

## CV 技術による身体動作の解析

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: 025-262-7428, ファックス: 025-261-2501

電子メール†: yamamoto@ie.niigata-u.ac.jp

あらまし: 動画像を使った動作解析の特徴は、身体に接することなく動作が測定できる点である。また、映像さえ残っていれば過去に遡って動作を解析することが可能である。本稿では、動画像による身体動作解析の現状や問題点を整理し、今後の課題について考えることにする。動作解析のプロセスは、4つのステップに分けることができる。まず、画像には身体以外にも様々な物体が写り込んでいることから、画像から身体像を抽出し、身体部位を検出する。身体が多関節モデルを導入し、モデルを身体像に一致させることにより、姿勢を測定する。この姿勢測定法は、順運動学手法と逆運動学手法に分けられる。動作は姿勢の変化分として、線形推定が可能である。姿勢は初期姿勢に変化分を累積することによっても得られるが、変化分の推定誤差を押える必要がある。得られた動作を編集することにより、新たな動作を生成することができる。

## Computer Vision-Based Human Motion Capturing

Masanobu Yamamoto

Department of Information Engineering, Niigata University

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

Tel: +81-25-262-7428, FaX: +81-25-261-2501

Email†: yamamoto@ie.niigata-u.ac.jp

**Abstract:** An computer vision-based motion capture has a remarkable advantage that it can measure a pose and motion of human body impermeable. This means that one can capture a human motion in movies or videos of the past. This paper surveies computer vision-based approaches for human motion capture, and proposes future directions. A process of these approaches is composed of four steps. The first step is to detect parts of human body from the image. The second step is model fitting to estimate the human pose. The model fitting can be established by direct or inverse kinematics approaches. The third step is to estimate the motion. Several constraints can determine the motion. The final step is to edit and replay the motion on computer graphics.

## 1 まえがき

コンピュータビジョン (CV) の研究は、1960 年代に始まった。研究対象が、積木の世界から自然界へと推移していったが、しばらくは静止したシーンが対象であった。1970 年代の終頃から、動きのあるシーンが対象とされはじめた。驚くべきことに、この頃既に身体動作解析が試みられていた [25][16][24][2][23]。しかし、実動画像を扱った研究は少なく、実動画像による身体動作解析が本格化するのには、1990 年代に入ってからである。この理由は色々考えられるが、一つは、機材の問題であろう。当時は動画像をコンピュータに入力できる機材は高価で、大量の動画像が処理できるほどの計算機パワーを持っていた機関も限られていた。今では動画像が高性能のコンピュータに容易に取り込めるようになってきた。もう一つは、身体を解析するためのニーズが広く認識されてきたことにもよる。身体動作解析は、人間とコンピュータとのインターフェース、医療や福祉、スポーツ、放送、セキュリティ、監視、交通、アミューズメント、エンターテインメント、等多くの分野で必要とされている。

身体動作解析は、CV で対象とされる以前に、様々な学問分野で取り上げられてきた。特に、バイオメカニクスや人間工学では、光学式、磁気式、機械式、など身体動作を測定する様々な方式が提案されている。これらの方式に、動作を CG 上で再現する機能を加えたシステムを、モーションキャプチャと呼んでいる。

これらのモーションキャプチャは、幅広く利用されているが、身体に特殊な装置を取付たり、予めマーカ張り付けることが多く、自然な状態で身体動作を測定することが難しい。そのため、身体に非接触に動作が測定できることが望まれている。

CV 技術による動作解析は、身体に接触することなく測定できるため、次世代のモーションキャプチャと考えられる。本稿では、動画像による身体動作解析の現状や問題点を整理し、今後の課題について考えることにする。

## 2 身体領域の抽出

### 2.1 画像上の領域抽出

画像を用いた測定法には広い用途が期待される反面問題点もある。その一つとして、一般に画像には対象としている身体以外にも背景にさまざまな物体が映りこんでいる。そのため、まず画像から身体像の領域を分離しなくてはならない。

連続する画像間で濃淡値の差をとれば、動いている身体の輪郭を抽出することができる。また、背景画像があらかじめ得られているならば、画像と背景画像の差をとるだけで、身体領域を抜き出すこと

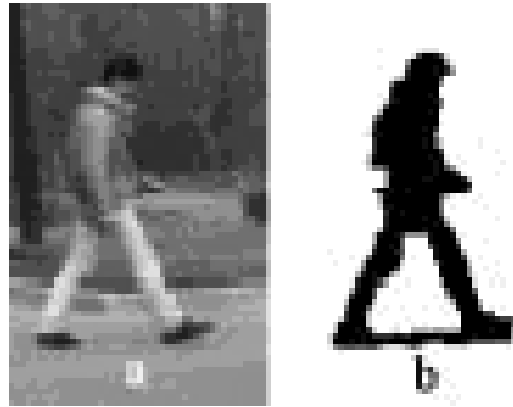


図 1: 身体像の抽出とシルエット. I.Haritaoglu ©IEEE 1998

もできる。ただし、これらの簡単な方法は、カメラが固定されている場合に限定されている。カメラも運動している場合には、オプティカルフローから身体特有の動き（例えば関節回りの回転運動）が利用できる [21]。得られた抽出画像には、人体の輪郭や領域以外にもノイズにより様々な領域が発生し身体と区別する必要がある。小さなノイズは、あらかじめ画像に平滑化処理を施しておいたり、差分画像に縮小と拡大を繰り返すことによって除去できる。このようにして人物像が抽出される。抽出された身体を図 1 に示す。人物像の領域を値 1、背景を 0 とすれば、身体像のシルエットを表す 2 値画像となる。

### 2.2 空間中の領域抽出

画像を使った測定法のもう一つの問題点は隠れの問題である。手や足などは動作中に胴体に隠されることがあり、逆に胴体も腕によって完全には観測することが難しい。隠れの問題を解決する直接的な方法としてはカメラの数を増やすことである。互いにキャリブレーションされたカメラ対から、身体上の 3 次元座標値を精度良く求めることができる [12]。しかし、複数のカメラの利用には、映像の同期やカメラキャリブレーションなどを必要とする。

### 2.3 領域中の部位の特定

身体像が領域として抽出されたならば、次は、そこから頭や手足のような身体各部を特定する。画像がカラーであるならば、顔や手の位置は肌の色を利用して決定することができる。直立や歩行あるいは走行しているならば、身体のプロポーションを使っておよその位置が分かる [5]。例えば、身体の高さを

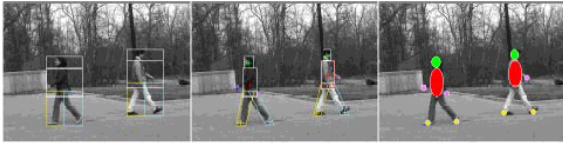


図 2: 身体部位の決定. I.Haritaoglu ©IEEE 1998

1としたとき、身体像の上から1/5は頭部、首から下1/2は胴体、足元から上1/2は足部とする。そのときの各部の位置が図2に示してある。また、シルエット中の先端部分は頭や手先、足先である可能性が高い。この情報も位置の決定には有用である。また、身体は足または手などの体の一部は、常に床などの環境に接しており、そこでは環境との相対速度がゼロとなる[10]。したがって、逆に相対速度がゼロの部位を足先または手先と認識することができる。

### 3 身体の姿勢の推定

画像から得られたデータは、身体の領域を表す画素の集合、あるいは、多視点画像からならば身体の形状を表す3次元座標値の集合である。このデータ構造は次のような欠点を持っている。

第1に、姿勢や動作を表すには、これらのデータ全てが必要というわけではなく、関節位置などのデータがあれば十分である。第2に、画像からのデータは時間の経過と共にどのように推移して行くかという情報が欠けている。第3に、画素データや形状データは、それぞれカメラ座標系やシーンの座標系で記述されていることである。このため、データを採取した環境や使用した機材によって、同じ動作でも全く異なったデータとして記述されてしまう。動作の比較や再現を効率良く行うためには、姿勢や動作の表現形式を統一しておく必要がある。

画像センサから得られたデータを簡潔に、しかも統一的形式で表すために、身体モデルを導入する。姿勢や動作の統一的形式として身体モデルに固有の座標系を用いる。このモデルを画像から得られたデータに照合させることによって姿勢を決定する。

#### 3.1 身体モデル

身体の構造を表すためのモデルとして、多関節モデルが一般的である。多関節モデルの最も簡単な例は、スティックモデルである。このモデルは、図3(a)に示されるように、スティック(棒)が端点で他のスティックとジョイント結合した構造を持ち、身体の骨格を表している。

図3(b)(c) (多面体、超楕円体モデル)の多関節

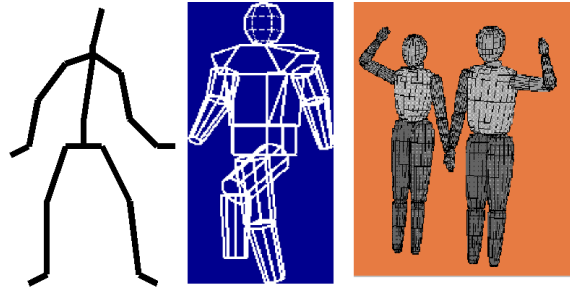


図 3: 身体モデル:(a) スティックモデル、(b) 多面体モデル、(c) 超楕円体モデル D.M.Gavrila ©IEEE 1996

モデルは、胴体や頭、手足のような身体各部を多面体あるいは超楕円体で近似したものである。これは、スティックモデルに肉を付けたものとも考えられる。

対象となる動作に応じてモデルをさらに精密に構成する場合もある。例えば、体のひねり運動では胴体を胸部、腹部、腰部に分ける必要があるし、背伸び運動では肩部を胴体から分離する必要がある。また、新体操やフィギュアスケートのような美しさを競う競技では手先の微かな動きも重要である。そのためには腕の先に手や指のモデルを付け加えることも必要となる。さらに、顔の表情も重要とあれば、皮膚の動きもモデルに加える必要がある。

一方、サッカーやアメリカンフットボールのようなフィールド競技では、選手間の関係プレーに基づくチームワークが重要となってくる。この場合は身体を平面上で移動する物体や一つの点のように簡略化しても良い。

#### 3.2 身体モデルの位置と姿勢

多関節モデルは身体各部の階層構造で構成されている。胴体を階層構造の最上位に置き、末端に向けて頭や手、足が継っている。階層構造内の直接の上下関係を親子関係ともよぶ。

この多関節モデルでは、最上位の胴体のみが並進と回転の6つの自由度を持っている。手足などの他の部分はそれらの親に関節でつながれているので、動きは関節位置を中心とした回転のみである。したがって、自由度は3である。図3のモデルでは、全自由度は33個になる。

腕の先に手、足首の先に足、さらにその先に指がモデルに加えられるとさらに自由度は多くなる。また、胴体も胸や腰や肩などに細かく分割できれば、微かな動きを表すことができるが、自由度の数はさらに多くなる。

しかし実際の姿勢を表すには、これらの自由度がすべて必要というわけではない。例えば、指の各関

節の動きは1自由度でもいいし、下腕や下脚は3自由度も要らない。上腕や上脚が2自由度あれば、1自由度で十分である。これらの場合には、自由度の総数は少なくなる。

姿勢を表す変数(パラメータ)は胴体の3次元座標値や関節角度である。各変数を  $q_i$  と表し、変数の集まりをベクトル  $\mathbf{q}$  とする。各変数に適当な値を与えれば、多関節モデルはそれに対応する姿勢を取ることができる。したがって、モデル上の点  $\mathbf{p}$  の座標値は、姿勢を表す変数( $\mathbf{q}$ )の関数である。式では

$$\mathbf{p} = \mathbf{f}(\mathbf{q}) \quad (1)$$

と書ける。ただし、この式では点の初期位置が省略されている。ここで、関数  $\mathbf{f}$  は一般に非線形ベクトル関数となる。

### 3.3 姿勢の推定

画像から図1のような身体のシルエット画像が得られたとする。この画像から身体の姿勢を推定するのがこの節の目的である。前項で導入した身体が多関節モデルを使って姿勢を推定しよう。

多関節モデルは、ロボットあるいはその腕部であるマニピュレータのモデルでもある。ロボット工学の分野では、マニピュレータの関節角度を与えることによって指先の位置と姿勢を計算する方法を順運動学手法とよんでいる。逆に、指先の位置と姿勢から関節角度を推定する方法を逆運動学手法とよんでいる。

身体の姿勢を推定する方法も、この順運動学と逆運動学の考え方を使って説明することができる。それぞれのアプローチを紹介しよう。

その前に一つ注意しておくべきことがある。それは、一つのカメラからの画像だけでは姿勢の決定が難しいという問題である。姿勢の決定には多方向からの観測が必要である。カメラの数が多くなるほど、得られた姿勢の信頼性が増してくる。また、死角も少なくなる。しかし、撮影システムとしては大掛かりになってくる。

#### 3.3.1 順運動学による姿勢推定

多関節モデルの位置と姿勢の変数に値を与えれば、順運動学手法によりモデルの位置と姿勢を計算することができる。カメラの位置とその向きが分かっているならば、得られたモデルの姿をカメラの投影面に表示させることができる。モデルの投影像と身体のシルエット画像を重ね合わせたとき、両者が一致しているならば、モデルの位置と姿勢は実際の身体の位置と姿勢を表していると言って良い。もし、一致していなければ、変数の値を変えて再度モデルの

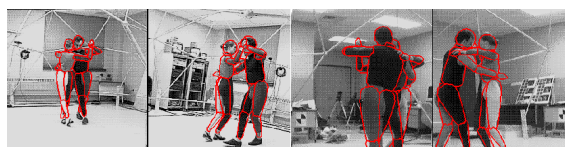


図4: アルゼンチンタンゴの追跡. D.M.Gavrila ©IEEE 1996

投影像とシルエット画像を重ね合わせる。この操作を両者が一致するまで繰り返す。

一致の度合いは、例えば、モデルの投影像とシルエット像を加えた領域を求め、それに対する重なった領域の比率によって評価できる。この比率は重なりが全くなければ0であり、重なり領域が増えるにしたがって値が大きくなり、完全に一致すれば、最大値1となる。一方、輪郭に着目した評価法もある。モデルの輪郭上の点に対してシルエット像の輪郭上で最も近い点を求める。この2点間の距離を輪郭全体にわたって平均した距離を一致度の評価値とする。モデルがシルエットと一致したとき平均距離は0となり、シルエットから離れるにしたがって大きくなる。

さて、位置と姿勢の探索問題であるが、多関節モデルは姿勢の自由度が大きいため、しらみつぶしの探索は不可能に近い。局所的な探索は局所解に陥る恐れがある。そこで、大域的な探索方法として、遺伝的アルゴリズム(GA)の使用[15]も試みられている。

一方、シルエット像の中で、頭や手足などの部位がどこにあるのか分かっているならば、部位ごとに探索を行えば良く、探索範囲は狭まる。また、求める位置と姿勢に近いものが与えられているならば、この位置・姿勢を初期値として、局所的な探索で解を得ることができる。

このような条件は、身体を追跡するときに満たされる。なぜならば、追跡は前の時刻に置かれたモデルを次の時刻の身体像と一致させるという手順の繰り返しであるので、前の時刻でのモデルが良い初期値となる。さらに、シルエット像中の部位は、このモデルの各部位の最近傍領域として得ることができるからである。

追跡例を図4に示す。これはアルゼンチンタンゴを踊っている踊り手の追跡である。踊り手のモデルは超楕円体で構成されている。一致の程度は、輪郭間の平均距離で評価している。モデルの位置と姿勢を小刻みに変えてシルエット像との平均距離を計算した。そのなかで最も平均距離が小さな位置・姿勢を解としている。この追跡では、最初に胴体、ついで手、最後に足の順に照合を行っている。

以上は、複数のカメラからの姿勢決定法であるが、

単眼カメラの場合には、対応する3次元モデルの姿勢を一意に決定できるとは限らない。一つの画像に対して複数の姿勢を候補に上げることができるが、決定を保留しておく。動画の前後のフレームでも対応付けを行ない、動作が滑らかに推移するような姿勢列を選択する[20]。この選択には、ビーム探索などを使って探索効率を上げることができる。

### 3.3.2 逆運動学による姿勢推定

マニピュレータの手先の位置が指定されたとき、その位置を実現するためにマニピュレータの関節角度を計算しなければならない。手先の位置と関節角度との関係は、身体モデルの関係式(1)と同様に非線形関係

$$p = f(q) \quad (2)$$

として表される。手先の位置  $p$  が与えられたとき、モデルの位置と姿勢は、未知数  $q$  に関する非線形方程式を解かねばならず、これが厄介な問題である。

身体の姿勢決定問題では、位置情報は画像から与えられる。身体のシルエット画像から腕や足の体軸が画面上の直線として抽出される。肘や膝の位置は体軸の交点から得られる。また、シルエット像の先端部の中心点から、頭や手や足先の位置が求まる。

モデル上での手先の位置を  $p$  としたとき、この投影はシルエットの手先位置に一致しなければならない。この結果、非線形方程式が一つ導かれる。肘や膝に対しても同じく非線形方程式が導かれる。また、モデルの腕や足などの体軸上の点  $p$  は、その投影位置がシルエット上の体軸に乗っていないなければならない。この関係からも、非線形方程式が導かれる。

非線形方程式が多数導かれるので、これらを連立させて解を求めることになる。解析的に解を得ることは難しいが、数値計算によって求めることにすれば様々な手法が利用できる。例えば、ガウス・ニュートン法を使えば適切な初期値からの繰り返し演算によって解に収束させることができる。

さて、ここで問題になるのは初期値の設定である。リンクの端点の位置、例えば上腕部の肘と肩、が与えられたとき、上腕の姿勢は2通り考えられる。どちらの姿勢を選択するかは、関節角の可動範囲を考慮して絞り込むことができる[1][3][9][22]。

さらに、数値解法では非線形方程式を線形化するが、得られた線形連立方程式が線形従属となる場合には、解の決定が困難である。これは、パラメータの数が過剰な場合と、姿勢がパラメータ空間中の特異点に陥った場合である。パラメータの数が過剰な場合は、多くは単眼視画像で起こるが、多視点画像でも特徴点の数が少なければ起こる可能性がある。例えば、頭部と手先しか位置が与えられない場合には、肘の姿勢

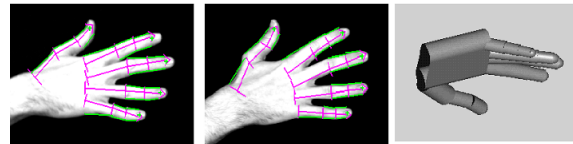


図 5: ステレオ画像からの手の姿勢決定. J.M.Rehg ©1994

を一意に決定することができない。冗長なパラメータには、任意の値を与えても構わないが、それが身体の姿勢を表している保証は無い。解を一意に決定するためには、さらに制約条件を必要とする。例えば、身体のバランスや動作の滑らかさ[1]や、身体パラメータの従属性[27]、などが利用されている。一方、特異点に陥った場合には、連立線形方程式の解が存在しない。特異点を脱出するためには、姿勢を摂動させて再計算する方法[13]や、リンクの伸縮を取り入れた特異点の無い身体モデル[11]も提案されている。

逆運動学手法によって、ステレオカメラから手の指の姿勢を求めた例を図5に示す。

## 4 動作の推定

身体の姿勢の推定を一コマごとに行えば、姿勢のデータ列が得られる。しかし、一コマごとのモデルの当てはめには、非常に多くの計算コストを必要とする。これに対して、姿勢の変化分は動画の変動から導かれる線形推定式を使って、容易に得ることができる[28][7]。したがって、最初のコマでモデルの当てはめを行い、初期姿勢を求め、その後は姿勢の変化分を推定し、初期姿勢に累積すれば計算コストを節約することができる。

### 4.1 動作の線形推定

モデルの位置及び姿勢  $q$  とモデル上の点  $p$  は、式(1)で関係付けられていた。この関係は非線形であるが、モデルの位置と姿勢の  $\delta q$  変化と、モデル上の点  $p$  の  $\delta p$  変位とは、ヤコビ行列を使って、次のような線形関係で表すことができる。

$$\delta p = J(q)\delta q \quad (3)$$

ただし、 $J(q)$  はヤコビ行列であり、 $J = \frac{\partial f}{\partial q}$  である。

変位  $\delta p$  は3次元移動ベクトルであり、これが与えられれば逆運動学により位置と姿勢の変化を推定することができる。しかし、画像からは3次元移動ベクトルを直接与えることはできず、次のように間接的に与える。

まず、空間中の3次元移動ベクトルは、画像上では2次元移動ベクトル  $\delta P = (\Delta X, \Delta Y)$  として観測され、両者は次式で関係付けられる。

$$\Delta P = \frac{1}{z} \begin{pmatrix} f & 0 & -X \\ 0 & f & -Y \end{pmatrix} \Delta p \quad (4)$$

動画像の濃淡値を  $E(X, Y, t)$  と表す。移動先でも濃淡値が変化しないとすれば、画面上の移動ベクトルは次の1次方程式に束縛されることが知られている [6]。

$$E_X \Delta X + E_Y \Delta Y + E_t = 0 \quad (5)$$

ここで、 $E_X, E_Y$  は画像の空間勾配であり、 $E_t$  は時間勾配である。それぞれ、エッジ検出オペレータや差分オペレータにより容易に計算することができる。

式(3),(4),(5)から、移動ベクトルを消去すると、

$$GJq + E_t = 0 \quad (6)$$

が得られる。ここで、 $G = (fE_X/z, fE_Y/z, -(XE_X + YE_Y))^T$  とする。モデルが身体に一致しているとするれば、モデルから身体上の点の3次元座標値を知ることができる。したがって、上式中の  $z$  は既知となり、上式は姿勢の変化分  $\delta q$  を未知数とする線形推定式となる。

身体(モデル)上の多数の点からそれぞれ線形推定式を導きそれらを連立させる。その結果、線形連立方程式を得ることができる。それを次のように記す。

$$A\delta q = b \quad (7)$$

ただし、 $A$  は係数行列、 $b$  は定数ベクトルである。

身体の位置と姿勢の変化は線形連立方程式(7)を解くことにより得られる。このときオプティカルフローは陽に求める必要はない。この連立方程式はカメラの視点ごとに導かれる。位置と姿勢の変化は、身体モデル固有の変数であるので、視点ごとに得られた連立方程式を、そのままさらに連立させることができる。すなわち、変化は大連立方程式

$$\begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_n \end{bmatrix} \delta q = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad (8)$$

を解いて得られる [19]。ここで添え字の数字はカメラの番号を表す。このようなことが可能なのは、姿勢をモデル固有の座標系で表しているからである。

カメラの数を増やせば、多方向からの観測情報を利用できるので、隠れの解消はもとより、冗長性を



図 6: 追跡開始フレームでの多視点画像照合. [19]



図 7: 投球動作の追跡結果 [19]

利用した信頼度の高い推定が可能となる [8]。複雑な動作の推定例として、野球投手のピッチング動作を推定した結果を示す [19]。ここでは、投手の周囲に4台のカメラを配置した。図6に初期時刻でモデルと身体像を一致させた結果を示す。モデルはワイヤーフレームで表示され、重ねて描かれている。図7にコマごとのモデルの位置を示す。

## 4.2 拘束による誤差の解消

姿勢の差分を累積する方法は、確かに計算コストが少なくすむ。しかし、差分の推定誤差も累積され長時間の測定は難しい。

この誤差を修正する方法としてキーフレーム法が提案されている [14]。この方法は、最初のコマだけではなく、最後のコマでもモデルを身体に当てはめておく。この当てはめによって得られた姿勢と、変化分の累積によって得られた姿勢が最終コマで一致するようにするのである。

姿勢の誤差を解消するためには、最後から一つ手前のコマで、当てはめによって与えられた最終姿勢になるようにモデルの姿勢を変化させればよい。これは可能ではあるが、姿勢の変化分が最後だけ異常に大きくなり、人間の動作としては不自然である。

自然な動作を得るためには、全コマでの姿勢の変化分を再計算する必要がある。その計算は、第1に動作が滑らかに推移するように、第2に累積姿勢が最終姿勢に一致するように、第3に差分の推定式(7)あるいは(8)を満たすように行われる。この計算は最終姿勢を途中の姿勢に伝播させる繰り返し演算により行われる。

動作の早さあるいは複雑さによっては、モデルの当てはめが最初と最終のコマだけでは足りず、間の幾つかのコマでも行う必要がある。このモデルの当てはめを行ったコマをキーフレームとよぶことにする。例として、円盤投げの動作を追跡した結果を示す。

キーフレーム法は累積誤差の直接的な修正法であ



図 8: 円盤投げの追跡 [14]

るが、修正のための繰り返し計算に膨大な時間がかかる。そのためこれはリアルタイム処理向きではない。繰り返し処理を必要としない誤差の解消法として、環境からの身体への拘束を利用する方法がある [26]。

身体は必ずその環境に接している。身体が面に接しているとき、接点の動きは面に拘束される。身体がある軌道に沿って動いている場合には、適当な 2 つの面に同時に拘束されているとする。さらに、空間中の一点に固定されている場合には、適当な 3 つの面に同時に拘束されているとする。これらの拘束条件を使えば、動作パラメータの自由度を制限し、ノイズの混入による誤差の拡大を防ぐことができる。

動作の推定問題は、多関節モデルをロボットと考えれば、マニピュレータの軌道計算問題と等価である。

## 5 動作の再生と編集

ビデオ映像は、それを撮影したカメラから見た動作の再現である。もし、動作中の身体形状が 3 次元的に測定されているならば、座標系の変換により、任意の視点から見た映像を作ることができる。

CG では、動作を測定したときと異なる状況下で、動作の再現が要求される。例えば、通常の歩行データから、障害物が存在する場合の歩行の生成などがある。このためには、動作データの編集を必要とするが、生成した動作が力学的な妥当性を持っていないとしない [17]。図 9 は、通常の幅跳びの動作データを力学データに変換した上で、ひねりを加えた幅跳びの動作を作り出している。

このことは、動作を位置や姿勢といった、見かけの記述だけではなく、動作の要因まで踏み込むことの必要性を意味している。

## 6 あとがき

動画像を使った身体動作の解析技術について現状を整理した。今後、この技術が広く使われてゆくためには、(1) 2 次元画像から身体 3 次元姿勢推定、(2) 動画像からの立体形状測定、(3) 動作や形状測定の自動化・高速化、(4) 動作要因の推定、などの課題に取り組む必要がある。

動画像解析技術は、既存のモーションキャプチャに比べて、優れた点が多くある。なかでも、決定的な

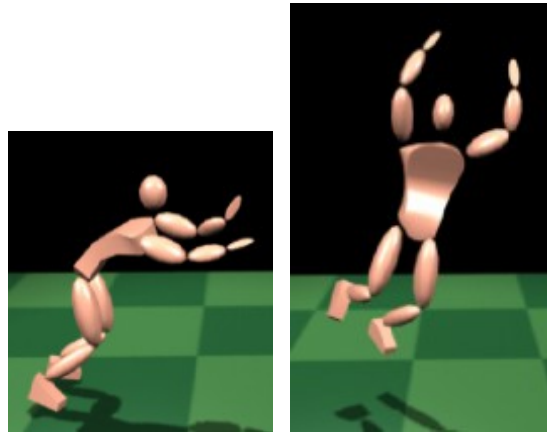


図 9: 左: 普通の幅飛び、右: 捻りを加えた幅飛び。Z.Popovic ©1999

のは、過去の映像からでも動作が測定できる点であろう。これは、映画の発明以来、現在に至るすべての映像が解析対象であることを意味している。この場合、解析のリアルタイム性は、必ずしも要求されないが、解析を待っている膨大な映像データを前にしたとき、処理の自動化・高速化は必要である。

謝辞 図の使用を許可していただいた、D.M.Gavrila, I.Haritaoglu, J.Rehg, Z.Popovic の各氏に感謝致します。

## 参考文献

- [1] 天谷賢治, 原 裕二, 青木 繁, "逆解析手法による 3 次元人体運動の再構成", 機械学会論文集(C 編), vol.63, no.608, pp.1167-1171, 1997
- [2] 浅田, 谷内田, 辻: 物体の 3 次元運動の解釈, 信学論, vol.J65-D, no.6, pp.606-613, 1982
- [3] C.Barron and I.A.Kakadiaris, "Estimating anthropometry and pose from a single image", IEEE CVPR, pp.669-676, 2000
- [4] Gavrila, D.M. and Davis, L.S.: 3-D model-based tracking of humans in action: a multi-view approach, *Proc. of IEEE conf. on Computer Vision and Pattern Recognition*, pp.73-80, 1996.
- [5] Haritaoglu, I., Harwood, D, and Davis, L.S.:  $W^4$ : Who? When? Where? What? Arealtime system for detecting and tracking people, *Proc. of the 3rd int'l Conf. on Automatic Face and Gesture Recognition*, pp.222-227, 1998

- [6] B.K.P.Horn and B.G.Schunck: Determining optical flow, *Artificial Intelligence*, vol.17, pp.185-203, 1981
- [7] 岩井儀雄, 八木康史, 谷内田正彦, "単眼動画画像からの手の3次元運動と位置の推定", *信学論*, vol.J80-D-II, no.1, pp.44-55, 1997
- [8] Kakadiaris, i.A. and Metaxas, D.: Model-based estimation of 3D human motion with occlusion based on active multi-viewpoint selection, *Proc. of IEEE conf. on Computer Vision and Pattern Recognition*, pp.81-87, 1996.
- [9] 亀田能成, 美濃導彦, 池田克夫, "シルエット画像からの関節物体の推定法", *信学論*, vol.J79-D-II, no.1, pp.26-35, 1996
- [10] M.K.Leung and Y.H.Yang: First sight: A human body outline labeling system, *IEEE PAMI*, vol.17, no.4, pp.359-377, 1995
- [11] D.D.Morris and J.M.Rehg, "Singularity analysis for articulated object tracking", *IEEE CVPR*, pp.289-296, 1998
- [12] Narayanan, P.J., Rander, P.W., and Kanade, T.: Constructing virtual worlds using dense stereo, *Proc. of the 6th int'l Conf. on Computer Vision*, pp.3-10, 1998
- [13] 大垣健一, 岩井儀雄, 谷内田正彦, "動きと形状モデルによる人物の姿勢推定", *信学論*, vol.J82-D-II, no.10, pp.1739-1749, 1999
- [14] 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画像からの人間動作の追跡, *電子情報通信学会論文誌*, Vol.J81-D-II, No.9, pp.2008-2018, 1998
- [15] 大谷淳, 岸野文郎: 遺伝的アルゴリズムを用いた多眼画像からの人物の姿勢のモデルベース推定, *映像情報メディア学会誌*, Vol.51, No.12, pp.2107-2115, 1997
- [16] J.O'Rourke and N.I.Badler: Model-based image analysis of human motion using constraint propagation, *IEEE PAMI*, vol.2, no.6, pp.522-536, 1980
- [17] Z.Popovic and A.Witkin: Physically based motion transformation, *Proc.SIGGRAPH99*, pp.11-20, 1999
- [18] Rehg, J.M. and Kanade, T.: Visual tracking of high DOF articulated structures: an application to human hand tracking, *Proc. of the 3rd European Conf. on Computer Vision*, Vol.2, pp.35-46, 1994.
- [19] 佐藤明知, 川田聡, 大崎喜彦, 山本正信: 多視点動画画像からの人間動作の追跡と再構成, *電子情報通信学会論文誌*, Vol.J80-D-II, No.6, 1997.
- [20] 島田伸敬, 白井良明, 久野義徳: 確率に基づく探索と照合を用いた画像からの手指の3次元姿勢推定, *電子情報通信学会論文誌*, Vol.J79-D-II, No.7, pp.1201-1217, 1996.
- [21] 玉木徹, 山村毅, 大西昇: オプティカルフローを用いた複雑背景下における人物の腕領域の抽出とパラメータ推定, *電気学会論文誌 C 分冊*, Vol.120-C, pp.1801-1808, 2000
- [22] C.J.Taylor, "Reconstruction of articulated objects from point correspondences in a single uncalibrated image", *IEEE CVPR*, pp.677-684, 2000
- [23] 築山, 白井: 連続する画像からの人の動きの検出, *計測自動制御学会論文集*, vol.19, no.1, pp.909-915, 1983
- [24] S.Tsuji, M.Osada and M.Yachida: Tracking and segmentation of moving objects in dynamic line drawing, *IEEE PAMI*, vol.2, no.6, pp.516-522
- [25] J.A.Webb and J.K.Aggarwal, "Visual interpreting the motion of objects in space", *IEEE Computer*, vol.14, no.8, pp.40-46, 1981
- [26] 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画像追跡, *映像情報メディア学会誌*, Vol.52, No.3, pp.331-336, 1998
- [27] 八木下勝利, 山本正信: 固有空間を利用した単眼視画像からの人体の姿勢推定, *電子情報通信学会技術報告*, PRMU99-85, 1999
- [28] 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.