

単眼人物動画の解析による モーションデータの自動取得に関する検討

武藤 麻矢 加藤 誠巳
(上智大学理工学部)

1. まえがき

人間の 3DCG アニメーション、運動学的フォーム解析、ロボットの遠隔操作など、モーションキャプチャシステムの応用範囲は多岐にわたっている。

しかし現在主流のシステムは、加速度センサや光学式や磁気式のマーカを使用しているため撮影範囲に限られ、被験者は装置やマーカを装着しなければならない制約があり動作も制限される。コストも高く、個人で利用するのは困難である。利用目的が多様化している現在、個人向けの簡易なシステムが求められている。

本稿では個人で手軽に使用できるソフトウェアベースのシステムについて検討を行った。既存のソフトウェアでは不可欠であった関節位置などの特徴点を手動で付ける作業や、撮影範囲やマーカ等の制限のない、誰にでも気軽に使用できる簡易なシステムを目的としている。

2. システムの概要

背景が固定されている動画が入力されると、ユーザが特徴点の入力などの特別な操作をする必要はなく、自動的に各部位を認識し追跡を行う。取得した追跡結果から、出力ファイル形式に必要な座標にパラメータ変換し、CG ソフトで人体のアニメーションに使用できるモーションデータを自動的に作成する。

3. システムの詳細

一般に CG で人間のモデルを扱う際は、身体の構造が決まっているので、骨を親子関係で連結して全身の構造を表した木構造の骨格(スケルトン)を CG キャラクタに設定する。すると親となるノードからの子ノードへの回転角度を制御することによって、姿勢を変化させることができる。本システムではスケルトン階層構造を使用し、首や肩など身体のパーツで分けた 17 のノードによる記述形式を採用した。

An Automatic Motion Data Acquisition by
Analyzing Monocular Moving Human Pictures
Maya MUTO, Masami KATO
Sophia University

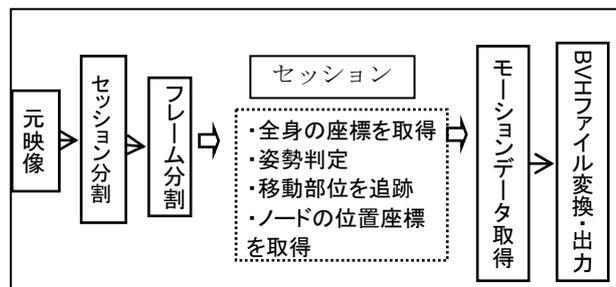


図1 システムの構成

システムの構成を図1に示す。ユーザはシステムに動画を入力し、アニメーションデータを作成したい範囲でセッションを区切る。システムは自動でシルエットから初期位置を検出し、用意した DB(データベース)からスケルトン階層構造や座標などの必要なデータをロードし、初期位置から各ノードを追跡する。その結果を出力形式に合わせてパラメータ変換し、一連の動作をアニメーションデータとする BVH ファイルを作成する。

3.1 動画の画像処理

入力は通常の民生用ビデオカメラやデジタルカメラの動画モードで撮影できる、320×240ピクセルで30フレーム/秒の動画像を用いる。入力した単眼動画像を Java Media Framework(JMF)を使用してフレーム毎に分割し、時系列順に並んだ静止画として扱い、各フレームに対して画像処理を行う。

3.2 シルエットからの姿勢認識^[1]

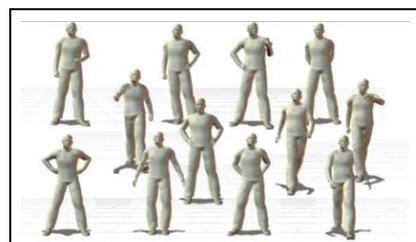


図2 姿勢サンプルの例

一般のソフトにおいては身体のパーツの位置を認識するためにはユーザが目で見て判断し、特徴点を入力する必要がある。これを自動的に認識させる手法として、姿勢サンプルのDBから姿勢の自動認識を行った。

画像からフレーム間差分を取ることによ

て人物領域のみを抜き出し、シルエット画像を得る。シルエットと姿勢サンプルのマッチングからそのフレームでの姿勢を判定し、パーツの位置とスケルトンの角度データを取得する。姿勢サンプルとは、シルエットと部位の位置座標とを対応させるために用意した二次元画像であり、三次元 CG 人体モデルを使用して様々な姿勢パターンにおいて生成した。図 2 に例を示す。各姿勢でのノードの回転角度とパーツの二次元座標を DB として用意する。マッチングにより DB 内で最も近い姿勢を選択し、その姿勢におけるデータを初期値として採用する。

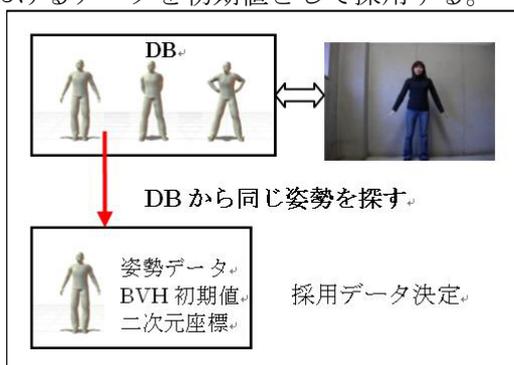


図 3 マッチングによる姿勢判定

3.3 動作の認識

3.2 で求められた各ノードの配置を初期位置として次フレーム以降で各ノードの追跡を行う。前フレームのノード座標を中心としてテンプレートを更新し、ブロックマッチングを行って次フレームでの位置を検出した。

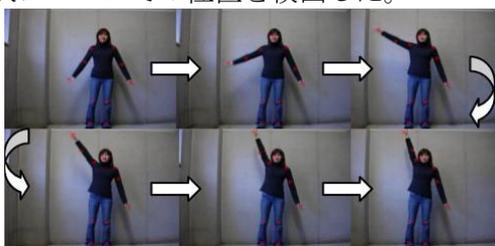


図 4 追跡結果例

3.4 ノード位置の推定^[2]

肩や手など色情報の変分が大きいノードに対し、単色の衣服領域内のノードでは取得しづらい。肘、膝に関しては衣服領域内に含まれる可能性が高いため、Inverse Kinematics (IK)^[3] に基づいた計算による関節の回転角度の推定を併用した。形状が特徴的な手と肩のノードは検出精度が高い。その二つのノードから IK を用いて間のノードの回転角度を計算し、肘のノードの存在範囲を推定する。同様に膝でも推定を行い、追跡結果を補完するのに使用する。

3.5 モーションデータ取得

出力形式は汎用性を持たせるため、多くの市販 CG 関連ソフトウェアで動作が保証されている、モーションデータを格納するファイルフォーマットの BVH 形式を用いた。BVH ファイルはスケルトン階層構造と、各フレームでの関節ノードを右手系オイラー角で記述した動作データから構成される。単眼の動画画像から取得できるデータは二次元座標である。これをスケルトンの回転角度データにパラメータを変換し、出力形式に格納するためのデータを計算する。

4. 実行結果



図 5 実行結果

本システムでの実行例を図 5 に示す。BVH ファイルの実行には人体 CG アニメーションソフトウェアの Poser 6J を用いた。

5. むすび

本稿では誰にでも扱える簡易なモーションキャプチャシステムの実現に関する検討について述べた。今後の課題として、三次元での正確な位置の推定や検出率の改善が挙げられる。

最後に、本学 e-LAB/マルチメディア・ラボの諸氏に謝意を表す。

参考文献

- [1] 亀田, 美濃: "シルエット画像からの関節物体の三次元姿勢推定法," 画像電子学会第21回大会, 1993.
- [2] J.Lee, S.Shin, : "A Hierarchical Approach to Interactive Motion Editing for Human-like Figures," Siggraph Computer Graphics Proceedings, 1999.
- [3] L.Wang, C.Chen: "A Combined Optimization Method for Solving the Inverse Kinematics Problem of Mechanical Manipulators," IEEE Transactions on Robotics and Automation, Vol.7, No.4, Aug 1991.