

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2008-514136

(P2008-514136A)

(43) 公表日 平成20年5月1日(2008.5.1)

(51) Int.Cl.	F I	テーマコード (参考)
H04N 7/26 (2006.01)	H04N 7/13 Z	5C059
G06T 7/00 (2006.01)	G06T 7/00 300F	5L096

審査請求 未請求 予備審査請求 未請求 (全 23 頁)

(21) 出願番号 特願2007-532627 (P2007-532627) (86) (22) 出願日 平成17年9月20日 (2005.9.20) (85) 翻訳文提出日 平成19年5月18日 (2007.5.18) (86) 国際出願番号 PCT/US2005/033733 (87) 国際公開番号 W02006/034308 (87) 国際公開日 平成18年3月30日 (2006.3.30) (31) 優先権主張番号 60/611,878 (32) 優先日 平成16年9月21日 (2004.9.21) (33) 優先権主張国 米国 (US)	(71) 出願人 507030416 ユークリッド・ディスカバリーズ・エルエルシー EUCLID DISCOVERIES, LLC アメリカ合衆国, マサチューセッツ州 O 1742, コンコード, スイート 212, モニュメント スクエア 30 (74) 代理人 100087941 弁理士 杉本 修司 (74) 代理人 100086793 弁理士 野田 雅士 (74) 代理人 100112829 弁理士 堤 健郎
--	---

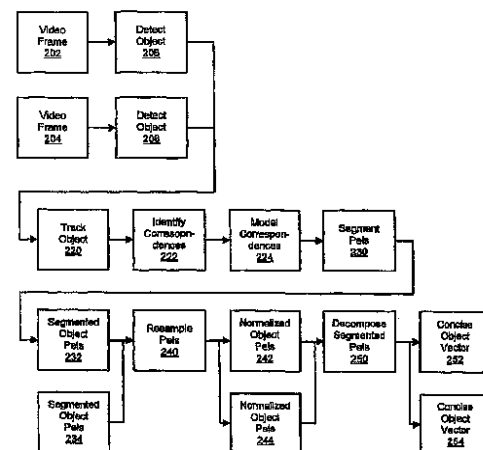
最終頁に続く

(54) 【発明の名称】 ビデオデータを処理する装置および方法

(57) 【要約】

ビデオデータを処理する装置および方法に関する。本発明はデータとデータの特定のパラメータ化のための近似モデルとの間の一致を評価するために使用できる、ビデオデータ表現を提供する。これにより、様々なパラメータ化技法の比較および特定データの連続的ビデオ処理の最適技法の選択を可能にする。この表現は膨大な処理の隙間を埋めるものとして、またはビデオデータを処理するフィードバック機構として、中間形式で利用できる。中間形式において利用される場合、本発明は、ビデオデータの記憶、強調、改良、特徴抽出、圧縮、コード化および伝送の処理において使用される。本発明は、頑健で有効な方法で顕著な情報を抽出するのに役立つと同時に、ビデオデータソースに一般的に係る問題に対処する。

【選択図】 図 2



【特許請求の範囲】**【請求項 1】**

複数のビデオフレームからビデオ信号データのコード化形式を生成するコンピュータ装置であって、

2 またはそれ以上のフレームの間でオブジェクトの対応する要素を識別する手段と、
これら対応する要素の相関関係をモデル化して、モデル化された相関関係を生成する手段と、

前記オブジェクトに関係する前記ビデオフレーム内の画素データを再サンプリングする手段であって、前記モデル化された相関関係を利用する再サンプリング手段と、

前記再サンプリングされた画素データの空間位置を復元する手段であって、前記モデル化された相関関係を利用する復元手段とを備え、

前記オブジェクトが1つ以上のオブジェクトであり、

前記再サンプリングされたデータがデータの間接形式であるコンピュータ装置。

【請求項 2】

請求項 1 において、前記オブジェクトが追跡方法によって追跡され、

前記ビデオフレームのシーケンスにおいてオブジェクトを検出するオブジェクト検出手段と、

前記ビデオフレームのシーケンスの2つまたはそれ以上のフレームを通して前記オブジェクトを追跡するオブジェクト追跡手段とを備え、

前記オブジェクト検出手段および前記オブジェクト追跡手段が、Viola / Jones 顔検出アルゴリズムを備えているコンピュータ装置。

【請求項 3】

請求項 1 において、前記オブジェクトがセグメント化方法を利用してビデオフレームからセグメント化され、

前記オブジェクトに関係する前記画素データを、前記ビデオフレームのシーケンスの他の画素データからセグメント化するセグメント化手段と、

前記復元した画素に関連するセグメント化データとともに組み立てて、オリジナルのビデオフレームを生成する生成手段とを備え、

前記セグメント化手段が時間積分を含むコンピュータ装置。

【請求項 4】

請求項 1 において、前記相関関係モデルが全体モデルに因子分解され、

前記相関関係の測定を全体動きのモデルに統合する統合手段を備え、

前記相関関係モデル化手段が、2次元アフィン型の動きモデルの解を求めるために、頑健なサンプリングコンセンサスを備え、

前記相関関係モデル化手段が、前記シーケンスの2つまたはそれ以上のビデオフレームのブロックを基礎とする動き検出から生成された有限差分に基づくサンプリング集団を備えたコンピュータ装置。

【請求項 5】

請求項 1 において、中間データがさらにコード化され、

前記正規化されたオブジェクト画素データを、コード化された表現に分解する分解手段と、

コード化された表現から、前記正規化されたオブジェクト画素データを再構成する再構成手段とを備え、

前記分解手段が主成分分析を含み、

前記再構成手段が主成分分析を含むコンピュータ装置。

【請求項 6】

請求項 1 において、

前記対応する要素の相関関係を識別するように、前記ビデオフレームのシーケンスの2つまたはそれ以上のフレームを通して多数のオブジェクトを追跡するオブジェクト追跡手段と、

10

20

30

40

50

前記画素データを再サンプリングするように、前記オブジェクトに関する前記画素データを、前記ビデオフレームのシーケンスの他の画素データからセグメント化するセグメント化手段と、

動きパラメータのコンパクトなパラメータ化を促進するように、動きパラメータを全体動きのモデルに統合する統合手段と、

前記正規化されセグメント化されたオブジェクト画素データを、コード化された表現に分解する分解手段と、

コード化された表現から、前記正規化されセグメント化されたオブジェクト画素データを再構成する再構成手段とを備えたコンピュータ装置。

【請求項 7】

請求項 5 において、フレームの非オブジェクト画素がオブジェクト画素と同一方法でモデル化され、

前記オブジェクトが他のオブジェクトが取り去られたときのフレームの残差の非オブジェクトであるコンピュータ装置。

【請求項 8】

請求項 5 において、前記セグメント化された画素および再サンプリングされた画素が、従来のビデオ圧縮および解凍処理に組み合わせられ、

前記再サンプリングされた画素を標準的なビデオデータとして従来のビデオ圧縮処理に供給する手段と、

モデルの相関関係データに対応するコード化済みビデオデータとともに記憶および伝送する手段とを備え、

前記分解手段および前記再構成手段が、前記従来のビデオ圧縮処理の圧縮効率の向上を可能にすることができる方法。

【発明の詳細な説明】

【関連出願】

【0001】

本出願は、2004年9月21日出願の米国特許仮出願第60/611,878号「主成分分析を利用するビデオ圧縮システムおよび方法 (System And Method For Video Compression Employing Principal Component Analysis)」の優先権を主張する。本出願は、2005年7月28日付の米国特許出願第11/191,562号の部分継続出願である。上記出願の全内容は参照により本明細書に引用したものとする。

【技術分野】

【0002】

本発明は一般にディジタル信号処理の分野、さらに詳細には、信号または画像データ、最も詳細には、ビデオデータの効果的な表現および処理のためのコンピュータ装置およびコンピュータによって実現される方法に関する。

【背景技術】

【0003】

本発明が属する従来技術の一般的なシステムの説明が図 1 に示されている。ここで、ブロックダイアグラムは典型的な従来技術のビデオ処理システムを示している。このようなシステムは一般に、以下のステージ、すなわち入力ステージ 102、処理ステージ 104、出力ステージ 106、および 1 つまたは複数のデータ記憶機構 108 を有する。

【0004】

入力ステージ 102 は、カメラセンサ、カメラセンサアレイ、距離計測 (range finding) センサといった構成要素、または記憶機構からデータを読み出す手段を有する。入力ステージは、人工および / または自然発生の現象の時間相関のあるシーケンスを表すビデオデータを提供する。データの顕著な要素はノイズまたは他の望ましくない信号によってマスクまたは汚染されている可能性がある。

【0005】

データのストリーム、アレイまたはパケットの形式のビデオデータが、予め定義された

10

20

30

40

50

転送プロトコルに従い、直接に、または中間記憶要素 108 を介して処理ステージ 104 に提供される。処理ステージ 104 は、専用のアナログもしくはデジタルデバイス、または中央処理装置 (CPU)、デジタル信号プロセッサ (DSP) もしくはフィールドプログラマブルゲートアレイ (FPGA) などのプログラマブル装置の形式をとり、所望の一連のビデオデータ処理操作を実行する。処理ステージ 104 は通常 1 つまたは複数の CODEC (符号器 / 復号器) を有する。

【0006】

出力ステージ 106 は、信号、表示、またはユーザもしくは外部装置に影響を与えることが可能な他の応答を生成する。一般に、出力装置が、インジケータ信号 (表示信号)、表示、ハードコピー、記憶装置において処理される処理データ表現を生成するために、または遠隔地にデータ伝送を開始するために用いられる。さらに、後続の処理操作における使用のために、中間信号または制御パラメータを提供するように用いられてもよい。

10

【0007】

記憶装置はこのシステムにおいて随意の要素として提示されている。用いられる場合、記憶要素 108 は、読取専用記憶媒体のような不揮発性、または動的ランダムアクセスメモリ (RAM) のような揮発性であってもよい。単一のビデオ処理システムが、入力ステージ、処理ステージおよび出力ステージに対して様々な関係を有する、複数タイプの記憶要素を有することは珍しいことではない。このような記憶要素の例は、入力バッファ、出力バッファおよび処理キャッシュである。

20

【0008】

図 1 のビデオ処理システムの主な目的は、入力データを処理して特定の用途に対して有意義な出力を生成することである。この目的を達成するために、ノイズ低減もしくは除去、特徴抽出、オブジェクトのセグメント化および / もしくは正規化、データのカテゴリ分類、イベントの検出、編集、データの選択、データの再コード化、ならびにトランスコード化などの処理操作が利用される。

【0009】

ほとんど制約されていないデータを生成する多くのデータソースは、人々、特に音響および視覚映像にとって重要な問題である。ほとんどの場合、これらのソース信号の基本的特性が、効率的なデータ処理目的に悪影響を与える。技術的仮定を導く際に用いられる単純な経験的および発見的方法から生じる誤りを持ち込むことなく信頼性の高い効率的な方法でデータを処理するには、ソースデータの本質的な多様性が障害となる。入力データが、狭く定義された特性セット (例えば、限定された記号値のセットまたは狭い帯域幅) に自然または故意に制限される場合、この多様性は用途に応じて軽減される。これらの全ての制約は、ほとんどの場合、商業的価値の低い処理技術をもたらす。

30

【0010】

信号処理システムの設計は、システムの意図する用途および入力として使用されるソース信号の期待される特性によって影響される。ほとんどの場合、要求される性能効率もまた重要な設計因子である。すなわち、性能効率は、利用可能なデータ記憶と比較した処理データ量、ならびに利用可能な計算能力と比較したアプリケーションの計算の複雑性によって影響される。

40

【0011】

従来のビデオ処理方法では、遅いデータ通信速度、大きな記憶容量条件、および妨害となる知覚呈示 (知覚を刺激するもの (perceptual artifact)) の形態で現れる、多くの非効率に苦しんでいる。ユーザが所望するビデオデータの使用および操作の方法は様々であり、また、特定の形式の知覚情報に対してユーザが生得的感受性を有するため、これらは重大な問題となる可能性がある。

【0012】

「最適な」ビデオ処理システムは、所望の一連の処理操作の実行において、効率的であり、信頼性が高く、頑健である。このような操作には、データの記憶、伝送、表示、圧縮、編集、暗号化、強調、カテゴリ分類、特徴検出および認識が含まれる。2 次的操作は

50

、他の情報源とこのように処理されたデータの統合を含む。このような処理システムの場合において同様に重要なのは、知覚呈示の混入を回避することによって、出力が人間の視覚に対応していなければならない。

【 0 0 1 3 】

ビデオ処理システムの速度、効率および品質が入力データのいずれかの特定の特性の特異性に強く依存しない場合、ビデオ処理は「頑健である (robust)」と説明される。頑健性はまた、入力の一つにエラーがあるときに操作を実行する能力に関係する。多くのビデオ処理システムは、アプリケーションの汎用的な集合への適用を可能にするだけの十分な頑健性を有しない。これらシステムは、システムの開発に使用された同一の狭く制約されたデータへの適用のみを提供する。

10

【 0 0 1 4 】

入力要素のサンプリングレートが検出現象の信号特性に適合しないことによって、顕著な情報が、連続値のデータソースの離散化において失われる可能性がある。また、信号強度がセンサの限界を超える場合も損失が発生し、飽和を生じさせる。同様に、入力データの全範囲の値が一連の離散値によって表現される場合、任意の量子化プロセスが実行されて入力データの精度が低下する際に情報が失われ、これによりデータの表現の精度が低下する。

【 0 0 1 5 】

集合体多様性は、データまたは情報ソースの集合におけるあらゆる不確定要素に触れる。視覚情報は一般に制限がないため、視覚情報のデータ表現は極めて大きい集合体多様性を有する。視覚データは、センサアレイ上に入射する光によって形成される、空間アレイシーケンスまたは時空間シーケンスを表現する。

20

【 0 0 1 6 】

視覚現象のモデル化において、ビデオプロセッサは、一般に、データの表現または読取りに、いくつかの一連の制限および/または構造を課す。この結果、このような方法は、出力の品質、出力に関する信頼性、およびデータに確実に実行される後続の処理タスクの種類に影響を与える、系統的誤り発生させる可能性がある。

【 0 0 1 7 】

量子化方法は、そのデータの統計的変動を保持することを試みる一方、ビデオフレーム内のデータ精度を低下させる。一般に、ビデオデータは、データ値の分布が確率分布に収集されるように解析される。また、データを空間周波数の混合データとして特徴付けるためにデータを位相空間に射影する方法もあり、これにより、精度の低下が拡散され、好ましい状態になる。これら量子化が集中的に利用されると、しばしば、知覚的に信じがたい色が発生させ、ビデオフレームの元の滑らかな領域に急峻な画像データを発生させる。

30

【 0 0 1 8 】

また、一般に、データの局所的な空間類似性を利用するために、差分コード化が用いられる。フレームの一部分におけるデータが、このフレームにおいて類似データの周辺で塊となっている傾向がある。また、後続のフレームにおいて同様の位置で塊となっている傾向がある。したがって、空間的に近接するデータに関するデータ表現は、量子化と組み合わせることができ、最終結果は、所定の精度に対して、差分表現がデータの絶対値を使用するよりも正確になる。例えば、白黒ビデオまたは低水準のカラービデオなどのように元のビデオデータのスペクトル分解能が制限される場合、この仮定はよく当てはまる。ビデオのスペクトル分解能が高くなると、同様の仮定が成り立たなくなる。これは、ビデオデータの精度を選択的に保護できないことに起因する。

40

【 0 0 1 9 】

残差のコード化は、表現誤差をさらに差分コード化して、元のデータの精度を所望のレベルの正確度に復元する点で、差分コード化に類似する。

【 0 0 2 0 】

これらの方法の多様性は、空間位相および空間スケールにおけるデータの相関関係を明らかにする別の表現にビデオデータを変換することを試みる。ビデオデータがこれらの方

50

法で変換されると、量子化および差分コード化の方法が変換データに適用されて、顕著な画像特徴の保持性の増加をもたらす。これら変換ビデオ圧縮技法のもっとも普及している2つの方法は、離散コサイン変換(DCT)および離散ウェーブレット変換(DWT)である。DCTにおける誤差はビデオデータ値の広範な変動として現れ、したがって、これらの誤った相関を局所化するために、DCTは一般にビデオデータのブロックにおいて使用される。この局所化による呈示(artifact)はブロックの境界に沿って現れることが多い。DWTについては、基底関数と特定のテクスチャの間に不一致が存在すると、より複雑な呈示が発生し、これにより「ぼけ」が引き起こされる。DCTおよびDWTの悪影響を打ち消すために、表現の精度を上げて、貴重な帯域幅を犠牲にして歪みを低減している。

10

【発明の開示】

【0021】

本発明は、コンピュータに実装されるビデオ処理方法であって、既存の最新のビデオ処理方法に、計算および解析の両方における利点を提供するビデオ処理方法である。本発明の方法の原理は、線形分解法、空間セグメント化方法および空間正規化の方法の統合である。ビデオデータを空間的に制約することによって、線形分解法の頑健性および適用性を大幅に向上する。これに加えて、データの空間セグメント化は、他の高変化データが解析対象データに空間的に近接する場合、誘引される非線形性を軽減する。

【0022】

詳細には、本発明は、信号データを効率的に処理して1つまたは複数の有益な表現を得る手段を提供する。本発明は、多くの一般に発生するデータセットの処理において有効であり、特に、ビデオおよび画像データの処理において有効である。本発明の方法は、データを解析し、このデータの1つまたは複数のコンパクトな表現を提供することによって、この処理およびコード化を容易にする。新しい、よりコンパクトなデータ表現のそれぞれは、ビデオデータのコード化、圧縮、伝送、解析、記憶および表示(これらに限定されない)を含む、多数の用途に対する、計算処理、伝送帯域幅および記憶容量の要件の低減を可能にする。本発明は、ビデオデータの顕著なコンポーネントを識別および抽出する方法を含み、データの処理および表現における優先順位付けを可能にする。信号のノイズおよび他の望ましくない部分は優先順位がより低いと識別され、したがって、さらなる処理を、より高い優先順位のビデオ信号の部分の解析および表現に集中させることができる。その結果、ビデオ信号は、以前に可能であったのよりも、よりコンパクトに表現される。正確度における損失は、知覚的に重要でないビデオ信号の部分に集中する。

20

30

【発明を実施するための最良の形態】

【0023】

ビデオ信号データでは、通常3次元シーンを2次元画像面へと射影として描く画像シーケンスに、ビデオフレームが組み立てられる。各フレームすなわち画像は、サンプル信号に対する画像センサ応答を表現する画素(pixel)から構成されている。多くの場合、サンプル信号は、2次元センサアレイによってサンプル化される特定の反射、屈折または放射された電磁エネルギーに相当する。連続的な逐次サンプリングによって、フレーム当たりの空間2次元と、ビデオシーケンスにおけるフレーム順序に対応する時間次元とを備える、時空間データストリームを得ることができる。

40

【0024】

本発明は、図2に示されているとおり、信号データを解析して顕著なコンポーネントを識別する。信号がビデオデータで構成されている場合、時空間ストリームの解析により、顔のような特定のオブジェクトであることが多い、顕著なコンポーネントを明らかにする。識別処理は、顕著なコンポーネントの存在および重要度を特定して、特定された顕著なコンポーネントのうちの最も重要な1つ以上を選択する。これは、ここで説明されている処理後または処理と同時に行為される、顕著性がより小さい他の顕著なコンポーネントの識別および処理を制限するわけではない。上述の顕著なコンポーネントはその後さらに解析され、変化するサブコンポーネントおよび不変のサブコンポーネントが識別される。不変

50

サブコンポーネントの識別は、コンポーネントの特定の外観のモデル化処理である、これにより、コンポーネントを所望の正確度に合成できるモデルのパラメータ化が明らかになる。

【 0 0 2 5 】

本発明の一実施形態では、前景のオブジェクトが検出され追跡される。オブジェクトの画素はビデオの各フレームから識別されてセグメント化される。ブロックに基づく動き検出（動きベクトル探索）が、複数フレーム中のセグメント化されたオブジェクトに適用される。次に、これらの動き検出が高次の動きモデルに統合される。動きモデルを利用して、オブジェクトのインスタンスを共通の空間構成に包み込む。この構成において、特定のデータに対しては、オブジェクトのより多くの特徴が一つにまとめられる。この正規化によって、複数フレームにわたってオブジェクトの画素値の線形分解が可能になり、コンパクトに表現される。オブジェクトの外観（アピアランス）に関する顕著な情報が、このコンパクトな表現に包含されている。

10

【 0 0 2 6 】

本発明の好ましい実施形態は、前景のビデオオブジェクトの線形分解を詳述する。オブジェクトは空間的に正規化され、これによりコンパクトな線形外観モデルを生成する。別の好ましい実施形態はさらに、空間的正規化の前に、ビデオフレームの背景から前景のオブジェクトをセグメント化する。

【 0 0 2 7 】

本発明の好ましい実施形態は、少しだけ動いてカメラに向かって話す人物のビデオに本発明を適用する。

20

【 0 0 2 8 】

本発明の好ましい実施形態は、空間的変換によって、鮮明に表現されるビデオ内の任意のオブジェクトに本発明を適用する。

【 0 0 2 9 】

本発明の好ましい実施形態は、特に、ブロックを基礎とした動き検出を用いて、ビデオの2または3以上のフレーム間での有限差分を求める。高次の動きモデルが、より効率的な線形分解を提供するために、有限差分から因子分解される。

【 0 0 3 0 】

< 検出および追跡 >

30

信号の構成要素の顕著なコンポーネントが決定されると、これらのコンポーネントは保持され、他のすべての信号コンポーネントは減少または除去される。顕著なコンポーネントの検出処理が図2に示されている。ビデオフレーム（202）が1つまたは複数のオブジェクト検出（206）処理によって処理され、1つまたは複数のオブジェクトが識別され、続いて追跡される。保持されるコンポーネントはビデオデータの中間的な形式を表現している。この中間データは、通常は既存のビデオ処理方法には利用されない方法を用いて、コード化される。中間データが複数の形式で存在するので、これらの中間的形式のいくつかをコード化するのに、標準のビデオコード化技術が用いられる。それぞれの例について、本発明は、最も効率的なコード化技法を決定して採用する。

【 0 0 3 1 】

40

好ましい一実施形態においては、顕著性の解析処理が、顕著な信号モードを検出して分類する。この処理の一実施形態は、強度がビデオフレーム内のオブジェクトの検出された顕著性に関係している応答信号を生成するように特に設計された空間フィルタの組み合わせを用いる。識別器が、ビデオフレームの異なる空間スケールで、異なる位置に、適用される。識別器からの応答強度が、顕著な信号モードの存在の可能性を示す。顕著性が強いオブジェクトが集中している場合、処理はそれを強い応答と識別する。顕著な信号モードの検出が、ビデオシーケンス内の顕著な情報に関する後続の処理および分析を可能にすることによって、本発明を特徴付ける。

【 0 0 3 2 】

1つまたは複数のビデオフレーム内の顕著な信号モードの検出位置が与えられると、本

50

発明は、顕著な信号モードの不変の特徴を分析する。さらに、本発明は、不変の特徴について、残差信号、すなわち「突出性の小さい」信号モードを解析する。不変の特徴の識別が、冗長な情報の低減および信号モードのセグメント化（すなわち分離）の基礎となる。

【 0 0 3 3 】

< 特徴点の追跡 >

本発明の一実施形態では、1つまたは複数のフレーム内の空間位置が、空間強度場勾配解析を通して決定される。これらの特徴は、「コーナー」として大まかに記載できる「線」の交点に対応する。このような実施形態はさらに、両方とも強いコーナーで相互に空間的に異なるコーナー（ここでは特徴点と称する）の組を選択する。さらに、オプティカルフロー推定の階層的な多重解像度を用いて、特徴点の時間的な並進変位を求めることができる。

10

【 0 0 3 4 】

図2において、オブジェクト追跡（220）処理が、オブジェクト検出処理（206および208）からの検出インスタンスを集め、さらに複数のビデオフレーム（202および204）にわたって検出された1つまたは複数のオブジェクトの特徴点の相関関係を識別する（222）。

【 0 0 3 5 】

特徴追跡の限定されない実施形態を利用することによって、ブロックを基礎とした動き検出などのより一般的な勾配解析を修正するために、特徴点を用いることができる。

【 0 0 3 6 】

20

別の実施形態は、特徴点追跡を基礎とする動き検出の予測を前もって処理する。

【 0 0 3 7 】

< オブジェクトを基礎とする検出および追跡 >

本発明の限定されない一実施形態では、頑健なオブジェクト識別器を用いてビデオフレーム内の顔を追跡する。このような識別器は顔に向けられた方位エッジに対するカスケード応答を基礎とする。この識別器では、エッジは一連の基本的なHaar特徴として定義され、これら特徴の回転は45°ごとである。カスケード識別器はAdaBoostアルゴリズムの変形形態である。さらに、応答計算は、エリア総和テーブルを使用して最適化される。

【 0 0 3 8 】

30

< 局所的位置合わせ >

位置合わせは、2つまたはこれ以上のビデオフレーム内で識別されたオブジェクトの要素間の相関関係の指定を伴う。これらの相関関係は、ビデオデータ内の時間的に別個の点におけるビデオデータ間の空間関係のモデル化の基礎となる。

【 0 0 3 9 】

特定の実施形態を説明し、これら実施形態に関連する、公知のアルゴリズムおよびこれらのアルゴリズムの発明派生物に関して、実行作業の低減を説明するために、本発明では、位置合わせの様々な限定されない手段が記述される。

【 0 0 4 0 】

時空シーケンス内の明白なオプティカルフローのモデル化の1つの手段は、2つまたはそれ以上のビデオデータのフレームから有限の場の生成を通して達成される。相関関係が空間および強度感知の両方において特定の一定の制約条件に適合する場合、オプティカルフロー場をわずかに予測できる。

40

【 0 0 4 1 】

図3に示すように、フレーム（302または304）が、おそらくは分解処理（306）または他の何らかのサブサンプル化処理（例えば、低域通過フィルタ）によって空間的にサブサンプル化される。これらの空間的に低減された画像（310および312）は、同様にさらにサブサンプル化されることができる。

【 0 0 4 2 】

< ダイヤモンド探索 >

50

ビデオフレームを重複しないようにブロックに分割すると仮定して、各ブロックに対する一致について前のビデオフレームを探索する。全域探索ブロックベース (FSBB) の動き検出によって、現在のフレーム内のブロックと比較する際、前のビデオフレーム内の最小誤差を有する位置を見出す。FSBBの実行は計算的に極めて負荷が大きく、多くの場合、局所的動きの仮定に基づいた他の動き検出方式に比べて優れた一致を得るとは限らない。ダイヤモンド探索ブロックベース (DSBB) の勾配降下動き検出は、各種サイズのダイヤモンド形状の探索パターンを用いて、ブロックの最高一致の方向に誤差勾配を反復的に移動するものであって、FSBBに対する一般的な代替方法である。

【0043】

本発明の一実施形態では、DSBBは、1つまたは複数のビデオフレームの間の画像勾配場の解析に用いられ、その値が後に高次動きモデルに因子分解される有限差分を生成する。

10

【0044】

ブロックに基づく動き検出が規則的メッシュの頂点の解析の同等物と捉えることができることを、当業者は認識するであろう。

【0045】

< 位相を基礎とする動き検出 >

従来技術では、ブロックを基礎とする動き検出は一般に、1つまたは複数の空間一致をもたらす空間探索として実現されていた。位相を基礎とする正規化相互相関 (PNCC) は、図3に示されているとおり、現在のフレームおよび前のフレームからのブロックを「位相空間」に変換し、これらの2つのブロックの相互相関を見出す。相互相関は、値の位置が2つのブロック間のエッジの「位相シフト」に対応する、値の場として表される。これらの位置はしきい値化によって分離され、その後、空間座標に変換して戻される。空間座標は別個のエッジ変位であり、動きベクトルに対応する。

20

【0046】

PNCCの利点は、ビデオストリームにおける利得 / 露出調整の許容差を可能にするコントラストマスキングを含むことである。また、PNCCは、空間を基礎とした動き検出量から多くの反復値を取得する単一ステップからの結果を可能にする。さらに、動き検出はサブピクセル精度である。

【0047】

本発明の一実施形態では、1つまたは複数のビデオフレーム間の画像勾配場の解析にPNCCを利用することにより、その値が後に高次の動きモデルに因子分解される有限差分を生成する。

30

【0048】

< 全体位置合わせ >

一実施形態では、本発明は、有限差分予測値の場から1つまたは複数の線形モデルを因子分解する。このようなサンプリングが発生する場を、ここでは、有限差分の母集団と称する。ここで記載される方法はRANSACアルゴリズムの予測値と同様な頑健な予測値を用いる。

【0049】

図4に示すように、全体の動きモデル化の場合においては、有限差分は母集団プール (404) に集められる並進動き検出 (402) である。この母集団プールは、これら動き検出のランダムサンプリング (410) および線形モデルのこれらサンプルからの因子分解 (420) によって、反復的に処理される。次いで、この結果を用いて、ランダム処理を通じて見出されたとおり、モデルに対する異常値を除外して、線形モデルをよりよく明確化するために集団 (404) を調節する。

40

【0050】

線形モデル予測アルゴリズムの一実施形態では、動きモデルの推定量は線形の最小2乗解に基づいている。この依存性により、推定量は異常値データによって狂わされてしまう。RANSACに基づいて、ここで開示されている方法は、データの部分集合の反復的な

50

予測によって異常値の影響に対抗し、データの重要な部分集合を記述する動きモデルを探求する、頑健な方法である。各探求によって生成されるモデルは、モデルを表すデータのパーセンテージについて試験される。十分な数の反復が行われている場合、モデルはデータの最大の部分集合に適合すると見なされる。

【 0 0 5 1 】

図 4 において考察され示されているとおり、本発明は有限差分の初期サンプリング（サンプル）および線形モデルの最小 2 乗予測を含む代替アルゴリズムの形態で、R A N S A C アルゴリズムを超える革新的方法を開示する。総合誤差が、解明された線形モデルを用いて母集団内の全サンプルに対して査定される。サンプルの残差が事前設定のしきい値に一致するサンプル数に基づいて、線形モデルにランク（順位）が割り当てられ、このランクは「候補コンセンサス」とみなされる。

10

【 0 0 5 2 】

最終基準が満たされるまで、初期サンプリング、解明およびランク付けが反復的に実行される。基準が満たされると、最大ランクの線形モデルが集団の最終コンセンサスとみなされる。

【 0 0 5 3 】

随意の改良工程では、候補モデルに最適適合する順に、サンプルの部分集合を反復的に解析し、1 つまたは複数のサンプルの追加が部分集合全体に対する残差誤差のしきい値を超えるまで部分集合サイズを増加する。

【 0 0 5 4 】

図 4 に示すように、全体のモデル予測処理（4 5 0）が、コンセンサスランク許容テスト（4 5 2）が満足されるまで繰り返される。ランクが達成されていない場合、線形モデルを明らかにするように努めて、有限差分の母集団（4 0 4）が、発見されたモデルに対して選別される。最適（最高ランク）の動きモデルが、処理 4 6 0 において解セットに加えられる。次いで、処理 4 7 0 においてモデルが再予測される。完了すると、母集団（4 0 4）が再選別される。

20

【 0 0 5 5 】

本発明の開示された限定されない実施形態はさらに、有限差分ベクトルの場として上述したベクトル空間のサンプリングの一般方法として一般化され、これにより、特定の線形モデルに対応する別のパラメータベクトル空間における部分空間の多様体を求めることができる。

30

【 0 0 5 6 】

全体位置合わせ処理の別の結果は、この処理と局所的な位置合わせ処理との間の差が局所的な位置合わせの残差を生じることである。この残差は局所モデルへの近似における全体モデルの誤差である。

【 0 0 5 7 】

< 正規化 >

正規化は、標準または共通の空間構成のために空間強度場を再サンプリングすることを意味する。これら相対的な空間構成がこのような構成の間の可逆空間変換である場合、画素の再サンプリングおよびこれに伴う補間もまた位相限界まで可逆性を有する。本発明の正規化方法は図 5 に示されている。

40

【 0 0 5 8 】

3 つ以上の空間強度場が正規化されるとき、中間の正規化の計算を保存することによって、計算効率が向上する。

【 0 0 5 9 】

位置合わせの目的、すなわち正規化のために画像を再サンプリングするのに使用される空間変換モデルは、全体モデルおよび局所モデルを含む。全体モデルは、並進から射影への増加する整列（increasing order）である。局所モデルは、基本的にはブロックによって、またはより複雑には区分線形メッシュによって決定される、隣接画素の補間を暗示する有限差分である。

50

【 0 0 6 0 】

オリジナルの強度場の正規化強度場へと補間によって、強度場の部分集合に基づいた P C A 外観モデルの線形性が向上する。

【 0 0 6 1 】

図 2 に示すように、オブジェクトの画素 (2 3 2 および 2 3 4) を再サンプリングして (2 4 0)、正規化されたバージョンのオブジェクトの画素 (2 4 2 および 2 4 4) を得ることができる。

【 0 0 6 2 】

< 3 次元の正規化 >

本発明の別の実施形態は、特徴点を三角形メッシュにテセレーション (tessellate (モザイク模様にする)) し、メッシュの頂点が追跡され、各三角形の頂点の相対位置が、これら 3 つの頂点に一致する平面に対する 3 次元の表面法線の予測に使用される。表面法線がカメラの射影軸に一致する場合、画像化された画素が、三角形に対応するオブジェクトのひずみが最も少ないレンダリングをもたらすことができる。表面法線に直交する傾向の正規化画像を生成することで、後続の外観ベースの P C A モデルの線形性を向上させる中間データタイプを保存する画素を生成することができる。

【 0 0 6 3 】

他の実施の形態は、全体の動きモデルを暗示的にモデル化するために、従来からのブロックを基礎とした動き検出を使用する。限定されない一実施形態では、この方法が、従来のブロックを基礎とする動き検出 / 予測によって記述される動きベクトルから全体のアフィン動きモデルを因子分解する。

【 0 0 6 4 】

< 段階的な幾何正規化 >

メッシュのエッジに不連続が位置したときに、暗示的に不連続をモデル化するために、テセレーションされたメッシュを整列させるように空間不連続の分類が使用される。

【 0 0 6 5 】

均質な領域の境界が、多角形輪郭によって近似される。輪郭は、オリジナルの頂点のそれぞれの優先度を求めるために連続的に近似される。共有される頂点に対する頂点優先度を保存するために、頂点の優先度が領域に渡って伝搬される。

【 0 0 6 6 】

画像の位置合わせが、強力な画像勾配で、高い優先度の頂点に偏る。得られる変形モデルは、画像化されたオブジェクトの形状に関連して空間不連続を保持する傾向にある。

【 0 0 6 7 】

< セグメント化 >

詳述するセグメント化処理によって識別される空間不連続性は、不連続のそれぞれの境界の幾何パラメータ化によって効率的にコード化され、これらを空間不連続モデルと称する。これらの空間不連続モデルは段階的方法でコード化され、コード化の部分集合に対応するより簡潔な境界記述を可能にする。段階的なコード化は、空間不連続の顕著な様相の大部分を保持しながら、空間形状の優先順位付けの頑健な方法を提供する。

【 0 0 6 8 】

本発明の好ましい実施形態は多重解像度のセグメント化解析と空間強度場の勾配解析とを組み合わせ、頑健なセグメント化を達成するためにさらに時間安定性の制約を用いる。

【 0 0 6 9 】

図 2 において、オブジェクトの特徴の相関関係が経時的に追跡 (2 2 0) されてモデル化 (2 2 4) されると、この動き / 変形モデルへの適合を、オブジェクトに対応する画素をセグメント化 (2 3 0) するために使用することができる。この処理を、ビデオ (2 0 2 および 2 0 4) において検出された多数のオブジェクト (2 0 6 および 2 0 8) について繰り返すことができる。この処理の結果は、セグメント化されたオブジェクト画素 (2 3 2) である。

10

20

30

40

50

【 0 0 7 0 】

本発明により利用される不変の特徴の解析の一形態は、空間不連続の識別に焦点が当てられる。これらの不連続は、エッジ、影、遮蔽、線、コーナー、または、1つもしくは複数のビデオの画像フレーム内の画素間の急激で識別可能な分離を発生させる、その他の可視特徴物として現れる。さらに、ビデオフレーム内のオブジェクトの画素が、互いに異なる動きではなく、オブジェクト自体に対してコヒーレントな動きを受ける場合にのみ、類似の色および/またはテクスチャのオブジェクト間の微細な空間不連続が現れる。本発明は、空間、テクスチャ（強度勾配）および動きのセグメント化の組み合わせを利用して、顕著な信号モードに関連する空間不連続を頑健に識別する。

【 0 0 7 1 】

< 時間セグメント化 >

並進動きベクトルの時間積分、すなわち高次動きモデルへの、空間強度場における有限差分の測定は、従来技術で説明されている動きセグメント化の形式である。

【 0 0 7 2 】

本発明の一実施形態では、動きベクトルの高密度場を生成して、ビデオ内のオブジェクト動きの有限差分を表現する。これらの導関数は、タイルの規則的分割によって、または空間セグメント化などの特定の初期化手順によって、空間的に一体にグループ化される。各グループの「導関数」は、線形の最小2乗推定量を使用して高次の動きモデルに統合される。次に、得られた動きモデルが、k平均クラスタ化技法を用いて動きモデル空間内のベクトルとしてクラスタ化される。導関数は、いずれのクラスタがそれら導関数に最も適合するかに基づいて分類される。次に、クラスタラベルが、空間分割の発現として空間的にクラスタ化される。この処理は空間分割が安定するまで続けられる。

【 0 0 7 3 】

本発明の別の実施形態では、所定の開口の動きベクトルが開口に対応する一連の画素位置に対して補間される。この補間によって定義されたブロックが、オブジェクト境界に対応する画素間を橋渡しする場合、得られる分類はブロックの特定の特異な対角領域分割（anomalous diagonal partitioning）である。

【 0 0 7 4 】

従来技術では、導関数を統合するのに使用される最小2乗推定量は異常値に極めて敏感である。この敏感さのために、反復が広範囲に発散する点に動きモデルのクラスタ化方法を大きく偏らせる動きモデルが生成される。

【 0 0 7 5 】

本発明においては、動きセグメント化の方法は、2つまたはそれ以上のビデオフレーム全体にわたる明らかな（目に見える）画素動きの解析によって空間不連続を識別する。この明らかな動きは、ビデオフレーム全体にわたる一貫性に関して解析され、パラメータの動きモデルに組み込まれる。このような一貫した動きに関連する空間不連続が識別される。時間変化は動きによって発生するため、動きセグメント化はまた、時間セグメント化と称することができる。しかし、時間変化はまた、局所変形、明るさの変化などといった、いくつかの他の現象によって引き起こされる可能性もある。

【 0 0 7 6 】

ここで説明された方法によって、正規化方法に対応する顕著な信号モードは、複数の背景差分法のうちの1つによって、周囲信号モード（背景または非オブジェクト）から識別され、かつ分離される。これらの方法は、各時刻において最小量の変化しか示さない画素として背景を統計的にモデル化する。変化は画素値の差として特徴付けられる。代わりに、動きセグメント化が、顕著な画像モードの検出された位置およびスケールを与えることによって達成される。距離変換を用いて、検出された位置からの各画素の距離を求めることができる。最大距離に対応する画素値が保持されている場合、背景の妥当なモデルを解明できる。言い換えれば、周囲信号は、信号差の測定を利用して時間的に再サンプル化される。

【 0 0 7 7 】

周囲信号のモデルを前提として、各時刻における完全な顕著信号モードの差を計算できる。これらの差のそれぞれを再サンプル化して、空間的な正規化信号差（絶対差）を得ることができる。次に、これらの差は相互に整列されて積算される。これらの差は顕著信号モードに対して空間的に正規化されているため、差のピークが、顕著信号モードの画素の位置にほぼ対応する。

【0078】

<非オブジェクトの解像度>

解明された背景画像を前提として、この画像と現在のフレームとの間の誤差は、空間的に正規化されて時間的に積算される。このような解明された背景画像は、「背景解像度」の章で説明される。

【0079】

次に、得られた積算誤差がしきい値処理され、初期の輪郭がもたらされる。輪郭は、その後、誤差の残差を輪郭変形に対してバランスさせるために空間的に伝搬される。

【0080】

<勾配セグメント化>

テクスチャのセグメント化方法、すなわち強度勾配セグメント化は、1つまたは複数のビデオフレーム内の画素の局所勾配を解析する。勾配応答はビデオフレーム内の画素位置近傍の空間不連続を特徴付ける、統計的測定値である。次に、いくつかの空間的クラスタ化方法のうちの1つを用いて、勾配応答同士を組み合わせることで空間領域を生成する。これらの領域の境界が、1つまたは複数のビデオフレームにおける空間不連続の識別に有効である。

【0081】

本発明の一実施形態では、コンピュータグラフィックスのテクスチャ生成からのエリア総和テーブルの概念が、強度場の勾配の計算を促進するために用いられる。連続的に合計された値の場が生成して、4つの追加操作に組み合わせられる4つの参照を通してオリジナルの場の任意の長方形の加算を促進する。

【0082】

別の実施形態は、画像に対して生成されるHarris応答を用いて、各画素の近辺が均質なエッジまたは均質なコーナーのいずれかとして識別される。応答値はこの情報から生成され、フレーム内の各要素についてエッジまたはコーナーの度合いを示す。

【0083】

<多重スケール勾配解析>

本発明の実施形態は、さらに、いくつかの空間スケールを通して画像勾配値を生成することによって、画像勾配サポートを制約する。この方法は画像勾配の修正に役立つことができる。例えば、異なるスケールにおける空間不連続が互いをサポートするために使用される。「エッジ」がいくつかの異なる空間スケールにおいて区別される限り、そのエッジは「顕著」である必要がある。より修正された画像勾配は、より顕著な特徴に相関する傾向にある。

【0084】

好ましい実施の形態においては、テクスチャの応答場が最初に生成されて、次にこの場の値が、k平均ビンニング/分割に基づいていくつかのビン(bin)に量子化される。次に、単一反復がwatershedセグメント化を適用できる値の間隔として各ビンを用いて、オリジナルの画像勾配値が段階的に処理される。このような手法の利点は、強力な空間バイアスで、均一性が相対的な意味で定義される点にある。

【0085】

<スペクトルセグメント化>

スペクトルセグメント化の方法は、ビデオ信号における白黒、グレースケールまたは色の画素の統計的確率分布を解析する。スペクトル識別器は、これらの画素の確率分布に対してクラスタ化操作を実行することにより構成される。次に、識別器を用いて、1つまたは複数の画素を確率クラスに属するとして分類する。次に、結果として得られた確率クラ

10

20

30

40

50

スおよびその画素はクラスラベルを与えられる。その後、これらのクラスラベルは明確な境界を有する画素の領域に空間的に関連付けられる。これらの境界は1つまたは複数のビデオフレーム内の空間的不連続を識別する。

【0086】

本発明は、空間分類に基づく空間セグメント化を利用して、ビデオフレーム内の画素をセグメント化する。さらに、領域間の対応は、空間領域と前のセグメント化における領域との重ね合わせに基づいて求められる。

【0087】

ビデオフレーム内のオブジェクトに対応するより大きな領域に空間的に結合される連続的な色領域からビデオフレームが大まかに構成される場合、色付けされた（またはスペクトルの）領域の識別および追跡によって、ビデオシーケンス内のオブジェクトの後続のセグメント化が容易になることが観測されている。

10

【0088】

< 背景セグメント化 >

ここで説明される発明はビデオフレーム背景のモデル化方法であって、検出されたオブジェクトと各ビデオフレーム内の各画素の間の空間距離測定の時間的最大に基づくモデル化方法を含む。オブジェクトの検出された位置を前提として、距離変換が適用され、フレーム内の各画素についてスカラー距離が生成される。各画素についてビデオフレームのすべてにわたる最大距離のマップが保持される。最大値が最初に割り当てられるか、あるいはその後新規かつ異なる値で更新されると、このビデオフレームの対応する画素が、「解明された背景」フレームに保持される。

20

【0089】

< 外観モデル化 >

ビデオ処理の共通の目的は、多くの場合、ビデオフレームのシーケンスの外観をモデル化して保存することである。本発明は、前処理を利用して、頑健で広範囲に利用可能な方法で適用される制限された外観のモデル化方法を実現することを目的としている。前述の位置合わせ、セグメント化および正規化は明らかにこの目的のためのものである。

【0090】

本発明は、外観変化（appearance variance）モデル化の手段を開示している。外観変化モデル化の主要な基本は、線形モデルの場合には、線形相関を利用するコンパクトな基準を示す特徴ベクトルの解析である。空間強度場の画素を表現する特徴ベクトルは外観変化モデルに組み込まれることができる。

30

【0091】

別の実施形態においては、外観変化モデルは画素のセグメント化された部分集合から計算される。さらに、特徴ベクトルは、空間的に重複しない特徴ベクトルに分離される。このような空間分解は空間的タイル法を用いて達成される。計算効率は、より全体的なPCA方法の次元数減少を犠牲にすることなく、これらの時間集合の処理によって達成される。

【0092】

外観分散モデルを生成すると、空間強度場の正規化を用いて、空間変換のPCAモデル化を現象することができる。

40

【0093】

< PCA >

外観変化モデルを生成する好ましい手段は、ビデオフレームをパターンベクトルと組み合わせる訓練マトリックスとするか、または訓練マトリックスに主成分分析（PCA）を組み合わせるかもしくは適用することである。このような展開が打ち切られると、結果として得られるPCA変換マトリックスは、ビデオの後続のフレームの解析および合成に用いられる。打ち切りレベルに基づいて、画素のオリジナルの外観の品質レベルの変更が達成される。

【0094】

50

パターンベクトルの構成および分解の特定的手段は当業者に公知である。

【0095】

顕著信号モードの周囲信号からの空間セグメント化およびこのモードにおける空間正規化を前提として、画素自体、すなわち結果として得られる正規化信号の外観は、画素の外観の表現の近似誤差とビットレートとの間の直接トレードオフを可能にする低いランクのパラメータ化を用いて、線形相関性があるコンポーネントに因子分解できる。

【0096】

図2に示すように、正規化されたオブジェクトの画素(242および244)をベクトル空間に射影でき、データの次元的にコンパクトなバージョンを生成するために、PCAのような分解処理(250)を使用して、線形相関関係をモデル化することができる(252および254)。

10

【0097】

<逐次(シーケンシャル)PCA>

PCAは、PCA変換を用いて、パターンをPCA係数にコード化する。PCA変換によって、より優れたパターンが表現されると、パターンをコード化するのに必要な係数がより少なくなる。訓練パターンの取得とコード化されるパターンとの間の時間の経過に伴ってパターンベクトルが劣化することを認識して、変換を更新することにより劣化に対抗する作用を助けることができる。新しい変換の生成に対する代わりとして、既存パターンの逐次更新が、特定の場合において計算的により有効である。

【0098】

20

多くの最先端のビデオ圧縮アルゴリズムは、1つまたは複数の他のフレームから1つのビデオフレームを予測する。予測モデルは一般的に、重ならないタイルへの各予測フレームの分割に基づいている。この重ならないタイルは、別のフレーム内の対応するパッチおよびオフセット動きベクトルによってパラメータ化される関連の並進運動に一致する。随意にフレームインデックスと結合されるこの空間的変位が、タイルの「動き予測」の変形を提供する。予測誤差が特定のしきい値を下回る場合、タイルの画素は残差のコード化に適し、圧縮効率における対応する利得が存在する。そうでなければ、タイルの画素は直接コード化される。この種類のタイルを基礎とする(代わりに、ブロックを基礎とする、と称される)動き予測方法は画素を含むタイルを並進させることによりビデオをモデル化する。ビデオの画像化現象がこの種類のモデル化に準拠する場合、対応するコード化の効率が向上する。このモデル化の制約は、ブロックを基礎とする予測に固有である並進運動に適合させるために、特定レベルの時間解像度すなわち1秒当たりのフレーム数が、動きのある画像化されるオブジェクトに持続すると仮定する。この並進モデルに関する別の必要条件是、特定の時間解像度に対する空間変位が制限されていることである。すなわち、予測が導き出されるフレームと予測されるフレームとの間の時間差が、比較的短い絶対時間量でなければならない。これらの時間解像度および動き制限は、ビデオストリーム内に存在する、ある一定の冗長ビデオ信号コンポーネントの識別およびモデル化を促進する。

30

【0099】

<残差を基礎とする分解>

MPEGビデオ圧縮では、現在のフレームは動きベクトルを使用して前のフレームの動き補償によって形成され、次いで補償ブロックに対して残差の更新を適用し、最終的に十分な一致を有しないいずれかのブロックが新しいブロックとしてコード化される。

40

【0100】

残差のブロックに対応する画素は、動きベクトルによって前のフレーム内の画素にマッピングされる。この結果は、残差の値を連続して適用することによって合成できるビデオを通る画素の時間経路である。これらの画素はPCAを用いて最適に表現される画素として識別される。

【0101】

<遮蔽を基礎とする分解>

本発明の別の改良は、ブロックに適用される動きベクトルが、画素を移動させることに

50

よって前のフレームからのいずれかの画素を遮蔽する（覆う）のか否かを決定する。各遮蔽の発生に対して、遮蔽画素を新しい層に分割する。また、履歴なしに出現する画素も存在する。出現した画素は現在のフレーム内のそれら出現画素に適合するいずれかの層に配置され、履歴の適合はその層に対して実施される。

【0102】

画素の時間連続性は様々な層への画素の接続および接合によって維持される。安定した層モデルが得られると、各層内の画素はコヒーレント動きモデルとの帰属関係に基づいてグループ化される。

【0103】

<サブバンドの時間量子化>

本発明の別の実施形態は、離散コサイン変換（DCT）または離散ウェーブレット変換（DWT）を用いて、各フレームをサブバンド画像に分解する。次に、主成分分析（PCA）がこれらの「サブバンド」ビデオのそれぞれに適用される。この概念は、ビデオフレームのサブバンド分解がオリジナルのビデオフレームと比較して、サブバンドのいずれかが1つにおける空間変化を少なくするというものである。

【0104】

動きのあるオブジェクト（人物）のビデオについては、空間変化がPCAによってモデル化される変化を左右する傾向にある。サブバンド分解は、いずれの分解ビデオにおける空間変化も減少させる。

【0105】

DCTについては、いずれのサブバンドに対する分解係数も、サブバンドのビデオに空間的に配置される。例えば、DC係数は各ブロックから取得され、オリジナルのビデオの郵便切手の変形のように見える、サブバンドのビデオに配置される。これは他のサブバンドのすべてに対して繰り返され、結果として得られるサブバンドビデオのそれぞれはPCAを使用して処理される。

【0106】

DWTでは、サブバンドはすでにDCTに対して説明されている方法で配置される。

限定されない実施形態において、PCA係数の打ち切りは変更される。

【0107】

<ウェーブレット>

データが離散ウェーブレット変換（DWT）を用いて分解されると、多重帯域通過データセットが低い空間解像度になる。この変換処理は、単一のスカラー値を得るまで、導き出されたデータに再帰的に適用される。分解された構造におけるスカラー要素は一般に、階層的な親/子方式で関連付けられる。結果として得られるデータは多重解像度の階層的な構造および有限差分を含む。

【0108】

DWTが空間強度場に適用されると、自然発生する画像現象の多くは、低い空間周波数のために、第1または第2の低帯域通過生成データ構造によってほとんど知覚損失なく表現される。高い周波数の空間データが存在しないかノイズと見なされるかのいずれかである場合、階層構造の打ち切りがコンパクトな表現を提供する。

【0109】

PCAを用いることによって少数の係数で正確な復元を達成できるが、変換自体は極めて大きい。この「初期の」変換の大きさを低減するために、ウェーブレット分解の組み立てゼロ・ツリー（EZT）構成を用いて変換マトリックスのより正確な変形を連続的に形成することができる。

【0110】

<部分空間分類>

本発明は、ベクトルの要素としてのいくつかの不連続なパラメータ化を表現する。これらのパラメータには、セグメント化されたオブジェクトの正規化された外観における画素、動きパラメータ、および2または3次元における特徴または頂点の構造的な位置（これら

10

20

30

40

50

に限定されない)が含まれる。これらベクトルのそれぞれはベクトル空間に存在し、この空間の形状の解析が用いられて、パラメータベクトルのコンパクトな表現が取得される。有利な形状は、ベクトルのクラスに集中した部分空間を有する。いくつかの部分空間が混合されて複雑な部分空間を生成すると、個々の単純な部分空間の識別が難しくなる。オリジナルのベクトルの何らかの相互作用(例えば、内積)を通して生成される、より高次元のベクトル空間におけるデータを調べることによって、このような部分空間の分離を可能にするいくつかのセグメント化方法が存在する。

【0111】

ベクトル空間のセグメント化の一方法は、多項式を表現するVeroneseベクトル空間へのベクトルの射影を含む。この方法は、従来技術において、一般化PCAすなわちGPCA技法として公知である。このような射影によって、多項式への法線が見出され、これら法線に関連するベクトルが一体にグループ化される。この技法の利用例は、時間にわたって追跡された2次元空間点相関関係の3次元構造モデルおよびこの3次元モデルの動きへの因子分解である。

【0112】

<ハイブリッド空間正規化圧縮>

本発明は、ビデオストリームを2つまたはそれ以上の「正規化」ストリームにセグメント化することの追加によって、ブロックを基礎とした動き予測コード化の仕組みの効率を拡張する。次に、従来のコーデック(CODEC)の並進動きの仮定を有効にできるように、これらのストリームが別個にコード化される。正規化ストリームをデコードすると、ストリームはこれらの適切な位置に非正規化され、一体に合成されて、オリジナルのビデオシーケンスが得られる。

【0113】

一実施形態においては、1つまたは複数のオブジェクトがビデオストリームにおいて検出され、その後、個々のオブジェクトのそれぞれに関係する画素が、非オブジェクトの画素を残してセグメント化される。次に、全体の空間動きモデルが、オブジェクトおよび非オブジェクトの画素に対して生成される。全体モデルが使用されて、オブジェクトおよび非オブジェクトの画素を空間的に正規化する。このような正規化は、ビデオストリームから非並進動きを効果的に取り除いており、遮蔽の相互作用が最小限にされている一式のビデオを提供している。これらは両方とも本発明の方法の有利な構成である。

【0114】

空間的に正規化された画素を有するオブジェクトおよび非オブジェクトの新しいビデオが、従来のブロックを基礎とした圧縮アルゴリズムへの入力として供給される。ビデオのデコードに関して、全体の動きモデルパラメータが用いられて、これらデコードされるフレームを非正規化し、オブジェクトの画素が非オブジェクトの画素に一体に合成されて、ほぼオリジナルのビデオストリームが生成される。

【0115】

図6に示すように、1つまたは複数のオブジェクト(630および650)に対する先に検出されたオブジェクトインスタンス(206および208)が、従来のビデオ圧縮方法(632)の別個のインスタンスでそれぞれ処理される。さらに、オブジェクトのセグメント化(230)から生じた非オブジェクト(602)も、従来のビデオ圧縮(632)を用いて圧縮される。これら別個の圧縮コード化(632)のそれぞれの結果は、各ビデオストリームに別個にそれぞれ対応している、従来方法でコード化された別個のストリーム(634)である。おそらくは伝送の後である、ある時点で、これら中間コード化されたストリーム(234)が、正規化された非オブジェクト(610)および多数のオブジェクト(638および658)の合成に解凍(636)される。画素が空間において相対的に正しく位置するように、これら合成された画素は、これらの非正規化バージョン(622、642および662)に非正規化(640)されることができる。これにより、合成処理(670)が、オブジェクトおよび非オブジェクトの画素を組み合わせることで完全なフレームの合成(672)とすることができる。

【 0 1 1 6 】

< ハイブリッドコーデックの統合 >

本発明の記述のとおり、従来のブロックを基礎とした圧縮アルゴリズムと正規化 - セグメント化の仕組みとの組み合わせにおいて、いくつかの進歩的な方法がもたらされる。第 1 に、特別なデータ構造および通信プロトコルが必要とされる。

【 0 1 1 7 】

主たるデータ構造は、全体の空間変形パラメータおよびオブジェクトセグメント化仕様マスクを含む。主たる通信プロトコルは、全体の空間変形パラメータおよびオブジェクトセグメント化仕様マスクの伝送を含む層である。

【 図面の簡単な説明 】

10

【 0 1 1 8 】

【 図 1 】 従来技術のビデオ処理システムを示したブロック図である。

【 図 2 】 ビデオを処理の主要なモジュールを示す、本発明の全体を示すブロック図である。

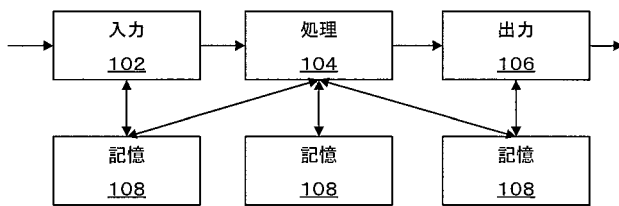
【 図 3 】 本発明の動き検出方法を示すブロック図である。

【 図 4 】 本発明の全体の位置合わせ方法を示すブロック図である。

【 図 5 】 本発明の正規化方法を示すブロック図である。

【 図 6 】 ハイブリッド空間正規化圧縮方法を示すブロック図である。

【 図 1 】

FIG. 1
Prior Art

【 図 2 】

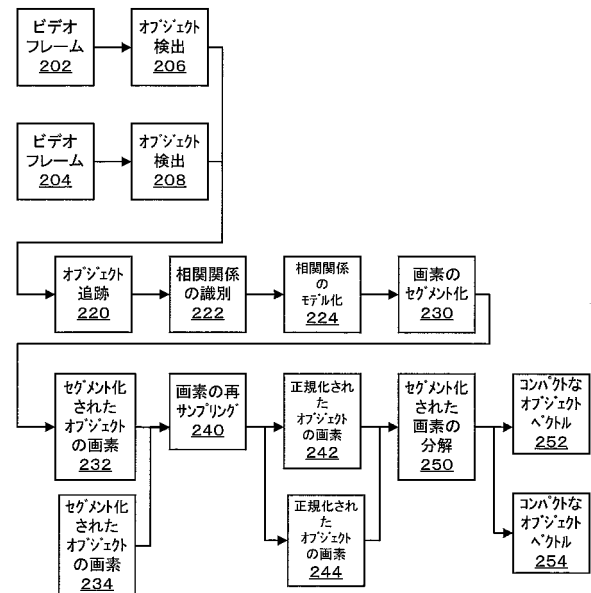


FIG. 2

【図 3】

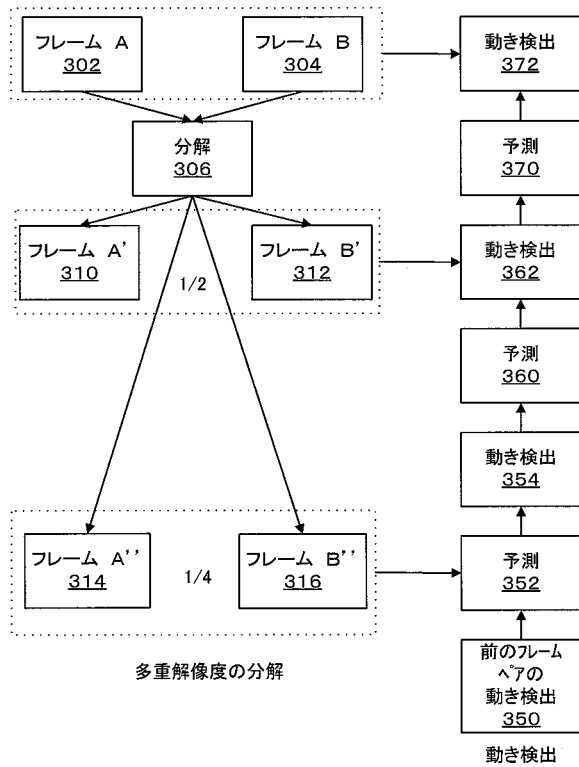


FIG. 3

【図 4】

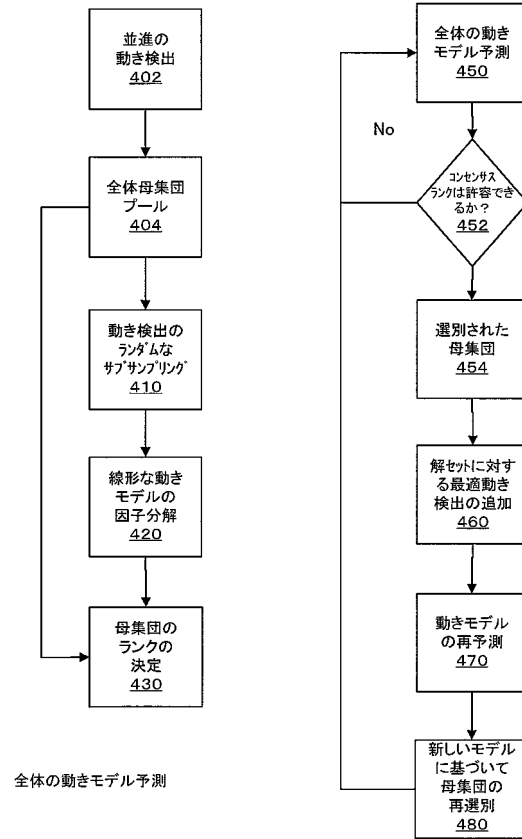


FIG. 4

コンセンサスランク許容の試験および処理

【図 5】

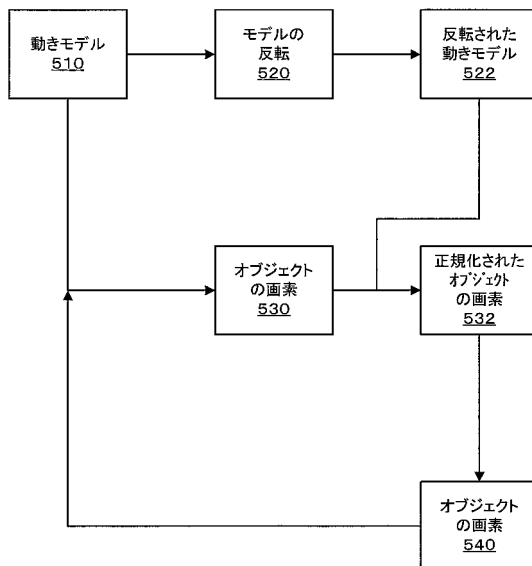


FIG. 5

【図 6】

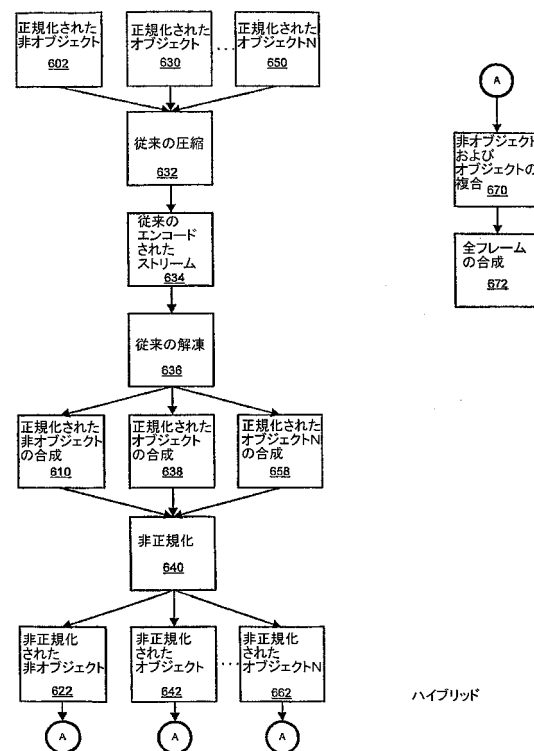


FIG. 6

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT		International application No. PCT/US05/33733
A. CLASSIFICATION OF SUBJECT MATTER IPC: G06K 9/00(2006.01);H04N 5/225(2006.01) USPC: 382/103;348/169 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) U.S. : 382/103; 348/169 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) Google Scholar		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 6,711,278 B1 (Gu et al.) 23 March 2004 (23.03.2004) col 1, line 13 - col 6, line 42).	1-5,7-8
Y	Viola, Paul and Michael Jones. "Rapid Object Detection using a Boosted Cascade of Simple Features". 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'01). IEEE: 2001. Volume 1. Pages 511- 518.	1-5,7-8
Y	US 2004/0135788 A1 (Davidson et al.) 15 July 2004 (15.07.2004), pages 8-9, paragraphs 112-113.	3
Y	US 2003/0235341 A1 (Gokturk et al.) 25 December 2003 (25.12.2003), page 5, paragraphs 64 and 69.	5,7-8
Y	US 2003/0194134 A1 (Wenzel et al.) 16 October 2003 (16.10.2003), Abstract.	7
A	US 2004/0013286 A1 (Viola et al.) 22 January 2004 (22.01.2004).	
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.		
* Special categories of cited documents:		
"A" document defining the general state of the art which is not considered to be of particular relevance	"I" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention	
"E" earlier application or patent published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone	
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art	
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family	
"P" document published prior to the international filing date but later than the priority date claimed		
Date of the actual completion of the international search 31 March 2006 (31.03.2006)	Date of mailing of the international search report 19 MAY 2006	
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US Commissioner for Patents P.O. Box 1450 Alexandria, Virginia 22313-1450 Facsimile No. (571) 273-3201	Authorized officer Damon Conover Telephone No. (572) 272-5448	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US05/33733

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2003/0103647 A1 (Rui et al.) 5 June 2003 (05.06.2003).	
A	US 6,625,310 B2 (Lipton et al.) 23 September 2003 (23.09.2003).	
A	Viola, Paul and Michael Jones. "Fast Multi-view Face Detection". Mitsubishi Electric Research Laboratories, Inc: July 2003.	

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US05/33733

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☒ Claims Nos.: 6
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of any additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW

(72)発明者 ペース・チャールズ・ポール

アメリカ合衆国, バーモント州 05763, ノース チッテンデン, スミス ファーム ロード
70

Fターム(参考) 5C059 KK01 KK08 MA00 MA23 MA24 MB03 MB12 MC11 NN01 NN21
NN27 NN40 PP12 PP26 PP27 PP28 TA61 TC02 TC12 TD02
TD03 TD05 TD09 TD12 TD13 TD15 TD16 TD17 UA05 UA12
UA22
5L096 AA06 CA04 HA09 JA11