

## 映画からの俳優の演技の測定とCGによる再演

八木下 勝利 古山 隆志 大久保 直人 山中 一 山本 正信

新潟大学工学部情報工学科

〒950-2181 新潟市五十嵐2の町8050

電話:025-262-7462, ファックス:025-261-2501

電子メール: yagisita@vision.ie.niigata-u.ac.jp

概要: 本論文では、動画像から人間の動作を解析し、それをアニメーションとして再構築する手法を示す。人間の動作の解析には時空間勾配法による手法を用いる。これにより普通の映画やビデオ映像などからでも対象の運動を測定できる。解析の精度向上のためにいくつかの拘束条件を利用している。解析された動作はデータベース化しておき、アニメーション作成時にはそのデータベースを利用してキャラクターの運動を再生する。データベースに無い動作を人間らしさや個性などが保存しながら生成する方法についても検討する。本手法を用いて複数の映画から解析した俳優の動作を一つのCGアニメーションとして再構成したので紹介する。

キーワード: 人間動作, 動画像追跡, コンピュータグラフィクス, 時空間勾配法, 拘束

## Producing a Computer Animation by Using Actors' Actions Captured from Old Movies

Katsutoshi Yagishita, Takashi Furuyama, Naoto Ookubo, Hajime Yamanaka  
and Masanobu Yamamoto

Department of Information Engineering, Niigata University

8050, Ikarashi 2-nocho, Niigata, 950-2181, Japan

Tel:+81-25-262-7462, Fax:+81-25-261-2501

E-mail: yagisita@vision.ie.niigata-u.ac.jp

**Abstract:** This paper presents a novel method for analysis and synthesis of human actions from movies. This method can capture human actions from movies by using the temporal-spatial gradient-based technique with action constraints from the scene. The captured actions are stored in an action database. We can produce a computer animation where the synthetic actors replay actions selected from the action database. Some techniques are described for action modifications according to varying the scene. We will present two actors playing together in a computer animation where the action of each actor has been captured from different old movie.

**Keywords:** Human motion, Visual tracking, Computer graphics, Gradient-based method, Constraint

## 1 はじめに

近年、映画のシーンやコンピュータゲームなどにおいて、実写に近いコンピュータグラフィクスが製作されている。これらの中には人間を対象としたものも多く、そのためいかに本物に近い人間の動きを作るかが重要な課題となる。

人間の動きを生成する方法の一つとして、シミュレーションによる方法がある。この方法では物理法則から運動を算出するが、物理法則だけでは人間らしい個性的な動きを生成するのは困難である。もう一つの方法として、実際の人間の動作を計測しそれを再現することが行われている。動作の計測には、モーションキャプチャやデータスーツなどが用いられている。しかしこれらの方法では人体にデバイスを取り付ける必要があったり、大がかりな装置が必要であったりして、対象となる人間に不要な負担を与え、その人本来の運動を測定することが難しい。また対象となる人物にスタジオ等で演じてもらわなくてはならないため、過去の動作や実際のスポーツシーンなどの運動の測定は不可能である。

本研究では、ビデオ映像や映画などのごく普通の動画像から運動を測定する。これにより人体に接触することなく運動の測定ができる。また映像さえ残っていれば過去における動作などでも測定が可能である。

様々な映像から俳優の動作を取り出したデータはデータベース化する。そしてそのデータベースから動作を選択し、様々なシーンをデジタル映画として作成する。今回はその一例として、二つの映画から取り出した俳優の演技を再構成し、新たなデジタル映画を作成した。

## 2 人間動作の動画像追跡

本節では動画像追跡法について説明する。まず2.1節で追跡対象である人体とその運動のモデル化について述べ、ついで2.2節で時空間勾配法を用いた追跡の原理と手順について述べる。

### 2.1 人体と運動のモデル化

人体は図1のような多関節ロボットモデルで表す。このモデルは人体を各関節で分解して、それぞれをソリッドモデルとして表したものである。各セグメントには図2のような親子関係を結んでおく。これによって、実際の人間と同じように、上位の部位(上腕)の運動により下位の部位(前腕)も移動することになる。さらに各セグメント毎に固有の座標系を用意し、それをそれぞれのセグメントに対応する関節の位置

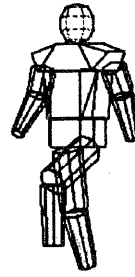


図 1: 多関節ロボットモデル

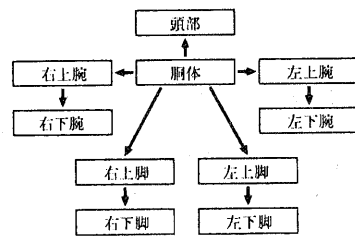


図 2: 親子関係

に合致させておく(図3)。各セグメントの運動はその座標系で表し、また座標系も共に移動するものとする。各座標系の位置・姿勢を並べたベクトルを  $q = (q_1, q_2, \dots)^T$  としたとき、その時間変化率  $\dot{q}$  は各座標系すなわち各関節の運動を表すパラメータとなる。これを運動パラメータと呼び、人体の運動をこのパラメータで記述する。

### 2.2 時空間勾配法による追跡

シーン及び投影面の座標系を図4のように設定する。カメラは焦点距離  $f$  の中心投影モデルとする。このとき、シーン中の点  $(x, y, z)$  とその投影点  $(X, Y)$  は以下の式で関係付けられる。

$$\begin{cases} X = f \frac{x}{z} \\ Y = f \frac{y}{z} \end{cases} \quad (1)$$

式(1)の両辺を時刻  $t$  で微分すると

$$\begin{cases} \dot{X} = \frac{f\dot{x} - X\dot{z}}{z} \\ \dot{Y} = \frac{f\dot{y} - Y\dot{z}}{z} \end{cases} \quad (2)$$

となる。

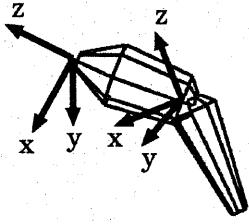


図 3: 物体座標系

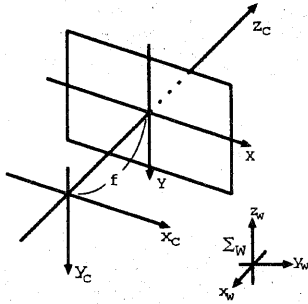


図 4: カメラ座標系

動画像の時刻  $t$  における点  $(X, Y)$  の明るさを  $E(X, Y, t)$  と表す。このとき、動画像上の速度ベクトル  $(\dot{X}, \dot{Y})$  は以下の時空間勾配式に従う [4]。

$$E_X \dot{X} + E_Y \dot{Y} = -E_t \quad (3)$$

式 (3) に式 (2) を代入することで

$$\mathbf{G}^T \dot{\mathbf{p}} = -E_t \quad (4)$$

となる。ただし

$$\mathbf{G} = \left( f \frac{E_X}{z}, f \frac{E_Y}{z}, -\frac{XE_X + YE_Y}{z} \right)^T \quad (5)$$

$$\mathbf{p} = (x, y, z)^T \quad (6)$$

である。

2.1節で人体はロボットモデルとして表された。ロボットモデル上の点  $\mathbf{p}$  の速度ベクトル  $\dot{\mathbf{p}}$  と、ロボットモデルの各関節の速度すなわち運動パラメータ  $\dot{\mathbf{q}}$  の関係は、ヤコビ行列  $\mathbf{J}$  を用いることで以下のように関係付けられる。

$$\dot{\mathbf{p}} = \mathbf{J} \dot{\mathbf{q}} \quad (7)$$

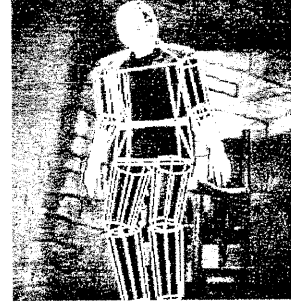


図 5: 画像とモデルの照合

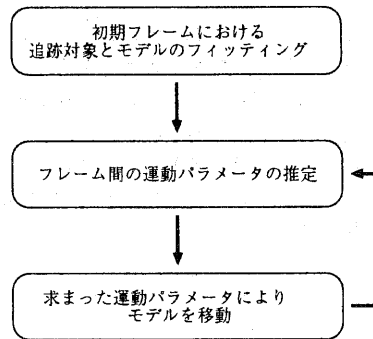


図 6: 動画像追跡の流れ

これを式 (4) に代入して

$$\mathbf{G}^T \mathbf{J} \dot{\mathbf{q}} = -E_t \quad (8)$$

を得る。これが、時空間勾配法による運動パラメータ  $\dot{\mathbf{q}}$  の推定式となる。

時空間勾配法による動画像追跡では、まず、図5のように先頭フレームで画像とモデルを合致させる。その後一フレーム毎に動画像中のモデル上の複数点から式 (8) を求め、それらを連立させて解くことで、次々と運動パラメータを求める。この流れを図6に示す。

### 3 拘束を利用した追跡

2節で説明した動画像追跡では、追跡の途中でモデルと実際の対象との間にずれが生じると、その後の追跡が著しく困難となる。そこでこのずれを避けるために、拘束を用いることを考える。

### 3.1 キーフレーム拘束

キーフレーム拘束とは、図5のような画像とモデルとの照合を最終フレームでも行い、最終フレームにおける追跡結果がその姿勢に収束するように拘束するものである。

$j$  番目の物体の、先頭フレームでの照合で与えられる親からみた位置・姿勢を  $T_{0,j} \cdot R_{0,j}$ 、第  $n$  フレームにおける追跡結果での位置・姿勢を  $\hat{T}_{n,j} \cdot \hat{R}_{n,j}$  とすると

$$\hat{T}_{n,j} = T_{0,j} + \sum_{i=1}^n S_{i,j} \quad (9)$$

$$\hat{R}_{n,j} = R_{0,j} \prod_{i=1}^n Q_{i,j} \quad (10)$$

と表される。ただし  $S_{i,j}$  は  $i$  番目のフレームにおける運動パラメータ  $\dot{q}$  の物体  $j$  の並進成分に対応する部分であり、 $Q_{i,j}$  は回転成分に対応する運動パラメータを回転行列で表したものである。ここで  $n$  番目のフレームにおける位置・姿勢がキーフレームとしてそれぞれ  $T_{n,j}$ 、 $R_{n,j}$  与えられているとすれば、キーフレーム拘束により

$$\hat{T}_{n,j} = T_{n,j} \quad (11)$$

$$\hat{R}_{n,j} = R_{n,j} \quad (12)$$

となる。

時空間拘束式 (8) とキーフレーム拘束式 (11) の両者の誤差を最小にするような解、つまり

$$L = \sum_{i=1}^n \|A_i \dot{q} - b_i\|^2 + \lambda \sum_{j=1}^m \{ \|\hat{R}_{n,j} - R_{n,j}\|^2 + \|\hat{T}_{n,j} - T_{n,j}\|^2 \} \quad (13)$$

を最小にする  $\dot{q}$  を求めることで、キーフレーム拘束と2節で述べた時空間勾配法とを融合した追跡となる。ただし  $A_i$ 、 $b_i$  は式 (8) を連立させたときの係数行列と定数ベクトルで、 $m$  は人体モデルのセグメントの数である。この解法は文献 [2] を参照されたい。

### 3.2 シーン拘束

人間などの運動はシーン中の他のものから拘束される。これを追跡に利用することで追跡の際のずれを抑えることを試みる。この拘束は、追跡対象上の点  $p$  の位置・速度・加速度をいくつかの平面によって表すことで表現する。これらの拘束式は全て

$$d^T \dot{q} = c \quad (14)$$

という一次方程式で表すことができる。拘束するものが曲面である場合には  $d$ 、 $c$  を拘束される点の位置

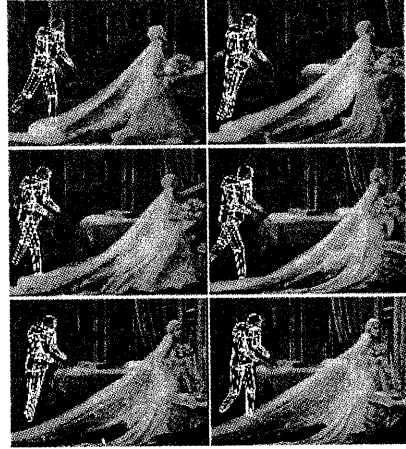


図7: 追跡結果 (フランケンシュタイン)

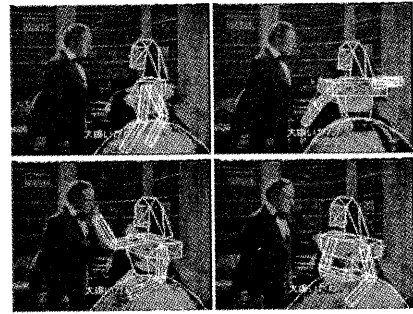


図8: 追跡結果 (風と共に去りぬ)

$p$  の関数とし、時刻によって拘束するものが変化する場合には  $d$ 、 $c$  を時刻  $t$  の関数とする。式 (14) を式 (8) とともに連立させることでシーン拘束を考慮した追跡が可能となる [3]。

## 4 追跡結果

2, 3節で述べた手法により追跡を行った結果を示す。図7は映画「フランケンシュタイン」<sup>1</sup>から怪物<sup>2</sup>が歩くシーンを追跡した結果である。46フレームの動画画像でキーフレームは4フレームに与えた。また直線歩行であるため速度に対するシーン拘束も用いている。図8は映画「風とともに去りぬ」<sup>3</sup>からスカー

<sup>1</sup>Frankenstein 1931, ©Universal Pictures Corporation

<sup>2</sup>Boris Karloff

<sup>3</sup>Gone with the wind, 1939, ©MGM/UA Home Enter-

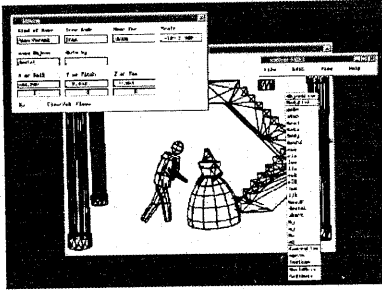


図 9: モデルの配置

レット<sup>4</sup>がアシュレー<sup>5</sup>を叩くシーンを追跡した結果である。高速な動きであるにもかかわらず、キーフレームにより追跡が可能となっている。

## 5 アニメーションの作成

追跡された動作はデータベース化されている。また人間をはじめ、シーンを構成する大道具や小道具などのモデルもデータベース化されている。それらのデータベースから作成するシーンに必要なデータを取り出して構成することでCGアニメーションが作成される。

### 5.1 モデル・カメラの配置

役者や大道具・小道具などのモデルやカメラの配置は、モデラ上でマウスを用いて設定する。配置の過程を図9に示す。このとき、アニメーションシステムにある物体間干渉チェック機能を利用して、モデル同士のめり込み等を事前に防ぐことが可能である。

### 5.2 動作のスケジューリング

各キャラクターの動作のスケジュールは、図10のようなデータファイルに記述しておく。これは個々のキャラクターの各動作について、スタート時刻とその動作名を記述する。このデータファイルをキャラクターの数だけ用意しアニメーションシステムに入力し全体の動きを構成する。

### 5.3 動作の変形

アニメーションに必要な動作を全て映画からの追跡によって取得するのは困難である。そこで、取得

tainment Inc. and Turner Entertainment Co.

<sup>4</sup>Vivien Leigh

<sup>5</sup>Leslie Howard

# No.	Data	Option
0	walk	1.2
50	stand	
70	hit	
85	turn	
100	walk	0.8

図 10: スケジューリング

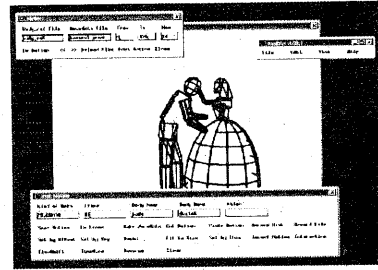


図 11: 動作の変形過程

された動作に対して変形を行う。また、シーン中のモデルの配置によっては、障害物をよけるように、あるいは対象物に近づくようになどの変形も必要となる。この変形の際には、元の動作に含まれる人間らしさ・個性が保存されることが望ましい。本研究では、単純にキーフレームを用いた変形の他にモーションワープ法[5]を用いることも行った。また物体間の干渉から生じる動作の変形については、幾つかの関数(クーロンの法則式など)を用意し、もともと追跡された運動にその関数を融合することで他の登場人物を避けて歩いたり、また追いかけていたりというシーンを自動的に作成することが可能となった。

動作の追跡や変形の際のエラーによってや、アニメーション中で使うモデルと追跡時に使うモデルが異なった場合には、例えば歩いている足が床面から浮いたり床面上を滑ったりといった不自然な動作となることがある。この場合には3.2節で述べたシーン拘束を動作に適用することで自然な動作に修正する。

以上の変形は、アニメーションシステム上で動作を確認しながら適用することができるようになっている。この操作の過程を図11に示す。

### 5.4 実際の作成

今回は、映画「フランケンシュタイン」と「風と共に去りぬ」からフランケンシュタインの怪物とスカーレットの動作を追跡し、「襲ってくる怪物をスカー

レットが叩く”というCGアニメーションを作成した。動画像追跡により実際に取得した動作は

- 怪物のうろつく動作  
(怪物が地下室で出口を見付けるためうろついている動作)
- 怪物が歩く動作  
(怪物が窓から部屋へ侵入し部屋の中を歩きまわっている動作, 図7)
- スカーレットの歩く動作  
(スカーレットが庭を歩いている動作)
- スカーレットが叩く動作  
(スカーレットがアシュレーを叩く動作, 図8)

である。以上の動作を組み合わせることにより、元の動作とは全く異なる意味を持ったアニメーションを作成する。アニメーションを

- 怪物がスカーレットに近づくシーン
- スカーレットが怪物を叩くシーン
- 怪物が逃げるスカーレットを追いかけるシーン

に分け作成した。以下に今回の作成手順を示す。

まず配置を行う。これは5.1で述べたとおりである。1番目のシーンに必要な怪物が歩く動作、スカーレットが立っている姿勢ともに取得しているので、変形はシーン拘束によって足の位置を床面に対して補正するのみであった。2番目のシーンでは、スカーレットが叩く動作は追跡されたものをそのまま利用し、怪物が叩かれた後に倒れ込む動作は追跡されていないため、キーフレームを与えて作成した。3番目のシーンは、スカーレットが歩く動作に合わせて怪物が歩く動作を修正することで追いかけるようにした。また足の位置の床面に対しての補正もしている。実際に作成されたアニメーションを図12に示す。

## 6 あとがき

本研究では、映画などの普通の動画像から人間の動作を取得し、それによってキャラクタの動作を生成しCGアニメーションを作成した。この方法により撮影用に演じてもらうことのできない人物の動きなどでもキャラクタに反映させる事が可能となる。今回は実際に2本の映画から俳優の動きを取り出し、それらを合成してアニメーションを作成した。

今後はアニメーションの作成過程における人手の介入を減らす必要がある。GUIの利用などによりアプリケーションは使い易いものとなっており、また物体間の干渉のテストや動作の変形の一部など自動

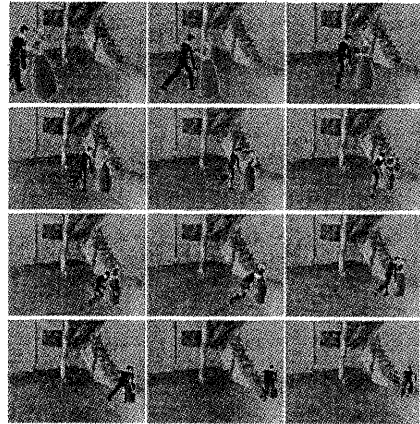


図 12: 作成されたアニメーション

化されている部分もあるが、さらに自動化を進め、シナリオを与えたら自動的にシーンが作成されるようなシステムが望ましい。また、動画像追跡の高精度化も課題である。

## 謝辞

動作解析のための映画の使用に関しましてご協力頂きましたユニバーサル・スタジオに感謝致します。

本研究は一部科研費重点領域研究「人工現実感」の支援を受けた。

## 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次元動画像追跡”, 信学論, J79-D-II, No.1, pp.71-83, (1996)
- [2] 大田佳人, 山際貴志, 山本正信: “キーフレーム拘束を利用した単眼動画像からの人間動作の追跡”, 信学論, J81-D-II, No.9, pp.2008-2018, (1998)
- [3] 八木下勝利, 山本正信: “シーン拘束を用いた人間動作の高精度動画像追跡”, 映情学誌, Vol.52, No.3, pp.331-336
- [4] B.K.P.Horn and B.G.Schunck: “Determining Optical Flow”, Artificial Intelligence, 17, 1-3, pp.185-203 (1981)
- [5] A. Witkin and Z. popovic: “Motion Warping”, Computer Graphics Proceedings (SEG-GRAPH 95),