

複数カメラによるリアルタイム・モデルベースド モーショントラッキング

森 敬充(九州大学) 倉爪 亮(九州大学) 村上 剛司(九州大学) 長谷川 勉(九州大学)

Real-time model-based motion tracking using multiple cameras

*Takamitsu MORI(Kyushu University), Ryo KURAZUME(Kyushu University),
Kouji MURAKAMI(Kyushu University), Tsutomu HASEGAWA(Kyushu University)

Abstract—This paper proposes a method for fast model-based human motion tracking by multiple cameras using 2D-Distance-Field. The 2D-Distance-Field represents the errors between captured images and projected images of a 3D model on the image plane and can be constructed rapidly using the Fast Marching Method. Using the 2D-Distance-Field and human model, 3D motion of human is estimated robustly in real time. This paper shows some experimental results including simultaneous motion tracking of multiple people.

Key Words: Motion capture, Motion tracking, Multiple cameras, Level Set Method

1. はじめに

人間とロボットが共生する環境において、ロボットが自身に搭載したセンサや計算機のみを用いて多様な状況で自律的に活動することは、センシング能力や計算能力の限界から実現困難である。これに対し、「ロボットタウンの実証的研究」プロジェクトでは、街中に分散配置された多数のセンサを用いて環境状況を認識し、その結果をロボットに伝えてロボットの行動を支援するシステムを開発している。本プロジェクトにおいて、多数のカメラで撮影された画像から、人間の位置や移動方向、手足の動きなどを計測する技術は、人間とロボットの安全なインタラクションを実現するための主要な検討課題の1つである。そこで本研究では、分散配置されたカメラを用いて人物動作を推定するシステムを開発する。

カメラ画像から人物動作を推定する手法として、モデルベースド・モーショントラッキング、すなわち人体の3次元モデルとカメラ画像上の人物像を比較することで人物の位置姿勢を推定する手法がある。本手法の多くは、まず適当な初期姿勢で人体モデルをカメラ画像上へ投影し、その投影像とカメラ画像との間でエッジや輪郭線などの画像特徴量を比較し、それらが最も良く一致する人体モデルの姿勢を探索するものである。しかし画像特徴量の比較に必要な計算コストが大きく、特に多数台のカメラを用いた場合、リアルタイム処理の実現には計算量の削減が大きな課題である。一方、円滑なロボット・人間のインタラクション支援を実現するには、動作推定の遅れは最小限にする必要がある。そこで本研究では、カメラ画像上で高速に構築可能な2次元距離場を用い、複数カメラから得られる姿勢パラメータの修正量を3次元モデル上で統合することにより、複数台のカメラを用いてもリアルタイム処理が可能なモーショントラッキング・システムを開発する。

これまでに、C.Sminchisescuらによってシルエットベースの2次元距離場を用いた手法 [1] が提案されているが、本研究では2次元距離場を粗密かつ輪郭線ベース

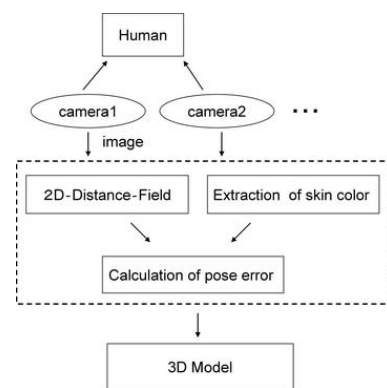


Fig.1 Proposed system

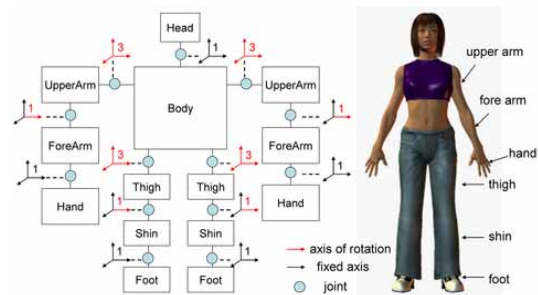


Fig.2 3D body model

で構築することで、より高速かつ高精度なモーショントラッキングを実現するものである。

2. 2次元距離場を用いたモーショントラッキングシステム

開発した手法は、背景差分により得られたカメラ画像上の人物の輪郭線に対して、2次元距離場を計算し、画像上に投影された3次元人体モデルの輪郭線上の距離値から姿勢修正量を求め、動作の推定を行う。システム構



original image contour image 2D-Distance-Field
Fig.3 Construction of 2D-Distance-Field

成を Fig.1 に示す .

2.1 3次元人体モデル

本研究で使用する3次元人体モデルは、Fig.2 に示すように、14個のリンクが13個の関節によって連結され、全21自由度を有している。なお、Fig.2において座標系の数字は各関節での自由度を示している。ただし今回の実験では、3次元人体モデルの位置姿勢は、胴体中心の3次元位置と胴体中心を通る床面に垂直な軸周りの回転角度、及び首、手首、足首を除く8個の関節（各腕2、各脚2、全16自由度）の回転角度によって決定されるものとする。

2.2 2次元距離場の構築と姿勢推定

提案手法において、2次元距離場 T とは、画面上の人物輪郭線 I_h からの最短距離を画面上の全ての点で求めたものである。画像上の点 $p = (x_p, y_p)$ での $T(p)$ は以下を満たす。

$$T(p) = 0, p \in I_h \quad (1)$$

$$|\nabla T(p)| = 1 \quad (2)$$

具体的な2次元距離場の構築方法について述べる。まず、カメラ画像から背景差分によって人物シルエット画像を取得し、これに Level Set Method[2] を適用して、人物輪郭線を抽出する。次に、抽出された輪郭線からの距離場 T を Fast Marching Method[2] によって構築する。2次元距離場の構築の様子を Fig.3 に示す。ただし、本手法では2次元距離場は輪郭線に近い領域（Distance Band）のみ詳細に構築し、輪郭線から離れた領域では粗に計算することで、構築処理を高速化している。

次に、構築された2次元距離場を用いた3次元人体モデルの位置修正量の計算手順を示す。ただし、使用した3次元人体モデルは、いわゆるサーフェスモデルであり、表面は微小な多数の3角パッチにより構成されている。

まず、3次元人体モデルを2次元距離場に投影する。ここで、3次元人体モデルの投影輪郭線上の点 $p_i = (x_{p_i}, y_{p_i})$ 、及びそれに対応した3角パッチを u_i とする。このとき、パッチ u_i に加えられる画面に平行な力 f_i 、及び3次元人体モデルの中心に加えられる力 F を以下のように求める。

$$f_{u_i} = \begin{bmatrix} T(p_i)T_x(p_i)/D(p_i) \\ T(p_i)T_y(p_i)/D(p_i) \end{bmatrix} \quad (3)$$

$$D(p_i) = \sqrt{T_x(p_i)^2 + T_y(p_i)^2} \quad (4)$$

$$F = \sum_{u_i} f_{u_i} \quad (5)$$

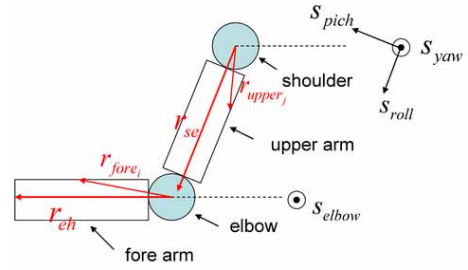


Fig.4 Arm model



frame 0 frame 1 frame 2
Fig.5 Examples of motion estimation

ただし、 T_x, T_y はそれぞれ T の x, y 方向の一次微分値である。求められた F を3次元人体モデルの中心に加えることで、位置を修正する。

次に、各関節の修正量を求める。ここでは簡単のため、Fig.4 に示す腕部分について考える。まず、3次元モデル前腕（手も含む）の投影輪郭線に対応するパッチを u_{fore_i} 、肘関節から u_{fore_i} へのベクトルを r_{fore_i}, u_{fore_i} に加えられる画面に平行な力を f_{fore_i} とし、肘関節周りに発生するモーメントを

$$M_{fore} = \sum_{u_{fore_i}} (r_{fore_i} \times f_{fore_i}) \quad (6)$$

として求める。これより、肘関節の回転軸方向を s_{elbow} とすると、肘関節の修正量は

$$\Delta\phi_{fore} = s_{elbow} \cdot M_{fore} \quad (7)$$

として求まる。

次に、3次元モデル上腕の投影輪郭線に対応するパッチを u_{upper_j} 、肩関節から u_{upper_j} へのベクトルを r_{upper_j}, u_{upper_j} にかかる力を f_{upper_j} 、肩関節から肘関節へのベクトルを r_{se} とする。これより、肩関節周りに発生するモーメントを

$$M_{upper} = \sum_{u_{upper_j}} (r_{upper_j} \times f_{upper_j}) + \sum_{u_{fore_i}} ((r_{fore_i} + r_{se}) \times f_{fore_i}) \quad (8)$$

として求める。さらに、肩関節の回転軸方向を $s_{roll}, s_{pitch}, s_{yaw}$ とすると、肩関節の修正量は

$$\Delta\phi_{roll} = s_{roll} \cdot M_{upper} \quad (9)$$

$$\Delta\phi_{pitch} = s_{pitch} \cdot M_{upper} \quad (10)$$

$$\Delta\phi_{yaw} = s_{yaw} \cdot M_{upper} \quad (11)$$

として計算できる。以上、Eq.(7), (9)~(11) を両腕両脚について求めることで、3次元人体モデルの姿勢を修正する。



Fig.6 Examples of body direction estimation

2.3 肌色情報の利用

手足が胴体などに重なって撮影された場合、本手法では手足の輪郭線を抽出することができず、2次元距離場だけでは姿勢修正量を正しく求めることができない。そこで、2次元距離場に加えて手足の肌色情報も利用する。

具体的には、3次元モデルをカメラ画像上に投影し、投影3次元モデルの手先足先位置から画像上の肌色位置へのベクトルを求め、これを3次元空間に逆投影した力を3次元モデルの手先足先に加えることで、各関節の修正量を求める。2次元距離場と肌色情報を用いた姿勢推定の様子を Fig.5 に示す。

2.4 人物の向きへの推定

次に、人物の向きへの推定手法について述べる。ただし、本システムでは3次元モデルの向きは胴体中心を通る床面に垂直な軸周りの回転角度 ϕ_{body} によって決定されるものとする。本節では人物向きへの推定手法として、1) 足先の向きを用いる方法と、2) 輪郭線を用いる方法の2つを提案する。

2.4.1 足先の向きからの推定

画像中の人物の足先の方向から体の向きを推定する手法について述べる。ただし本手法では、足首より下の足先が、前述の肌色抽出処理によりあらかじめ正確に切り出されていると仮定する。

まず、足の肌色領域を囲む最小矩形を求める。次に、その長辺を足の向きとし、両足の向きの合成ベクトルを画像中の人物の向きとして推定する。推定の様子を Fig.6 に示す。

2.4.2 輪郭線からの推定

人物輪郭線と3次元モデルの投影輪郭線を用いた向きの推定手法について述べる。本手法では、2.2章と同様に2次元距離場から3次元モデルの胴体、上腕、前腕の各3角パッチに加えられる力を求め、その合成モーメントを計算して人物の向きを推定する。

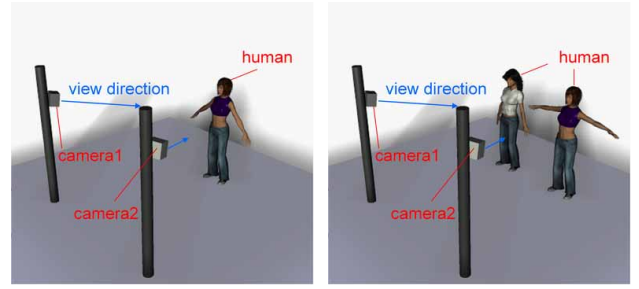
まず、3次元モデルの胴体、上腕、前腕(手も含む)の投影輪郭線に対応するパッチを u_k 、胴体中心から u_k へのベクトルを r_k 、Eq.(3) から求められる u_k にかかる力を f_k とし、胴体中心周りに発生するモーメントを

$$M_{body} = \sum_{u_k} (r_k \times f_k) \quad (12)$$

として求める。これより、胴体中心の床面に垂直な回転軸方向を s_{body} とすると、向きの修正量は

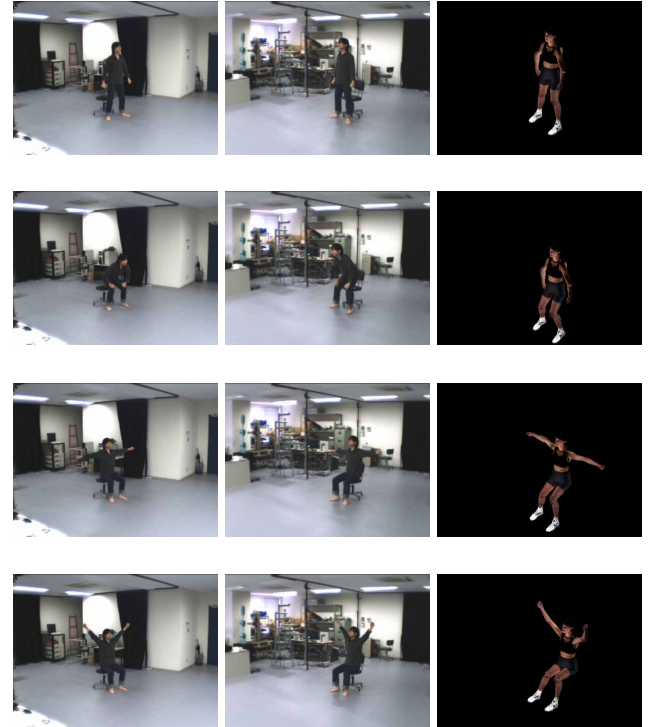
$$\Delta\phi_k = s_{body} \cdot M_{body} \quad (13)$$

として求まる。



(a) (b)

Fig.7 Experimental setup



camera1 camera2 3D model

Fig.8 Tracking results for sitting motion by two cameras

3. トラッキング実験

3.1 多視点画像によるリアルタイム・モーショントラッキング

複数カメラを用いた動作推定実験を行った。実験で使ったカメラは Dragonfly2(Point Grey Inc.) であり、画像の取得から動作推定まで全てオンラインで行った。画像の大きさが 320×240 の時の処理時間は、2次元距離場の構築が 8.7ms、姿勢修正量計算が 3.0ms である。実験環境を Fig.7(a) に、椅子に座る動作に対する推定結果を Fig.8 に示す。

3.2 複数人物動作の同時推定

2人の人物動作を同時に推定する実験を行った。実験環境を Fig.7(b) に、実験結果を Fig.9 に示す。これより、2人の人物に対しても、1人の場合と同様に動作推定が可能であることを確認した。なお、処理時間は、2

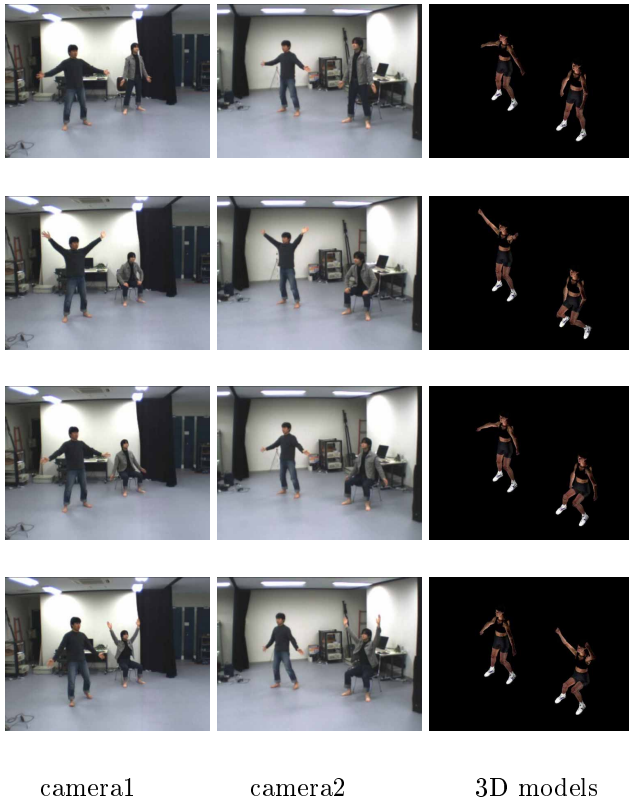


Fig.9 Tracking results of two people by two cameras

次元距離場の構築が 6.5ms, 姿勢修正量計算が 4.2ms である。実験環境が違うため単純な比較はできないが, 対象が 1 人の場合と比較して処理時間に大きな差が無いことがわかる。提案手法は, 一度 2 次元距離場を構築すれば, 追跡する輪郭数が増えた場合でも, 姿勢修正量の計算時間はほとんど変化しないという特徴があり, 複数人物動作の同時推定に特に適していると考えられる。

3.3 向きの推定実験

これまでの実験では, 対象が方向転換しないと仮定し, 人物の向きの推定は行っていない。そこで, 方向転換を含む歩行動作に対する人物方向の推定実験を行った。実験環境は Fig.7(a) と同様に人物 1 人に対してカメラ 2 台を配置している。足の向きからの推定結果を Fig.10 に, 輪郭線からの推定結果を Fig.11 に示す。ただし, 足先の白い直線は推定された胴体の向きを示している。これより, 両手法とも静止・移動時の方向転換に応じて, 胴体の向きの推定が可能であることがわかる。しかし, 足の向きからの推定は, 足先が正確に抽出される必要があり, 実用上は問題が多いと考えられる。

4. まとめ

本稿では, 輪郭線ベースの粗密 2 次元距離場を用いたリアルタイム・モーショントラッキング法を提案した。また実験により, 歩行動作や椅子に座る動作などが推定可能であること, また複数人の動作推定にも適用可能であることを確認した。さらに, 肌色情報及び 2 次元距離場を用いることにより, 人物の向きの推定が可能であることを確認した。

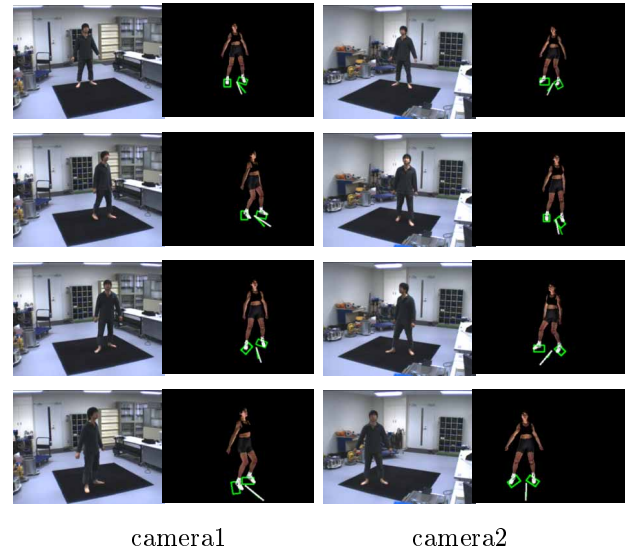


Fig.10 Estimation of body direction by foot direction

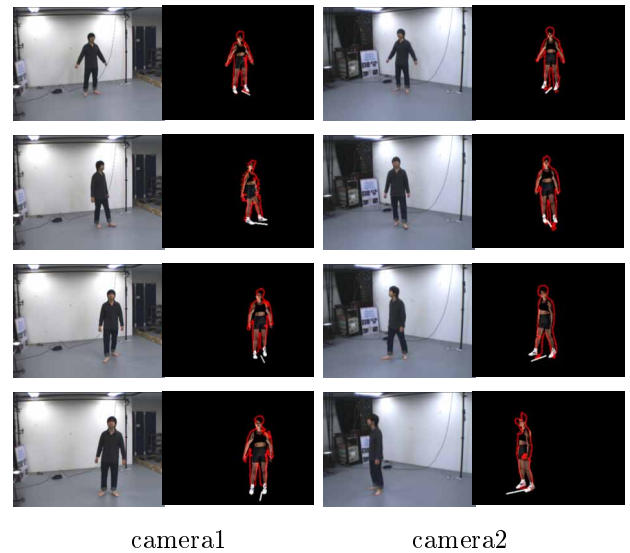


Fig.11 Estimation of body direction by contour lines

謝辞

本研究は, 文部科学省の平成 18 年度科学技術振興調整費による「科学技術連携施策群の効果的・効率的な推進」の一環として実施したものであり, また一部は文部科学省科研費補助金基盤 B 「ロボット群と環境固定センサ群の相互支援による人間・ロボット共生環境の情報構造化 (18360124)」および「5 感センサネットワークによる被災建造物内 RT システムの開発 (19360119)」の助成を受けて行われました。

- [1] C. Sminchisescu and A. Telea: "Human Pose Estimation From Silhouettes: A Consistent Approach Using Distance Level Sets", WSCG International Conference on Computer Graphics, Visualization and Computer Vision, 2002
- [2] J. Sethian: "Level Set Methods and Fast Marching- Methods, second edition", Cambridge University Press, UK, 1999