

ボリュームデータの細線化とグラフマッチングを用いた事例ベース人体姿勢推定

田中 秀典[†] 中澤 篤志^{†,††} 竹村 治雄^{†,††}

特殊なマーカや計測装置の装着が不要なビジョンベースの人体姿勢推定（モーションキャプチャ）手法において、人体部位同士の接合等によるトポロジ変化に対応できる手法はほとんど提案されていなかった。そこで我々は、あらかじめ用意した様々なトポロジを持つ人体形状の特徴のデータベースを用いて安定的に姿勢推定を行う手法を提案する。まず、複数のカメラから人物の動きを撮影し視体積交差法によってボリュームデータを求め、細線化処理によってトポロジを保持したままスケルトンを得る。次にスケルトンを属性付きグラフで表現し、あらかじめ用意した人体の取りうる様々なトポロジの事例を同様にグラフ化したものとのマッチングを行ってスケルトンの各部に対応する人体部位を推定する。そしてスケルトンにおいて曲率の高くなる位置を時系列のスケルトンから求め、その位置を各部位における関節の位置として推定する。8台のカメラを用いた動作実験によって従来手法では困難だった体のトポロジが変化する動作に対しても本手法が対応できることが確認できた。

Example-Based Approach for Human Pose Estimation Using Volume Data and Graph Matching

HIDENORI TANAKA,[†] ATSUSHI NAKAZAWA^{†,††}
and HARUO TAKEMURA^{†,††}

In this paper, we propose a novel example based marker free motion capture method. First, volume data of actor's body and its skeleton are obtained from multiple camera views at each frame through visual hull based method and thinning process. Then, examining the structure of the skeleton, we can acquire an attributed graph and comparing the acquired graph and the graphs in the Model Graph Database, correspondences between body parts and skeleton parts are obtained. Finally, we estimate the position of the joints examining the curvature of the skeleton. The experiment result shows the validity our approach.

1. はじめに

人体の姿勢・動作を計測するモーションキャプチャ手法は、コンピュータグラフィックスやVR、リハビリ医療、ロボティクス、デジタルアーカイブ等の幅広い分野において利用されている。しかし従来の手法では計測対象の人物に光学マーカや磁気センサなどを装着する必要があった。そこで近年カメラ映像を用いた、マーカやデバイスの装着が不要なモーションキャプチャ手法がコンピュータビジョンの分野において研究されている^{1)~5)}。

多くの研究では関節構造をもった人体モデルを導入し、カメラ映像中の特徴とのフィッティングによる人体モデルのトラッキングを行うというアプローチを用

いている。人体の各部位を表すモデルとしては円柱や楕円体などの3Dプリミティブが用いられており、カメラ映像中の特徴としてはシルエット、オプティカルフロー、輪郭線などが用いられる。関節パラメータの推定はこれらのモデルと画像特徴との誤差の最小化問題として定式化されている。さらに、Mikiらはこのフレームワークに円柱、球、楕円体から構成される人体モデルと、複数のカメラ映像から得られた3次元形状（ボリュームデータ）を取り入れた⁶⁾。関節の初期パラメータは四肢の伸ばされた状態から推定され、トラッキングは拡張カルマンフィルタによって行われる。一方 Caillette らは色付きのBlob群からなる人体モデルと、Voxel Coloringによって得られる色付きのボリュームデータを使用した⁷⁾。Blob群とボリュームデータのマッチングはEMアルゴリズムによって行われる。Aguiar らはさらに一般的な手法として、人体関節構造の知識を用いない手法を提案した⁸⁾。この手法では多数の楕円体がボリュームデータにフィッティングされ、時系列間での楕円体のマッチングを行う

[†] 大阪大学大学院 情報科学研究科
Graduate School of Information Science and Technology, Osaka University
^{††} 大阪大学 サイバーメディアセンター
Cybermedia Center, Osaka University

ことで関節構造の推定が行われる。Kehl らの手法⁹⁾は人体モデルとして関節に運動して変形するメッシュモデルを用い、ボリュームデータの表面と Stochastic Meta Descent Optimization によってフィッティングされる。

これらの 3D 関節モデルとのフィッティングを行う トップダウンのアプローチでは、関節パラメータの収束計算のための計算コストが高い、初期パラメータの推定が手動、トラッキング失敗からの復帰が困難といった問題が起こる。フィッティングの際には多数の点群で表されるボリュームデータとモデル形状の類似度を繰り返し計算する必要がある。またトラッキングを行うには、あるフレームのパラメータ推定はその前の推定結果に依存する形となるため、あるフレームにおけるエラーによって以降のフレームすべての推定にエラーが生じる。また、初期パラメータを与える必要があるため、これらの手法においては関節の初期パラメータは撮影された映像に応じて手動で与えられるか、逆に測定対象となる人物が初期フレームにおいて特定の姿勢をとることで適切な初期パラメータを与えている。

これらの問題を回避するために、ボトムアップな手法も提案されている。これらの手法ではボリュームデータを直接解析し、スケルトンと呼ばれる 3 次元形状の中心を通る軸から関節位置が推定される。このボトムアップな手法は各フレームについて独立に処理が行われるため、トップダウンなアプローチで生じる問題が回避される。

Chu らの手法¹⁰⁾ではボリュームデータは Isomap¹¹⁾と呼ばれる手法を用いて姿勢に対して形状が不変な 3 次元部分空間に投影され、その空間内でスケルトンが抽出される。関節の位置は元の 3 次元空間においてスケルトンの曲率を解析することで推定される。Isomap によって投影された空間では四肢と頭はそれぞれ放射状に伸ばされるため、各部位のセグメンテーション及びスケルトンの抽出は容易に行われる。しかしこの投影には大きな計算コストがかかることと、ループ構造を持つボリュームデータに関して処理が行えないという問題があった。ループなどを含むトポロジは例えば手を腰に当てる、両手を握るなどの姿勢において頻繁に現れる。よって、よりロバストな推定のためにこれらのトポロジの変化について考慮する必要がある。また、得られたスケルトンが体のどの部位に対応するかといった問題については考慮されていなかった。

Brice らはボリュームデータを床面に対して水平方向に“スライス”することでスケルトンを抽出する手法を提案した¹²⁾。姿勢の推定方法はこの水平方向に依存した非常にヒューリスティックなルールによって行われるため、直立した姿勢でなければ対応できない。

Sundaresan らは Laplacian Eigenspace をボリュームデータのセグメンテーションに用いた。ボリューム

はセグメント毎にスプライン曲線にフィッティングされ、2 種類のトポロジを持つグラフ構造によってスプライン曲線に部位する部位が判定される。しかしこの手法では四肢と体の接触の無いものと腰に手の触れたものの 2 種類のトポロジについてのみしか対応がなされていなかった。

我々の提案する手法は Sundaresan らの手法に似たボトムアップのアプローチを用いる。ボリュームデータに対して齋藤らの 3 次元細線化処理¹³⁾を施すことでスケルトンを抽出し、スケルトンの時系で列の曲率を解析することで関節位置を推定する。スケルトンに対応する体の部位は事例ベースの判定法を用いる。すなわち様々な姿勢をとったときのスケルトンを属性付きグラフで表したものをデータベースとして保存しておき、入力されたスケルトンを同様にグラフ化したものとマッチングさせることでデータベース内で最も類似したグラフを選択する。このときにデータベース内のグラフのノードには対応する部位がラベリングされており、マッチングの結果から対応部位を判定することができる。本手法には従来手法と比較して以下のよう長所がある。

- ボトムアップなアプローチを用いることで、関節パラメータの初期化やトラッキングの失敗時の復帰に関する問題を回避できる。
- 収束アルゴリズムを用いないため、モデルベースのアプローチや他の空間に投影するアプローチに比べて計算コストが低い。
- 部位の判別には事例ベースの方法を用いてヒューリスティックなルールを排除したため、一般化が容易である。ノイズの含まれるデータや複数の人間、他の関節構造を持つデータに関しても本手法を適用することができる。

以降、2 章では提案手法の詳細について述べる。次に 3 章にて実験について述べ、最後に 4 章でまとめと今後の課題について述べる。

2. 提案手法

提案手法の流れを図 1 に示す。本手法はボリュームデータの取得 (2.1 節)、3 次元細線化 (2.2 節)、モデルグラフデータベースを用いた部位の判定 (2.3 節)、曲率からの関節位置の推定 (2.4) から構成される。まず、時系列のボリュームデータが 8 台のカメラからの映像をもとに Visual Hull アルゴリズムによって取得される。次に、ボリュームデータの中心を通る軸 (スケルトン) が 3 次元細線化処理によって抽出される。スケルトンからそのトポロジを表現する属性付きグラフが生成され、あらかじめ用意された様々なトポロジを持つモデルグラフのデータベース (MGDB) とマッチングされることで、入力スケルトンと各部位の対応関係が得られる。最後に関節位置の推定のために時系列

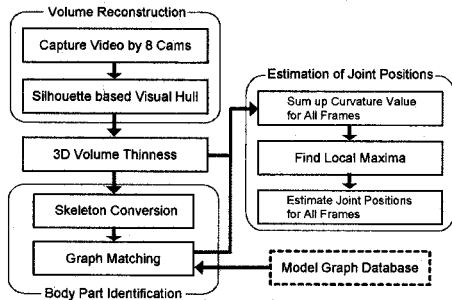


図1 提案手法の流れ。ボリウムデータの取得，3次元細線化，部位の判定，関節位置の推定の4つの処理から成る



図2 ボリウムデータとそのスケルトン

におけるスケルトンの部位ごとの曲率が評価される。以降に各処理の詳細を述べる。

2.1 ボリウムデータの取得

姿勢推定の対象となる人物のボリウムデータシーケンスを取得するために，提案手法では Visual Hull アルゴリズム¹⁴⁾を使用した。複数のキャリブレーション済みのカメラから対象人物が撮影され，人物の領域(シルエット)を得るためには背景差分と2値化処理が行われる。なお背景差分には床に映る影を除去するために CIELab 色空間で L の値に軽い重みをつけた距離を用いる。最後に得られたシルエットに対して Visual Hull アルゴリズムを適用し，各フレームにおける対象人物のボリウムデータを得る(図2)。

2.2 3次元細線化

取得されたボリウムデータに3次元細線化処理を行うことで，人体形状のスケルトンを得る。3次元の細線化を行う手法はレベルセット法やリープグラフなどがあるが，本提案手法では計算の簡易さから齋藤らのアルゴリズム¹³⁾を採用した。まず，それぞれのボクセルに対してその深さ(元の形状における表面までの最短距離)を計算する。次に深さの小さいものから，すなわち形状の表面のボクセルから順に消去可能性判定が行われ，消去可能なものが削除される。消去可能性は隣接する近傍26個のボクセルの配置を元に，元の形状のトポロジを保存するように判定され，スケルトンの縮退の防止のための端点(近傍に1つしかボクセルが残っていないもの)の保存とともに削除可能と判断されたボクセルが逐次削除される。形状の表面から順に削除が行われることによって，最終的に元のボリウムの中心を通る太さ1の3D曲線の集合がスケルトンとして抽出される(図2)。

以上の処理の結果，体の部位を表さない短い曲線

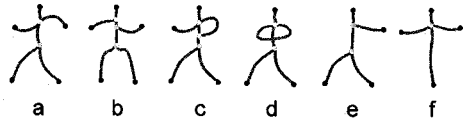


図3 人体スケルトンのトポロジ。(a): 2本の腕が1点で胴体に接続している。(b): 異なる点で2本の腕が胴体に接続。(c,d): ループ構造，(e,f): 完全に接触し，スケルトンが融合している。

(ヒゲ)が得られることがある。これはボリウムの表面にできる突起に起因し，本手法ではこれをまず曲線の長さに関する閾値処理で削除する。さらに閾値処理では除去できないヒゲには，我々は MGDB 内にこのようなヒゲを含むグラフを追加することで対処した。これについては後述する。

2.3 モデルグラフデータベースを用いた部位判定

この処理ではスケルトンに含まれる各曲線(枝)と人体の体の部位の対応が判定される。提案手法では人体のトポロジの変化を考慮して部位の判定を行う。本節ではまず人体の取りうる様々なトポロジについて述べ，次に入力されたスケルトンに対して部位の判定を行う方法を述べる。

2.3.1 人体形状のトポロジの種類

体の部位に接触の無い状況を考えると，理想的にはスケルトンは6つの枝(頭，胴体，右腕，左腕，右足，左足)から構成され(図3-a)，各枝と部位の対応は容易に判定される。しかし実際には，以下のような理由によってスケルトンのトポロジは異なるものとなる。

- (1) 両腕の枝が胴体の枝に一点で交わらない(図3-b)。
- (2) 複数の部位が接触する(図3-c,d)。
- (3) 複数の部位が全体で接触し，融合する(図3-e,f)。

これらの場合に関節位置を推定するには，2つに分かれた同じ部位を表す枝の結合，及び結合した2つの異なる部位の枝の分離が必要となる。我々の手法では，これら接触によって部位と1対1で対応しない枝を考慮した部位の判定を行う。ここで，様々なトポロジを持つグラフに対してヒューリスティックに対応部位を判定するルールを構築することは困難であるため，統一的なルールで部位の判定を行うために我々は次に述べる事例ベースの手法を採用した。

2.3.2 部位判定アルゴリズム

部位判定にはグラフマッチングを用いた事例ベースの判定法を用いる(図4)。あらかじめ，人体の異なるトポロジを持ついくつかのスケルトンを用意し，それらを元に属性付きグラフ(サンプルグラフ)を作る。サンプルグラフの各ノードには元のスケルトンが対応している人体の部位を手動で与え，これを MGDB に保存する。入力されたスケルトンも同様に属性付きグラフ(入力グラフ)に変換され，サンプルグラフとマッチングする。このマッチングによって最も類似す

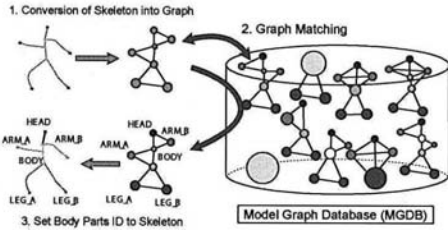


図4 MGDBを用いた部位判定. スケルトンはグラフに変換される. このグラフはMGDB内のグラフをマッチングされ, マッチング結果から対応する部位が判定される.

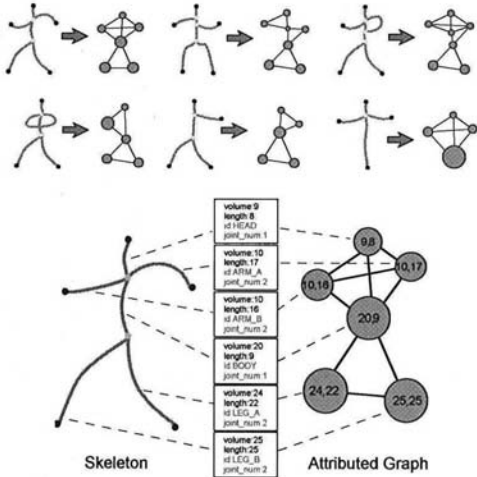


図5 スケルトンのグラフ表現. 各枝はグラフのノードで, 枝同士の接続関係はリンクで表現される. それぞれのノードは長さ, 体積の2つの属性値を持つ, また, MGDB内のグラフのノードには対応する部位のID及び関節の個数が保存される.

るサンプルグラフと入力グラフのノード間の対応が得られ, これを元に入力グラフの部位が判定される.

2.3.2.1 スケルトンのグラフ表現

スケルトンはいくつかの曲線(枝)とそれらを結ぶ接続点から構成される. 提案手法では, 以下のようにしてスケルトンをグラフ表現に変換し, 人体のトポロジ及び各部の特徴を表現する. 各枝はノードで表され, それをつなぐ接続点がリンクで表される(図5). また, ノードには長さ及体積の2つの属性が与えられる. 長さ属性は枝の長さをスケルトン全体の枝の長さの合計で正規化したものを, 体積属性は元のボリュームデータにおける枝周囲の体積を表す. 体積属性は枝を構成する各点について, 2.2節で述べた深さ属性から計算され, 長さと同様全体の体積で正規化される. これらの正規化は1つのMGDBを異なる体型の人物にも適用可能にするために行われる.

2.3.2.2 モデルグラフデータベース

サンプルグラフは実際の人体ボリュームのスケルト

表1 部位を表すID及びその部位の中に見れる関節個数

FULL(4)	HEAD(2)
BODY(1)	LOWBODY(1)
ARM_A(3)	ARM_B(3)
LOWARMA(1)	LOWARMB(1)
LEG_A(3)	LEG_B(3)
HEAD_BODY_ARMS(2)	HEAD_UPARM(3)
HEAD_UPARMS_BODY(2)	HEAD_BODY_UPARMS(2)
BODY_ARM(1)	BODY_UPARM(1)
BODY_LEGS(3)	ARM_LOOP(6)
NOISE(1)	

ンから得られる. これらのグラフの各ノードには部位のID及び部位に含まれる関節の個数(表1)が手動で与えられ, MGDBに保存される. 部位の接触などによってある部位が複数の枝に分割されている場合は, それぞれの枝に同じIDを割り当てる.

このスケルトンのグラフ表現はトポロジの変化ごとに異なるが, トポロジの変化しない範囲での姿勢(関節の角度)に対しては不変である. 従って, 膨大な人体の姿勢のパターンに対してごく少数のグラフでスケルトンのトポロジを表現できる. 論理的にはサンプルの個数は人体形状の取りうるトポロジの個数と同数で十分である. しかし, 我々はマッチングの精度を上げるために1つのトポロジに対していくつかのサンプルグラフをMGDB内に用意し, ほとんど起こりえないトポロジに関しては除いた.

2.3.2.3 グラフマッチング

グラフマッチングは入力グラフとMGDB内のグラフに対して行われる. 本手法ではEdit-Distanceベースのグラフマッチングアルゴリズム¹⁵⁾を用いた. Edit-Distanceは2つのグラフの間の非類似度を表し, 一方のグラフを他方グラフに変換するのに必要な編集操作にかかるコストの和の最小値として定義される. 本手法ではノードとエッジの削除及び属性変更の操作を編集操作として定義し, 以下のようなコスト関数を定めた.

$$\text{cost}(\text{del_node}(n)) = 1.0 \cdot n.\text{vol} \quad (1)$$

$$\text{cost}(\text{del_edge}(e)) = k_0 \quad (2)$$

$$\text{cost}(\text{sub_node}(n1, n2)) = k_1 |n1.\text{vol} - n2.\text{vol}| + k_2 |n1.\text{len} - n2.\text{len}| \quad (3)$$

$$\text{cost}(\text{sub_edge}(e1, e2)) = 0 \quad (4)$$

$n.\text{vol}$ と $n.\text{len}$ はノード n の体積属性, 長さ属性を表す. $k_0 \sim k_2$ は定数である.

ノードの体積属性が大きいものほどボリュームデータ内で大きな構造をもつものから作られたノードであると考えられるため, ノードの削除には体積に比例するコストを与える. また, ノードの属性変更にかかるコストは, それぞれのノードに対応する部位同士の非類似性を表す. ここでは体積と長さ両方の差を考慮し, その線形和でコストを表現する. 係数となる k_1 と k_2 は3.1節で述べるように実験的に決定した. ここで定

数 k_1 と k_2 に関しては k_0 や 1.0 (ノードの削除の重み係数) より十分に小さい値とし、トポロジの等しくかつ最も類似するグラフが選択されるようにした。エッジは属性を持たせていないので、属性変更のコストを 0、削除のコストを k_0 と定数にした。エッジの削除はトポロジの変化を引き起こすため、 k_0 には十分な大きなコストを与えた。従って、グラフ間のマッチングコストは主にノード間の属性変更のコストとなる。

グラフマッチング問題は NP 完全として知られているが、高速に処理を行うアルゴリズムがいくつか提案されている。本手法では Messmer らのアルゴリズム¹⁵⁾ を用いた。このアルゴリズムは複数のモデルグラフに対して、共通に含まれる部分グラフを抽出して包含関係をもとに木構造を構成し、この木構造を辿りながらマッチングを行うことでマッチングを高速化している。よって本提案手法のように複数のモデルグラフとのマッチングを行うのに適している。

グラフマッチングの結果、入力グラフと MGDB 内で最も類似したサンプルグラフとのノード間の対応が得られる。この対応関係をもとにサンプルグラフに与えられた部位情報が入力グラフに得られ、元のスケルトンの枝に対応する部位が判定される。

2.4 関節位置の推定

対応する部位の判明したスケルトンに対して、スケルトンの各枝から関節位置を推定する。

まずは接触によって分離された同じ部位を表す枝の接続及び接続された異なる部位を表す枝の分離を行い、腕や胴体などの部位ごとに 1 本ずつの枝が対応するよう変換する。例えば、胴体に対して両腕が別の点に接続されている場合は胴が 2 本の枝に分割されるが、このとき 2 つの枝を結合する (図 6-左)。接触によってループ構造が生じた場合は先端に当たる部分を切り離す (6-右)。ある部位の先端が他の部位に接触する接触関係については、この結合・分離によって枝と部位が 1 対 1 に対応付けられる。また、この際ヒゲとして認識された枝は以降の処理で無視する。

ここで、体の側面に腕が密着するような姿勢については、枝が完全に他の部位の枝と融合してしまうためスケルトンを用いるアプローチでは解決できない。よって本手法ではそのような姿勢については部位の判定のみを行い、関節位置推定に関しては枝と部位との対応が 1 対 1 になるもののみを扱う。

関節位置の推定は、それぞれの枝の時系列での曲率を解析することで行われる。具体的には以下のようなステップで行われる。

- (1) 枝の長さがすべて同じになるよう正規化し、枝の端からの各相対位置における曲率を計算する。
- (2) 同じ部位を表す枝に関して、時系列で同位置に当たる部分の曲率を積算する (図 10)。
- (3) 積算された曲率について、極大となる位置をその部位における関節の相対的な位置として、ID

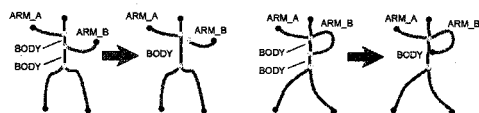


図 6 スケルトンの結合・分離。左：複数に分離された枝の結合。右：ループ構造の分離。

ごとに割り当てた関節の個数分だけ抽出する。

- (4) 各フレームのスケルトンについて、先ほどの相対的な関節位置に対応する関節位置を求める。

全フレームの曲率を足し合わせることで、シーケンス中で曲がっていたすべての関節の情報が得られる。この曲率の極大値をもとに推定した関節位置を用いて、全フレームで一貫した関節の構造を得ることができる。

3. 実験

本提案手法を用いて実験を行った。

実験環境として内部で運動する人物を中央を周囲の天井から見下ろす形に配置された 8 台のカメラが撮影するブルーバックのスタジオを用いた。カメラはすべてキャリブレーション済みで、同期信号によって最大 33fps で解像度 XGA の映像がそれぞれに接続されたコンピュータに送られる。3 人の体型の異なる人物 (被験者 1, 2, 3), 及び 2 人の人物 (被験者 1, 4) が同時にスタジオ内で動作している様子を撮影した。なお、撮影以降の処理はすべて一台の PC を用いてオフラインで行い、ボリュームデータは一辺 2cm の立方体として復元した。

3.1 MGDB とマッチングの係数

MGDB には被験者 1 を撮影して得られたグラフから 31 種類のトポロジを含む合計 41 個のサンプルグラフを選び、それぞれのノードには手動で対応する部位 ID を与えた。

グラフマッチングのマッチングコスト (式 3) に関する係数 k_1 と k_2 を決定するために、用意した MGDB と被験者 1 のグラフ 352 フレーム分で k_1 と k_2 を変化させながらグラフマッチングを行った。被験者 1 のグラフにもサンプルグラフ同様に部位の情報を手動で与え、部位の判定が正しく行われたフレーム数の割合 (正解率) を評価した。結果、 $k_1 : k_2 = 1.0 : 0.83$ のときに最も良い正解率 (92.4%) が得られた。

3.2 実験結果

スケルトンに対する部位の判定結果を図 7, 図 8 に示す。

図 7 は被験者 1 のボリュームデータから作成されたスケルトンの部位判定の結果である。上段はカメラから得られた画像の 1 つであり、中段には部位判定の結果が枝ごとに色づけされたスケルトンで表されている。下段は入力グラフにマッチしたサンプルグラフである。結果より、我々の手法で手と胴体の接触や両足

の融合などによるスケルトンのトポロジの変化に対応して部位の判定が行えた。また、ヒゲによるトポロジの変化に対しても MGDB 内に対応するサンプルグラフが用意されていることで対応できることが確認された（図 7 下段右）。

図 8 は被験者 2, 3 のスケルトンの部位判定の結果である。この実験では被験者 1 の映像から作成された MGDB 及びグラフマッチングの係数を用いた。成功率に関しては被験者 2 で 94.6%, 被験者 3 については 93.3% となった。長さや体積を正規化することにより、異なる人物に対しても部位判定が正しく行われることがわかった。

図 9 はシーン中に 2 人の人物がいる場合の部位判定結果である。MGDB に 2 人が接触したときのサンプルグラフも追加すると、このように 2 人の人物が接触している場合でも部位の判定が行える。

図 10 は被験者 1 の 291 フレームのスケルトンにおける曲率を積算したものである。この積算した曲率の極大点より、腕、足における各関節の相対的な位置が決定される。例えば腕の ID を持つ枝の極大点は端から順に手首、肘、肩の関節を、足の ID を持つ枝の極大点は足首、膝、股関節を表すものとみなされる。図 11 に各フレームのスケルトンに対してこの相対的な位置から関節位置を求めたものを色付き折線で示す。各関節の位置が正しく推定できていることがわかる。

処理時間に関しては、ボリューム復元に 5.75s, 細線化処理に 480ms グラフへの変換とグラフマッチングに 156ms, 関節位置の推定に 40ms を要した。

3.3 考 察

実験結果より、提案手法は人体形状のトポロジが変化するような姿勢・動作に対しても正しく部位の判定が行えることがわかった。用意した MGDB は別の人物に対しても 90% 以上の成功率で部位判定が行える。時系列を考慮することでよりよい結果が得られるため、この成功率は十分な値と考えられる。また、提案手法はボリュームデータのノイズにも強い。細線化処理で得られたスケルトンには閾値処理では取り除けない長さのヒゲが含まれる場合があるが、実験では、部位判定のプロセスにおいて MGDB 内のヒゲを含むサンプルがマッチすることでヒゲを認識し、以降の処理で無視することができることが示された。

一方で提案手法には以下のような問題点も見られた。まずは部位判定における頭と腕の混同が挙げられる。頭と腕は多くの場合、胴体に対してトポロジ的に同等な接続関係にある（図 3-a）。さらに上腕を表す枝が胴体と融合している場合、腕を表す枝と頭を表す枝の長さ・体積属性は近い値となる。この問題を解決するため、グラフの属性に長さや体積以外にボリュームデータの形状に関する特徴を加えて識別力を高めることを考えている。

我々のグラフベースのアプローチは非常に一般的で、

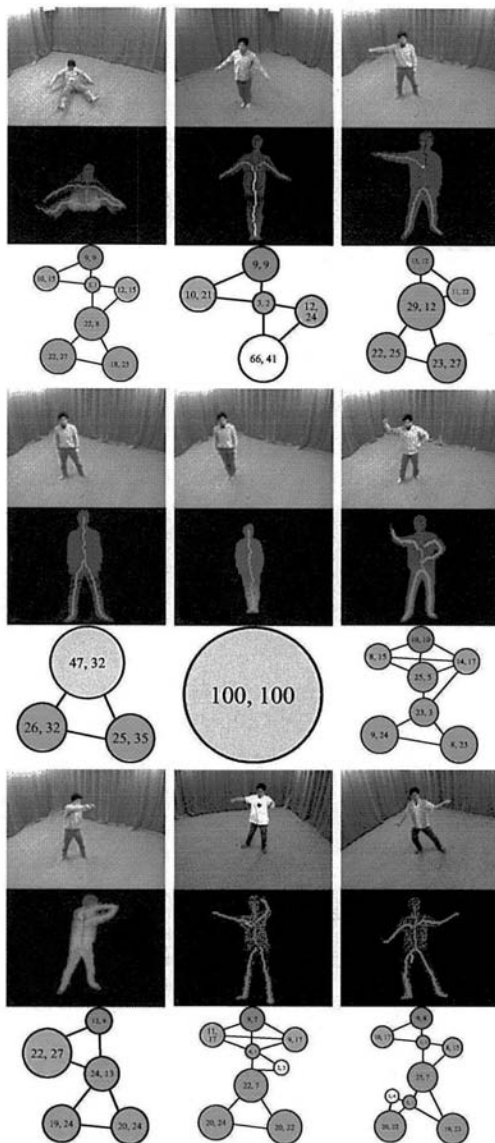


図 7 被験者 1 の部位判定の結果。上段：カメラ画像。中段：部位別に色分けされたスケルトン。下段：入力グラフにマッチしたサンプルグラフ。白色のノードはヒゲを表す。

簡単に拡張できる。実験で示したように、提案手法は複数の人物が接触しているシーンに対しても、サンプルグラフを用意し MGDB に加えることで 1 人のときと同様に処理が行える。このようなシーンの扱いは従来のトップダウン型のアプローチでは探索すべきパラメータの組み合わせが非常に多くなるため困難であった。

処理時間に関しては、1 フレームあたり約 6.4s で

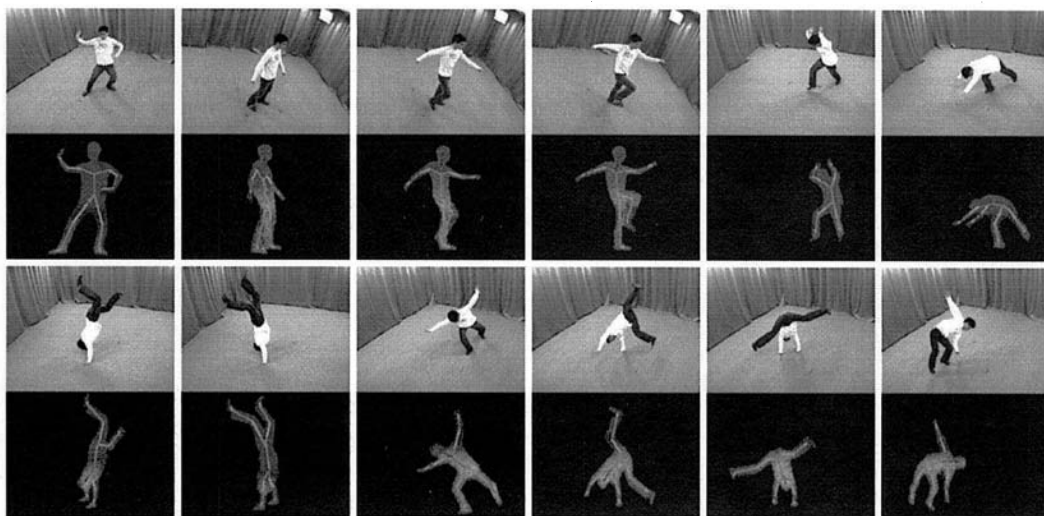


図 11 関節位置の推定結果. 関節間を色の異なる折線で示した.

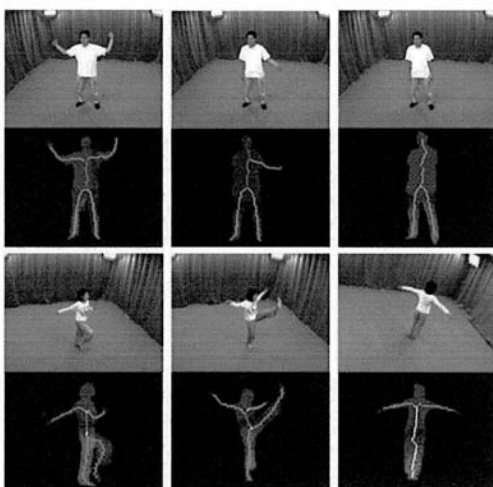


図 8 被験者 2, 3 の部位判定の結果. 上段: カメラ画像. 下段: 部位別に色分けされたスケルトン.

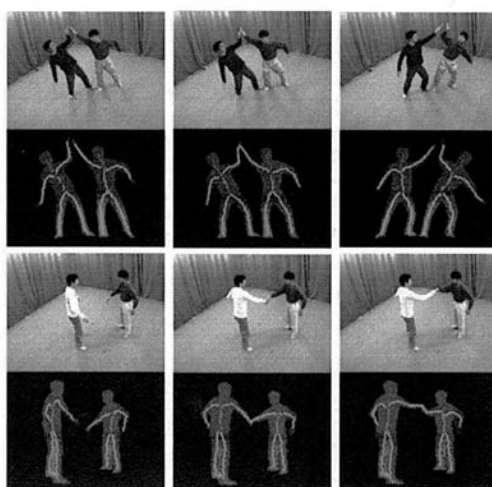


図 9 2 人の被験者での部位判定結果. 上段: カメラ画像. 下段: 部位別に色分けされたスケルトン.

あった. そのうち細線化と部位の判定, 関節位置推定にかかった時間は 0.7s 程度で, 処理時間の 90% (5.8s) はボリュームデータの生成にかかる時間であった. 残りの処理もモデルフィッティングを行うトップダウンな手法や別の空間にボリュームデータを投影してスケルトンを抽出する方法と比較して高速に処理が行えると考えられる. また, ボリュームデータの復元に関してはより高速な既存手法を用いた速度の改善を予定している.

4. 結論と今後の課題

本稿では, ボリュームデータを用いたマーカの不要なモーションキャプチャ手法を提案した. 提案手法はボリュームデータをボトムアップに解析するため, 姿勢に関する初期パラメータが不要であり, 3次元細線化処理でスケルトンを抽出するため高速に処理が行える.

我々は姿勢によってトポロジの変化する人体スケルトンと人体部位の対応付け問題に対して, 事例ベー

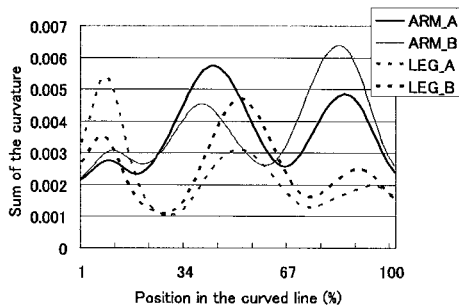


図 10 291 フレームのスケルトンの各位置における曲率の積算値

スの手法を用いてこれを解決した。様々な人体のトポロジをグラフとしてデータベースに保存しておき、スケルトンから得られた入力グラフとマッチングすることによって部位の判定が行われる。また、実験より時系列でのスケルトンの曲率の解析によって関節構造位置の推定が行われ、正しく関節位置が推定されることが示された。関節位置推定はそれぞれのフレームについて独立に行われるため、あるフレームで起こったエラーが他のフレームの推定に影響を及ぼすといったトラッキング手法における問題が回避された。さらに、提案手法は複数人物に対応できるように容易に拡張可能である。実験結果により、2人の場合のトポロジに関しても1人の場合と同様にサンプルグラフを用意するだけで部位の判定が可能であった。

しかし、提案手法で用いたグラフ構造には位置や向きなどの情報が含まれておらず、左右の手足など同じ形状で対称な位置にある部位を区別することが困難である。グラフ構造にこれらの情報を加えることで解決できると考えられる。また、推定された関節位置の精度評価と従来手法との処理時間の比較も今後の課題である。

参 考 文 献

- 1) D.M. Gavrilu. The visual analysis of human movement: A survey. *Computer Vision and Image Understanding*, Vol.73, No.1, pp. 82–98, 1999.
- 2) J.K. Aggarwal and Q. Cai. Human motion analysis: A review. *Computer Vision and Image Understanding*, Vol.72, No.3, pp. 428–440, 1999.
- 3) T.B. Moeslund and E. Granum. A survey of computer vision-based human motion capture. *Computer Vision and Image Understanding*, Vol.81, No.3, pp. 231–268, 2001.
- 4) L. Wang, W. Hu, and T. Tan. Recent developments in human motion analysis. *Pattern*

- Recognition*, Vol.36, No.3, pp. 585–601, 2003.
- 5) J.J. Wang and S.Singh. Video analysis of human dynamics—a survey. *Real-Time Imaging*, Vol.9, No.5, pp. 321–346, 2003.
- 6) I.Mikić, M.Trivedi, E.Hunter, and P.Cosman. Human body model acquisition and tracking using voxel data. *International Journal of Computer Vision*, Vol.53, No.3, pp. 199–223, 2003.
- 7) F.Caillette and T.Howard. Real-Time Markerless Human Body Tracking with Multi-View 3-D Voxel Reconstruction. In *Proc. of British Machine Vision Conference*, Vol.2, pp. 597–606, 2004.
- 8) E.Aguiar, C.Theobalt, M.Magnor, H.Theisel, and H.P. Seidel. M3: Marker-free model reconstruction and motion tracking from 3d voxel data. In *Pacific Graphics*, pp. 101–110, 2004.
- 9) M.B.R. Kehl and L.V. Gool. Full body tracking from multiple views using stochastic sampling. In *Proc. of CVPR'05*, Vol.2, pp. 129–136, 2005.
- 10) C.-W. Chu, O.C. Jenkins, and M.J. Mataric. Markerless kinematic model and motion capture from volume sequences. In *Proc. of CVPR'03*, Vol.2, pp. 475–482, 2003.
- 11) J.B. Tenenbaum, V.Silva, and J.C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, Vol. 290, No. 5500, pp. 2319–2323, 2000.
- 12) M.Brice and B.Sáida G.Erwan. Human model and pose reconstruction from multi-views. In *Proc. of ICM'05*, pp. 788–795, 2005.
- 13) T.Saito, K.Mori, and J.Toriwaki. A sequential thinning algorithm for three dimensional digital pictures using the euclidean distance transformation and its properties. *The transactions of the Institute of Electronics, Information and Communication Engineers. D-II*, Vol. J79-D-II, No.10, pp. 1675–1685, 1996.
- 14) A. Laurentini. The visual hull concept for silhouette-based image understanding. *IEEE Trans. on PAMI*, Vol.16, No.2, pp. 150–162, 1994.
- 15) B.T. Messmer and H.Bunke. A new algorithm for error-tolerant subgraph isomorphism detection. *IEEE Trans. on PAMI*, Vol.20, pp. 493–504, 1998.