセクションのセットがいったんランク付けされると、ビデオランク付けモジュール 222 は、ランク付けされたビデオファイル又はビデオセクションのサムネイル画像表現を表示することができる。このことに応じて、サムネイル画像表現の選択が、メディアプレーヤ 122 に対応ビデオセクションを再生させ、あるいは、他のアプリケーションに共有及び又は編集のために対応ビデオセクションを提供させることができる。10

【0062】

ユーザインターフェースモジュール 224 は、ユーザがインターフェース 204 を介してビデオ解析エンジン 102 のモジュールと相互作用することを可能にする。例えば、ユーザインターフェースモジュール 224 は、ユーザが重要度解析のためのビデオファイルを選択し、ビデオファイル内で識別される人間の顔に情報でタグ付けし、肯定的特徴スコアバイアスかけのために人間の顔を強調し、及び／又はサムネイル画像表現を介してメディアプレーヤ 122 による再生のためのビデオファイル及びビデオセクションを選択することを可能にする。いくつかの実施形態において、ユーザは、ユーザインターフェース 224 をさらに使用して、ビデオ解析エンジン 102 による解析のためにビデオファイルの低レベル特徴のうち 1 つ以上又は高レベル特徴のうち 1 つ以上を選択することができる。20

【0063】

データストア 226 は、様々なモジュールにより使用されるデータを記憶することができる。少なくともいくつかの実施形態において、データストア 226 は、ビデオファイル 228、ランク付けされたビデオファイル 230、ランク付けされたビデオセクション 232、及び／又は、ランク付けされたビデオファイル 230 及びランク付けされたビデオセクション 232 に関連付けられたメタデータ 234 を記憶することができる。他の実施形態において、データストア 226 は、ビデオファイル及びビデオセクションをランク付けするのに使用されるビデオファイル又はビデオセクションに関連付けられたデータ（例えば、重要度スコア）を記憶することができる。データストア 226 は、モジュールにより生成されるさらなる帰結及び値、例えば、ホモグラフ変換マトリクス、特徴スコア、ビデオ重要度スコア、セクション重要度スコア等をさらに記憶することができる。30

【0064】

いくつかの実施形態において、1 つ以上のさらなるアプリケーションがコンピューティング装置 104 上にインストールすることができる。こうしたアプリケーションにはビデオ編集アプリケーションが含まれてもよく、ビデオ編集アプリケーションは、元のビデオファイルのうちの精選ビデオセクションから新しいビデオファイルを編成するのに使用される。例えば、上記ビデオ編集アプリケーションは、ユーザが、一緒にデジタル的に組み合わせられるべき特定のスコア閾値を超えるセクション重要度値を有するビデオセクションを選択して、ハイライトビデオファイルを作成することを可能にすることができます。アプリケーションにはオンライン共有アプリケーションがさらに含まれてもよく、オンライン共有アプリケーションは、ユーザがビデオファイル、ビデオセクション、又はハイライトビデオをオンラインで投稿することを可能にする。さらなる実施形態において、1 つ以上の他のアプリケーションがコンピューティング装置 104 上にインストールされて、ビデオファイル及びビデオセクションについてデータストア 226 に記憶されたデータにアプリケーションインターフェースを介してアクセスしてもよい。上記アプリケーションはデータにアクセスして、他の方法で解析結果を使用し得る。換言すると、ビデオ解析エンジン 102 は、より低いレベルのサービスとして機能して、データを上記アプリケーションに提供することができる。4050

【0065】**例示的な処理**

図4～図7は、ビデオファイルのルールに基づく重要度解析を実行する様々な例示的処理を説明する。各例示的処理において動作が記載される順序は、限定とみなされることは意図されず、説明される動作のうち任意数が、任意の順序で及び／又は並列に組み合わせられて各処理を実施してもよい。さらに、図4～図7の各々における動作は、ハードウェア、ソフトウェア、及びこれらの組み合わせにおいて実装され得る。ソフトウェアの文脈において、動作はコンピュータ実行可能命令を表し、該命令は、1つ以上のプロセッサにより実行されるときに、列挙された動作を1つ以上のプロセッサに実行させる。一般に、コンピュータ実行可能命令には、特定機能が実行され又は特定抽象データタイプが実装されることを引き起こすルーチン、プログラム、オブジェクト、コンポーネント、データ構造、及び同様のものが含まれる。

10

【0066】

図4は、ルールに基づくビデオ解析を使用してビデオファイル内の特徴を解析してビデオファイルのビデオセクションをその重要度に基づいてランク付けする一例示的処理400を例示するフロー図である。ブロック402において、ビデオ解析エンジン102は、ビデオファイル、例えばビデオファイル108(1)などをデコードして、ビデオフレームと該ビデオフレームに関連付けられたオーディオデータとを取得することができる。ビデオデコーダモジュール210は、様々なコーデックを使用してビデオファイルをデコードすることができ、例えば、H.264、MPEG-2、MPEG-4等である。上記関連オーディオデータは、ビデオフレームと同じ継続時間、例えば、1秒の1/30又は1秒の1/60を有し得る。しかしながら、いくつかの実施形態において、ビデオデコーダモジュール210は、関連オーディオデータを取得することなくビデオフレームを取得するように構成されてもよく、あるいは、逆もまた同様である。

20

【0067】

ブロック404において、ビデオ解析エンジン102は、ビデオフレームのビデオ特徴又はオーディオデータのオーディオ特徴のうち少なくとも1つを解析して、特徴スコアを取得することができる。ビデオ解析エンジン102は、上記解析を、1つ以上の局所的ルールに基づいて実行することができる。様々な実施形態において、解析される1つ以上の特徴には、高レベル特徴及び／又は低レベル特徴が含まれ得る。例えば、低レベル特徴には、露出品質、彩度品質、ビデオフレームの揺れ、平均明度、色エントロピー、及び／又はビデオフレーム間におけるヒストグラム差などの特徴が含まれ得る。高レベル特徴には、ビデオフレーム内で検出される顔の数量、位置、及び／又は顔面特徴などの特徴が含まれ得る。

30

【0068】

ブロック406において、ビデオ解析エンジン102は、ビデオフレームの特徴スコアをビデオフレームのメタデータとして記憶することができる。様々な実施形態において、ビデオ解析エンジン102は、メタデータをデータストア226に記憶することができる。こうしたメタデータは、ビデオフレームの解析の再発を、同じビデオフレームを伴う対応ビデオファイル又はビデオセクションの重要度の将来の決定の間、低減させ、あるいは除外することができる。

40

【0069】

判断ブロック408において、ビデオ解析エンジン102は、解析すべきビデオファイルのさらなるフレームが存在するかどうかを決定することができる。換言すると、ビデオ解析エンジン102は、ビデオファイルのビデオフレーム及び関連オーディオデータのすべてがデコードされたかどうかを決定することができる。ビデオ解析エンジン102が、解析すべきさらなるフレームが存在すると決定する場合(判断ブロック408において“yes”)、処理400はブロック402にループバックし得る。ブロック402において、ビデオ解析エンジン102は、後続ビデオフレーム及び関連データをビデオファイルから取得することができる。しかしながら、ビデオ解析エンジン102が判断ブロック4

50

08において、ビデオファイルのさらなるビデオフレームが解析されることにならないと決定する場合（判断ブロック408において“no”）、処理400はブロック410に進み得る。

【0070】

ブロック410において、ビデオ解析エンジン102は、少なくとも1つの大局的ルールを1つ以上の特徴結果に適用することができる。例えば、最も高いグループ重要度スコアを有するグループ内に顔が出現する人は、主要な人物と考えられることができる。こうした情報に基づいて、高レベル解析モジュール214は大局的ルールをビデオフレームに適用することができ、したがって、主要人物を示すビデオフレームの顔重要度スコアがさらに評価されることがある。

10

【0071】

ブロック412において、ビデオ解析エンジン102は、ビデオファイルの各ビデオフレームのすべての特徴スコアを組み合わせて、各ビデオフレームの対応フレーム重要度スコアを導出することができる。例えば、ビデオ解析エンジン102は、各ビデオフレームの正規化された特徴スコアのセットを平均して、各ビデオフレームのフレーム重要度スコアを取得することができる。

【0072】

ブロック414において、ビデオ解析エンジン102は、ビデオファイルのメタデータを記憶することができる。メタデータには、ビデオファイルのビデオフレームのビデオフレーム重要度スコア及び／又は各ビデオフレームの特徴スコアが含まれ得る。

20

【0073】

ブロック416において、ビデオ解析エンジン102は、ビデオフレームのビデオフレーム重要度スコアに基づいて、ビデオファイルをビデオセクションに区分することができる。いくつかの実施形態において、ビデオ解析エンジン102は、ウィンドウマスクの算出を使用して、ビデオファイルをビデオセクションに区分することができる。他の実施形態において、ビデオ解析エンジン102は、ビデオファイル内に捕捉されたモーションデータのズーム及び垂直並進モーションのゼロ交差を使用して、ビデオファイルをビデオセクションに区分することができる。

【0074】

ブロック418において、ビデオ解析エンジン102は、ビデオファイルの各ビデオセクションのセクション重要度値を算出することができる。様々な実施形態において、ビデオ解析エンジン102は、セクション重要度値を、ビデオセクション内のビデオフレームの正規化されたフレーム重要度スコアに基づいて生成することができる。いくつかの例において、重要度算出モジュール218は、ビデオセクションのセクション重要度値の生成の間、1つ以上の特徴スコアに対するバイアスかけをさらに適用することができる。

30

【0075】

ブロック420において、ビデオ解析エンジン102は、ビデオセクションをそのセクション重要度値に従ってランク付けすることができる。例えば、ビデオセクションは、最も高い重要度から最も低い重要度へランク付けされることができ、あるいは、逆もまた同様である。ビデオセクションがいったんランク付けされると、ビデオ解析エンジン102は、ビデオセクションのサムネイル画像表現を表示することができる。サムネイル画像表現の選択は、メディアプレーヤ122が対応ビデオセクションを再生すること、又は他のアプリケーションが対応ビデオセクションを共有及び／又は編集のために提供することを引き起こすことができる。別法として、ビデオ解析エンジン102は、ランク付けデータを別のアプリケーションに提供することができ、したがって、該アプリケーションは、ランク付けされたビデオセクションのサムネイル表現を表示することができる。

40

【0076】

図5は、ルールに基づくビデオ解析を使用してビデオファイルの特徴を解析してビデオファイルをその重要度に基づいてランク付けする一例示的な処理500を例示するフロー図である。ブロック502において、ビデオ解析エンジン102は、ビデオファイルの集

50

合、例えばビデオ集合 106 から、ビデオファイルを取得することができる。様々な実施形態において、ビデオ解析エンジン 102 は、ビデオファイルをユーザからの選択入に基づいて取得することができる。

【0077】

ブロック 504 において、ビデオ解析エンジン 102 は、ビデオファイルのビデオ重要度値を、ビデオファイル内のビデオフレームのビデオフレーム重要度スコアに基づいて算出することができる。ビデオ解析エンジン 102 は、図 4 に示される処理 400 のブロック 402～412 において説明されたとおり、ビデオファイルの各フレーム重要度スコアを計算することができる。様々な実施形態において、ビデオ解析エンジン 102 は、ビデオフレーム重要度スコアを平均して、ビデオファイルのビデオ重要度値を導出することができる。別の実施形態において、ビデオ解析エンジン 102 は、最も高い重要度を有するビデオファイルの 1 つ以上のビデオセクションのビデオフレーム重要度スコアを平均して、ビデオファイルの重要度値を導出するように構成されてもよい。いくつかの実施形態において、ビデオファイルのビデオ重要度値の算出は、フレームに関連付けられた 1 つ以上の重要度スコアにバイアスをかけることをさらに含むことができる。10

【0078】

判断ブロック 506 において、ビデオ解析エンジン 102 は、さらなるビデオファイルが解析されるべきであるかどうかを決定することができる。ビデオ解析エンジン 102 は、上記の決定を、ユーザから受信する選択入に基づいて行うことができる。ビデオ解析エンジン 102 が、解析されるべきさらなるビデオファイルが存在すると決定する場合（判断ブロック 506 において “yes”）、処理 500 はブロック 502 ヘループバックし得る。ブロック 502 において、ビデオ解析エンジン 102 は、さらなる解析のためにビデオファイルの集合から別のビデオファイルを取得することができる。20

【0079】

しかしながら、ビデオ解析エンジン 102 が、解析されるべきさらなるビデオファイルがないと決定する場合（判断ブロック 506 において “no”）、処理 500 はブロック 508 に進み得る。ブロック 508 において、ビデオ解析エンジン 102 は、解析されるビデオファイルを、対応するビデオ重要度値に基づいてランク付けすることができる。例えば、ビデオファイルは、最も高い重要度から最も低い重要度へランク付けされることができ、あるいは、逆もまた同様である。ビデオファイルがいったんランク付けされると、ビデオ解析エンジン 102 は、ビデオファイルのサムネイル画像表現を表示することができる。これにおいて、サムネイル画像表現の選択は、メディアプレーヤ 122 が対応ビデオフレームを再生すること、又は、他のアプリケーションが対応ビデオセクションを共有及び／又は編集のために提供することを引き起こすことができる。別法として、ビデオ解析エンジン 102 は、ランク付けデータを別のアプリケーションに提供することができ、したがって、該アプリケーションは、ランク付けされたビデオファイルのサムネイル表現を表示することができる。30

【0080】

いくつかの実施形態において、ビデオ解析エンジン 102 は、ビデオファイルのビデオセクション及び／又はビデオファイルを、ビデオフレームの高レベル特徴について取得される特徴スコアに基づいてランク付けするように最初試みることができる。こうした実施形態において、ビデオ解析エンジン 102 は、ビデオフレーム内の高レベル特徴の不十分な存在に起因して上記最初の試みが失敗するとき、ビデオフレームの高レベル特徴と低レベル特徴との双方についての特徴スコアを取得することに頼ってランク付けを作り出すことができる。40

【0081】

図 6 は、ビデオフレームの顔重要度スコアを計算する一例示的な処理 600 を例示するフロー図である。ブロック 602 において、高レベル解析モジュール 214 は、ビデオフレームに対して顔検出を実行して、1 つ以上の顔を検出することができる。検出された顔は、ビデオフレームを捕捉したカメラに対面しているか、又はカメラに対して横向きであ50

り得る。上記検出に基づいて、高レベル解析モジュール214は、検出された顔のリストを、ビデオフレーム内におけるその位置、各顔によりカバーされるビデオフレームのエリア、及び検出における信頼度を示す各顔の検出信頼度スコアと共に生成することができる。

【0082】

ブロック604において、高レベル解析モジュール214は、顔追跡を実行して、1つ以上の顔を追跡することができる。様々な実施形態において、高レベル解析モジュール214は、複数のビデオフレームにわたり人間の顔を追跡することができる。こうして、高レベル解析モジュール214は、ビデオファイル内に存在する顔のセットを確認し、さらに、各顔がビデオファイル内に出現する頻度を追跡することができる。

10

【0083】

ブロック606において、高レベル解析モジュール214は、顔面特性に基づくスコアバイアスがビデオフレームに割り当てられるべきかどうかを決定することができる。様々な実施形態において、上記決定は、1つ以上の顔面特性がビデオフレーム内で存在し又は欠如するかどうかに基づいて行われることができる。ゆえに、高レベル解析モジュール214が、顔面特性に基づくスコアバイアスが割り当てられるべきであると決定する場合（判断ブロック606において“yes”）、処理600はブロック608に進み得る。しかしながら、高レベル解析モジュール214が、顔面特性に基づくスコアバイアスが割り当てられるべきでないと決定する場合（判断ブロック606において“no”）、処理600はブロック610に直接進み得る。

20

【0084】

ブロック608において、高レベル解析モジュール214は、ビデオフレーム内の少なくとも1つの顔の顔面特性に基づいて、スコアバイアスをビデオフレームに割り当てることができる。顔面特性には、顔サイズ、顔の動き、及び／又は特定の顔面特徴の存在又は欠如（例えば、笑っているか否か、閉じられた／開かれた目等）が含まれ得る。例えば、最小サイズ閾値よりもより小さいか又は最大サイズ閾値よりもより大きいサイズであるビデオフレーム内の顔が、高レベル解析モジュール214による否定的スコアバイアスの割り当てを結果としてもたらし得る。別の例において、特定の人の顔を示すビデオフレームの重要度スコアが、該顔の、ビデオフレームの中心に対する、軸（例えば、x軸又はy軸）に沿った距離に従って、肯定的又は否定的にバイアスをかけられてもよく、したがって、中心に対してより近い顔がビデオフレームのより高い重要度を結果としてもたらし、逆も又は同様である。否定的スコアバイアスは、ビデオフレームの顔重要度スコアを減少させる重み因子であり得る。さらなる一例において、高レベル解析モジュール214は、笑っている及び／又は開いた目を有する各顔について、肯定的スコアバイアスを割り当ててもよい。

30

【0085】

ブロック610において、高レベル解析モジュール214は、少なくとも1つの顔に対して、顔認識又は顔グループ化のうち少なくとも1つを実行することができる。顔グループ化において、高レベル解析モジュール214は、追跡される顔をグループ化して、異なるビデオフレーム上で検出される顔が同じ人間に属するかどうかを決定することができる。顔認識において、高レベル解析モジュール214は、顔認識アルゴリズムを使用して、ビデオフレーム内で検出される各人の顔を人の既知のアイデンティティにマッチさせることができる。

40

【0086】

ブロック612において、高レベル解析モジュール214は、顔追跡データ又は顔認識データのうち少なくとも1つに基づいて、1以上の主要人物を識別することができる。例えば、各グループ内の顔の数は、ビデオファイル内における顔の行き渡りに正比例する。さらに、より多くの行き渡りは、その顔を有する人間のより高い重要度を示し、逆もまた同様である。このことに応じて、最多数の顔を有するグループに属する顔が、高レベル解析モジュール214によって、主要人物に属するとして決定されることがある。別の例

50

において、主要人物に属するとして指定された顔が、顔認識アルゴリズムによってビデオフレーム内に存在するとして検出されるとき、主要人物が識別されてもよい。

【0087】

ブロック614において、高レベル解析モジュール214は、ビデオフレーム内の主要人物の各存在について、ビデオフレームに肯定的スコアバイアスを割り当てることができる。肯定的特徴スコアは、ビデオフレームについて算出される顔重要度スコアを高めることができる。肯定的スコアバイアスは、ビデオフレームの顔重要度スコアを増加させる重み因子であり得る。

【0088】

ブロック616において、高レベル解析モジュール214は、ビデオフレームの顔重要度スコアを計算することができる。顔重要度スコアは、ビデオフレーム内の各顔のサイズ及び／又は動きに比例して算出されることができる。顔重要度スコアの計算は、肯定的及び／又は否定的スコアバイアスに基づいて算出されることができる。

【0089】

図7は、特徴点の動きを解析することによってビデオ内の重要セクションを決定する一例示的な処理700を例示するフロー図である。ブロック702において、モーション解析モジュール216は、ビデオファイル、例えばビデオファイル108(1)などの、ビデオフレームを取得することができる。ビデオデコーダモジュール210が、モーション解析モジュール216による解析のためにビデオファイルからビデオフレームを解析していくてもよい。

10

【0090】

判断ブロック704において、モーション解析モジュール216は、ビデオファイルの終わりが到達されたかどうかを決定することができる。モーション解析モジュール216が、ビデオファイルの終わりが到達されていないと決定する場合(判断ブロック704において“no”)、処理700はブロック706に続き得る。

20

【0091】

ブロック706において、モーション解析モジュール216は、ビデオフレーム内の特徴点を検出することができる。様々な実施形態において、モーション解析モジュール216は、ビデオフレームをダウンサンプリングし、より小さい次元のダウンサンプリングされた画像のピラミッドを作成することができる。それから、ダウンサンプリングされた画像がモーション解析モジュール216によって比較されて、ダウンサンプリングされた画像間における共通点(すなわち、特徴点)を決定することができる。

30

【0092】

判断ブロック708において、モーション解析モジュール216は、ビデオフレームがビデオファイルの最初のビデオフレームであるかどうかを決定することができる。このことに応じて、モーション解析モジュール216が、ビデオフレームが最初のビデオフレームであると決定する場合(判断ブロック708において“yes”)、処理700はブロック702にループバックし得る。ブロック702に戻ると、モーション解析モジュール216は、ビデオファイルの別のビデオフレームを取得することができる。しかしながら、モーション解析モジュール216が、ビデオファイルがビデオの最初のビデオフレームでないと決定する場合(判断ブロック708において“no”)、処理700はブロック710に進み得る。

40

【0093】

ブロック710において、モーション解析モジュール216は、ビデオフレーム内の特徴点を、ビデオファイルの前のビデオフレーム内の特徴点のさらなるセットにマッチさせることができる。様々な実施形態において、モーション解析モジュール216は、最大数の特徴点がマッチするように2つの隣接するフレームを整列させる変換を決定することによって、マッチングを実行することができる。少なくとも1つの実施形態において、上記変換は、ロバストパラメータ推定の一実装であるジオメトリックマッチングを用いて実行されることができる。

50

【 0 0 9 4 】

ブロック 712において、モーション解析モジュール216は、ビデオフレーム間における特徴点の動きを説明するホモグラフィ変換マトリクスを算出することができる。少なくともいくつかの実施形態において、モーション解析モジュール216は、ランダムサンプリング及びコンセンサス(RANSAC)アルゴリズムを使用して、ホモグラフィ変換マトリクスを取得することができる。

【 0 0 9 5 】

ブロック714において、モーション解析モジュール216は、ホモグラフ変換マトリクスから、ビデオフレームのモーションデータを計算することができる。様々な実施形態において、モーション解析モジュール216は、ホモグラフ変換マトリクスを用いて、ズーミング及び垂直並進成分の大きさ及び方向を抽出することができる。上記大きさ及び方向は、2つの隣接フレーム内における、記録したカメラの意図的動作を表す。その後、処理700はブロック702にループバックすることができ、したがって、モーション解析モジュール216は、処理のためにビデオファイルの別のビデオフレームを取得することができる。

10

【 0 0 9 6 】

判断ブロック704に戻ると、モーション解析モジュール216が、ビデオファイルの終わりが到達されたと決定する場合(判断ブロック704において“yes”)、処理700はブロック716に続き得る。ブロック716において、モーション解析モジュール216は、ビデオフレームのモーションデータの移動平均と、各ビデオフレームのフレーム動き揺れとを算出することができる。ビデオファイルの1つ以上のビデオフレームのモーションデータが、上記算出の前に組み合わせられることができる。様々な実施形態において、モーション解析モジュール216は、ビデオフレームの動きベクトルのモーションデータ移動平均を算出することによって、ビデオフレームを記録したカメラの揺れ動きを決定し、モーションデータ移動平均から上記カメラの意図的動きを減算することができる。モーションデータ移動平均の算出は、局所的分散を抑制し、意図的動きを表す長期トレンドを維持する。換言すると。

20

【 0 0 9 7 】

ブロック718において、モーション解析モジュール216は、ゼロ交差ゾーンを確認して(ascertain)、モーションデータの局所的なピーク及び谷情報、すなわち極大及び極小を見つけることができる。局所的ピーク及び谷情報は、ビデオファイルのシーン移行点を示し得る。

30

【 0 0 9 8 】

ブロック720において、モーション解析モジュール216は、局所的ピーク及び谷情報に基づいて、ビデオファイルをビデオセクションに区分することができる。さらに、モーション解析モジュール216は、各ビデオセクションについて、ズーム及びパンモーションの大きさを組み合わせ、平均することができる。ビデオセクションのズーム及びパンモーションの大きさの平均により表される加速の量がモーション解析モジュール216によって使用されて、ビデオセクションにモーション重要度スコアを割り当ててもよい。その後、モーション解析モジュール216は、最も高いモーション重要度スコアを有する1つ以上のビデオセクションをビデオファイルの重要セクションとして指定することができる。

40

【 0 0 9 9 】

本明細書に説明される手法の使用は、ユーザがビデオファイルをユーザに対する重要度に基づいてランク付けすることを可能にすることができる。こうしたランク付けに基づいて、ユーザは、どのビデオファイルを保持すべきか、及びどのビデオファイルを削除すべきかを決めることができる。いくつかの例において、ユーザは、ビデオファイルのランク付けをさらに使用して、特定のビデオファイルをオンライン共有ウェブサイトに投稿すべきかどうかを決定することができる。本明細書に説明される手法は、ビデオファイルの重要なセクションを表すサムネイル画像表現をさらに提示することができ、したがって、ユー

50

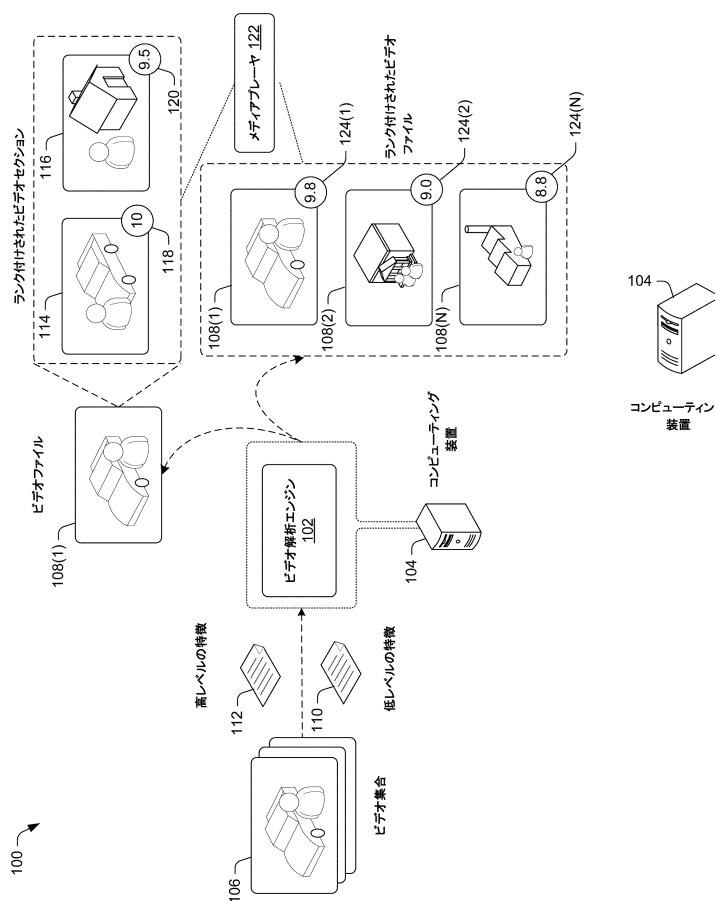
ザは、一見しただけで、ビデオファイルのうち興味深い部分を言うことができる。上記情報は、ビデオファイルを編集することにおいてユーザを支援して、ビデオファイルのコンテンツ品質を向上させ、あるいは個別セクションを強調することができる。

【0100】

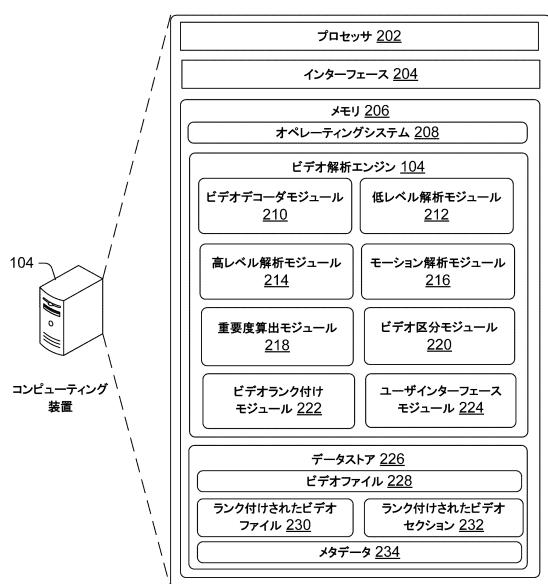
結び

最後に、様々な実施形態が構造的特徴及び／又は方法論的動作に固有の言語において説明されたが、別記の表現に定義される対象事項は必ずしも説明された特定の特徴又は動作に限定されないことが理解されるべきである。むしろ、上記特定の特徴及び動作は、請求される対象事項を実施する例示的形態として開示されている。

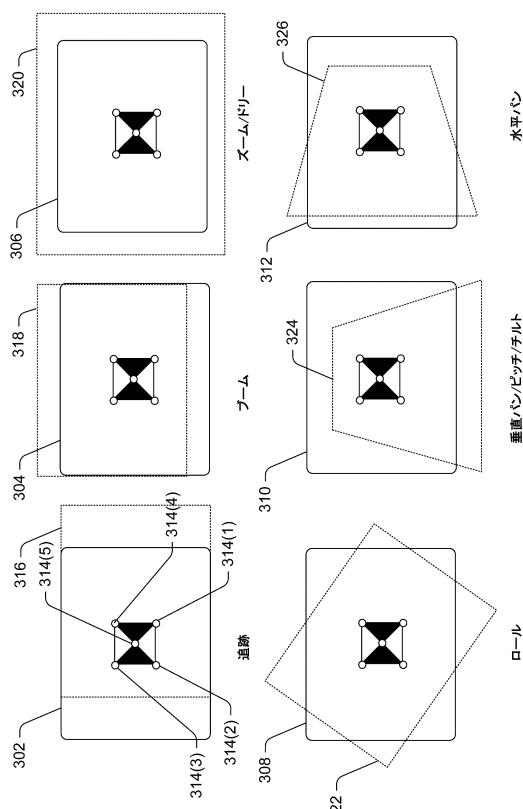
【図1】



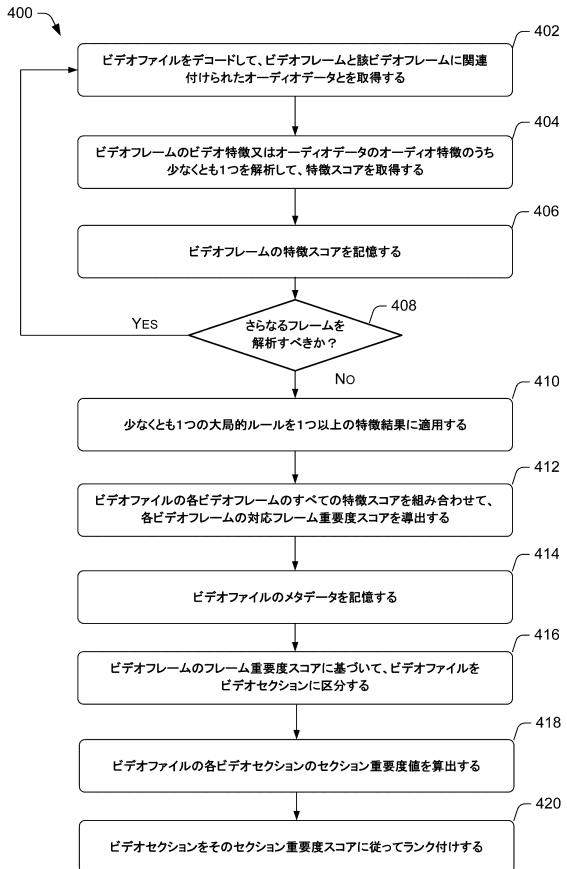
【図2】



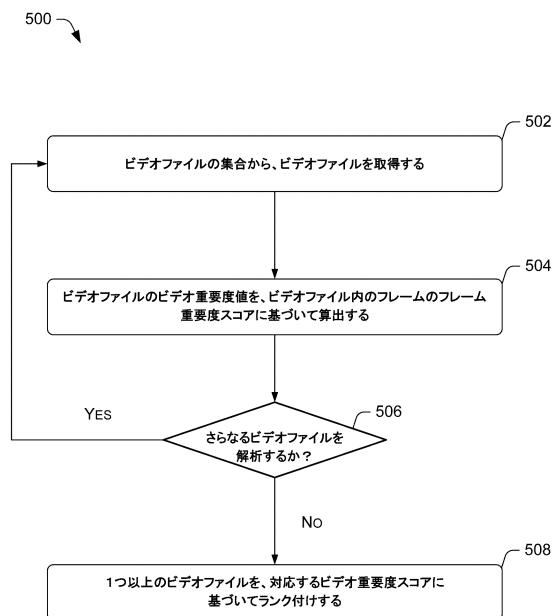
【図3】



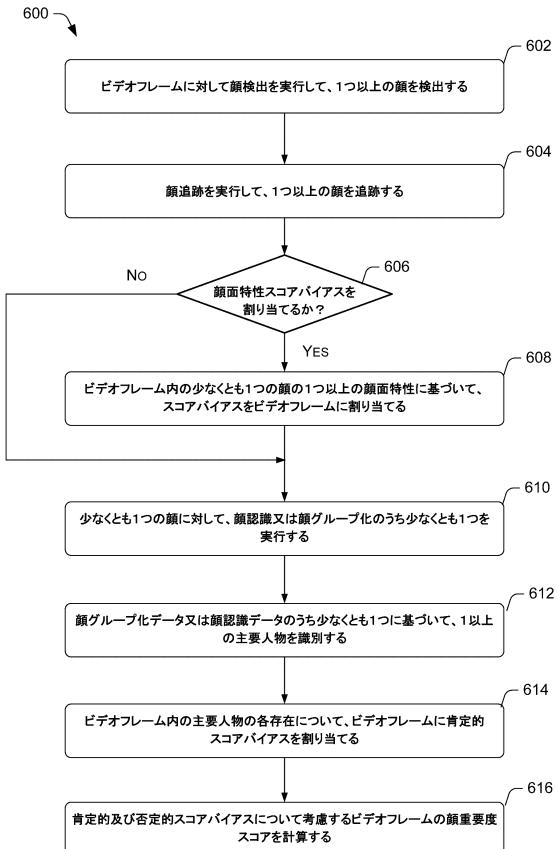
【図4】



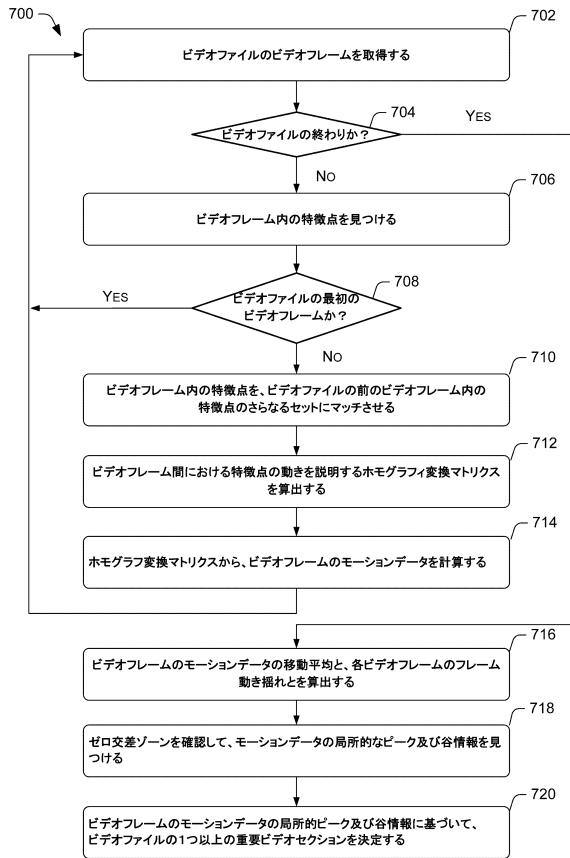
【図5】



【図6】



【図7】



フロントページの続き

(72)発明者 スリ , ニティン

アメリカ合衆国 98052-6399 ワシントン州 レッドモンド ワン マイクロソフト
ウェイ マイクロソフト テクノロジー ライセンシング, エルエルシー エルシーエー - インタ
ーナショナル パテンツ (8/1172) 内

(72)発明者 ワン , ツオンジー

アメリカ合衆国 98052-6399 ワシントン州 レッドモンド ワン マイクロソフト
ウェイ マイクロソフト テクノロジー ライセンシング, エルエルシー エルシーエー - インタ
ーナショナル パテンツ (8/1172) 内

(72)発明者 メヘンデル , オムカ-

アメリカ合衆国 98052-6399 ワシントン州 レッドモンド ワン マイクロソフト
ウェイ マイクロソフト テクノロジー ライセンシング, エルエルシー エルシーエー - インタ
ーナショナル パテンツ (8/1172) 内

(72)発明者 アイヴォリー , アンドリュー エス .

アメリカ合衆国 98052-6399 ワシントン州 レッドモンド ワン マイクロソフト
ウェイ マイクロソフト テクノロジー ライセンシング, エルエルシー エルシーエー - インタ
ーナショナル パテンツ (8/1172) 内

(72)発明者 スプロール , ウィリアム ディー .

アメリカ合衆国 98052-6399 ワシントン州 レッドモンド ワン マイクロソフト
ウェイ マイクロソフト テクノロジー ライセンシング, エルエルシー エルシーエー - インタ
ーナショナル パテンツ (8/1172) 内

審査官 西谷 憲人

(56)参考文献 特開2011-040860 (JP, A)

米国特許出願公開第2010/0070523 (US, A1)

米国特許出願公開第2014/0023348 (US, A1)

特開2010-109592 (JP, A)

特開2005-318548 (JP, A)

特表2007-519053 (JP, A)

国際公開第2013/186958 (WO, A1)

特開2010-062621 (JP, A)

米国特許出願公開第2008/0019661 (US, A1)

特開2010-103878 (JP, A)

米国特許出願公開第2011/0150275 (US, A1)

米国特許出願公開第2007/0169613 (US, A1)

米国特許出願公開第2013/0179112 (US, A1)

米国特許出願公開第2009/0115864 (US, A1)

(58)調査した分野(Int.Cl. , DB名)

H04N 5/14

G06T 7/246

H04N 5/76