

Bottom-UpとTop-Down アプローチの統合による 単眼画像からの人体3次元姿勢推定

3D Human Pose Estimation Integrating Bottom-Up and Top-Down Approach from Monocular Image

大西克則¹

Katsunori Onishi

滝口哲也²

Tetsuya Takiguchi

有木康雄²

Yasuo Ariki

神戸大学大学院工学研究科¹

Graduate School of Science and Technology, Kobe University

神戸大学自然科学系先端融合研究環²

Organization of Advanced Science and Technology

1 はじめに

本論文では、単眼画像から人体の3次元姿勢を推定する手法を提案する。画像特徴には、物体の大まかな形状を表すことが可能な HOG 特徴を用いる。また3次元姿勢は、人体を多関節モデルで表現しその各関節角で記述する。推定は、重回帰分析を用いた Bottom-up アプローチと、パーティクルフィルタを用いた Top-down アプローチを統合することで、より高精度な推定を行う。

2 特徴量

人体3次元モデル特徴では、人体の構造に基づいて各関節の角度を特徴量 x とする。本論文では、web で公開されているモーションキャプチャデータ [1] を用いた。

入力画像に対しては、あらかじめ背景差分法を用いて人物領域を切り出す。その際、画像サイズの正規化も行い、HOG 特徴 [2] を抽出する。

3 姿勢推定方法

人体の姿勢推定問題は、大きく分けて二つのアプローチがある。一つは画像を観測し、3D 姿勢を推定する Bottom-Up アプローチであり、もう一つは逆向きの推定を行う Top-Down アプローチである。本論文では、Bottom-Up で大まかな姿勢を推定し、Top-Down でより精度を向上させる、双方向から推定する手法を提案する。

3.1 重回帰分析

Bottom-Up アプローチでは、重回帰分析を用いて姿勢推定を行う。画像から得られた HOG 特徴ベクトル z と人体3次元モデルの特徴ベクトル x の関係を、次式で近似する。

$$x = Rz + \varepsilon \quad (1)$$

ε は残差ベクトルである。 R の値は n 個の学習データ $\{(y_i, x_i) \mid i = 1 \dots n\}$ を用いて、最小自乗推定法により決定する。

3.2 パーティクルフィルタ

Top-Down アプローチでは、パーティクルフィルタを用いて繰り返し推定を行うことで、真値に収束させる。

時刻 t における姿勢を状態量 x_t 、同時刻の画像特徴を z_t 、時刻 1 から t までの画像特徴を $Z_t = (z_1, \dots, z_t)$ とする。 x_t の事後確率分布 $p(x_t|Z_t)$ はベイズの定理により、式 (2) のようになる。

$$p(x_t|Z_t) = k_t p(z_t|x_t) p(x_t|Z_{t-1}) \quad (2)$$

事後確率は N 個のサンプル群とその尤度評価によって得られる重みを用いて、離散的に近似する。

重回帰分析により推定された姿勢を初期値とし、その周辺に粒子をサンプリングする。その後は時刻 $t-1$ の尤度に基づき、粒子をリサンプリングし状態遷移を繰り返す。複数回繰り返した後、最も尤度の高かった状態を出力する。尤度評価は、各粒子の状態量からシルエット画像を生成し、入力画像に対するシルエットと画素毎の一致度を用いる。

4 実験

3D 姿勢には、CMU のデータ [1] を用いた。その中から、歩く (316 フレーム)、走る (148 フレーム)、ボールを蹴る (801 フレーム)、の3種類の動作を学習データとして用いた。各動作に対して45度毎に8視点から見た画像をCGで作成し、一方向あたり1265、全方向で10120フレームを用いた。

図1に実験結果を示す。Bottom-Up では推定できていなかった部分も、Top-Down から推定することで、良好な結果が得られていることが確認できる。



(a) 入力画像



(b) Bottom-Up 推定



(c) Top-Down 推定

図1 実験結果

5 まとめ

本論文では、Bottom-Up、Top-Down の統合により、推定精度を向上させる手法を提案し、実験によりその有効性を示した。しかし、Top-Down により精度が悪化する場合がある。その原因として、尤度評価に用いる特徴量が考えられる。今後は、その特徴量と、Bottom-Up からの推定方法を改善していく予定である。

参考文献

- [1] CMU Human Motion Capture DataBase. Available online at <http://mocap.cs.cmu.edu/>.
- [2] N.Dalal and B.Triggs, "Histograms of Oriented Gradients for Human Detection," CVPR, 886-893, 2005.