

(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2008-521347

(P2008-521347A)

(43) 公表日 平成20年6月19日(2008.6.19)

(51) Int.Cl. F I テーマコード (参考)
H04N 7/32 (2006.01) H04N 7/137 Z 5C059

審査請求 未請求 予備審査請求 未請求 (全 29 頁)

| | |
|---|---|
| (21) 出願番号 特願2007-543165 (P2007-543165) (86) (22) 出願日 平成17年11月16日 (2005.11.16) (85) 翻訳文提出日 平成19年7月11日 (2007.7.11) (86) 国際出願番号 PCT/US2005/041253 (87) 国際公開番号 W02006/055512 (87) 国際公開日 平成18年5月26日 (2006.5.26) (31) 優先権主張番号 60/628,819 (32) 優先日 平成16年11月17日 (2004.11.17) (33) 優先権主張国 米国 (US) (31) 優先権主張番号 60/628,861 (32) 優先日 平成16年11月17日 (2004.11.17) (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 |
|---|---|

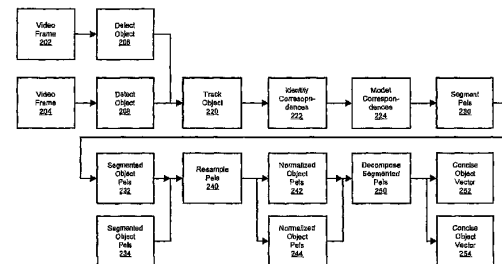
最終頁に続く

(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関係する問題に対処する。

【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、元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母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 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。

10

20

30

40

50

【請求項 7】

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を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、

前記圧縮および解凍処理が、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、コンピュータ装置。

【請求項 8】

10

請求項 1 において、前記相関関係モデルが局所変形に因子分解され、

前記オブジェクトに対応する 2 次元メッシュオーバーレイ画素を定義する手段と、

局所動きのモデルへの相関関係測定手段とを備え、

前記メッシュ定義手段は規則的な頂点グリッドおよびエッジに基づいており、

前記相関関係測定が、2 つ以上のビデオフレーム間のブロックを基礎とした動き検出から生成される有限差分に基づく頂点変位を含む、コンピュータ装置。

【請求項 9】

請求項 8 において、前記頂点は離散画像特徴に相当し、

前記オブジェクトに相当する顕著な画像特徴を特定する手段を備え、

前記特定手段は、画像勾配 Harris 応答の解析を行う、コンピュータ装置。

20

【請求項 10】

離散線形部分空間に存在するデータベクトルを分離するコンピュータ装置であって、

データベクトルセットにおいて部分空間セグメント化を実行する手段と、

陰関数表現のベクトル空間において接線ベクトル解析の適用によって、部分空間セグメント化基準を制約する手段とを備え、

前記部分空間セグメント化方法が GPCA であり、

前記陰関数表現のベクトル空間が Veronese マップであり、

前記接空間制約が Legendre 変換である、コンピュータ装置。

【発明の詳細な説明】

【関連出願】

30

【0001】

本出願は、2004 年 11 月 17 日出願の米国特許仮出願第 60/628,861 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)」、および 2004 年 11 月 17 日出願の米国特許仮出願第 60/628,819 号「ビデオデータの処理および符号化装置および方法 (Apparatus and Method for Processing and Coding Video Data)」の優先権を主張する。本出願は、また、2005 年 7 月 28 日出願の米国特許出願第 11/191,562 号の部分継続出願である、2005 年 9 月 20 日出願の米国特許出願第 11/230,686 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】

40

【0002】

本発明は一般にデ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。

50

【 0 0 0 4 】

入力ステージ 1 0 2 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。

【 0 0 0 5 】

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、予め定義された転送プロトコルに従い、直接に、または中間記憶要素 1 0 8 を介して処理ステージ 1 0 4 に提供される。処理ステージ 1 0 4 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 1 0 4 は通常 1 つまたは複数の CODEC (符号器 / 復号器) を有する。

【 0 0 0 6 】

出力ステージ 1 0 6 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。

【 0 0 0 7 】

記憶装置はこのシステムにおいて随意的要素として提示されている。用いられる場合、記憶要素 1 0 8 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。

【 0 0 0 8 】

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。

【 0 0 0 9 】

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。

【 0 0 1 0 】

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。

【 0 0 1 1 】

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

10

20

30

40

50

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。

【 0 0 1 2 】

「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、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。

10

【 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。

【 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。

20

【 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。

30

【 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、系統的誤りを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を発生させ、ビデオフレームの元々は滑らかであった領域に急峻な画像データを発生させる。

40

【 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

50

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。

【 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。

【 発明の開示 】

【 0 0 2 1 】

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、空間正規化に相当するデータの空間セグメント化は、空間正規化のみから誘引される利益を増大するように作用することができる。

【 0 0 2 2 】

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。

【 発明を実施するための最良の形態 】

【 0 0 2 3 】

ビデオ信号データでは、通常3次元シーンを2次元画像面への射影および画像化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(例えば、電磁エネルギー、音響エネルギーなど)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。

【 0 0 2 4 】

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。

10

【 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（アピランス）に関係する顕著な情報が、このコンパクトな表現に包含されている。

20

【 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。

30

【 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 >

40

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 】

50

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。

【0032】

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。

【0033】

<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。

【0034】

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）。

【0035】

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限定するために、特徴点を用いることができる。

【0036】

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。

【0037】

<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。

【0038】

<局所的な位置合わせ>

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。

【0039】

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。

【0040】

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

10

20

30

40

50

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。

【 0 0 4 1 】

図 3 に示すように、フレーム (3 0 2 または 3 0 4) が、おそらくは分解処理 (3 0 6) 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 (3 1 0 および 3 1 2) 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 >

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 (F S B B) 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。F S B B 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 (D S B B) 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、F S B B に対する一般的な代替方法である。

【 0 0 4 3 】

本発明の一実施形態では、D S B B 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。

【 0 0 4 4 】

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。

【 0 0 4 5 】

< 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 (P N C C) 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。

【 0 0 4 6 】

P N C C 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、P N C C 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。

【 0 0 4 7 】

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 P N C C 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。

【 0 0 4 8 】

< 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 R A N S A C アルゴリズムの予測値と同様な頑健な予測値を用いる。

【 0 0 4 9 】

10

20

30

40

50

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母集団(404)を調節する。

【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。

10

【0051】

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、RANSAC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。

20

【0052】

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母集団の最終コンセンサスとみなされる。

【0053】

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。

【0054】

図4に示すように、全体のモデル予測処理(450)が、コンセンサスランク許容テスト(452)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(404)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460において解セットに加えられる。次いで、処理470においてモデルが再予測される。完了すると、母集団(404)が再選別される。

30

【0055】

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。

40

【0056】

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。

【0057】

<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

50

正規化方法は図 5 に示されている。

【0058】

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。

【0059】

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、隣接画素の補間を意味する有限差分である。

【0060】

元の強度場の正規化強度場への補間によって、強度場の部分集合に基づいた PCA 外観モデルの線形性が向上する。

【0061】

図 2 に示すように、オブジェクトの画素 (232 および 234) を再サンプリングして (240)、正規化されたバージョンのオブジェクトの画素 (242 および 244) を得ることができる。

【0062】

< 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、後続の外観を基礎とする PCA 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。

【0063】

他の実施の形態は、全体の動きモデルを暗示的 (陰関数表現的) 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、従来から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。

【0064】

図 9 は、全体正規化と局所正規化の組合せ方法を示す。

【0065】

< 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。

【0066】

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、元の頂点のそれぞれの顕著な優先度を求めるために、連続的に低い精度で近似される。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。

【0067】

本発明の一実施形態において、多角形分解方法によって、均質な領域分類に付随した境界の優先順位付けが可能になる。スペクトル類似性のような均質性基準に従って画素が分類され、分類ラベルは空間的に領域に連結される。さらに好ましい限定されない実施形態においては、空間連結性を判別するために、4 または 8 の連結性基準が適用される。

【0068】

好ましい実施形態においては、次に、これら空間領域の境界は多角形に離散化される。全ての均質領域の全ての多角形の空間オーバレイは、次にテセレーションされて、仮のメッシュに結合される。元のメッシュの知覚の顕著性 (perceptive saliency) の多くを保

10

20

30

40

50

持する、単純なメッシュ表現を示すために、仮のメッシュの頂点は、いくつかの基準を用いて分解される。

【 0 0 6 9 】

好ましい実施形態においては、本明細書の別の箇所において記載する画像の位置合わせ方法が、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 7 0 】

好ましい実施形態においては、領域境界を絞り込むのに、動的輪郭が用いられる。多角形領域の動的輪郭は、1回の繰り返しを伝播してもよい。異なる領域の各動的輪郭頂点の「変形」または動きは平均化操作に組み合わされて、暗示されたメッシュの制約された伝播が可能になる。この暗示されたメッシュに対しては、これら動的輪郭頂点が全て属している。

10

【 0 0 7 1 】

好ましい実施形態においては、頂点には、異なる領域の輪郭の一部である隣接頂点に対するメッシュ内の隣接頂点数が割り当てられる。これら別の頂点是对立していると定義される。頂点が1の数を有する場合、対立頂点を有さず、保持される。2つの隣接する反対の頂点が両方とも1の数を有する場合（これら2つの頂点が異なる多角形内にあり、互いに隣接することを意味する）、一方の頂点は他方の頂点に帰着する。1の数の頂点が、2の数を有する近隣の多角形頂点と対立する場合、1の数の頂点は2の数の頂点に帰着し、この2の数である頂点の数が1になる。したがって、もう一つ別の近隣の反対の頂点が存在する場合、この頂点は再度帰着される。この場合、元の頂点の数を保存しておくことが重要である。これは、頂点が帰着された場合、元の頂点の数に基づいて帰着の方向をとることができるからである。つまり、頂点aが頂点bに帰着した後は頂点bは頂点cに帰着しない。代わりに、頂点cが頂点bに帰着する必要がある。これは、頂点bが既に1つの帰着に使用されているからである。

20

【 0 0 7 2 】

好ましい実施形態においては、T分岐ポイントが特に処理される。これらT分岐ポイントは、隣接する多角形にポイントを有しない多角形のポイントである。この場合、各多角形頂点は画像ポイントマップに最初にプロットされる。このマップは頂点の空間位置および多角形識別子を特定する。次に、各多角形の周縁がトラバースされ、他の多角形からの隣接頂点があるか否かがチェックされる。他の領域から近接する頂点が存在する場合、それら頂点はそれぞれ現多角形からの近接する頂点を既に有するか否かがチェックされる。有しない場合、現多角形の頂点として現ポイントが追加される。この追加のチェックによって、他の多角形の孤立した頂点がT分岐ポイントを生成するのに用いられることが保証される。これに対して、有する場合、この領域が既に一致頂点を有する箇所に、新たな頂点を追加するのみである。したがって、近接する頂点がこの現領域に対立しない場合のみ、対立するベクトルが追加される。さらなる実施形態においては、マスク画像を利用することで、T分岐検出の効率が良くなる。多角形頂点が順次対象となり、頂点の画素が多角形頂点に属すると特定されるように、マスクが更新される。次に、多角形の周縁の画素がトラバースされ、多角形頂点と空間的に一致する場合は、これら画素は現多角形内の頂点として記録される。

30

40

【 0 0 7 3 】

好ましい実施形態では、1つまたは複数の重複する均質な画像勾配領域によって、スペクトル領域が再配置されて別の均質なスペクトル領域も重複すると、先に再配置された全ての領域に、現在再配置されている領域と同一のラベルが付与される。したがって、要するに、スペクトル領域が2つの均質領域によって重複されると、これら2つの均質領域によって重複される全てのスペクトル領域は、同一ラベルを取得する。これより、1つのスペクトル領域は2つの均質領域の代わりに1つの均質領域によって実際に覆われているかのようである。

【 0 0 7 4 】

50

本発明の一実施形態においては、隣接マージ基準を見出すために、領域リストよりも領域マップを処理するのが有利である。さらなる実施形態では、不均質領域を用いる識別器を訓練するためにスペクトルセグメント化識別器が改良されてもよい。さらに、cannyエッジ検出のようなエッジの使用に基づく別のセグメント化を付加して、次に最初の多角形セットを特定するために有効な輪郭を供給することによって、均質領域を大幅に区別化できる。

【0075】

< 局所正規化 >

本発明は、時空間ストリームにおける画素が「局所的な」方式で位置合わせされる方法を提供する。

10

【0076】

このような局所化された方法は、幾何学的なメッシュの空間適用を採用する。この方法は、これにより、画像化された事象の局所変形と関連付けて明らかな画像明るさの恒常性の曖昧さを解決する際に、画像化された事象（特に、画像化されたオブジェクト）における局所化されたコヒーレンシーが組み入れられるように画素を解析する方法を提供する。

【0077】

このようなメッシュは、画像平面における表面変形の区分線形モデルを提供するために、局所正規化の手段として採用される。画像化された事象は、多くの場合、ビデオストリームの時間解像度がビデオの動きに比べて高い場合、区分線形モデルに相当する。モデル仮定（model assumption）の例外が、様々な技法で取り扱われる。これら技法としては、トポロジカルな制約、隣接頂点条件、ならびに画素の均質性および画像勾配領域の解析である。

20

【0078】

一実施形態においては、頂点が特徴点に相当する三角形要素から構成されるメッシュを生成するために、特徴点を用いられる。他のフレームの相当する特徴点は、局所変形モデルを生成するために、三角形の補間「ワーピング」および同様に画素の補間「ワーピング」を暗示する。

【0079】

図7は、このようなオブジェクトメッシュの生成を示す。図8は、このようなオブジェクトメッシュを局所的にフレームに正規化するために用いることを示す。

30

【0080】

好ましい実施形態においては、三角形マップが生成され、これにより、マップの各画素の元となる三角形を特定する。さらに、各三角形に対応するアフィン変換が、最適化ステップとして予め計算される。さらに、局所変形モデルを作成すると、元の画素の座標を決定してサンプル化するために、固定画像（前回の画像）が空間座標を用いてトラバースされる。

【0081】

別の実施形態においては、全体変形の後に局所変形が実行される。上述の先に開示された明細書においては、2つまたはそれ以上のビデオフレームにおいて画素の空間的な正規化に全体位置合わせ方法が用いられる処理として、全体正規化が記載されていた。全体正規化されたビデオフレームの結果は、さらに局所正規化される。これら2つの方法の組合せによって、局所正規化が解決を導く全体の絞込みとなる。これにより、局所方法の解決されなければならない曖昧さを大幅に削減できる。

40

【0082】

別の限定されない実施形態においては、特徴点、すなわち「規則的メッシュ」の場合には頂点が、これら特徴点の近傍における画像勾配の解析を通して限定される。この画像勾配は、直接、またはHarris応答のような間接的な計算を通して計算される。さらに、これらポイントは、空間制約および画像勾配の低下に関連する動き予測エラーによってフィルタリングされる。限定されたポイントが、多くのテセレーション技法の1つによってメッシュの基礎として用いられ、要素が三角形であるメッシュをもたらす。三角形ごと

50

に、ポイントおよびその残差動きベクトルに基づいてアフィンモデルが生成される。

【0083】

好ましい実施形態においては、三角形アフィンパラメータのリストが維持される。このリストは繰り返され、すなわち変更が逐次追加されて、現在 / 前回のポイントリストが構築される（頂点探索マップを用いて）。現在 / 前回のポイントリストは変換を予測するのに用いられるルーチンに渡される。この変換は、その三角形についてのアフィンパラメータを計算する。アフィンパラメータ、すなわちモデルは、三角形アフィンパラメータリストに保存される。

【0084】

さらなる実施形態においては、本方法が三角形識別子画像マップをトラバースする。ここで、マップの各画素は、画素が要素であるメッシュにおける三角形についての識別子を含む。三角形に属する各画素に対して、その画素についての相当する全体変形座標および局所変形座標が計算される。次に、相当する画素をサンプル化してその値を相当する「正規化」位置に適用するために、これら座標が用いられる。

【0085】

さらなる実施形態においては、画像勾配の検索に起因する密度および画像強度対応強さに基づいて、空間制約がポイントに適用される。動き検出が画像強度の残差のノルムに基づいて行われた後に、ポイントが分類される。次に、ポイントは空間制約に基づいてフィルタリングされる。

【0086】

さらなる実施形態においては、スペクトル空間セグメント化が採用されて、小さい均質空間領域が、強度および / または色と類似する空間関連性（affinity）に基づいて、近傍の領域と合併される（マージされる）。均質テクスチャ（画像勾配）の領域との重ね合わせに基づいて、スペクトル領域を共に組み合わせるために、均質マージングが用いられる。さらなる実施形態は、次に中心周辺（center-surround）ポイントを用いる。これらポイントは、メッシュの頂点をサポートするための限定された対象ポイントのように、小さい領域が大きい領域に囲まれているかのようである。さらに限定されない実施形態においては、中心周辺ポイントは、 3×3 、 5×5 または 7×7 画素内に境界ボックスが存在する領域として画定される。また、この境界ボックスについての空間画像はコーナー形状である。この領域の中心はコーナーとして分類され、さらに、この位置を有利な頂点位置としてみなす。

【0087】

さらなる実施形態においては、水平方向および垂直方向の画素有限差分画像が、各メッシュエッジの強さを分類するために用いられる。空間位置と一致する有限差分をエッジが多数有する場合、このエッジは画像化現象の局所変形に極めて重要であるとみなされる。エッジの有限差分の合計の平均間に大きい導関数差分（derivative difference）を有する場合、領域エッジが量子化ステップではなくテクスチャ変更エッジに相当する可能性が高い。

【0088】

さらなる実施形態においては、空間密度モデル終了条件が、メッシュ頂点の処理を最適化するために採用される。検出矩形が発生する空間エリアのほとんどを包含する十分な数のポイントが検査されると、処理は終了する。終了はスコアを生成する。この処理に入る頂点および特徴点は、このスコアによって分類される。存在ポイントにポイントが空間的に極めて近い場合、またはポイントが画像勾配におけるエッジに一致しない場合、このポイントは廃棄される。これらに該当しない場合、ポイントの近傍における画像勾配は下降し、勾配の残差が限界を超えると、このポイントも廃棄される。

【0089】

< 規則的メッシュ正規化 >

本発明は、前述の規則的メッシュを利用する局所正規化方法を拡張する。このメッシュは、検出オブジェクトに一致した位置およびサイズを有している根本の画素に関係なく構

10

20

30

40

50

築される。

【0090】

検出オブジェクトを所与として、空間フレーム位置および顔のサイズを示すスケールが、顔の発生領域にわたって規則的メッシュを生成する。好ましい実施形態においては、長方形メッシュの輪郭を描くために、重ならないタイルセットを用いて、三角形メッシュ要素を有する規則的メッシュを生じさせるためにタイルの対角線区分化を実行する。さらに好ましい実施形態においては、タイルは従来のビデオ圧縮アルゴリズム（例えば、MPEG-4 AVC）におけるタイルに相当している。

【0091】

好ましい実施形態においては、前述のメッシュの頂点は、訓練に用いられる特定のビデオフレームにおけるこれら頂点を取り囲む画素領域の解析を介して優先順位付けされる。このような領域の勾配の解析が、局所画像勾配（例えば、ブロックを基礎とした動き検出）に依存する各頂点の処理に関して正確さを提供する。

10

【0092】

多数のフレームにおける頂点位置の一致は、画像勾配の単純な下降によって見つけれられる。好ましい実施形態においては、これは、ブロックを基礎とした動き検出によって達成される。本実施形態においては、極めて正確な頂点によって、極めて正確な一致が可能となる。極めて正確な頂点一致からの推定される曖昧な画像勾配を解決することにより、暗示的に正確さの低い頂点一致となる。

【0093】

好ましい実施形態においては、規則的メッシュが矩形追跡の発生を大きく変える。タイルは 16×16 に生成され、対角線上に切られて三角形形状のメッシュを生成する。これら三角形の頂点は動き推定される。動き検出はそれぞれポイントを有するテクスチャの型によって決まる。テクスチャは、コーナー、エッジおよび均質の3つのクラスに分割され、テクスチャはまた頂点の処理順序を定義する。コーナー頂点は、近傍の頂点検出を用いる。すなわち、近傍ポイント（利用可能な場合）の動き検出が、予測動きベクトルに用いられ、動き検出がそれぞれに適用される。重大誤差が最小（lowest mad error）の動きベクトルが、この頂点動きベクトルとして用いられる。コーナーに用いられる探索方法は全て（幅広い、小さい、および原点）である。エッジに対しては、再び最短の近傍動きベクトルが予測動きベクトルとして用いられ、最も誤差量の少ないものが用いられる。エッジについての探索方法は、小さいおよび原点である。均質については、近傍ベクトルが探索され、最も誤差の少ない動き検出が用いられる。

20

30

【0094】

好ましい実施形態においては、各三角形頂点に対する画像勾配が計算され、クラスおよび大きさに基づいて保存される。したがって、コーナーはエッジの前であり、このエッジは均質の前である。コーナーについて、強いコーナーは弱いコーナーよりも前であり、エッジについて、強いエッジは弱いエッジよりも前である。

【0095】

好ましい実施形態においては、各三角形についての局所変形は、その三角形に関する動き検出に基づいている。各三角形はその三角形に対して推定されたアフィンを有する。三角形がトポロジ的に反転しない場合、つまり縮退しない場合、推定三角形の一部である画素が、得られた推定アフィンに基づいて、現画像をサンプル化するのに用いられる。

40

【0096】

<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。

【0097】

50

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。

【0098】

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(220)されてモデル化(224)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(230)するために使用することができる。この処理を、ビデオ(202および204)において検出された多数のオブジェクト(206および208)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(232)である。

【0099】

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。

【0100】

<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。

【0101】

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。

【0102】

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。

【0103】

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。

【0104】

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。

【0105】

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。

【0106】

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。

【0107】

< 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。

【0108】

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。

【0109】

< 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。

【0110】

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されて、4つの追加操作に組み合わせられる4つの参照を通して元の場の任意の長方形の加算を促進する。

【0111】

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。

【0112】

< 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。

【0113】

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

10

20

30

40

50

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、単一の繰り返し 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、元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。

【0114】

< 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。

【0115】

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。

【0116】

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。

【0117】

< 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。

【0118】

< 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。

【0119】

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。

【0120】

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。

【0121】

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

化を減少することができる。

【 0 1 2 2 】

< P C A >

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て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 (P C A) 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 P C A 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、画素の元の外観の品質レベルの変更が達成される。

【 0 1 2 3 】

パターンベクトルの構成および分解の特定の手段は当業者に公知である。

【 0 1 2 4 】

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。

【 0 1 2 5 】

図 2 に示すように、正規化されたオブジェクトの画素 (2 4 2 および 2 4 4) 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、P C A のような分解処理 (2 5 0) を使用して、線形相関関係をモデル化することができる (2 5 2 および 2 5 4) 。

【 0 1 2 6 】

< 逐次 (シーケンシャル) P C A >

P C A は、P C A 変換を用いて、パターンを P C A 係数にコード化する。P C A 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、特定の場において計算的により有効である。

【 0 1 2 7 】

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は、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。

【 0 1 2 8 】

< 残差を基礎とする分解 >

M P E G 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

10

20

30

40

50

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。

【0129】

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特定される。

【0130】

< 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）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。

10

【0131】

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。

【0132】

< 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元のビデオフレームと比較して、サブバンドのいずれか1つにおける空間変化を少なくするというものである。

20

【0133】

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。

【0134】

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、元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。

30

【0135】

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。

【0136】

< 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。

40

【0137】

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。

50

【0138】

P C A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込みゼロ・ツリー（E Z T）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。

【0139】

< 部分空間分類 >

当業者には理解されるとおり、離散的にサンプル化された事象データおよび派生データは、代数的ベクトル空間に対応するデータベクトルのセットとして表現される。これらベクトルは、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（これらに限定されない）を含む。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もよく、サンプル化されたベクトル、つまりパラメータのベクトルのコンパクトな表現が取得される。有利な形状条件は、コンパクトな部分空間を形成するパラメータベクトルによって類型化されている。1つまたは複数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。元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。

【0140】

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P C AすなわちG P C A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、グループ化され、さらに、元の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。

【0141】

G P C A技法は、定義されたように適用された場合には不完全であり、データベクトルがほとんどノイズなく生成された時にのみ結果を生成できる。先行技術では、G P C Aアルゴリズムを導入するのに、監督的なユーザの指導を仮定している。これにより、G P C A技法の潜在力が大幅に制限されてしまう。

【0142】

本発明は、G P C A方法の概念的基礎を拡張して、ノイズおよび混合した余次元（mixed co-dimension）がある多重部分空間の識別およびセグメント化を頑健に取り扱う。これより、最先端技術では、G P C A技法に監督が不要であるという改良がもたらされる。

【0143】

先行技術では、Veroneseマップの多項式の法線ベクトルを、これら法線ベクトルの接空間を意識せずに、処理する。本発明の方法は、Veroneseマップを通常見つける法線ベクトルの空間に直交する接空間を見つけるように、G P C Aを拡張する。この「接空間」、すなわちVeroneseマップの部分空間は、次に、Veroneseマップの因子として用いられる。

【0144】

接空間は、平面波展開法、および位置と接空間座標の間のLegendre変換によって、特定される。この接空間座標は、形状的オブジェクトの表現、特にVeroneseマップの多項式への法線のタンジェントの表現の二重性を明らかにするものである。離散Legendre変換が、法線ベクトルに相当する制約された導関数形式を定義するように、凸解析によって適用される。これは、ノイズが存在する法線ベクトルの計算によって、データベクトルをセグメント化するのに用いられる。この凸性解析がG P C Aに組み入れられて、より頑健なアルゴリズムを供する。

【0145】

本発明は、G P C Aを適用する際に、繰り返し因子分解法を利用する。特に、先行技術

における導関数に基づく実装が、全く同一のここで記述するGPCA方法によって分類されたデータベクトルの組合せを絞り込むために拡張される。繰り返し適用することで、本技法が、Veroneseマッピングにおける候補の法線ベクトルを頑健に見つけるのに用いられる。因子化ステップで、絞り込まれたベクトルセットに関連する元のデータが、元のデータセットから取り除かれる。残されたデータセットが、この革新的なGPCA技法で同様に解析されてもよい。この革新的な方法が、監督されない方法においてGPCAアルゴリズムを用いるのに不可欠である。図11は、データベクトルの再帰的な絞込みを示す。

【0146】

GPCA技法の新規な拡張は、Veronese多項式ベクトル空間において多重根が存在する場合に、極めて効果的である。さらに、従来技術では縮退する場合、すなわちVeroneseマップにおける法線がベクトル空間軸に平行である場合でも、本方法は縮退しない。

10

【0147】

図10は、基本的な多項式近似および差分の方法を示す。

【0148】

<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、元のビデオシーケンスが得られる。

20

【0149】

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。

【0150】

30

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、ほぼ元のビデオストリームが生成される。

【0151】

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。

40

【0152】

50

< 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 5 3 】

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 】

【 0 1 5 4 】

【 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 7 】 局所正規化に用いられる本発明のメッシュ生成方法を示すブロック図である。

【 図 8 】 局所正規化に用いられる本発明の正規化方法に基づくメッシュを示すブロック図である。

【 図 9 】 本発明の全体正規化および局所正規化の方法の組合せを示すブロック図である。

【 図 10 】 本発明の基本的な多項式近似および差分の方法に基づく G P C A を示すブロック図である。

【 図 11 】 本発明の再帰的 G P C A 絞込み方法を示すブロック図である。

【 図 1 】

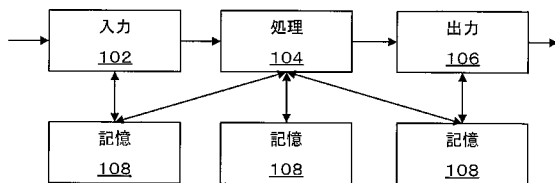


FIG. 1
Prior Art

【 図 2 】

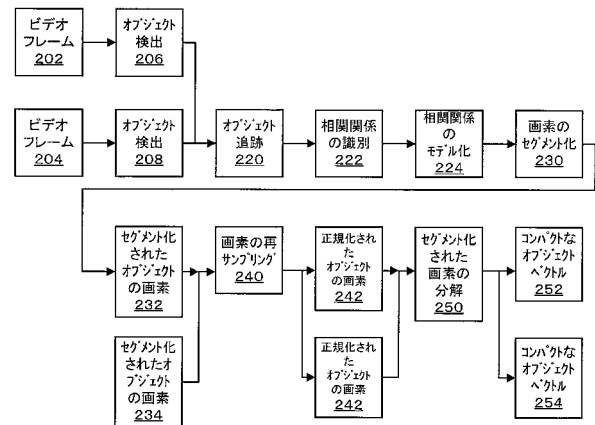


FIG. 2

【図 3】

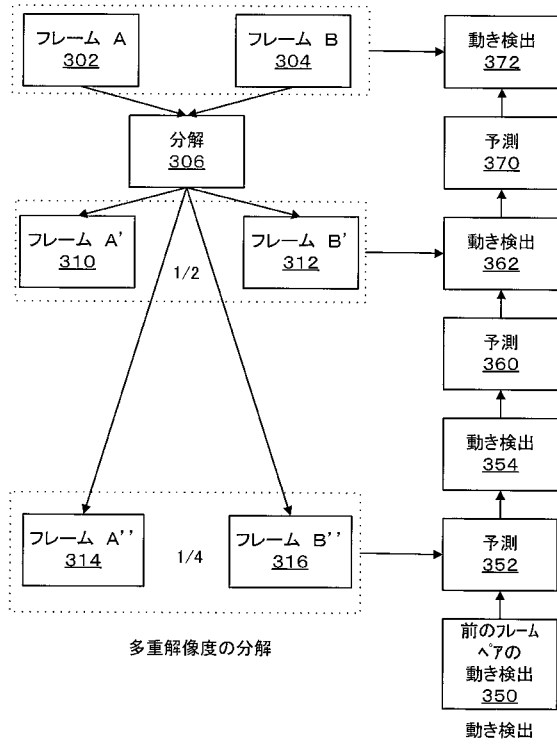


FIG. 3

【図 4】

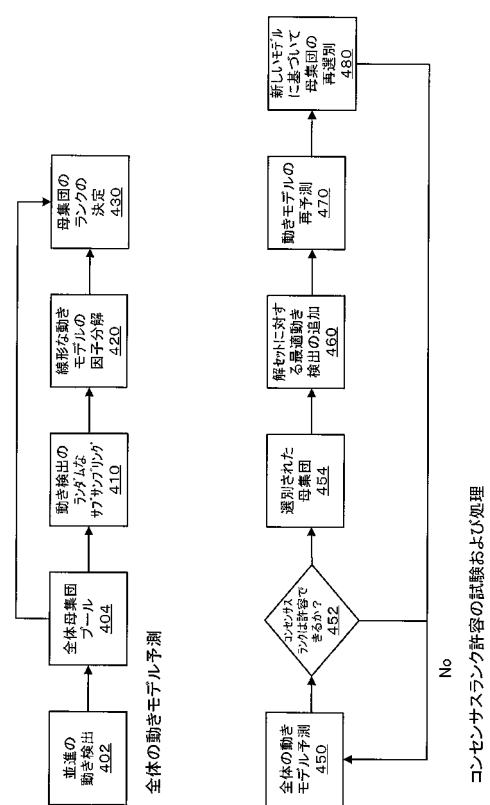


FIG. 4

【図 5】

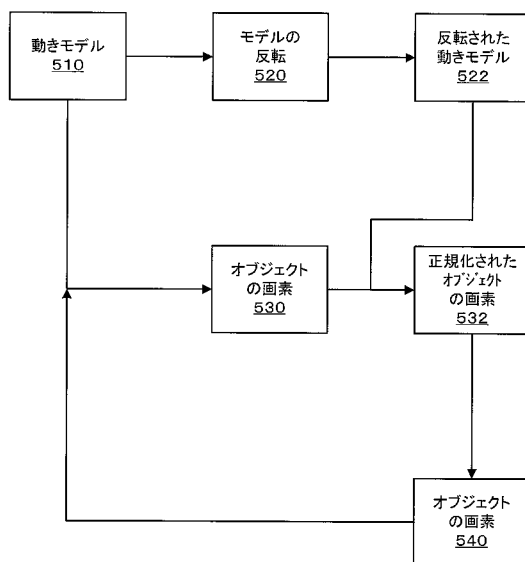


FIG. 5

【図 6】

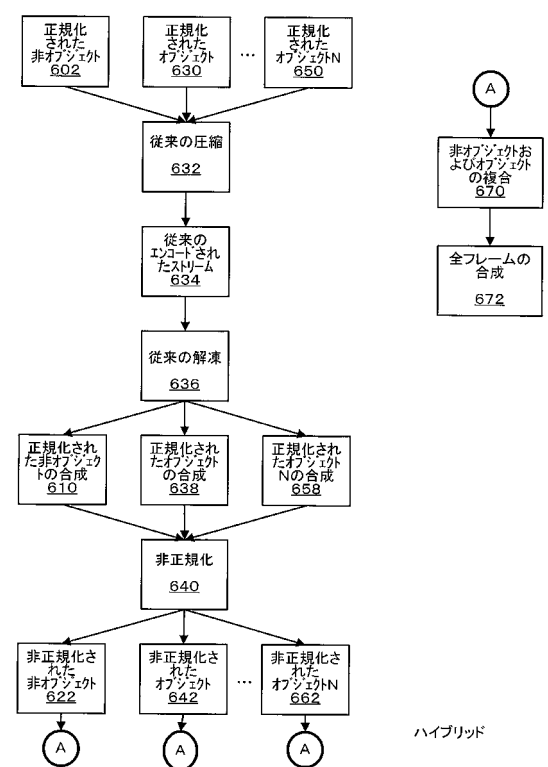


FIG. 6

【図 7】

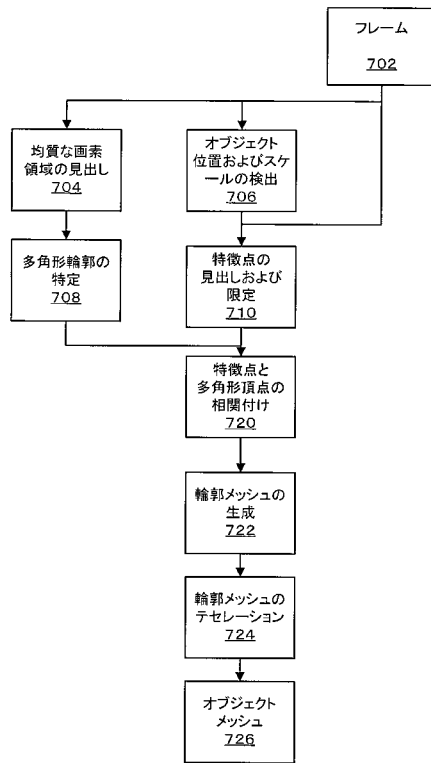


FIG. 7

局所正規化メッシュ生成

【図 8】

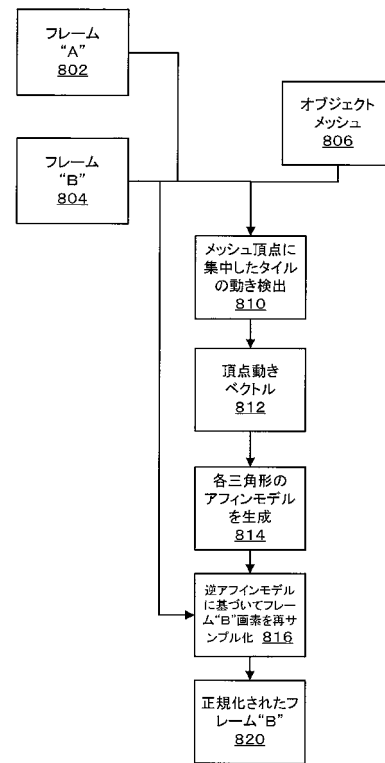


FIG. 8

局所正規化メッシュを基礎とした正規化

【図 9】

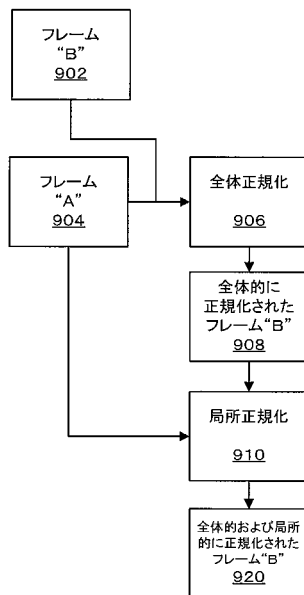


FIG. 9

全体正規化と局所正規化の組合せ

【図 10】

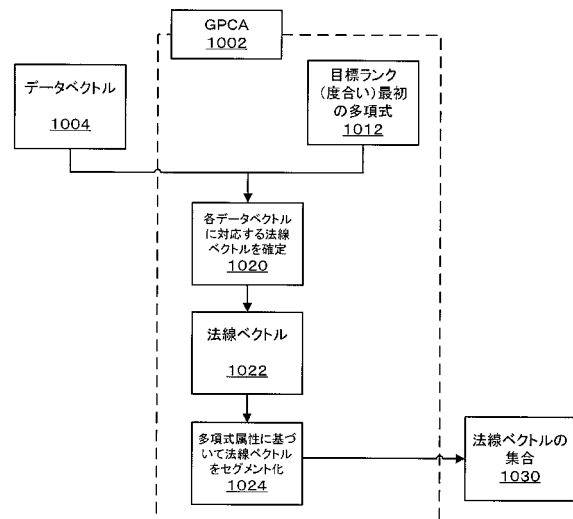
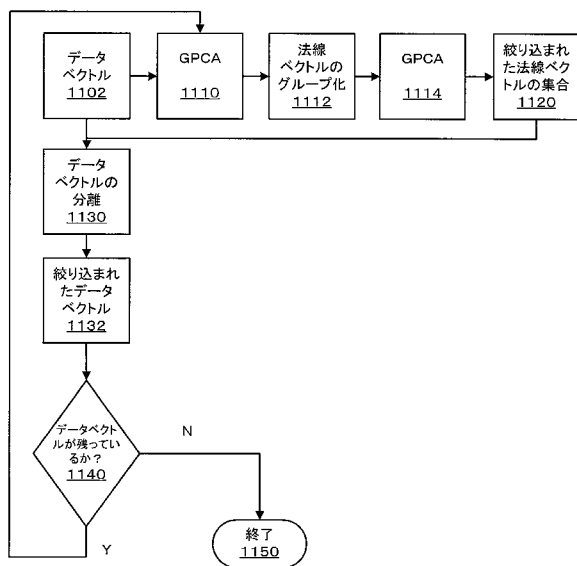


FIG. 10

GPCA—基本的な多項式近似および差分

【図 11】



GPCA—再帰的なGPCA絞り込み

FIG. 11

【 国際調査報告 】

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US05/41253

| A. CLASSIFICATION OF SUBJECT MATTER IPC: G06K 9/36(2006.01),9/46(2006.01),9/00(2006.01) USPC: 382/236,103 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. : 382/236,103 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) column 1, line 13 - column 6, line 42. | 1-9 |
| 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. | 1-9 |
| 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 |
| Y | US 2003/0194134 A1 (Wenzel et al.) 16 October 2003 (16.10.2003), Abstract. | 6 |
| <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" |
| "E" | earlier application or patent published on or after the international filing date | "X" |
| "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" |
| "O" | document referring to an oral disclosure, use, exhibition or other means | "&" |
| "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 24 August 2006 (24.08.2006) | | Date of mailing of the international search report 25 OCT 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 <i>Renai</i> Telephone No. (571) 272-5448 |

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US05/41253

C. (Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|------------|--|-----------------------|
| Y | Toklu, Candemir, A. Murat Tekalp, and A. Tanju Erdem. "Simultaneous Alpha Map Generation and 2-D Mesh Tracking for Multimedia Applications". Proceedings of the International Conference on Image Processing: 1997. 26-29 october 1997. Volume 1. Pages 113-116. | 8 |
| Y | Urban, Martin. "Harris Interest Operator". http://cmp.felk.cvut.cz/cmp/courses/dzo/resources/lecture_harris_urban.pdf 28 January 2003. | 9 |
| Y | Vidal, Rene, Yi Ma, and Shankar Sastry. "Generalized Principal Component Analysis (GPCA)". Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition: 2003. Volume 1. Pages 621-628. | 10 |
| Y | US 2004/0107079 A1 (MacAuslan) 3 June 2003 (03.06.2004), paragraphs 6 and 9. | 10 |
| Y | US 6,661,004 B2 (Aumond et al.) 9 December 2003 (09.12.2003), column 5, lines 11-13. | 10 |

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, KN, 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

(74)代理人 100144082

弁理士 林田 久美子

(72)発明者 ウェイス・ジョン

アメリカ合衆国, マサチューセッツ州 0 2 4 7 6 , アーリントン, グランド ビュー ロード
4 9

(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 0 5 7 6 3 , ノース チッテンデン, スミス ファーム ロード
7 0

F ターム(参考) 5C059 LB05 MB02 MB04 MB12 UA02