

逐次モンテカルロ法を用いた単眼画像列からの姿勢推定*

江本 光晴[†]
広島市立大学[†]

林 朗[‡]
広島市立大学大学院[‡]

末松 伸朗[§]
広島市立大学大学院[§]

神原 利彦[¶]
広島市立大学大学院[¶]

1. はじめに

視覚監視, 福祉ロボット, ヒューマン・インタフェースなどの分野において, コンピュータが人間の動作や姿勢を認識することは, 重要である. しかし, コンピュータで認識させることは難しい問題である. なぜなら, 人間の構造は複雑で多くの自由度を持ち, 人間の姿勢を画像平面へ投影したのものにはオクルージョン (自己隠蔽) が発生するからである.

これら問題を解決する一般的な手法として, 人体の構造モデルを用いて姿勢推定を行う手法がいくつか提案されている. これら手法は, 事前に人体の構造モデルを構築しておき, シルエット画像にモデルを当てはめることで姿勢の推定を行っている. そのため, 精巧な人体構造モデルが必要であり計算的負担が大きいという欠点がある.

そこで, 本研究では人体モデルを必要としない新しい推定手法を提案する. 本手法では, 姿勢と画像を位相と視点の関数として学習し, 逐次モンテカルロ法を用いて姿勢の推定を行う. また, カメラの視点は任意であり, 連続的に変化する視点でも姿勢の推定を行う.

2. 手法の概要

我々は単眼画像列から人体の姿勢推定を行うのにあたって, 人体モデルを必要としない学習アプローチを採用する. モーション・キャプチャのデータから作成した画像を訓練データとし, 姿勢と画像を位相と視点の関数として学習し, 逐次モンテカルロ法を用いて, 画像に対応する位相と視点を導出する. そして, その位相と視点から姿勢の推定をする.

画像から姿勢の推定を実現するには, 人体の持つ自由度の多さ, 異なる姿勢が視点によってはよく似た画像を与える曖昧さ, 自己隠蔽, 姿勢の空間での確率分布の複雑さ, 時間的空間的な解像度不足など, 克服しなければならない問題が多くある. 岩村 [1] は, 画像と姿勢の関係を画像と姿勢がなす高次元直積空間における混合ガウス分布で表現し, これらの問題点を解決するように試みた. しかし, この表現方法では, 分布パラメータを推定するのに, 大量の訓練データを必要とするうえ, 高精度の姿勢推定ができないことがわかった.

そこで, 本研究では, 姿勢と画像 (モーメント) の関係を以下のように表現する.

1. 周期運動を仮定し, 人体の姿勢を周期運動の位相の関数として表現する.

*Pose Estimation from monocular image sequence using sequential monte carlo

[†]Mitsuharu Emoto · Hiroshima City University

[‡]Akira Hayashi · Graduate School, Hiroshima City University

[§]Nobuo Suematu · Graduate School, Hiroshima City University

[¶]Toshihiko Kanbara · Graduate School, Hiroshima City University

2. 画像の特徴量 \mathcal{F} を位相と視点の関数 f として表現する.

$$\mathcal{F} = f(\text{phase}, \text{view})$$

なお, この関数は訓練データから生成する.

このとき, 姿勢推定を以下の手順で行う.

1. 画像が得られる.
2. 与えられた画像に対応する (位相, 視点) の候補を逐次モンテカルロ法を用いて求める.
3. 候補列の事後確率から (位相, 視点) を推定する.
4. (位相, 視点) から姿勢を復元する.

3. 逐次モンテカルロ法

本研究では, 事後確率分布の推定に逐次モンテカルロ法を用いる. 逐次モンテカルロ法は, 事後確率分布を直接近似するのではなく, m 個の重み付けされたサンプルの集合で事後確率分布を近似し, 推定をするアルゴリズムである [2].

逐次モンテカルロ法は各時刻 t において以下の手順を m 回繰り返し, 生成された m 個のサンプルにより事後確率分布の推定を行う. また, サンプルの初期集合 ($t = 0$) は, 一様分布からサンプリングされ, 均一な重み ($1/m$) で重み付けされる. なお, x_t は時刻 t の (位相, 視点) からなる状態空間のサンプルである.

1. 時刻 $t-1$ の状態の事後確率分布から x_{t-1} をサンプリングする.
2. 状態遷移の条件に従って, サンプル x_{t-1} から時刻 t のサンプル x_t を求める. ここでの状態遷移の条件は, 「動作はほぼ等速度で行われる」, 「向きは急に変わらない」になる.
3. 画像の位相と視点の関数を用いてサンプル x_t を平均とする確率密度分布を生成し, 与えられた画像 (モーメント) の確率密度を求め, サンプル x_t の重み付けを行う.

4. 画像列からの姿勢推定

4.1 訓練データ

モーション・キャプチャのデータから, 全体座標系における 3D マーカー位置からなるベクトルの集合 \mathcal{P}_{3d} が得られる. 3D マーカー位置を人体の中心 (腰) に高さを設定した正射影カメラを用いて画像平面に投影することで, 2D マーカー位置からなるベクトルの集合 \mathcal{P}_{2d} が得られる. この投影は, 3次元モデルの鉛直下方向の中心軸の周りに $\{\theta_i | \theta_i = \pi/16, 0 \leq i \leq 31\}$ 方向から行う. ここで, i は視点位置 i であり, 人間の進行方向に正対

するカメラの角度を 0, 視点位置 0 とし, 回転方向は時計回りである。

また, 3D マーカー位置が与えられると CG 技術を使ってレンダリングし, シルエット画像が得られる。

各方向の動作 1 周期分のシルエット画像から画像特徴ベクトルを求める。画像特徴ベクトルはモーメント特徴量であり, 本研究では並進およびスケールに不変な 0 次から 3 次までのモーメント特徴量を用いる。

4.2 姿勢の推定

動作の種類がわかっている場合での姿勢推定を考える。その動作の画像の位相と視点の関数を作成し逐次モンテカルロ法に与え, 与えられた画像列から求めた画像特徴ベクトルを入力とする。そして, 事後確率を最大にする (位相, 視点) を決定し, 推定を行う。

4.3 動作の分類

動作の種類がわからない場合での動作分類の問題を考える。動作を分類するため, 各動作ごとに画像の関数を作成し, 逐次モンテカルロ法に用いる。さらに, サンプルをそれぞれの動作ごとの状態空間に生成させ, パラレルに推定していく。そして, 事後確率を最大にする動作を選び, 動作の分類を行う。

4.4 動作の変化

「歩いてから走る」など動作が変化する場合での動作の推定を考える。4.3 節の方法では, 事後確率分布が 0 または極端に小さい動作が生じ, 次の時刻において, その動作での状態空間でサンプリングがされなくなり, それら動作が候補として取り上げられなくなるということが考えられる。そこで, 4.3 節に加え, それぞれのサンプル集合に新たなサンプルを追加することで, 動作の変化に対応させる。そして, 事後確率を最大にする動作を選び, 推定を行う。

5. 実験

5.1 姿勢の推定

訓練データとテストデータは同じものを使い, 動作「歩く」について実験を行った。図 1 に実験を 30 回行ったときの推定誤差 (正解と推定の (位相, 視点) 間のユークリッド距離) を示す。また, 推定した結果から復元した画像列を図 2 に示す。

図 1 から, 最初のフレームで大きくずれてしまった場合もあった。しかし, 常に誤推定されてはいないので, 上手く推定されているといえる。

5.2 動作の分類

「歩く」「スキップ」「走る」の 3 つの動作を訓練データとし, 動作「歩く」をテストデータとして実験を行った。図 3 に各動作の事後確率の推移を示す。なお, 実線が「歩く」, 破線が「スキップ」, 点線が「走る」である。

図 3 から, フレーム 1 で「スキップ」の事後確率が「歩く」より大きい, 次のフレームから「歩く」の事後確率が最大になっていることがわかる。ゆえに, 上手く分類されているといえる。

5.3 動作の変化

「歩く」「スキップ」「走る」の 3 つの動作を訓練データとし, 「歩いたのちスキップし, そして走る」という動

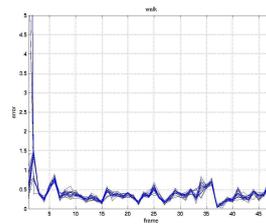


図 1: 推定誤差

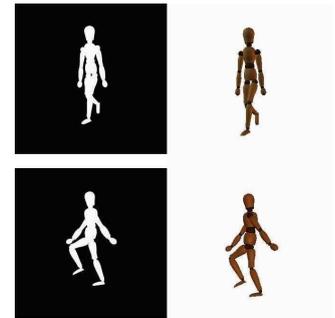


図 2: 姿勢推定の結果

作をテストデータとして実験を行った。図 4 に各動作の事後確率の推移を示す。なお, 実線が「歩く」, 破線が「スキップ」, 点線が「走る」である。

図 4 から「歩く」「スキップ」「走る」の順で事後確率が最大になっていることがわかる。上手く推定されているといえる。

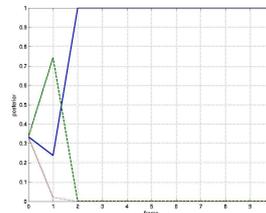


図 3: 動作の分類: 実線が「歩く」, 破線が「スキップ」, 点線が「走る」である

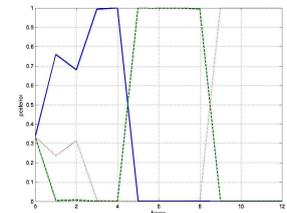


図 4: 動作の変化: 実線が「歩く」, 破線が「スキップ」, 点線が「走る」である

6. 結論および今後の課題

本研究では単眼画像列からの人体の姿勢推定を試みた。人体モデルを利用して姿勢を推定する研究は多いが, 我々は人体モデルの不要な学習アプローチを採用し, 姿勢と画像の関係を学習した。

画像から姿勢推定するには, 自由度の多さや曖昧さなど多くの問題がある。我々は, 姿勢を周期運動の位相の関数, 画像を位相と視点の関数とし, 逐次モンテカルロ法を用いて, 事後確率を最大とする (位相, 視点) を決定し, 推定を行う。実験により, 提案手法の有効性を示した。

今後の課題として, 同一人物での服装, 歩き方の違いや個人差などのノイズの評価をどう扱うかが挙げられる。

参考文献

- [1] 岩村篤樹. 学習を用いた単眼画像列からの姿勢の推定. 修士論文, 広島市立大学大学院 情報科学研究科, 2003.
- [2] Sebastian Thrun, Dieter Fox, Wolfram Burgard, and Frank Dellaert. Robust Monte Carlo Localization for Mobile Robots. Artificial Intelligence Journal, 2001.