

6C-3

パターンマッチングと時空間勾配法の併用による
モーションキャプチャ

塩谷 厚志†

†新潟大学大学院自然科学研究科

山本 正信†

†新潟大学工学部情報工学科

1. はじめに

CG キャラクタに動作を与えるために、モーションキャプチャ(MoCap)が利用されることがある。キャラクターに人物と同じ動作を与えることができる半面、光学式MoCapなどの接触式MoCapシステムは専用機器やスタジオが必要となる。一方、市販のビデオカメラで撮影した動画像から、人物の動作を測定することができれば、より簡単にCGアニメーションの作成を行うことができるようになる。動画像を使って時空間勾配法[1]をベースに身体の動作を測定するために非接触式MoCapシステムの研究・開発[2]を行ってきたが、2.で示すように、まだユーザの負担が大きい。

本論文では、ユーザの作業量を減らし、より簡単に使用できるMoCapシステムを提案する。

2. 非接触式モーションキャプチャの問題点

2.1 初期フレーム照合と時空間勾配法

ビデオ映像からの動作測定法では、身体の構造を表す多関節モデルを利用することが一般的となっている。図1(a)に多関節モデルの例を示す。このモデルは全部で16個の部位から構成されており、Waistが木構造の最上位となっている。図1(b)に各部位の接続関係を木構造で表している。矢印の向きは親から子への向きを示す。モデルを画面上の身体に当てはめるとき(図1(c))、モデルから身体的位置・姿勢を知ることができる。動画像の最初のフレームでモデルを当てはめ、初期位置・姿勢を得た後、時空間勾配法から他のフレームでの位置・姿勢を求める。

時空間勾配法[1]とは動画像から運動パラメータ x を推定するための一手法である。運動パラメータとはモデルの各部位の並進と回転を表す変数である。身体上の点 P_c と変位 ΔP_c と運動パラメータはヤコビ行列 J を使い次式で表せる。

$$\Delta P_c = Jx \tag{1}$$

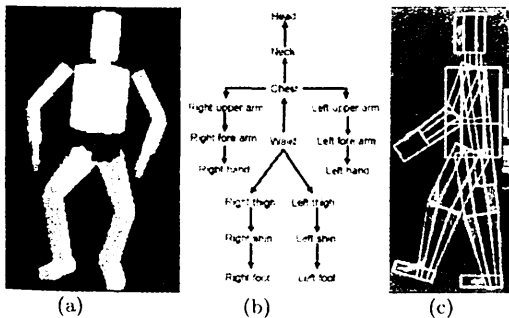


図1: 身体が多関節モデル

時刻 t における投影面上の点 (X, Y) の濃度値を $E(X, Y, t)$ とする。投影面上の移動ベクトル ΔP は動画像の空間勾配 (E_X, E_Y) と時間勾配に束縛される。

$$\nabla E \Delta P + E_t = 0 \tag{2}$$

空間変位 ΔP_c の投影 ΔP を式(2)に代入すると次式が得られる。

$$\nabla E Jx + E_t = 0 \tag{3}$$

式(3)から運動パラメータ x の最小2乗解を求める。

運動パラメータから姿勢の変位が分かり、それを初期姿勢に累積することで全フレームでの身体の姿勢を知ることができる。ところが、推定誤差も累積されるため、累積から得た姿勢は実際の姿勢から時間とともに大きくずれてしまう。

2.2 ずれの修正法と問題点

ずれを少なくするために、測定開始フレームと終了フレームでモデルを当てはめ、中間フレームでの位置と姿勢を線形補間より求める。姿勢の線形補間は、開始姿勢と終了姿勢の差を軸回りの回転で表し、回転角度をフレーム数で案分する。得られた位置・姿勢はパラメータ空間内に軌道を描く。この軌道を初期軌道とする。初期軌道は必ずしも実際の身体の動作軌道に一致しない。両者が一致するように時空間勾配法で初期軌道を修正する。

しかし、フレーム数が多くなった場合、位置・姿勢を開始・終了フレームで与えるだけでは不十分(図3(a)、図4(a))で、いくつかの中間フレームでも、位置・姿勢を与えることになる。位置・姿勢が与えられたフレームをキーフレームと呼ぶ。キーフレームの数が多くなった場合、ユーザの負担は大きくなる。

3. 初期軌道の生成

動作軌道とほぼ同じ初期軌道を作ることができれば、キーフレームの多用によるユーザの負担を減らすことができる。そこでテンプレートマッチングにより、画像中の顔、手、足など部位の追跡を行い、逆運動学から初期動作を作る。

3.1 テンプレートマッチング

$M \times M$ 画素の探索画像 I 上で $N \times N$ 画素のテンプレート画像 T を移動させ照合を行う。ここで $N < M$ である。 (a, b) は入力画像内におけるテンプレート画像の左上位置を示す。このとき式(4)に示すテンプレート画像と探索画像との残差 $R(a, b)$ が最小になる位置を求める。

$$R(a, b) = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} |I(a+m, b+n) - T(m, n)| \tag{4}$$

この処理は残差逐次検定法により効率化が図れる。

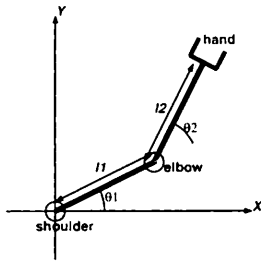


図 2: 2 自由度の腕の略図

3.2 逆運動学

逆運動学は、リンクの長さやエンドエフェクタの位置から、各関節の角度を求める手法である。人物の動作がカメラの光軸に垂直な平面のみで行われるとすると、足や手は2自由度となる(図2)。 θ_1 、 θ_2 を導出する2自由度の逆運動学の式は

$$\cos \theta_2 = \frac{P_x^2 + P_y^2 - l_1^2 - l_2^2}{2l_1l_2} \quad (5)$$

$$\theta_2 = \text{atan2}(\pm \sqrt{1 - \cos^2 \theta_2}, \cos \theta_2) \quad (6)$$

$$k_c = l_1 + l_2 \cdot \cos \theta_2 \quad (7)$$

$$k_s = l_2 \cdot \sin \theta_2 \quad (8)$$

$$\theta_1 = \text{atan2}\left(\frac{-k_s \cdot P_x + k_c \cdot P_y}{k_c^2 + k_s^2}, \frac{k_c \cdot P_x + k_s \cdot P_y}{k_c^2 + k_s^2}\right) \quad (9)$$

となる。解は2つ求まるが $\theta_2 > 0$ として解を1つに定める。ここで、 P_x と P_y はエンドエフェクタの位置(図2のhandの位置)、 l_1 と l_2 は各リンクの長さである。 l_1 、 l_2 の長さやshoulderの位置は人物に合わせたモデルから得られる。

4. 実験

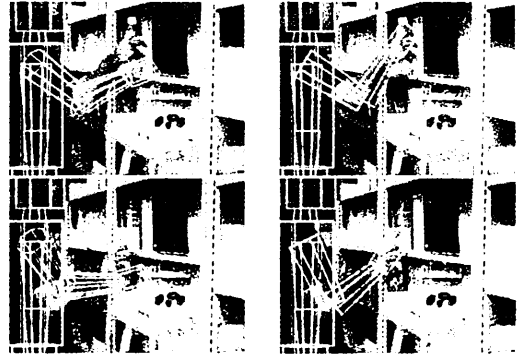
開始・終了フレームの位置・姿勢だけでは測定が難しい動画像に対して、本手法による初期軌道を与えて動作測定を行った。動作はカメラの光軸に垂直な平面のみで行われるものとした。

図3は棚の上にある物体を棚の下に移動させる動作である。39フレームで構成されている。図3(a)は線形補間による初期軌道を使用した結果、図3(b)は本手法による初期軌道を使用した結果である。

図4は右から左への歩行動作である。153フレームで構成されている。図4(a)は線形補間による初期軌道を使用した結果、図4(b)は本手法による初期軌道を使用した結果である。なお、実験2では右腕は測定対象外とした。

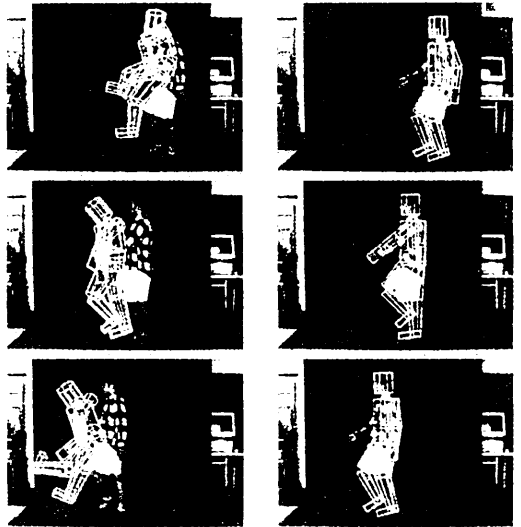
5. まとめ

非接触式 MoCap システムに初期軌道を与えて、モデルを手動で身体に合わせることによるユーザの負担を軽減する手法を提案した。測定結果も、2次元の動作に対しては良好な結果を得ることができた。今後は、3次元の動作に対応できるよう拡張する。



(a) (b)

図 3: 実験結果 1



(a) (b)

図 4: 実験結果 2

参考文献

- [1] 山本正信, 川田聡, 近藤拓也, 越川和忠: 「ロボットモデルに基づく人間動作の3次元画像追跡」, 電子情報通信学会論文誌, Vol.J79-D-II, No.1, pp.71-83, 1996.
- [2] 山本正信: 「ドリフト修正機能を有する動画からの身体動作測定法」, 電子情報通信学会論文誌, Vol.J88-D-II, No.7, pp.1153-1165, 2005.