シルエット画像を用いた人体の動作推定 (改定版)

A Human Motion Estimation Method Using a Silhouette Image Sequence

亀田 能成　　美濃 導彦　　池田 克夫

KAMEDA Yoshinari　MINOH Michihiko　IKEDA Katsuo

京都大学工学部

Faculty of Engineering, Kyoto University

1　はじめに

人間の身体の姿勢や運動を画像から認識する際の問題は、対象である様々な姿勢を取り得るので、画像のみからでは必要な情報が得られないことである。このため、予め対象となる身体の形状モデルを用意し、画像とのモデルマッチングにより認識を行う手法が一般的である。

筆者らはコンピュータグラフィックスの技法を用いて、人間の身体に代表される多関節物体の形状を正確にモデル化し、それを用いて単眼視のシルエット画像から姿勢推定を行う方法を提案している。この方法では、関節の動作する範囲と形状とがモデルに記述できるので、マッチング処理が簡潔になり、シルエットのみから姿勢推定が可能である。シルエット画像１枚のみでは隠蔽等によって一部の部位の姿勢推定が不可能であり、そうした部位についてはマッチングに矛盾を来さない可動範囲を推定している [1, 2]。

本稿では、上述の方法の動画像への拡張について述べる。人間の身体が運動を行う場合、慣性が生ずる。これを採り入れることで、各フレームにおける探索範囲の設定を行うと共に、運動の一貫性を考慮した手法を提案する。

2　人体モデル

人間の身体を表現する人体モデルは、15 個の剛体部品とそれらを結ぶ関節によって木構造で表現できる。関節を被覆する表面形状は関節角度によって規定される。全ての関節角度を決定すれば、人体モデルの姿勢が一意に定まる。

3　アルゴリズム

マッチング処理は、動画像の各フレームを二値化したシルエット画像に対して行う。

探索範囲の設定

人体は各部位とも慣性を持つと見なされるので、短い時間内では各関節は軸回りに一定の角速度で回転運動を行うと考えられる。実際には各角速度は徐々に変化していくので、その回転角加速度の度合によってどれだけの関節角度範囲を探索すべきかが決定される。

$$|r_p^a(t+\Delta t) - (r_p^a(t) + v_p^a(t)\Delta t)| \leq \frac{1}{2}\Delta v_{max}(\Delta t)^2 \qquad (1)$$

ただし、$r_p^a(t)$ は時刻 t における関節 p の軸 a の角度、$v_p^a(t)$ は同様にその角速度を意味する。Δv_{max} は、時間 Δt 内に関節が等角加速度運動すると仮定したときに得られる最大速度である。

マッチング処理

シルエット画像とのマッチング処理はフレーム毎に行う。第 i フレームの姿勢推定結果 $r_p^a(t_i)$ をもとに、第 $i+1$ フレームにおける関節角度 $r_p^a(t_{i+1})$ の探索範囲を式 (1) から求める。なお、関節角速度 $v_p^a(t_i)$ はそれまでのフレームの処理結果から求める。第 1 フレームについては、別の手法 [2] を用いて関節角度を求める。また、第 1 フレームでは関節角速度はその値を $0[deg/sec]$ とし、代わりに Δv_{max} を大きめの値に設定する。

各フレームにおけるマッチング処理は、人体モデルの木構造を辿ることにより行う。根ノードに対応する部位から葉ノードに向けて、シルエットと各部位の投影領域との重なりを考慮した評価関数をもとにそれぞれの関節角度を決定する。

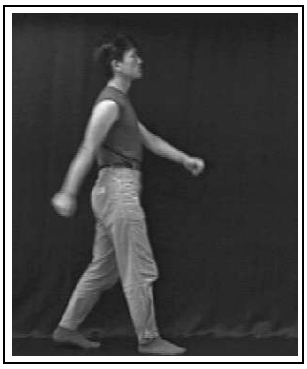

図 1 第 5 フレーム

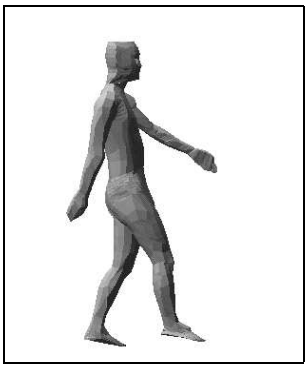

図 2 推定結果

4　実験

男性の歩行運動を 30 フレーム取り込み、実験を行った。結果の一部を図 1、図 2 に示す。実際には図 1 のような濃淡画像ではなく、二値画像を使用している。式 (1) において Δv_{max} は $7.2 \times 10^2[deg/sec]$ に設定した。

人体モデルの根ノードには頭部が対応するが、今回の実験では頭部抽出は顔領域に注目した別処理を行っている。なお、頭部の抽出処理にも慣性が考慮されている。

5　おわりに

本稿では，人間の身体を撮影した動画像に対し、シルエット画像を用いて運動推定を行う方法について提案し、実験結果を示した。今後の課題としては、隠蔽や重なりが比較的長く続く場合に対応するすることや、モデルの形状が実際の対象と異なる場合のマッチング評価法の改善などが挙げられる。

参考文献

[1] 亀田、美濃、池田: “シルエット画像からの関節物体の姿勢推定能力について、” 第 47 回情処全大, 2, pp.151-152, 1993.

[2] KAMEDA, MINOH, IKEDA: “Three Dimensional Pose Estimation of an Articulated Object from its Silhouette Image,” ACCV'93, pp.133-136, 1993.