

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4836065号
(P4836065)

(45) 発行日 平成23年12月14日(2011.12.14)

(24) 登録日 平成23年10月7日(2011.10.7)

(51) Int.Cl.		F I			
G06T	1/00	(2006.01)	G06T	1/00	340B
G06T	3/00	(2006.01)	G06T	3/00	500A
G06T	7/20	(2006.01)	G06T	7/20	200Z

請求項の数 5 外国語出願 (全 17 頁)

(21) 出願番号	特願2005-100997 (P2005-100997)	(73) 特許権者	393031586
(22) 出願日	平成17年3月31日 (2005.3.31)		株式会社国際電気通信基礎技術研究所
(65) 公開番号	特開2006-285358 (P2006-285358A)		京都府相楽郡精華町光台二丁目2番地2
(43) 公開日	平成18年10月19日 (2006.10.19)	(74) 代理人	100099933
審査請求日	平成20年2月28日 (2008.2.28)		弁理士 清水 敏
		(72) 発明者	ペドラム・アザド
			京都府相楽郡精華町光台二丁目2番地2
			株式会社国際電気通信基礎技術研究所内
		(72) 発明者	アレス・ウデ
			京都府相楽郡精華町光台二丁目2番地2
			株式会社国際電気通信基礎技術研究所内
		(72) 発明者	ゴードン・チェン
			京都府相楽郡精華町光台二丁目2番地2
			株式会社国際電気通信基礎技術研究所内

最終頁に続く

(54) 【発明の名称】 エッジトラッキング方法及びそのコンピュータプログラム

(57) 【特許請求の範囲】

【請求項1】

対象物の一連の2D画像中においてエッジをトラッキングする方法であって、

前記対象物の3D形状を3D空間内のエッジにより記述するモデルを準備するステップと、

前記対象物のモデルのエッジを前記画像の画像平面に投影するステップと、

前記画像平面上に投影されたエッジの各々の周囲の予め定められた区域内で前記画像のエッジ画素の組を探索するステップと、

前記投影されたエッジの各々のエッジ画素の組について回帰直線を求めることにより、前記エッジ画素の組の各々に線分を当てはめるステップと、

前記投影されたエッジの各々について、前記線分を規定する所定のパラメータの組でその線分を表すステップと、

前記線分を用いて対象物のモデルを更新するステップとを含み、

前記モデルは、前記モデルのとり得る複数個の形状と、前記モデルが当該形状を取る確率とから求められた平均形状で記述され、

前記更新するステップは

前記モデルの形状に対し、前記確率により規定されるパーティクルフィルタを適用することにより、前記モデルの形状の新たな組を予測するステップと、

前記モデルの形状の新たな組の各々のエッジの組と、前記エッジ画素の組に対して求められた前記線分との間で所定の誤差尺度を計算するステップと、

10

20

前記新たな組の各形状に、前記計算するステップで計算された関連の誤差尺度に従って、前記対象物が各形状である確率を割当てるステップと、
それぞれの確率で重み付けされた新たな形状の総和を計算することによって、新たなモデルを計算するステップとを含む、エッジトラッキング方法。

【請求項 2】

前記探索するステップが、

前記投影されたエッジの各々について探索点の組を決定するステップと、

前記投影されたエッジの各々の前記探索点の各々について、投影されたエッジの両側において、前記投影されたエッジと交差する方向の予め定められた距離内に存在する、画像の高コントラスト点を探索するステップとを含む、請求項 1 に記載の方法。

10

【請求項 3】

前記投影されたエッジと交差する前記方向は、前記投影されたエッジと直交する方向である、請求項 2 に記載の方法。

【請求項 4】

前記決定するステップが、

前記投影されたエッジの各々について、一定の距離をあけて前記探索点の組を決定するステップを含む、請求項 2 または請求項 3 に記載の方法。

【請求項 5】

コンピュータ上で実行されると、請求項 1 ~ 請求項 4 のいずれかに記載の全てのステップを当該コンピュータに実行させる、コンピュータプログラム。

20

【発明の詳細な説明】

【技術分野】

【0001】

この発明はモーションキャプチャシステムに関し、特に、高フレームレートに適する、一連の画像中の対象物のエッジをトラッキングすることに関する。

【背景技術】

【0002】

マーカなしで人体のモーションキャプチャをするという考えは、なんら付加的な構成を必要とせず、画像シーケンスのみの操作によって人の動きをキャプチャしようというものである。このようなシステムのためのハードウェアは、コンピュータに加えて、1台又は2台以上のカメラのみからなり、このため手ごろな価格で容易に設定できる。

30

【0003】

しかしながら、ヒューマノイドロボットの能動型の頭部にこれを適用するためには、いくつかの制限に対処しなければならない。比較的短距離に配置された2台のカメラという限定に加えて、能動型の頭部は潜在的に動く可能性があり、このため前景のセグメント化にオプティカル・フローまたは類似のアルゴリズムが利用できないことを念頭に置く必要がある。最大の問題は、少なくとも30Hz、最適な結果を達成するためには好ましくは60Hzまたはそれ以上のフレームレートでリアルタイムに計算を行わなければならないことである。

【0004】

40

モーションキャプチャの問題の一般的な定義は、入力画像の各々について、又は画像のタプルについて、基となる関節で連結された3Dの人体モデルの正確な形状を見出すこと、というものである。主な問題は、人体モデルの自由度(Degree Of Freedom: DOF)の数が増えるにつれて、探索空間が指数関数的に増大することである。人間の体の現実的なモデルは少なくとも25DOFを有する。DOFがこれほど高いと、探索空間の次元が非常に高くなる。

【非特許文献1】A. ブレーク及びM. イサード、「アクティブ輪郭」、スプリンガー、1998年(A. Blake and M. Isard. Active Contours. Springer, 1998.)

【非特許文献2】J. ドイチャー、A. ブレーク、及びI. レイド、「アニールされたパーティクルフィルタリングによる関節により連結された物体のモーションキャプチャ」、

50

コンピュータビジョン及びパターン認識 (CVPR)、2126-2133 ページ、ヒルトンヘッド、USA、2000年 (J. Deutscher, A. Blake, and I. Reid. Articulated body motion capture by annealed particle filtering. In Computer Vision and Pattern Recognition (CVPR), pages 2126-2133, Hilton Head, USA, 2000.)

【非特許文献3】J. ドイチャー、A. ダビソン、及びI. レイド、「接続された物体のモーションキャプチャに関連する高次元探索空間の自動的パーティショニング」、コンピュータビジョン及びパターン認識 (CVPR)、669-676 ページ、カウアイ、USA、2001年 (J. Deutscher, A. Davison, and I. Reid. Automatic partitioning of high dimensional search spaces associated with articulated body motion capture. In Computer Vision and Pattern Recognition (CVPR), pages 669-676, Kauai, USA, 2001.)

【非特許文献4】M. イサード及びA. ブレーク、「濃縮-視覚的トラッキングのための条件付き密度伝搬」、コンピュータビジョン国際ジャーナル、29(1):5-28、1998年 (M. Isard and A. Blake. Condensation - conditional density propagation for visual tracking. International Journal of Computer Vision, 29(1):5-28, 1998.)

【非特許文献5】J. マコーミック、「視覚的トラッキングのための確率モデル及び確率的アルゴリズム」、オックスフォード大学、博士論文。UK、2000年 (J. MacCormick. Probabilistic models and stochastic algorithms for visual tracking. PhD thesis, University of Oxford, UK, 2000.)

【非特許文献6】J. マコーミック及びM. イサード、「パーティション化されたサンプリング、関節により連結された物体、及びインターフェース品質の手のトラッキング」、コンピュータビジョンヨーロッパ会議 (ECCV)、3-19 ページ、ダブリン、アイルランド、2000年 (J. MacCormick and M. Isard. Partitioned sampling, articulated objects, and interface-quality hand tracking. In European Conference Computer Vision (ECCV), pages 3-19, Dublin, Ireland, 2000.)

【非特許文献7】H. サイデンブラー、「単眼のビデオシーケンスにおける3Dの人の動きの確率論的トラッキングと再構築」、王立工科大学博士論文、ストックホルム、スウェーデン、2001年 (H. Sidenbladh. Probabilistic Tracking and Reconstruction of 3D Human Motion in Monocular Video Sequences. PhD thesis, Royal Institute of Technology, Stockholm, Sweden, 2001.)

【発明の開示】

【発明が解決しようとする課題】

【0005】

従って、この発明の目的の1つは、人間の体等の動く対象物を、一連の画像でトラッキングすることを可能にする高フレームレートに適した方法を提供することである。

【課題を解決するための手段】

【0006】

この発明の一局面に従えば、対象物の一連の画像中においてエッジをトラッキングする方法は、対象物の形状を記述するモデルを準備するステップと、対象物のモデルのエッジを画像の画像平面に投影するステップと、画像平面上に投影されたエッジの各々の周囲の予め定められた区域内で画像のエッジ画素の組を探索するステップと、投影されたエッジの各々のエッジ画素の組について線分を当てはめるステップと、投影されたエッジの各々について、線分を規定する所定のパラメータの組でその線分を表すステップと、線分を用いて対象物のモデルを更新するステップとを含む。

【0007】

上述の方法によれば、計算量が実質的に減じられ、一連の画像中での対象物の画像トラッキングを高フレームレートで行なうことができる。

【0008】

10

20

30

40

50

好ましくは、探索するステップは、投影されたエッジの各々について探索点の組を決定するステップと、投影されたエッジの各々の探索点の各々について、投影されたエッジの両側において、投影されたエッジと交差する方向の予め定められた距離内に存在する、画像の高コントラスト点を探索するステップとを含む。

【0009】

より好ましくは、投影されたエッジと交差する方向は、投影されたエッジと直交する方向である。

【0010】

さらに好ましくは、決定するステップが、投影されたエッジの各々について、一定の距離をあけて探索点の組を決定するステップを含む。

【0011】

上述のモデルは、モデルのとり得る複数個の形状とそれぞれに関連する確率とで記述されてもよい。更新するステップは、モデルの形状の確率分布関数のうち、選択されたもののパーティクルにノイズを加えることにより、モデルの形状の新たな組を予測するステップと、モデルの形状の新たな組の各々のエッジの組と、エッジ画素の組との間で所定の誤差尺度を計算するステップと、形状の新たな組の各形状に、計算するステップで計算された関連の誤差尺度に従って重みを割当てするステップと、それぞれの重みで重み付けされた新たな形状の総和を計算することによって、新たなモデルを計算するステップとを含む。

【0012】

この発明の別の局面は、コンピュータ上で実行されると、上述の方法のいずれかの全てのステップをコンピュータに実行させるコンピュータプログラムに関するものである。

【発明を実施するための最良の形態】

【0013】

1.はじめに

この実施例では、ベイズのフレームワークによるパーティクルフィルタリングを用いて現在の形状の確率分布を計算する。濃縮アルゴリズムとしても知られるパーティクルフィルタリングは、非特許文献4で紹介されているように、一般に輪郭トラッキングのために（非特許文献1、5、6）、特に非特許文献2及び7において示されるように人体のモーションキャプチャのために、適用可能な頑健な技術であることがわかっている。

【0014】

視覚的トラッキングにおいて、パーティクルフィルタは以下のように動作する。 x_t をt番目のフレームの状態ベクトルとし、 y_t を観測ベクトルとし、 $Y_t = \{y_1 \dots y_t\}$ を y_t の履歴とする。状態の推定は、確率密度分布(pdf) $p(x_t | Y_t)$ を推定することとして定式化される。パーティクルフィルタでは、 $p(x_t | Y_t)$ は以下で示すディスクリートのサンプルのいくつかの組として表される。

【0015】

【数1】

$$\{(s_t^{(i)}; \pi_t^{(i)})\} (i = 1 \dots N)$$

ここで $s_t^{(i)}$ ($i = 1 \sim N$) は、状態空間 x_t におけるランダムでディスクリートのサンプルを示す。サンプル $s_t^{(i)}$ は対応の重み $\pi_t^{(i)}$ に対応する確率を有する。従って、 $p(x_t | Y_t)$ で非ガウス pdf を近似することが可能である。

【0016】

パーティクルフィルタでは、各サンプルは推定ステップと観測ステップとで更新される。推定するステップでは、サンプル s_{t-1} が、重み $\pi_{t-1}^{(i)}$ に比例する確率に従って、以下の組から選択される。

【0017】

【数2】

$$\{(s_{t-1}^{(i)}; \pi_{t-1}^{(i)})\}$$

10

20

30

40

50

その後新たなサンプル $s_t^{(i)}$ が動力学モデル $p(x_t | s_{t-1} = s_{t-1})$ から生成される。観測ステップでは、現時点での観測ベクトル y_t が与えられると、観測された密度 $p(y_t | x_t = s_t)$ に基づいて各サンプル $s_t^{(i)}$ の尤度が評価され、尤度に比例する対応の重み $w_t^{(i)}$ が計算される。最後に、現在の状態が、全てのサンプルを組合せることによって推定される。

【0018】

パーティクルフィルタでは、探索空間が大きくなるとより多くのパーティクルを必要とする。この問題に対処する1つの戦略は、被験者の動きの範囲を、例えば被験者が歩いていると仮定することによって限定し、それによって形状空間の大きさを減じることである。一般に、このようなアプローチはいずれも、結果として得られるトラックの一般性を大いに限定する。このため、このアプローチは好ましくない。別の方策は、分解、すなわち階層的な探索を行なうことによって形状空間の大きさと探索空間のサイズとの線形関係を仮定することである。このような方法の短所は、ある種の動き、例えば体軸周りの回転等では、体の全ての部分を同時に観察することによって得られるような重要な情報が失われることである。

10

【0019】

パーティクルの数を減じる一般的かつ有効な方法は、非特許文献2、3で提示されたシミュレーテッドアニリングの考え方である。しかし、最終的に得られたシステムは部屋の角の固定位置で3台のカメラを用いるものであり、1GHzのクロック信号で動作する最先端のプロセッサでも1フレームを処理するのに平均で15秒を要する[非特許文献3]

20

【0020】

この実施例は高度に最適化された尤度関数を備えた人体のモーションキャプチャシステムに関するものであり、これは従来の人体のモーションキャプチャシステムのいずれに対しても2倍から3倍の速度を達成する。この実施例は高速キャプチャリングに寄与するパーティクルフィルタによるエッジ検出に基づくものである。このアプローチは上述の戦略等、他の最適化のいずれに対しても独立して適用可能である。

【0021】

さらに、この実施例で採用されるアルゴリズムは直線エッジでモデル化され得るいかなる種類の対象物にも利用可能であり、またわずかな修正を加えれば楕円等の他の基本要素を組入れることもできる。このアプローチの唯一の制限は、これが60Hz（またはそれ以上）の高フレームレートに依存することである。しかし、以下で示すように、リアルタイムの人体のモーションキャプチャではいずれにせよ高フレームレートが必要とされるので、これは問題にならない。

30

【0022】

2. 基本アルゴリズム

人体モデルの形状 s の投影エッジと現在の入力画像 z とが与えられると、人体のモーションキャプチャシステムのエッジの手がかりのための尤度関数 $p(z | s)$ は、投影されたエッジの組を与える形状が適切な形状である尤度、すなわちエッジ画像に最も合致する尤度を計算する。ここで人体モデルの形状のエッジを「投影する」とは、エッジを予め定められた画像平面に投影し、画像平面上の投影されたエッジの画像の座標を計算することを意味する。

40

【0023】

図1を参照して、基本的な技法は、投影されたエッジ20の全体にわたり、一定の距離をあけて探索点を決定し、それぞれの探索点から、投影されたエッジ20に直交する（両側の）方向で、固定された探索距離内の高コントラスト特徴点30を探索すること、すなわちカメラ画像中でエッジ画素を見出すことである[非特許文献4]。この目的のために、カメラ画像は通常、階調の勾配を用いたエッジ検出器により前処理され、画像エッジ22が生成される。

【0024】

50

尤度は、二乗誤差合計 (Sum of Squared Difference: SSD) に基づいて計算される。表記の便宜上、全てのエッジは $M = L / \Delta$ 個の範囲に分けられて1つの連続したスプライン内に含まれるものとする。ただし、 L は現在の画像中の投影されたエッジの全ての長さの合計を表す。 m 番目の点についてエッジとしての特徴が見出された距離を d_m とし、 μ は、エッジとしての特徴を見出すことができない場合に適用される一定の最大誤差を表す。従って、尤度関数は以下の式で与えられる。

【0025】

【数3】

$$p(\mathbf{z} | \mathbf{s}) \propto \exp \left\{ -\frac{1}{2\sigma^2 M} \sum_{m=1}^M f(d_m, \mu) \right\} \quad (1)$$

10

ただし $f(v, \mu) = \min(v^2, \mu^2)$ である。

この実施の形態では、わずかに修正した尤度関数を用いて、より良好な結果を得る。その尤度関数とは、エッジごとの平均誤差を計算し、すべての誤差の平均誤差を計算するというものである。この修正した尤度関数を以下の式で表す。

【0026】

【数4】

$$p(\mathbf{z} | \mathbf{s}) \propto \exp \left\{ -\frac{1}{2\sigma^2} \sum_{k=1}^E d(e_k) \right\} \quad (2)$$

20

ここで e_k は特定の投影されたエッジを表し、 E は投影されたエッジの総数を表す。さらに、

【0027】

【数5】

$$d(e) = \frac{1}{M_e} \sum_{m_e=1}^{M_e} f(d_m, \mu) \quad (3)$$

30

キャニーのエッジ検出器 (Canny Edge Detector) 等の高レベルエッジ検出器は、一般に、1画素幅の線を生成するので、ある特定の画素がエッジの一部であるかどうかのテストを複雑にするか、又は不適當にする。ドイツャーは非特許文献2で、階調の勾配を用いたエッジ検出マスクを用い、この結果をしきい値処理してスプリアスエッジを除去し、これに続いてガウス平滑化マスクを適用して画素マップを作成することを提案しており、このマップでは各画素値はそのエッジにどの程度近接しているかに関連づけられている。こうすることによって、画素があるエッジの一部であるか否かのテストと、そのエッジへの近接度の計算とが1回の動作で解決できる。最大値を見出すためとループ管理のために必要とされるさらに3つの動作を考慮すると、基本的動作の数は以下の式で表される。

40

【0028】

【数6】

$$T(L) = 8 \cdot \delta \cdot \frac{L}{\Delta} \quad (4)$$

3. 新たなアルゴリズム

この実施の形態の戦略は、処理ステップをもう一歩進めて、対象エッジの幾何学的な表現を抽出することである。このために、この実施の形態では、オン-ザ-フライで(すぐさま)のエッジ検出という基本的な考え方を用いる。これは、人体モデルによって与えられる情報をエッジ検出アルゴリズムに組込むというものである。主な違いは、この実施の

50

形態ではパーティクルの各々に対してアルゴリズムを行なうのではなく、これを画像の前処理に用いることである。その後このアルゴリズムは、人体モデルの投影されたエッジに対応する点の各組に当てはまる線分（回帰直線）を求めることで拡張される。

【0029】

少なくとも30Hz、好ましくは60Hzまたはそれ以上の高フレームレートで画像が獲得されると仮定すれば、このアプローチはかなり高速であるばかりか、オン・ザ・フライのエッジ検出より良好な結果を達成する。一見したところ、これは驚くべきことであるように思われる。というのも、理論的には、オン・ザ・フライのエッジ検出は、エッジを検出するために人体モデルの予測された形状を用いて観測を行なうために、予測を用いることによる最強の情報基盤を有するからである。しかし、欠点は、比較的大きな動きの予測ではエッジの同定誤りにつながるおそれがあることである。フレームレートがこの発明の新たなアプローチに十分なほど高くなければ、このアプローチはうまくいかない。一般的な低フレームレートの効果はセクション4で論じる。

10

【0030】

3.1 直線の当てはめ

人体モデルをエッジ検出に組込む主な利点は以下の2つである。

【0031】

- ・階調の勾配を、予測されるエッジと垂直に計算できる。

【0032】

- ・検出されたエッジ画素と投影されたエッジとのマッピングが得られる。

20

【0033】

検出された画素と投影されたエッジとのマッピングが、人体モデルの投影されたエッジの各々についてわかっているので、アルゴリズムで特定された画素の組に対して回帰直線を見出すことができる。この組に対して最小二乗という意味で最適回帰直線を求めた後、この組の全ての画素について反復処理し、計算された直線までのそれらの距離を予め規定されたしきい値と比較することによって、アウト라이어（異常値）をフィルタリングによって除去する。最終的な結果はこのフィルタリングされた画素の組に対する回帰曲線を求めることで計算される。

【0034】

この実施の形態では、画像の階調の勾配は予測されるエッジと垂直な方向で計算される。しかしながら、この方向は、必ずしも予測された方向と垂直でなくてもよい。高コントラストの特徴点の探索は、予測されたエッジと交差する方向に行なえばよい。

30

【0035】

3.2 誤差尺度

直線の対の組が得られた後、尤度の計算のために誤差尺度を定義することになる。理論的には、2Dで2つの直線間の距離を定義できるのは、2直線が平行の場合のみである。しかし、図2で示すように、線Eと線分/AB40又は42（ここで「/」は線分を表すものとし、本来文字ABの上に表記すべきものである。）との尺度を以下のように定義できる。基本的アルゴリズムによって計算される誤差尺度は、ディスクリートの点での観測値に基づいて計算された、直線Eと線分/AB40又は42との平均距離である。この平均距離は、図2に例示されるように、総面積を長さ/AB'で除したものに等しい。従って、誤差尺度は以下のように計算される。

40

【0036】

【数7】

$$d(e, \overline{AB}) = \frac{|d(e, A)| + |d(e, B)|}{2} \quad (5)$$

本発明者らは、この尺度について可能な組合せ2つを試みた。Eを画像から抽出された直線と定義し、/ABを投影されたエッジとする組合せと、その逆とである。後者の場合、抽出された直線の各々について端点A及びBを定義しなければならない。これは図3に

50

示すように、フィルタされた点の組のうち最初と最後の点を画像のエッジを表す直線に直交投影することでなされる。本発明者らは、このようにすることが最良の結果を達成することを見出した。なぜなら、これは画像エッジの長さの推定値を尤度関数に組込んでいるからである。長さの推定が悲観的であるため、結果として得られる誤差尺度は楽観的になる。実際、これはエッジの幾何学的表現に対して行なわれる式(3)で表される基本アルゴリズムの正確なエミュレーションである。式(2)で適用される新たな関数は最終的に以下のように定義される。

【0037】

【数8】

$$d'(e) := d(e, \overline{AB_e}). \quad (6)$$

10

別の可能な誤差尺度は、図2に示される線Eと線分/ABとの間の面積であろう。

【0038】

図3は誤った形状に誤差尺度を適用した例を示す。点 A_1 と B_1 との間、及び点 A_2 と B_2 との間の白の点は、検出されたエッジ画素を特定し、線60及び62は回帰直線を示し、線50及び52は人体モデルから投影されたエッジを示す。入力画像はガウスカールによる 3×3 平滑フィルタで平滑化されている。

【0039】

上述のエッジ検出は、プロップ(かたまり状の画像)のエッジ検出に拡張可能である。プロップモデルを楕円または円で表す場合、探索点は楕円線上の一定距離だけ離れた点として決定され、各探索点において、投影された楕円線の両側のある距離内で画像エッジが探索される。その後、検出された画像エッジ点に楕円を当てはめることができる。この場合、誤差は楕円の重心間の距離によって測定できる。

20

【0040】

3.3 有効性

前処理ステップが成功し、エッジが正確に抽出されれば、本形態の新たな誤差尺度はより正確なので、より有効なアルゴリズムが得られる。エッジを成功裏に抽出するためには、以下の2つの条件が満たされなければならない。

【0041】

- ・画像中のエッジはほぼ直線でなければならない。

30

【0042】

- ・フレームレートは十分高くなければならない。

【0043】

通常を着衣を仮定すれば、第一の条件が満たされることは、発明者らのテストで示された。この仮定は適切なものである。なぜなら、基本アルゴリズムでさえ、あまりにゆったりとした着衣ではトラックの性能が悪くなるからである。テストでは、30Hzのフレームレートで十分であり、これで現実的な速度の動きがキャプチャされる。フレームレートの問題は、セクション4で詳細に検討する。この実施の形態で採用される新たなアルゴリズムがより効果的である理由は以下のように説明できる。

【0044】

- ・直線を抽出するのでノイズが抑制される。

40

【0045】

・直線に基づいて誤差尺度を計算するので、完璧でないエッジであっても正確な尤度が計算される。

【0046】

- ・フレームレートに比例する速度のみを許すことで、誤ったエッジの検出が避けられる。

【0047】

これらの考察は、実際に正しいことが検証された。テストでは、この実施の形態で採用された新たなアルゴリズムの性能が基本アルゴリズムより劣ることは決してなく、実際の

50

ところは、多くの場合より良好な結果を達成した。

【0048】

3.4 効率性

このセクションでは、基本アルゴリズムの効率を新たなアルゴリズムのそれと理論上で比較する。この考察は、セクション6で提示するテスト結果で検証する。最大解像度は、 $\alpha = 1$ を選択することで達成される。以下の考察では、 $\alpha = 1.5$ を選択する。投影された全エッジの長さの合計は、ここで提示された画像シーケンスでは $L = 1500$ (画素) である。

【0049】

基本アルゴリズムでは、セクション2で概要を述べた計算を、フレームレート f で各フレームの各パーティクルについて行なわなければならない。これは、 N 個のパーティクルでは $N f T (1500) = 1.8 \cdot 10^5 N f$ 回の演算を、毎秒行なうことを意味する。計算に要する作業については、画像の前処理は考慮していない。なぜなら、画像処理は新たなアルゴリズムでも同様に行なわなければならないからである。新たに必要となったエッジ抽出の作業は、セクション6の結果で検証されるように、無視できるものである。新たなアルゴリズムでは、各パーティクルに対し、式(5)を評価する必要がある。パーティクル1個当たりの演算数は約30であり、そのほとんどが浮動小数点演算である。総計算作業量は

$$T' = 90 \quad (7)$$

個の基本演算に相当するものであり、これは毎秒 $60 N f$ の作業量を意味する。基本アルゴリズムで必要とされる基本演算数と比較すると、検討された例ではこれは $T (1500) / T' = 2 \cdot 10^3$ 倍の高速化を意味する。しかし、システム全体では、実際の高速化はこれを下回る。なぜなら、この最適化された実現例では、人体モデルの運動と結果として得られる2Dの輪郭との計算もまた、各パーティクルに対して行なう必要があり、これがボトルネックとなるからである。尤度関数の計算とシステムとの高速化の実際については、セクション6で示す。

【0050】

4. 低フレームレートの影響

入力画像シーケンスのフレームレートが低いと、どのようなマーカ無しの人体のモーションキャプチャシステムにとっても不利となる。以下では、同じ動きを第1のフレームレート f_1 と、 $f_2 < f_1$ である第2のフレームレート f_2 でキャプチャし、この結果、画像シーケンス V_1 及び V_2 がそれぞれ得られたと仮定する。この場合、 V_2 での連続画像の関節部の各々での角度変化は、 V_1 と比べて f_1 / f_2 倍だけ大きい。総探索空間は DOF の数 d によって指数関数的に増大するので、探索空間のサイズ $s(V_1)$ と $s(V_2)$ との関係は以下で与えられる。

【0051】

【数9】

$$\|s(V_2)\| = \left(\frac{f_1}{f_2}\right)^d \|s(V_1)\| \quad (8)$$

これは、このシステムにおいてフレームレートを半減すると 2^{28} 倍のパーティクルが必要となる、という意味ではなく、定性的な結論として、2倍のフレームレートで動作させれば実際の探索空間を大いに減じることができる、という意味である。従って、パーティクルの数が同じでフレームレートが2倍ならば、人体のモーションキャプチャシステムの精度と解像度とはかなり高くなる。

【0052】

高フレームレートのもう1つの大きな利点は、誤ったエッジを検出する危険性が当然のことながら低くなることである。一般に、コンピュータビジョンでは、より小さい探索空間は同時に2つの長所を持つ。すなわち、より高い「効率」と、より高い「頑健性」とである。例えば、立体視の場合にエピポーラジオメトリが関連処理をかなり高速化するだけ

10

20

30

40

50

でなく、誤った対応の確率を減じると同様、人体のモーションキャプチャシステムでも、高フレームレートの結果として得られる、より小さい探索空間は、誤った形状の確率を減じるとともに、より頑健なシステムを結果としてもたらす。

【 0 0 5 3 】

人体のモーションキャプチャシステムのエッジの手がかりに関する良い目安は、連続した画像間での、体部品の主軸に対し直交する方向の距離が、この体部品の幅の半分以下である、というものである。そうでないと、形状を誤る確率が比較的高くなる。なぜなら、体部品の2個のエッジのうち1個がすでに、間違っただけのエッジと対応付けられやすくなっているからである。この実施の形態で用いられる最適化されたアルゴリズムでは、2倍から3倍の高速化が達成されるが、エッジは予測とは「独立に」抽出され、そのため、この条件が満たされることが厳格に求められる。結論として、フレームレートが高ければ、より効率的で、より精度が高く、より頑健なシステムが得られると言え、その差は単に線形の倍率 f_1 / f_2 だけでは説明できない。

10

【 0 0 5 4 】

5 . 実現例

5 . 1 構造

図4はこの発明の一実施の形態に従った人体のモーションキャプチャ (Human Motion Capture : HMC) システム 1 2 4 を示す。図4を参照して、HMC システム 1 2 4 は、ロボットの頭部に比較的短い距離をおいて装着された2台のカメラからの2つの画像 1 2 0 及び 1 2 2 を受け、最も確からしい人体モデルのパラメータ 1 2 6 を出力する。2台のカメラを使うこと、すなわち画像対のシーケンスを処理することで、リアルな3Dの動きを抽出することが可能である。

20

【 0 0 5 5 】

図4を参照して、HMC システム 1 2 4 は、画像 1 2 0 及び 1 2 2 を平滑化するとともにノイズを除去するための一対の 3×3 ガウス平滑化モジュール 1 4 0 と、エッジ (図3に示すエッジに当てはまる線の A 及び B の対) 及び肌の色の重心を抽出し、着衣の色のマップと深度計算マップとを計算するための一対の前処理モジュール 1 4 2 と、前処理モジュール 1 4 2 によって抽出され計算された情報に基づく動力学モデルを用いて、前回の確率密度関数のパーティクルにノイズを加えることにより、新たな人体モデルの形状を予測するための新形状組予測モジュール 1 4 4 と、形状を、前処理モジュール 1 4 2 から出力された情報と比較することにより、右と左のカメラにつきそれぞれ予測された形状 s_i の各々について、確率密度関数 $\{ (s_i , \pi_i) \}$ を計算する、一対の尤度計算モジュール 1 4 6 とを含む。

30

【 0 0 5 6 】

HMC システム 1 2 4 はさらに、尤度計算モジュール 1 4 6 によって計算された確率密度関数 $\{ (s_i , \pi_i) \}$ を記憶するための PD (確率密度) 記憶部 1 4 8 と、人体モデルの平均形状

【 0 0 5 7 】

【 数 1 0 】

$$\sum_{i=1}^N s_i \pi_i \quad (\text{ただし } N \text{ は新たな形状の数 (パーティクルの数) である定数})$$

40

を、PD記憶部 1 4 8 に記憶された確率密度関数に基づいて計算し、計算された平均形状を人体モデルパラメータ 1 2 6 として出力するための平均形状計算モジュール 1 5 0 と、次のフレームの前処理のために、平均形状計算モジュール 1 5 0 によって計算された平均形状

【 0 0 5 8 】

【数 1 1】

$$\sum_{i=1}^N s_i \pi_i$$

を記憶するための平均形状記憶部 1 5 2 とを含む。平均形状記憶部 1 5 2 は前処理モジュール 1 4 2 に接続され、前処理モジュール 1 4 2 はこの平均形状を利用して、次の画像 1 2 0 及び 1 2 2 を前処理する。同様に、新形状組予測モジュール 1 4 4 は、P D 記憶部 1 4 8 に記憶された確率密度関数を利用して、次の画像フレームのための新たな形状の組を予測する。

【0 0 5 9】

図 5 は前処理モジュール 1 4 2 の詳細を示す。図 5 を参照して、前処理モジュール 1 4 2 は、形状 s_i の各々について、カメラ画像 1 2 0 及び 1 2 2 内において、人体モデルから投影されたエッジから距離 d 以内でエッジ画素を見出し、これらのエッジ画素に当てはまる線分を求め、この線分の 2 つの端点に関する情報を出力するためのエッジ抽出モジュール 1 6 0 と、各画像内で、予め定められた肌の色を持つプロップの座標の重心を抽出するための肌色重心抽出モジュール 1 6 2 と、動きをキャプチャしようとするユーザの着衣の色のマップを計算するための着衣色マップ計算モジュール 1 6 4 と、2 つの画像 1 2 0 及び 1 2 2 に基づいて、画像の深度マップを計算するための深度マップ計算モジュール 1 6 8 とを含む。これらのモジュール 1 6 0、1 6 2、1 6 4 及び 1 6 8 の出力は全て、新形状組予測モジュール 1 4 4 に与えられる。

【0 0 6 0】

図 6 は図 5 に示すエッジ抽出モジュール 1 6 0 を実現するコンピュータプログラムの制御構造を示す。図 6 を参照して、ステップ 1 8 0 で、人体モデルの平均形状から投影された全てのエッジについて、ステップ 1 8 2 ~ 1 9 0 が繰返される。これらのステップが全てのエッジについて繰返されたあと、制御はこのルーチンから出る。

【0 0 6 1】

ステップ 1 8 2 で、処理すべき投影エッジの全体にわたり、一定の距離 d をあけて探索点が決定される。ステップ 1 8 4 で、ステップ 1 8 6 が全ての探索点について繰返される。ステップ 1 8 6 では、探索点を起点とし、投影されたエッジに直交する方向（の両方）の一定の探索距離 d 内で、高コントラスト特徴点を探索する。この動作により、カメラ画像中のエッジ画素が見出されることとなる。

【0 0 6 2】

ステップ 1 8 6 が全ての探索点について繰返されると、制御はステップ 1 8 8 に進む。ステップ 1 8 8 では、ステップ 1 8 6 で見出されたエッジ点に当てはまる線分（回帰直線）を求める。ステップ 1 9 0 で、この線分の 2 つの代表点 A 及び B（図 3 を参照）が決定される。

【0 0 6 3】

5. 2 動作

この実施例の H M C システム 1 2 4 は以下のように動作する。図 4 を参照して、2 つの画像 1 2 0 及び 1 2 2 が与えられると、ガウス平滑化モジュール 1 4 0 が各画像を平滑化し、ノイズを除去する。このように処理された画像は前処理モジュール 1 4 2 に与えられる。

【0 0 6 4】

前処理モジュール 1 4 2 では、図 5 を参照して、先のフレームサイクルで得られた、人体モデルの平均形状

【0 0 6 5】

【数 1 2】

$$\sum_{i=1}^N s_i \pi_i$$

10

20

30

40

50

について、エッジ抽出モジュール160がこの形状から投影されたエッジを計算し、投影されたエッジの各々全体にわたり、一定の距離をあけて探索点を決定し、探索点を起点として、投影エッジに直行する方向で一定の探索距離内の高コントラストな特徴を持つ点を探索し、これらエッジ画素に当てはまる線分を求め、線分の2つの代表点を決定し、これらの点の座標を出力する。

【0066】

同様に、肌色重心抽出モジュール162は画像中で肌の色を持つブロップの重心を出力し、着衣色マップ計算モジュール164は予め定められた着衣の色で画素のマップを計算し、深度マップ計算モジュール168は両方のカメラの画像から、深度マップを計算する。これらのモジュールの出力は全て、新形状組予測モジュール144に与えられる。

10

【0067】

新形状組予測モジュール144は、PD記憶部148に記憶された最後の確率密度関数 $\{(s_i, \pi_i)\}$ のパーティクルにノイズを加えることにより、動力的モデルを用いて新たな人体モデル形状の組を予測する。この新たな人体モデル形状の組は尤度計算モジュール146に与えられる。

【0068】

予測された形状の各々について、尤度計算モジュール146は前処理モジュール142から出力された情報と、対象となるモデル形状から投影された画像とを比較し、エッジ、肌の色重心、着衣色マップ、及び深度マップの誤差に基づいて確率を計算し、その後これらの確率を乗ずることにより、対象となるモデル形状の全体の確率を計算し、形状 s_i の各々について、確率密度関数 $\{(s_i, \pi_i)\}$ を出力する。

20

【0069】

PD記憶部148はこの確率密度変数を記憶する。

【0070】

平均形状計算モジュール150は、平均形状

【0071】

【数13】

$$\sum_{i=1}^N s_i \pi_i$$

30

を、平均形状計算モジュール150に記憶された確率密度関数に基づいて計算し、この平均形状を人体モデルパラメータ126として出力する。平均形状記憶部152は平均形状計算モジュール150の出力を記憶して、次のフレームサイクルの動作に備える。

【0072】

6. 結果

本発明者らは、この実施例の新たなアルゴリズムの効率を、基本アルゴリズムを利用したシステム及びオン・ザ・フライのエッジ検出に基づくアルゴリズムと比較した。結果を表1に示す。

【0073】

【表 1】
表 1

	新アルゴリズム	基本アルゴリズム	処理しながらのエッジ検出
ガウス平滑化及びグレースケール変換	3 ms	3 ms	3 ms
前計算(エッジ検出)	4 ms	4 ms	-
人体モデル更新	25 ms	25 ms	25 ms
尤度関数の評価	2 ms	4440 ms	6280 ms
総計	34 ms	4472 ms	6308 ms

表 1 は 2 . 8 G H z クロック信号で動作する最新のコンピュータでの効率を比較したものである。テストでは、1000個のパーティクルを用いた。値は1フレームの処理時間である。表 1 を参照して、この実施の形態に係るアルゴリズムは、2000倍以上の高速化を達成できる可能性がある。しかし、システム全体の現在の実現例では、人体モデルの運動及びその結果得られる輪郭の計算がボトルネックであり、このため、実質的な高速化は130倍となっている。

【0074】

60Hzでキャプチャされたビデオシーケンスのテストでは、この実施の形態で利用されたアルゴリズムは基本アルゴリズムを利用したシステム及びオン-ザ-フライのエッジ検出に基づくアプローチに比べ、さらに良好な結果を達成した。すなわち、この実施の形態で計算された尤度はより適切であると思われた。これを理論で証明するのは困難である。しかし、結果として得られた、計算された平均形状を投影したものを含むビデオ画像は、この有効性を明らかに示している。処理された8個のフレームのシーケンスを図7に示す。

【0075】

図7は60Hzでキャプチャされた8枚の画像のシーケンスを例示する。左から右へ、1列目が画像1-4、2列目が画像5-8である。人体モデルの計算された平均形状の投影されたエッジが白のエッジで表示されている。30Hzでキャプチャされたビデオシーケンスも、このシステムがこのような低フレームレートでも動作することを示した。

【0076】

7. 結論

この明細書で、発明者らは、パーティクルフィルタを用いた、エッジトラッキングのための、高度に最適化された尤度関数を備えた人体のモーションキャプチャシステムを提示した。用いられたアルゴリズムを詳細に説明し、効率と有効性とを検討し、通常のアプローチと理論的に比較し、結果を実際に検証した。この実施の形態で用いられるアルゴリズムの唯一の付加的な条件である、少なくとも30Hz、好ましくは60Hz以上という高いカメラフレームレートは、制限というよりはむしろ、いかなるリアルタイムの人間モーションキャプチャシステムでも、理にかなった選択というべきであろう。

【0077】

セクション6で示したように、この実施の形態で用いられるアルゴリズムでは、尤度関数を通常のアプローチに比べ2000倍も速く、オン-ザ-フライのエッジ検出に基づくアプローチに比べ3000倍も速く計算する。ここで、システム全体の有効な高速化は通常のアプローチに比べ約130倍であると述べた。この実施の形態で利用した尤度関数では、さらなる高速化のためのボトルネックは人体モデルとなる。

【 0 0 7 8 】

本発明者らが知る限りでは、この実施の形態で用いたアルゴリズムによってはじめて、マーカ無しでリアルタイムの人体のモーションキャプチャが実現可能なものとなった。このアルゴリズムは、ヒューマノイドロボットの能動型の頭部に適用するための人体のモーションキャプチャシステムに重要な役割を果たすであろう。

【 0 0 7 9 】

今回開示された実施の形態は単に例示であって、本発明が上記した実施の形態のみに制限されるわけではない。本発明の範囲は、発明の詳細な説明の記載を参酌した上で、特許請求の範囲の各請求項によって示され、そこに記載された文言と均等の意味および範囲内でのすべての変更を含む。

【 図面の簡単な説明 】

【 0 0 8 0 】

【 図 1 】 人体モデルから投影されたエッジから、どのように画像エッジが求められるかを概略的に示す図である。

【 図 2 】 線分と線の誤差尺度とを概略的に示す図である。

【 図 3 】 画像の 2 つのエッジに当てはめられた線分と人体モデルから投影された 2 本の線との結果を示す図である。

【 図 4 】 この発明の一実施例に従った人体のモーションキャプチャシステム 1 2 4 のブロック図である。

【 図 5 】 図 4 に示す前処理モジュール 1 4 2 の詳細図である。

【 図 6 】 図 5 に示すエッジ抽出モジュール 1 6 0 を実現するコンピュータプログラムの制御フロー図である。

【 図 7 】 この発明の実施例によって処理された 8 個のフレームシーケンスを示す図である。

【 符号の説明 】

【 0 0 8 1 】

5 0 及び 5 2 投影されたエッジ

6 0 及び 6 2 回帰直線

1 2 0 及び 1 2 2 画像

1 2 4 H M C システム

1 2 6 人体モデルパラメータ

1 4 0 ガウス平滑化モジュール

1 4 2 前処理モジュール

1 4 4 新形状組予測モジュール

1 4 6 尤度計算モジュール

1 4 8 P D 記憶部

1 5 0 平均形状計算モジュール

1 5 2 平均形状記憶部

1 6 0 エッジ抽出モジュール

1 6 2 肌色重心抽出モジュール

1 6 4 着衣色マップ計算モジュール

1 6 8 深度マップ計算モジュール

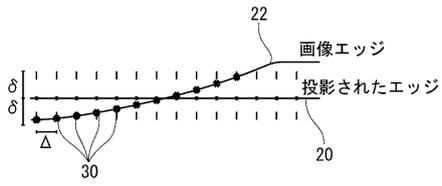
10

20

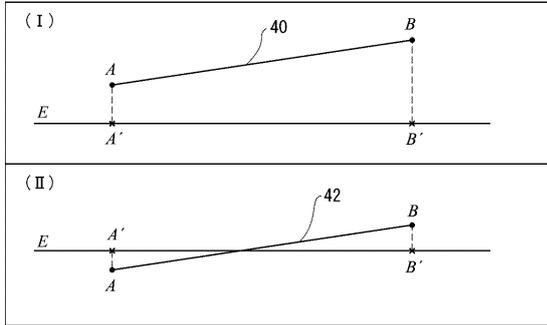
30

40

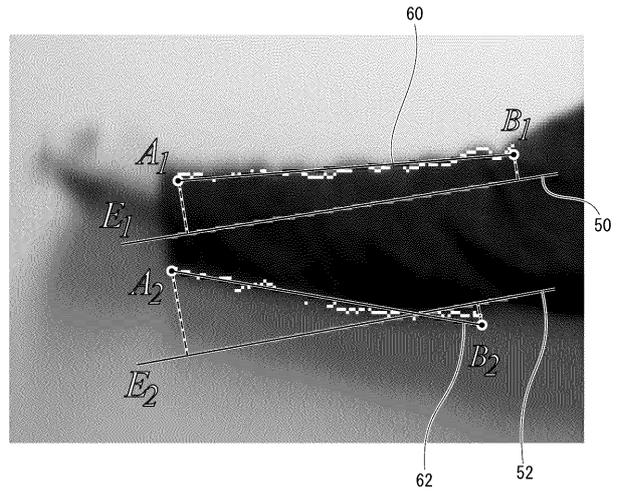
【図1】



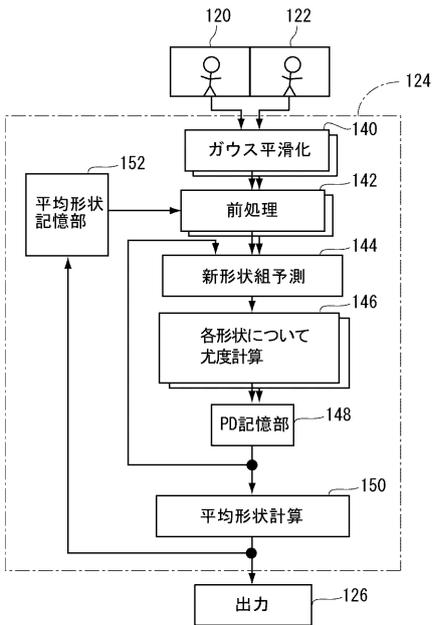
【図2】



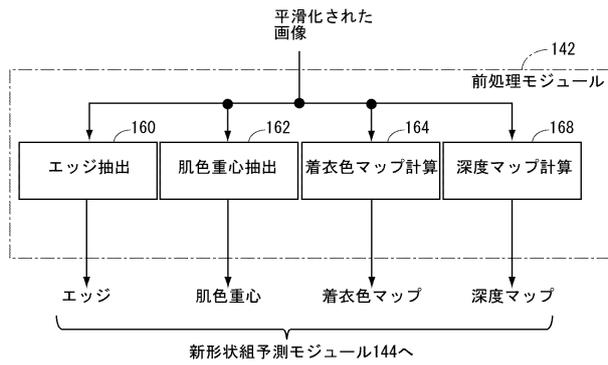
【図3】



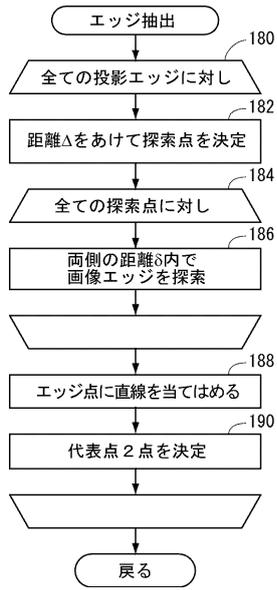
【図4】



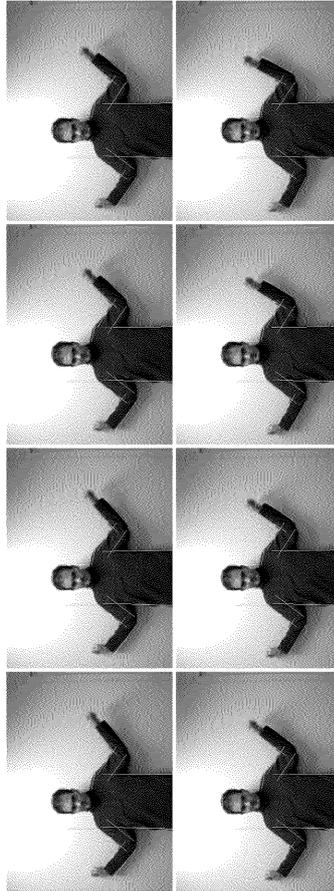
【図5】



【図6】



【図7】



フロントページの続き

審査官 秦野 孝一郎

(56)参考文献 特開2001-266159(JP,A)
特開2003-216951(JP,A)
特開2001-005973(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06T 1/00

G06T 3/00

G06T 7/20