

多視点カメラを用いた3次元人体動作の実時間推定

谷口 倫一郎

米元 聡

九州大学
大学院システム情報科学研究院
rin@limu.is.kyushu-u.ac.jp

九州産業大学
情報科学部設置準備室
yonemoto@ip.kyusan-u.ac.jp

実生活環境において誰もが利用可能な仮想環境インタフェースを構築するためには、機器の装着を強いられることなく、しかも実時間で人間の3次元動作を計算機へ入力する手段が必要となる。本研究では、センサとしてカメラを用いて人間の頭部や手、足等の部位の3次元位置を追跡し、それらから人間の3次元動作を推定する手法について提案する。人体部位の追跡は色領域 (blob) をベースに行い、複数視点からの観測結果を融合して部位の3次元位置を安定に求めている。しかし、この手法だけでは人間の姿勢推定に必要な肘、膝の位置が得られないので、これらの位置を求めるために逆運動学を実時間で近似的に解く方法を導入し、人間の3次元姿勢を実時間で推定することを可能にしている。

A Vision-based Real-time Human Motion Capture System Using Multiple Cameras

Rin-ichiro Taniguchi

Satoshi Yonemoto

Department of Intelligent Systems
Kyushu University
rin@limu.is.kyushu-u.ac.jp

Kyushu Sangyo University
yonemoto@ip.kyusan-u.ac.jp

This paper presents a real-time human motion capture using skin-color-based blobs and human motion synthesis based on real-time inverse kinematics. Our purpose is to do seamless mapping of human motion in the real world into virtual environments. In general, virtual environment applications such as 'smart' interface require real-time human motion tracking system without special devices or markers. However, since such vision-based human motion tracking system is essentially unstable and can only acquire partially observable information, we have to establish a robust estimation mechanism of visual features from the observable information. In this paper, we demonstrate a real-time and on-line interaction system using multiple cameras which realizes human figure motion synthesis from limited perceptual cues.

1 はじめに

近年は *SmartRoom*[1] など、接触型の器具を用いることなく、ビデオカメラを用いて人間の行為を知覚するような知的なサイバースペース生成の研究が進んでいる。これは非接触な方式を実現することが、コンピュータビジョンの達成すべき研究課題であるばかりでなく、仮想空間への入口となる計測空間と日常の生活空間とをシームレスに繋げるために有効と考えられているからであり、本研究もこのようなセンサを装着することによる制約を打破する非接触型の知覚系の構築を目指すものである。

1.1 関連研究

ビデオカメラで撮影した映像より人間の動作を推定する非接触型の方法は、従来より盛んに研究が進められてきた。特に、人体に関する詳細な形状・動き情報を推定する方法としては、オフラインで獲得された画像系列に対しモデルベースで解析する方法 [2][3] がある。しかし処理時間が膨大であることや、モデル像と画像のマッチングを行なうために必要な初期情報は通常オフラインでしか得ることができないという制約により、実時間、オンラインでの適用には課題が残っていた。一方、近年の計算機能力の向上により、オンラインでの画像獲得およびその解析がようやく可能となり、今日では画像を詳細に解析するアプローチとは異なり、比較的単純かつ高速な画像解析手法を用い、実時間性を重視したものが開発され始めている。ほぼ実時間で動作する人体動作解析システムには大別して色情報を用いるもの、輪郭情報を用いるものの2つのアプローチがあり、具体的には以下のようなものが提案されている。

色情報を用いる手法 *Pfinder*[1] は、一般的な領域クラスタリング手法を用いて2次元の blob 特徴、すなわち人体を推定するための画像特徴としての領域を検出する。blob 特徴とは、均一な色領域の重心およびその形状のモーメントで定義される特徴である。しかし、人体の部位を領域に対応付けるためのアルゴリズムや、シーンの制約条件が不明であり、人体を自動に推定するためには何らかの事前情報を必要とする。実際のアプリケーションでは、抽出した2次元の blob 特徴をもとに、2視点からのステレオ計測により3次元情報の推定を行なっている。比較的安価なシステム構成でありながら実時間に近い速

さで動作可能であり、非接触方式による比較的精度の低い抽出結果でも多くのアプリケーションが十分に機能することを示した。

最近 *Pfinder* をベースに、システムに物理法則に基づいた制御機構を導入することで、抽出結果の不安定さを回避し、リアリティの高い上半身動作の再現を達成することが可能な方法が開発された [4]。予め想定した動作に対しては良好な再現結果を得ることができるが、全身動作のように多くの動作表現を行うためには工夫が必要である。

シルエット輪郭情報を用いる手法 シルエット輪郭から人体情報を解析し、その行動、人数などを認識する手法 [5]。このシステムは人間のモニタリングシステムであり、人体モデルの再現を目的としていないため、連続的な人体動作の再現には適していない。

1.2 人体動作推定の問題

上述の関連研究からも分かるように、人体動作の推定問題は、なるべく制約のない、リアリティの高い推定を実時間で行う方式の開発が課題であると言える。一般に、人体動作の推定問題は、図1で示すような階層からなっている。第一段階は、画像から人間を抽出するのに必要な画像特徴を抽出するレベルである。第二段階はそれらの画像特徴を統合し、人体の3次元的な姿勢を抽出するレベルであり、基本的には物理的な情報を取り扱う世界である。現在までの研究のほとんどは、この段階までにとどまっている。第三段階は、第二段階で得られた結果を基に、さらに詳細な情報や時系列情報を獲得・統合して、行動や意図、感情などの抽象化された概念を抽出するレベルである。第三段階を実現するためにも、第二段階でリアリティの高い人体動作を推定・再現する必要がある。本研究では、以下の課題を実現するシステムを構築する。

- 少ない画像特徴からリアリティの高い人体動作を再現する。
- 実時間で動作可能であり、ユーザ介入の後処理を必要としないなどオンライン性がある。
- 計測範囲を拡張し、様々な姿勢を認識できるよう、多視点情報を利用する。

また、本システムでは前述の *Pfinder* システムと同様に、画像特徴として色領域 blob を用いる。その理由は、画像上で手先や足先の追跡が他の画像特

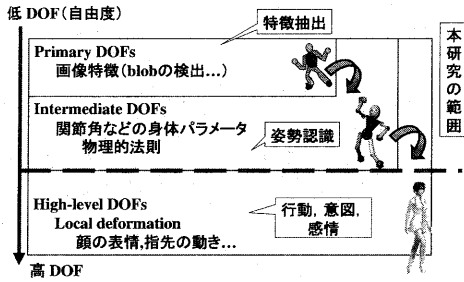


図 1: 人体動作のモデル

微に比べ安定で、抽出精度の人体姿勢依存性が低いと考えるからである。

2 Blob 特徴の追跡

以下では、手や顔に相当する色領域 blob を追跡し、その 3 次元位置を求める方法について述べる。

2.1 領域の検出

背景差分と外接矩形の検出 前処理として、背景差分を施し、人体領域を囲む外接矩形を検出する。

色識別 本手法では、入力画像中に観測される肌色領域は手、顔のいずれかの領域であると解釈される。色識別の方法に、照明変化に比較的ロバストな方法として、パラメトリックな色モデルによる識別手法を用いている [6]。各画素の色特徴 (r, g, b) が、次のような濃淡値 i をパラメータとした 2 次形式で表せると仮定する。

$$\begin{aligned} R &= R_2 i^2 + R_1 i, \\ G &= G_2 i^2 + G_1 i, \\ B &= B_2 i^2 + B_1 i \end{aligned} \quad (1)$$

ここで、6 つのモデルパラメータ係数 R_1, \dots, B_2 は、予め学習用実画像から推定したものである。

色識別においては、以下の式により、各画素について観測される色特徴 (r, g, b) とモデル色特徴間のマッチングを行い、しきい値処理を行う。

$$error = (\hat{r} - \hat{R})^2 + (\hat{g} - \hat{G})^2 \quad (2)$$

ここで、 $\hat{r} = r/(r+g+b)$, $\hat{g} = g/(r+g+b)$, $\hat{R} = R/(R+G+B)$, $\hat{G} = G/(R+G+B)$ 。

2.2 Blob の定義

領域特徴の 1 次モーメントまで考慮したガウシアン blob 特徴は、楕円形状モデル $e = (\mu, \Sigma)$ である (平均 μ , 共分散 Σ)。しかし、肌色領域のように、同じ色特性の領域を複数個追跡する場合、このモデルでは画素単位での識別に誤りが生じる恐れがある。領域追跡を正確に行うためには、より正確な領域形状、領域境界を把握する必要がある。

そこで、本研究では、blob 特徴に正確な領域境界を表現可能なように局所的な変形 d を追加する。

$$p(s) = e(s) + d(s) \quad (3)$$

ここで、 s は楕円周のパラメータを表す。以下、「blob」を、式 (3) の p と定義する。

2.3 画素の識別

初期ラベリング まず、外接矩形内の領域を各ガウシアン blob 領域へとクラスタリングする。どの部位に相当するか決定する尺度としては、前フレームまでに推定された 2 次元重心位置 (あるいは予測位置) との近接性のみを用いる。これにより、初期の推定位置 (μ', Σ') が求まる。

変形楕円による輪郭境界の探索 次のステップは、初期推定位置 (μ', Σ') を用いて領域境界を決定する、すなわちどの画素が各 blob に属するかを厳密に決定することである。境界探索のための探索領域を、初期楕円の周上に配置する (図 2(上) 参照)。各探索点における局所窓 W から、尤もらしい領域境界を決定する。通常、境界を決定する尺度としてエッジ勾配などが用いられるが、本手法では、衣服のエッジの影響も考えられるため、以下のような色の分離境界を定める分離度 $J(p)$ をその尺度としている。

$$\begin{aligned} & \text{if } \sum_{(x_i, y_i) \in W} n(x_i, y_i) > 0 \\ J(p) &= |N_W/2 - \sum_{(x_i, y_i) \in W} p(x_i, y_i | skin) + \epsilon|^{-1} \end{aligned} \quad (4)$$

ここで、 $p(x_i, y_i | skin)$ は肌色の尤度、 N_W は W 内の画素数、 $n(x_i, y_i) = \mathbf{p}_i^T \cdot \mu' \mathbf{p}$ (図 2(下) 参照)。したがって、 $d(s) = \arg \max J(\cdot)$ 。

推定した変形楕円内で対応画素の重心 μ を計算することにより、blob の 2 次元位置を求める。

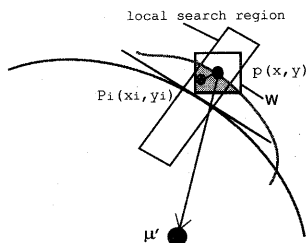
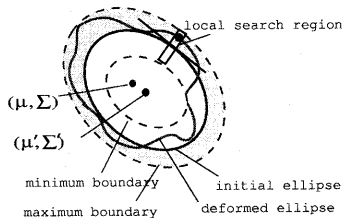


図 2: (上) 変形楕円モデル. (下) 境界探索の領域.

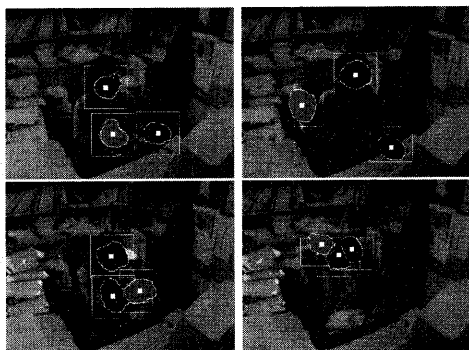


図 3: 推定した変形楕円領域の例.

2.4 3次元位置の復元

2 視点の画像から blob 特徴が観測される時、その 3 次元位置をステレオ視により計算することが可能である。予め得られるキャリブレーション情報により、カメラ座標原点と blob 中心位置を結ぶ視線を各視点について計算し、それら視線の交差位置を求める。

視線 1 が、 $\mathbf{T}_1 = \mathbf{o}_1 + t_1 \mathbf{d}_1$ と表現され (t_1 は媒介変数、 \mathbf{o}_1 はその視点のカメラ座標原点、 \mathbf{d}_1 はその単位視線ベクトル)、視線 2 が、 $\mathbf{T}_j = \mathbf{o}_2 + t_2 \mathbf{d}_2$ (t_2

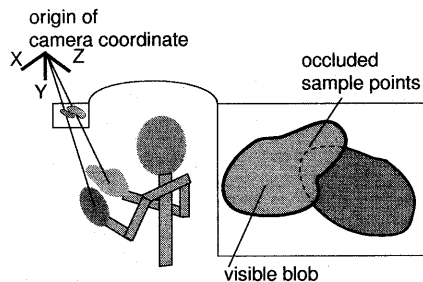


図 4: 隠れの検出.

は媒介変数、 \mathbf{o}_2 はカメラ座標原点、 \mathbf{d}_2 は単位視線ベクトル) と表現される時、視線 1 上にあり、視線 2 から最小距離である点 \mathbf{T} は、 t_1 をパラメータとして求めると以下ようになる。

$$\mathbf{T} = \mathbf{o}_1 - \frac{(\mathbf{d}_1 \times \mathbf{m}_2, \mathbf{o}_1 \times \mathbf{m}_2 - \mathbf{n}_2)}{\|\mathbf{d}_1 \times \mathbf{m}_2\|^2} \mathbf{d}_1 \quad (5)$$

ただし

$$\mathbf{m}_j = \frac{\mathbf{d}_j}{\sqrt{1 + \|\mathbf{o}_j \times \mathbf{d}_j\|^2}},$$

$$\mathbf{n}_j = \frac{\mathbf{o}_j \times \mathbf{d}_j}{\sqrt{1 + \|\mathbf{o}_j \times \mathbf{d}_j\|^2}}.$$

この点 \mathbf{T} が 3 次元 blob 位置である。

2.5 隠れの検出

人体領域の追跡においては、顔領域が一定位置にあり、手領域が顔領域を隠す状況や、片方の手領域がもう一方の手領域を一時的に隠し、横切るといような状況が起こり得る。この場合、blob 領域の可視性が重要である。推定したモデルの 3 次元予測位置を再投影することにより、どの blob が可視であるかを判定することが可能である。このモデルベースの予測を「画像生成による解析」[2]と呼ぶ。具体的には、まず、カメラ座標原点からの距離を Z バッファ法により計算し、次フレームでの blob の可視性を求める。それをもとに、隠れが起こっている blob についてはその影響を最小限にするため、画素識別の際に隠れた境界の推定を行わず、初期楕円弧に境界を固定するように制限する(図 4 参照)。

3 人体モデルの動き生成

3.1 推定 blob の人体モデルへのあてはめ

知覚モジュールにより得られるのは胴体重心・頭部・両手・両足の計6つの3D blob位置(以後, これを知覚データと呼ぶ)のみである. 輪郭解析などを行い肘・膝などの関節位置を推定する方法[7][8]もあるが, 不安定な結果しか得られず推定可能な条件も厳しくなる. したがって, 限られた少数の知覚データを用いて人体モデルの姿勢を生成するため, 以下の逆運動学(インバースキネマティクス)を用いてこの問題を解決する. なお人体モデルの表現としては, 知覚データより求める位置に両膝・両肘を加えた部位数14, 自由度23(詳細は後述)を考えた多関節構造モデルとする(図5).

3.2 実時間逆運動学

3.2.1 従来の手法

逆運動学問題は一般に, ロボティクス, コンピュータグラフィクス[9]の分野で用いられる. 解法のアプローチとしては, 解析的に解く方法[10], 数値解析で解く方法[11]があり, また, ヤコビ法, 最適化法などが提案されている. また, 腕のみであるが, 実際の人体の動きデータを収集しそれらから導出した線形関係式を用いるユニークな方法などもある[12]. ロボティクスおよびその関連分野では, 生体に近い腕の到達運動など研究されているが, 実際は動力学と併用されることが多く, 質量, 弾性など対象に関する多くの事前情報が必要とされる. 一方, コンピュータグラフィクスの分野でも近年, 実時間での逆運動学問題解法が開発されつつある. しかしながら, 利用用途がモデリングツールであり, ユーザとのインタラクションを前提としているため, 時系列データを必要としない最適化手法を用い, 局所解に陥った場合はユーザが修正するようなアプローチが用いられている.

3.2.2 提案手法

我々の目的においては逆運動学に関し, 上述の一般的手法と異なる以下の性質が望まれる.

- 実時間で両腕両脚の計4つの接続リンクに関する逆運動学について同時に解が得られる

- 知覚モジュールにより入力されるゴール(3-D blob)位置は精度が悪いということを前提とする
- 連続的かつ自然な人体の動きを与える

本論文では, 上記の性質を考慮したアプローチとして解析的に解く方法を提案する. まず逆運動学におけるゴールとして様々なものが考えられるが, 我々のような目的の場合, 知覚モジュールから入力される3-D blob位置が正確でないことが起こりうる. 特に, ゴール位置が解の存在範囲を越えて設定された場合, 連続性が満たされなくなる. よって, 実際のゴールを方向およびその距離(絶対値)と定義することで対処する. つまり, ゴールの方向は必ず満たし, その距離は設定ゴール位置が解の存在範囲内である時のみ3-D blob位置に一致させる(図5の左上参照). 一般に位置姿勢表現はロール・ピッチ・ヨーなどの回転・並進自由度が用いられ, それらが多関節構造の相互依存関係にあるため, このままでは解析的に解くことは困難である. そこで解析的に解けるよう自由度を適切に分ける方法が必要である.

3.2.3 解法

2つの可変リンクに関する運動学方程式を以下のように定義する.

$$\begin{aligned} \mathbf{g}_i = & \mathbf{T}^b \mathbf{R}^b(0, R_y^b, R_x^b) \mathbf{R}^{b'}(R_z^{b'}, 0, 0) \\ & \mathbf{T}^{l_1} \mathbf{R}^{l_1}(R_z^{l_1}, R_y^{l_1}, 0) \mathbf{R}^{l_1'}(R_z^{l_1'}, 0, R_x^{l_1'}) \quad (6) \\ & \mathbf{T}^{l_2} \mathbf{R}^{l_2}(0, 0, R_x^{l_2}) \mathbf{t}^e \end{aligned}$$

ここで, $\mathbf{g}_i = (g_x, g_y, g_z, 1)^T$ は, ゴール $i(1, \dots, 4)$ を表す位置ベクトル, \mathbf{T}^b および $\mathbf{R}^b, \mathbf{R}^{b'}$ は胴体に関する相対並進, 回転行列, \mathbf{T}^{l_1} および $\mathbf{R}^{l_1}, \mathbf{R}^{l_1'}$, \mathbf{T}^{l_2} および \mathbf{R}^{l_2} はリンク1, 2に関する相対並進, 回転行列である('のものは解析的に解くために分けられた回転行列を表す). \mathbf{t}^e はリンク2の末端のゴール位置への局所座標系での並進ベクトルである. また $\mathbf{R}(R_z, R_y, R_x)$ の回転自由度 R_z, R_y, R_x はそれぞれロール・ピッチ・ヨーを表す.

すると, 逆運動学方程式の解析解としては, その幾何学的な性質より

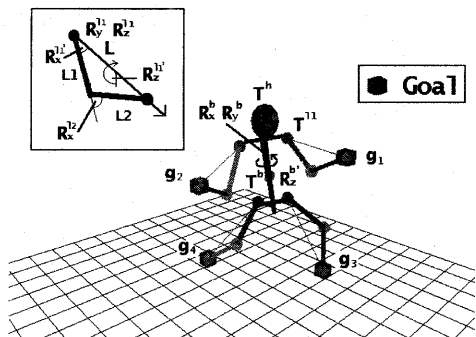


図 5: 人体モデルの構造およびゴール方向の定義

$$R_y^{l1} = -\arccos\left(\frac{g_z - wT_z^{l1}}{\|g - wT^{l1}\|}\right)$$

$$R_z^{l1} = -\arctan\left(\frac{g_y - wT_y^{l1}}{g_x - wT_x^{l1}}\right)$$

$$R_x^{l1} = \arccos\left(\frac{L_1^2 + L_2^2 - L^2}{2L_1L}\right)$$

$$R_x^{l2} = \arccos\left(\frac{L_1^2 + L_2^2 - L^2}{2L_1L_2}\right) - \pi$$

$$(|L_1 - L_2| \leq L \leq L_1 + L_2)$$

が得られる。ここで、 L_1, L_2, L はそれぞれリンク 1, リンク 2 の長さ, リンク 1 の原点からゴールまでの距離を表し, wT^* はワールド座標を表す(図 5 参照)。

また、胴体の方向に関しては、胴体位置 wT^b および頭部位置 wT^h により $wT^h - wT^b$ の方向に体軸が平行になるように決定する。

$$R_x^b = -\arcsin\left(\frac{wT_y^h - wT_y^b}{\|wT^h - wT^b\|}\right)$$

$$R_y^b = -\arctan\left(\frac{wT_x^h - wT_x^b}{wT_z^h - wT_z^b}\right)$$

なお、 R_z^{l1} は(6)式により、両腕両脚のそれぞれに設定されるパラメータであり、腕に関しては、肘の上げ下げ、脚に関しては膝の開き具合に対応する、より人間らしい動きを与えるための回転角である。これらは逆運動学の解からは直接得られないが、人体の姿勢表現には必要なパラメータであり、ここでは実際の測定値から実験的な定数値を当てはめた。また、 R_z^b は人体の向き(パン)を定める回転角であり、これも逆運動学の解には直接影響を与えない、

人体の姿勢表現には必要なパラメータであるが、人体の向き(パン)を 6 つの blob 位置のみから推定することは難しいため固定値にしており、その自動推定法の開発は今後の課題である。

本手法の特徴をまとめると以下のようになる。

- 解析的な方法であるため実時間計算可能
- 少数の知覚データから人体モデルの姿勢を定める目的に特化しているため 2 つの可変リンクのみ適用可能
- 逆運動学の解には直接現れないが、人体の姿勢には影響を与えるパラメータを設定できるため多くの動作表現が実現可能

4 実験結果

4.1 システム構成

実時間・オンラインでの動作を実現するため、本手法を PC クラスタ上に実装した。本実験においては、PC クラスタを、高速ネットワーク Myrinet で接続された最大 8 台の PC(700MHz × 2 の CPU) と 6 台の IEEE1394 ベースのデジタルカラービデオカメラ SONY DFW-V500 により構成している。

以下の実験では、2~6 台の PC が画像獲得・2 次元画像処理を行い、1 台の PC が 2 次元画像処理を行った PC からの情報を受信し、3 次元の復元処理、人体モデルの動作生成、人体モデルの描画を行う。さらに 1 台の PC が 2 次元画像処理を行う PC に同期信号を送信し、PC 間の同期を管理する(図 6)。入力画像のサイズは 320 × 240 画素(YUV:422)で 30fps で動作可能である¹。

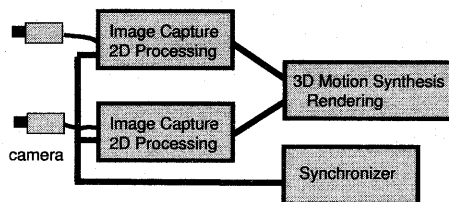


図 6: 多視点動画画像処理の PC クラスタへの実装。

¹DFW-V500 の仕様上、外部トリガにより同期をとると、最大 15fps の速度でしか画像を取り込めないが、システムの性能としては 30fps のスループットが得られている

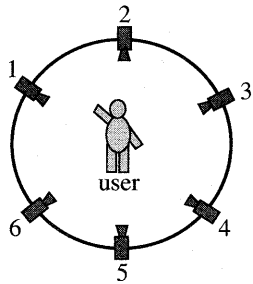


図 7: 6 視点カメラの均等配置.

4.2 多視点カメラの利用による計測範囲の拡張

本手法においては、人体動作の計測は正面像の対に対してのみ適用可能である。したがって、多視点カメラを用いて正面像の選択を行うことができれば、計測範囲の拡張を行うことが可能となる。しかしながら、高精度に人体の向き(パン角)を求めるには胴体を正確に測定する必要があり、現在のアプローチでは困難である。そこで本実験では簡単に、両足先のなす角度により、正面像の視点対を選択することを考える。すなわち、以下の手順で視点選択を行う。

- 鉛直上方からの仮想カメラ座標を想定し、左右足先の投影点 p_l , p_r を結ぶ 2 次元ベクトルと、各視点のカメラ座標原点の投影点 p_c 、ワールド座標原点の投影点 p_w を結ぶ 2 次元ベクトルを求める。
- それらを正規化し、ベクトルのなす角度を計算することにより、最適なすなわち直角に近い角度の隣り合う視点対を選択する。

本実験では、6つのカメラを同等の視野を保ちつつ60度の均等間隔で円周上に配置し(図7)、直立状態の動作に関しては、どの向きでも正面像を観測可能な観測環境を想定した。図8に、その入力画像および人体モデルの再構成像を示す。また、選択された視点対を太枠で示している。再投影による blob の予測位置を用いているので、滑らかな視点の切替えを行うことができ、推定位置の変動の影響は少ないことが確認できた。6つのカメラ程度でも、あらゆる人体の向きについての安定な正面像を選択可能であった。

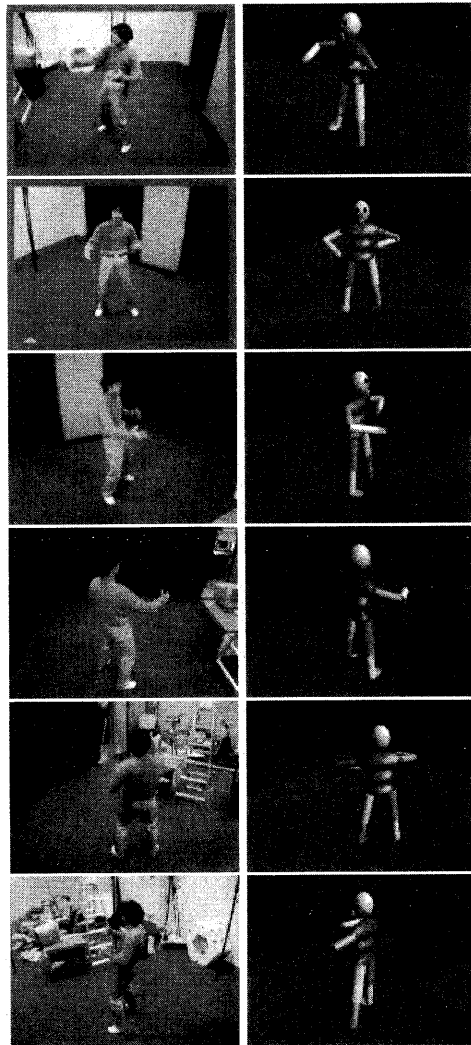


図 8: オンライン視点選択による全身動作の再構成の例(左: 入力画像(視点1-6). 右: 同一視点からの人体モデル再構成像).

4.3 デスクトップインタフェースへの適用

小規模なアプリケーションとして、カメラ2台を用いた上半身の動作の再現を行うデスクトップ型の3次元操作用インタフェースを構築した(図9)。このアプリケーションでは、PCは2D処理に2台、3Dと表示に1台、同期制御に1台の計4台のPCを用いている。このようなアプリケーションでの利用においては、手と顔の重なりや交差が問題となるが、

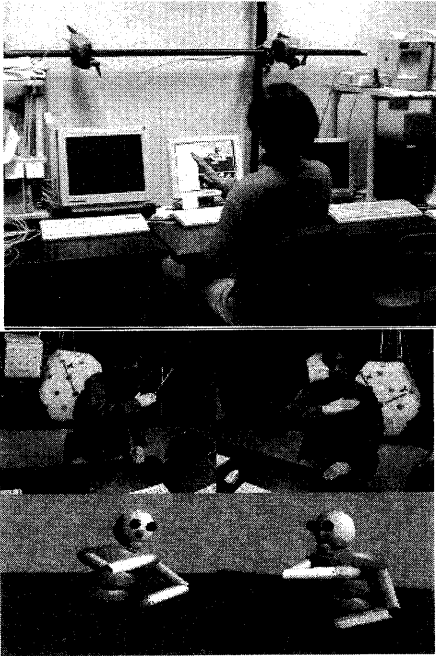


図 9: デスクトップインタフェースへの適用: (上) システム構成。 (下) 入力画像と同一視点からの推定モデル像。

本手法により, ある程度誤対応を防ぐことができた。

5 おわりに

本論文では多視点カメラを用いた非接触型の実時間人体動作推定システムを提案した。特に, 複数 blob 存在下での安定な追跡方法, 小数の知覚データに対する実時間逆運動学の導入による人体動作の再生方法について述べた。本システムはオンラインかつ実時間で動作するため, 仮想空間とのインタラクションなど様々なアプリケーションへの適用が可能である。特殊なマーカなど装着不要な高度なモーションキャプチャシステムであり, しかも全周に配置した多視点カメラを用いることによって, 計測範囲の拡張を行ったので, その応用範囲は広いと考えられる。本システムに力学的あるいは感性的な動作フィルタリングを施すことによって, より人間らしい動作の生成や, 知覚データに大きく影響されない安定な動作の再現が可能となると考えられるが, これは今後の課題である。

謝辞

本研究の一部は, 日本学術振興会未来開拓学術研究推進事業「分散協調視覚による動的 3 次元状況理解」プロジェクト (JSPS-RFTF 96P00501) の補助を受けて行った。

参考文献

- [1] C. Wren, A. Azarbayejani, T. Darrell, A. Pentland: Pfunder: Real-Time Tracking of the Human Body, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.19, No.7, pp.780-785, 1997.
- [2] S.Yonemoto, N.Tsuruta, R.Taniguchi: Tracking of 3D Multi-Part Objects Using Multiple Viewpoint Time-varying Sequences, Proc. of 14th Int. Conf. on Pattern Recognition, pp.490-494, 1998.
- [3] C.Bregler, J.Malik: Tracking People with Twists and Exponential Maps, Proc. of Conf. on Computer Vision and Pattern Recognition, pp.8-15, 1998.
- [4] C.Wren, A.Pentland: Understanding Purposeful Human Motion, Proc. of 4th Int. Conf. on Automatic Face and Gesture Recognition, 2000.
- [5] I.Haritaoglu, D.Harwood, L.S.Davis: W⁴: A real-time system for detecting and tracking people, Proc. of Conf. on Computer Vision and Pattern Recognition, p.962, 1998.
- [6] 米元, 有田, 谷口: 多視点動画画像処理による実時間全身モーションキャプチャシステム—視覚に基づく仮想世界とのインタラクション—, 映像情報メディア学会誌, Vol.54, No.3, pp.409-416, 2000.
- [7] Y.Azoz, L.Devi, R.Sharma: Reliable Tracking of Human Arm Dynamics by Multiple Cue Integration and Constraint Fusion, Proc. of Conf. on Computer Vision and Pattern Recognition, pp.905-910, 1998.
- [8] 岩澤, 海老原, 竹松, 坂口, 大谷: Shall We Dance? の構築, 信学技報, PRMU98-114, pp.15-22, 1998.
- [9] C.Welman: Inverse Kinematics and Geometric Constraints for Articulated Figure Manipulation, MSc Thesis, CS, Simon Fraser University, 1993.
- [10] D.Tolani: An Inverse Kinematics Toolkit for human modeling and simulation, PhD Thesis, University of Pennsylvania, 1998.
- [11] J.Zhao, N.Badler: Inverse Kinematics positioning using nonlinear programming for highly articulated figures, Transactions, on Computer Graphics, Vol.13, No.4, pp.313-336, 1994.
- [12] Y.Koga: Planning Motions with Intentions, Proc. of SIGGRAPH'94, pp.24-29, 1994.