

## 骨格線モデルと円筒多関節モデルを併用した人体の姿勢推定

宮本 新† 鷺見和彦† 松山隆司†

**概要** カメラによる撮影から取得された人物の三次元形状に基いた姿勢推定は、動作を入力とした新たなインターフェース等への応用が期待されている。一般的には問題を容易にするために人体の構造に基いた人体モデルを仮定し、入力となる三次元形状に対して最適なフィッティングとなる人体モデルの姿勢を探索することで推定する。視体積交差法等の手法で取得されるボクセルデータには、物体の構造に関する情報が含まれないため、これを入力として人体モデルとのフィッティングをする場合、局所解に陥る可能性がある。そこで本研究では、ボクセルデータから取得される骨格線モデルと、人体の構造に基いてトップダウンに構成される円筒多関節モデルとを併用した姿勢推定手法を提案する。骨格線モデルから一つの体節に相当する部分を探索し、それを元に円筒の初期軸方向を決定することで、フィッティング時の円筒多関節モデルの探索空間が小さくなる。また、円筒によって定義される骨格構造に基いて、骨格線モデルの該当部分を人体骨格に近い構造へと修正する。このような双方向の処理を各体節ごとに適用することで、誤りが少なく高速な姿勢推定が可能となる。

### Human Pose Estimation from 3D Object Skeleton using Articulated Cylindrical Human Model

ARATA MIYAMOTO, † KAZUHIKO SUMI † and TAKASHI MATSUYAMA †

**Abstract** Human motion capture has become a large research area in recent years due to its usefulness for a number of applied fields. Many pose estimation methods using multi-viewpoint images employ human models given a priori, and try to find the best pose which fits to the input data such as 3D volume of the object given by shape-from-silhouette method. However, 3D volume of the object is a set of voxels (3D cube) and does not have the information about body structure. On the other hand, 3D skeleton of the object extracted from 3D volume represents global structure of the object, but it is difficult to extract it robustly. In this paper, we propose a method which integrate articulated cylindrical human model and skeleton information of the object to realize more robust pose estimation. We first estimate the initial pose of the body parts based on the skeleton, and refine its pose so as to fit to the 3D volume. Then, we correct the skeleton based on the human model structure. We iterate this cross-validation process for each body parts. We show some experiments to demonstrate the effectiveness of our method.

#### 1. 序 論

##### 1.1 背 景

人体の動作解析および姿勢推定は、コンピュータビジョンに関する研究において活発に取り組まれている分野の一つである。その応用範囲は広く、人体動作を入力とした新たなインターフェースの開発、計算機上に構築された仮想環境で人間が作業を行う VR への利用など多岐に渡る。

撮影時にマーカ等の特徴点として利用する姿勢推定と比べて、純粋な撮影映像のみを用いる手法では、着衣への制約が少なく、被検者に拘束感を与えることがないため、より自然な状況で撮影を行うことが可能と

なる。

ところが、実環境での撮影から復元される三次元形状は、計算機内部においては単なる多数の点の集まりとして表現されており、そこには関節の位置や体節の長さのような、人体の構造に関する情報は含まれていない。したがって、マーカのような特徴点を含まない純粋な映像から姿勢を推定する場合、問題はより複雑なものとなる。

##### 1.2 人体モデルに基いた姿勢推定

姿勢推定の問題を容易にするために、一般的には人体モデルが利用される。人体モデルを用いたアプローチとは、人体の寸法や各関節の接続関係といった、身体の幾何的構造に関する情報を持つモデルを、前提知識として計算機内部に用意しておくという方法である。また、骨格の周りを取り囲む幾何学的な形状を定義することにより、人体のボリュームが表される。こ

† 京大大学院情報学研究科  
Graduate School of Informatics, Kyoto University

れには、人体の各部位を円筒により表現した単純なモデル<sup>1)</sup>から、実際の人の皮膚形状までを考慮に入れた精密なモデル<sup>2)</sup>まで、その研究の目的に応じて種々の表現方法が用いられる。

モデルを用いた姿勢推定においては、モデルと入力画像との比較により、現在の姿勢がどの程度入力となる三次元形状に対して相応しいかを評価する。この場合、姿勢推定は三次元形状と比較した際に評価値が最大となる人体モデルの姿勢を探索する処理へと帰着される。

ところが人体モデルは関節の数に応じた多数の自由度を持っており、モデルが取り得るあらゆる姿勢について評価値を求めることは、処理に要する時間の観点から考えると非現実的である。そこで何らかの方法により、モデルの持つ状態空間における探索範囲を限定する必要がある。

これを解決するために一般的に用いられるのは、ある初期姿勢を開始点として、現在の姿勢よりも良い評価が得られる姿勢へと、反復的に更新するという手法である。このような処理を一定回数、もしくはある一定の精度の解が得られるまで繰り返し、最終的に得られたモデルの姿勢を推定結果として出力する。

またモデルの持つ自由度を様々な制約条件によって削減し、あらゆる姿勢における評価値の計算を実時間で可能とする手法<sup>3)</sup>もある。この場合、上記のような予測処理は必要なくなるが、撮影環境やユーザの動作に対して様々な前提条件が必要となる。

人体モデルの姿勢を反復的に更新する手法では、処理開始前におけるモデルの初期姿勢をどのように決めるかということが問題となる。これに対して、既存の研究ではモデルの初期姿勢を以下のように単純な方法により決定している。

- 初期フレーム
  - 既知の姿勢を動作の開始点とする<sup>4)</sup>。
  - 利用者自らが初期姿勢を計測する<sup>5)</sup>。
- 2フレーム目以降
  - 前フレームでの推定により求められた姿勢を次フレームでの初期姿勢とする。

上述の方法では、初期フレームにおける初期姿勢は利用者が決定しなければならず、利便性の面から考えるとあまり好ましくない。また2フレーム目以降の処理では、各フレームごとに人体モデルの姿勢が逐次更新される。そのためあるフレームにおいて局所解に陥ってしまった場合、以降のフレームにおいて正確な姿勢推定が不可能となるという問題がある。

### 1.3 研究の目的

上述のように、人体モデルの初期姿勢の決定は、推定姿勢に対する信頼度にも影響を及ぼす重要な要素となる。実環境での撮影から三次元形状を取得する代表的な方法としては、視体積交差法<sup>7)</sup>が挙げられる。視体積交差法から得られるボクセルデータは、単に三次

元空間中での座標が定義された点の集まりにより、三次元物体の占める領域を定義するデータであるため、人体の関節位置のような特徴点を取得することが困難である。そのため、ボクセルデータのみを入力とした姿勢推定アルゴリズムでは、前述のような単純な初期姿勢の決定方法に頼らざるを得ない。

そこで本研究では、三次元形状をグラフ構造により記述した骨格線モデルを利用して、人体モデルの初期姿勢を決定する手法を提案する。これにより、時系列データ中の各フレームに対して個別に姿勢推定をすることが可能となるため、局所解に陥るといった問題が解決できる。また、初期フレームでの初期姿勢の決定も可能となる。

以下、第2章では、骨格線モデルと本研究で用いる人体モデルの特徴について述べる。第3章では、本研究における提案手法の原理について述べる。また、第4章では、評価実験を行い、その結果に対する考察をする。最後に第5章では本論に対する結論を述べる。

## 2. 物体形状の記述形式

本手法では骨格線モデルを利用することでおおまかな姿勢を予測し、人体モデルの自由度に関する探索範囲を削減するために利用する。ここでは、骨格線モデルの生成の原理とその性質について述べる。また、本手法で用いる人体モデルである円筒多関節モデルの設計についても説明する。

### 2.1 骨格線モデルによる三次元形状の記述

物体の形状に対する記述形式には、その抽象化の度合いに応じたいくつかの種類が存在する。このうち骨格線モデルは、図1に示すように、物体の形状をいくつかの主要なノードと、それらを接続するエッジによるグラフ構造で表現したものである。

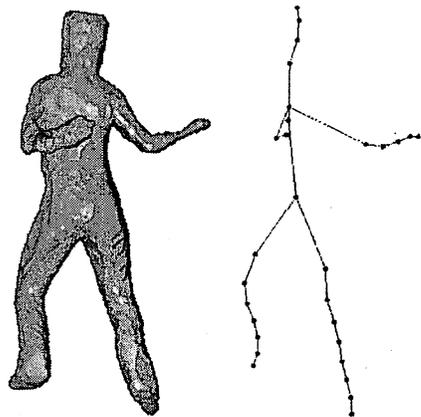


図1 三次元形状の骨格線モデルによる表現

### 2.1.1 リーブグラフによる物体形状の骨格線記述

本研究では、対象形状の骨格線モデルによる記述方法としてリーフグラフ<sup>8)</sup>と呼ばれるグラフ構造を用いる。リーフグラフは三次元物体の表面  $S$  上で定義された実連続関数  $\mu(v)$  ( $v$  は  $S$  上の任意の点) の値に基いて、 $S$  を部分領域に分割し、各部分領域の代表点同士の接続関係により対象の形状をグラフ構造で表したものである。 $\mu$  の定義としてはいくつかの種類が考えられるが、本研究では測地距離に基いた関数を用いる。表面  $S$  上の2点  $v, p$  に対する測地距離  $g(v, p)$  とは、 $v$  から  $p$  まで物体の表面を辿った場合の最短距離である。このとき物体表面  $S$  上の点  $v$  は、

$$\mu(v) = \int_{p \in S} g(v, p) dS \quad (1)$$

と定義される。すなわち、 $S$  上の点  $v$  から、 $S$  上のすべての点までの測地距離を加算したものが、点  $v$  での  $\mu(v)$  の値となる。図2は人体形状における  $\mu$  値の分布を示したものである。

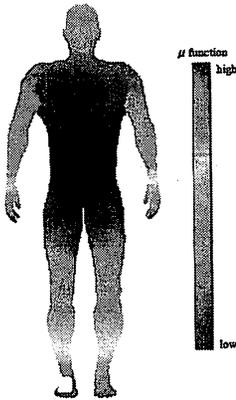


図2 人体形状における  $\mu$  値の分布  
文献9)より引用

### 2.1.2 骨格線モデルの特徴

三次元形状をボリュームで記述したボクセルデータに比べると、グラフ構造による記述である骨格線モデルは、より三次元物体のトポロジカルな特徴を把握し易いという利点を持つ。人間の姿勢はいくつかの関節の位置、あるいは骨格の位置により定義されており、骨格線モデルを利用することで、そのような姿勢を決定する特徴点を比較的容易に取得することが可能となる。

ただし、骨格線モデルは実環境での撮影からボトムアップに構成されるデータであり、人体の幾何的構造に基いていない。例えば、人物が腰に手を当てる、あるいは両足を密着させるなど、異なる体節間の接触を含むような姿勢を取った場合、三次元表面上に余分な

経路が生じるため、 $\mu$  値が変化する。そのため、同じ人物を対象とした骨格線モデルであっても、ループの発生、異なる体節の相当する枝同士との融合などにより、幾何的構造が絶えず変化する。

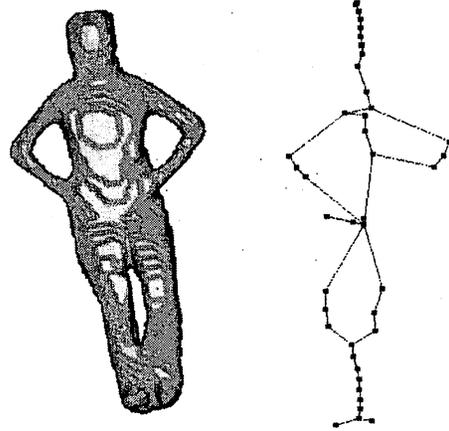


図3 ループや枝同士の融合が存在する骨格線モデル

### 2.2 円筒多関節モデルによる人体姿勢の表現

本研究では人体モデルとして円筒多関節モデルを用いる。このモデルでは図4のように11本の連結された骨格により人体の骨格構造を定義する。このとき各骨格を軸として、既知の長さを持つ円筒により三次元空間中での各体節のボリュームを定義する。各円筒間の従属関係は固定されており、この人体モデルにおいては各円筒の軸方向を指定することで姿勢が一意に決定される。

本研究で用いるモデルでは、円筒の軸周りの回転を考えない。したがって、各円筒は軸方向に関して2つの自由度を持つ。また、円筒の軸方向とは別に、入力となる三次元形状の座標系と人体モデルの座標系との位置の対応を取るためのベクトルを決定する必要がある。

#### 2.2.1 円筒の軸方向に対する評価関数

人体モデルを用いた姿勢推定手法では、入力となる人体形状データと人体モデルとをある評価関数により比較した際に、最大の評価値が得られるような姿勢へと人体モデルをフィッティングすることとなる。

本手法では、円筒多関節モデルとのフィッティングには視体積交差法により得られる三次元ボクセルデータを用いる。ボクセルデータは三次元空間上における点の集合により物体の占める領域を定義した記述である。そこで、円筒多関節モデルを構成するある円筒  $c$  について、その軸方向が  $v$  であるとき、この円筒内に含まれるボクセルの総数をその軸方向に対するフィッティングの評価値  $E$  とする。すなわち評価値  $E$  は軸方向  $v$  の関数として、

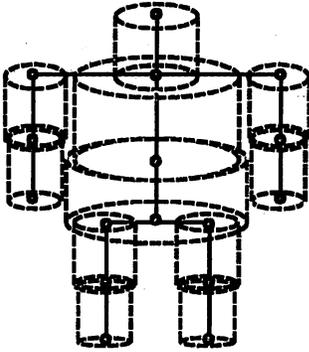


図4 円筒多関節モデルの構造

$$E = I(c(\mathbf{v})) \quad (2)$$

と定義される。そして、円筒  $c$  は評価値  $E$  が最大となるような軸方向、

$$\mathbf{v}_f = \max_{\mathbf{v}} I(c(\mathbf{v})) \quad (3)$$

へとフィッティングする。

### 3. 骨格線モデルと円筒多関節モデルを併用した人体姿勢推定の原理

骨格線モデルは実環境での撮影からボトムアップに構成されるデータ構造でありながら、物体形状のトポロジカルな構造を把握しやすいという特徴を持つため、人体の姿勢推定という骨格構造の空間中での配置を求める問題との相性が良い。円筒多関節モデルは人体の骨格構造に基いた前提知識であり、姿勢推定を円筒多関節モデルと入力三次元形状とのフィッティングという形式へと還元することで、問題を容易にすることが可能となる。

ただし、円筒多関節モデルは多くの自由度を持つため、モデルの取り得るあらゆる姿勢を全て調べることは計算機の計算コストの点で非現実的である。また、フィッティングの評価関数には多くの局所解が含まれるため、まちがった姿勢において評価値が最大となる可能性もある。そこで骨格線モデルを利用すれば、円筒の軸方向のある程度予測することで、探索空間の削減が可能となる。

一方、骨格線モデルには人体の骨格構造に関する知識は含まれない、すなわちどの枝が人体のどの体節に相当するかという情報を含まない。円筒多関節モデルは人体の各体節の寸法とその接続関係についての情報を含むため、これを利用して各枝と体節との対応を取ることができる。このように、骨格線モデルと円筒多関節モデルのそれぞれの長所を利用して、お互いの短所を補うことで、信頼性の高い姿勢推定アルゴリズムを目指す。

#### 3.1 本手法における姿勢推定アルゴリズムの構成

本手法では、円筒多関節モデルにおける円筒同士の従属関係に従って、ある関節を開始点としたツリー構造において、根ノードに近い側の円筒から順次フィッティングをして、その軸方向を決定する。一連の処理の流れは図5のようになる。

各円筒の処理については、まず円筒の長さを元に骨格線モデルを探索し、その円筒が表す体節に相当する部分骨格線を骨格線モデル上から見つける。複数の体節によって共有されるノードが検出された場合には、ノードの分割やエッジの再接続により骨格線モデルを修正する。この部分骨格線を元に対応する円筒の初期軸方向を決定し、フィッティングする。上記の処理を、すべての円筒に対して適用した後の円筒多関節モデルの姿勢を、最終的な姿勢推定結果として出力する。

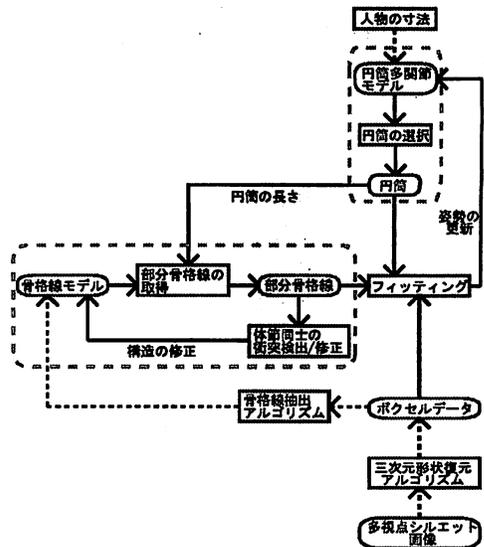


図5 本手法における姿勢推定処理の流れ

#### 3.2 処理開始関節の決定

円筒の従属関係に従って各円筒を順次フィッティングする場合、円筒多関節モデル中で処理の開始点となる関節が、骨格線モデル上のどのノードと対応するかを決定する必要がある。

式1で定義される関数  $\mu$  の値は、測地距離の総和に基いている。したがって、三次元形状の表面上において  $\mu$  値が最小となるのは、他の点までの平均的な測地距離が最も短い点であるため、三次元形状の中心付近に存在する。また、対象となる三次元物体が突起部分を持つ場合、そのような突起内において先端部分は、他の点までの平均測地距離が最大となる。したがって、図2のように人体において  $\mu$  値が最小となる点は、身体を中心である腰の辺りに存在する。また、

頭、手、足の先端部分において、 $\mu$  値はそれぞれ極大となる。

そこで骨格線モデルを、図 6 のように  $\mu$  値が最小となる分割領域に相当するノード  $n_r$  を根とし、 $\mu$  値が極大となる分割領域に相当するノード  $n_l$  を葉とするツリー構造とみなす。

このとき、円筒多関節モデルにおける腰関節  $p_{wast}$  について、 $p_{wast} = n_r$  とすることにより根ノード  $n_r$  と対応付けることが可能である。これにより、図 6 中央のように腰関節  $p_{wast}$  を開始点とした円筒の従属関係に基づいて、各円筒を人体の先端に向かう順序でフィッティングすることができる。

ただし、異なる体節同士が接触した姿勢では、三次元形状の表面に余分な経路が生じることにより  $\mu$  値の分布が変化するため、根ノード  $n_r$  の位置は僅かに変動する。

また同様に、人体の先端部分の関節、

$p_k (k \in \{head, r\_torso, l\_torso, r\_hand, l\_hand\})$  について、 $p_k = n_l$  と定めることで葉ノード  $n_l$  と対応付けることができる。この場合、図 6 右のようにそれぞれの先端関節  $p_k$  を開始点とし、各円筒を人体の中心、すなわち腰関節へと向かう順序で各円筒をフィッティングする方法が考えられる。

葉ノード  $n_l$  は姿勢が変化しても、人体に占める相対的な位置は変化しないが、複数存在するため、5つの人体の先端部分の関節  $p_k$  と一対一に対応付けることができない。また、異なる体節間の接触が存在する姿勢においては、骨格線モデル上にループが生じることが原因で、特定の先端関節に対応する葉ノードが存在しない場合もある。

以上のことから本手法では、まず腰関節  $p_{wast}$  を開始点として円筒多関節モデルの全身をフィッティングする。これにより推定された5つの先端部分の関節  $p_k$  と、最も距離が近い葉ノード  $n_l$  同士をそれぞれ対応付ける。その後、葉ノード  $n_l$  と対応を取ることができた先端関節  $p_k$  を一つ選び、その先端関節を開始点とした順序で各円筒を再度フィッティングする。

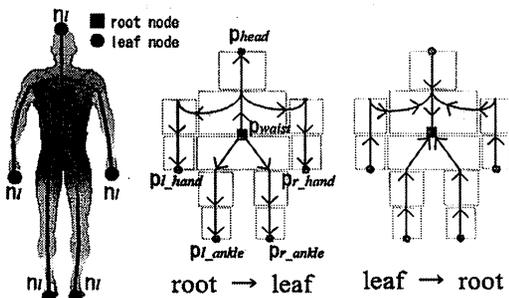


図 6 骨格線モデルの根ノードと葉ノードと人体関節との位置関係

### 3.3 部分骨格線の取得

ここで、図 7 に示すように、ある円筒によって表される体節に相当する部分領域を、骨格線モデル  $R$  上から部分骨格線として獲得する方法を述べる。

まず、円筒多関節モデル中からフィッティングの対象となる円筒  $c$  を選択する。円筒  $c$  において、円筒同士の従属関係で親に当たる円筒に近い側の関節を  $p_o$ 、遠い側の関節を  $p_e$  とする。円筒  $c$  と骨格線モデル上の部分骨格線との対応が取られている場合、関節  $p_o$  および  $p_e$  と骨格線モデル上のノード  $n_o$  と  $n_e$  の間でそれぞれ対応が取られている。このとき、骨格線モデル上の2つのノード  $n_o$  と  $n_e$  間の経路上に存在するノードとエッジの集合が、円筒  $c$  に対応する部分骨格線となる。

また関節  $p_o$  は、円筒間の従属関係において親に相当する円筒の関節の一つと一致する。したがって、根ノードに近い側の円筒から順次フィッティングをする場合、関節  $p_o$  の位置は既に決定されており、円筒  $c$  の持つ自由度はもう一方の関節  $p_e$  に関してのみとなる。そこで、固定点  $p_o$  に相当する骨格線モデル上のノード  $n_o$  を開始点として骨格線モデルを探索し、可動点  $p_e$  に相当するノード  $n_e$  を見つけ出す。

円筒  $c$  の長さ  $l$  は一定であるため、 $p_e$  の取り得る位置は、 $p_o$  を中心とする半径  $l$  の球面上  $S$  に限られる。したがって、ノード  $n_e$  が存在する位置としては、球面  $S$  と骨格線モデル  $R$  の交点  $r_i (i = 1, 2, \dots)$  のうちいずれか一つとなる。そこで、円筒  $c$  の軸方向を  $n_o$  から  $r_i$  へと向かうベクトルの向きとしたとき、円筒に含まれるボクセル数を最大とする交点  $r_i$  を、ノード  $n_e$  とする。すなわち、式 2 で表される評価関数  $E$  を用いると  $n_e$  は、

$$n_e = \max_{r_i} E(c(r_i - n_o)) \quad (4)$$

により決定される。

骨格線モデル上のこの位置に便宜的にノードを挿入し、これを  $n_e$  とする。ノード  $n_o$ 、 $n_e$  間の最短経路を  $N'$  としたとき、円筒  $c$  に対応する部分骨格線  $R_{part}$  は以下のように表される。

$$R_{part} = (N', E') \quad (5)$$

$$N' = \{n_o (= n_o), n_1, \dots, n_k (= n_e)\} \quad (6)$$

$$E' = \{(n_o, n_1), (n_1, n_2), \dots, (n_{k-1}, n_k)\} \quad (7)$$

#### 3.3.1 骨格線モデルにおける体節同士の衝突検出とその修正

骨格線モデル中にループや、枝同士の融合が生じている場合には、図 8 左のように一つのノードを複数の異なる体節同士が共有することとなる。部分骨格線はそれぞれ一つの体節に相当しており、ある体節に相当する部分骨格線に、別の体節に相当する部分骨格線上のノードが含まれないように、骨格線モデルを修正

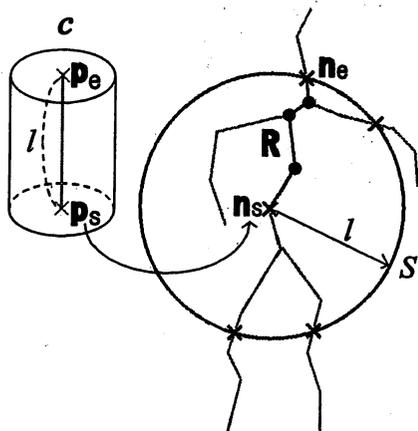


図7 リーブグラフ上からの部分骨格の取得

する必要がある。

そこで、各円筒が表わす体節に相当する部分骨格線を取得する過程において、すでにある体節に含まれることが判明しているノードは値  $T$ 、そうでないノードは値  $F$  となるフラグを各ノード中に保持することで、両者を区別する。開始時にはすべてのノードのフラグを値  $F$  で初期化し、処理の過程である体節に含まれていると判断されたノードは、フラグの値を  $T$  に変える。

異なる体節に共有されるノード  $n$  について、そのフラグ  $F_n$  は最初に  $F_n \leftarrow F$  で初期化されており、処理の過程である体節  $l_1$  に含まれることが判明した時点でフラグは  $F_n \leftarrow T$  へと書き換えられる。さらにその後、ノード  $n$  が別の体節  $l_2$  にも含まれることが判明した場合、すでに  $n$  は体節  $l_1$  に含まれており  $F_n = T$  となっているため、体節  $l_1$  と体節  $l_2$  がノード  $n$  の部分で衝突していることが検出できる。

ここで、ある体節に相当する部分骨格線の経路上に、フラグの値が  $T$  であるノードが含まれる場合、同じ座標値にそのノードの複製を作成する。このように、複数の体節に共有されるノードについて必要な数だけ複製を作成することで、図8のように異なる体節間の衝突を回避するように修正することが可能となる。

円筒  $c$  に相当する部分骨格線  $R_{part}$  が、式の様で得られているとしたとき、 $R_{part}$  に含まれる各ノード  $n_i (i = 1, 2, \dots, k)$  について、そのフラグ  $F_i$  の値に応じて以下のような処理を行う。

- step1.  $i \leftarrow 2$  とする。
- step2.  $i > k$  ならば終了する。
- step3.  $F_i = T$  ならば step4. へ、そうでなければ step6. へ進む。
- step4. ノード  $n_i$  と同じ座標にノード  $n'_i$  を作成する。
- step5.  $F_{i-1} = T$  ならば、エッジ  $(n'_{i-1}, n'_i)$  を作成

する。  $F_{i-1} = F$  ならば、エッジ  $(n_{i-1}, n_i)$  を削除して、エッジ  $(n_{i-1}, n'_i)$  を作成する。  $i \leftarrow i+1$  として step2. へ戻る。

- step6.  $F_{i-1} = T$  ならば、エッジ  $(n_{i-1}, n_i)$  を削除し、エッジ  $(n'_{i-1}, n_i)$  を作成する。  $i \leftarrow i+1$  として step2. へ戻る。

これにより、図9のように各体節に相当する部分骨格線が、骨格線モデル上で互いに独立した経路として存在し、他の部分骨格線と共有するノードを持つことがない部分骨格線  $R'_{part}$  が取得される。

また、新たに得られた部分骨格線  $R'_{part}$  に含まれる各ノードのフラグをすべて  $T$  とする。すなわち、

- $i = 1, 2, \dots, k$  について、  $F_i = F$  ならば  $F_i \leftarrow T$  とする。  $F_i = T$  ならば、ノード  $n_i$  のコピーである  $n'_i$  のフラグを  $F'_i \leftarrow T$  とする。

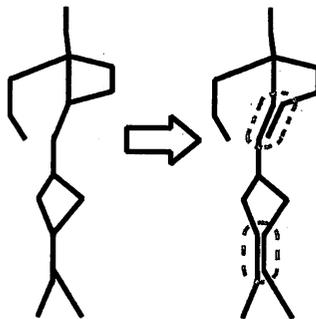


図8 異なる体節間で共有されるノードが存在する骨格線モデル

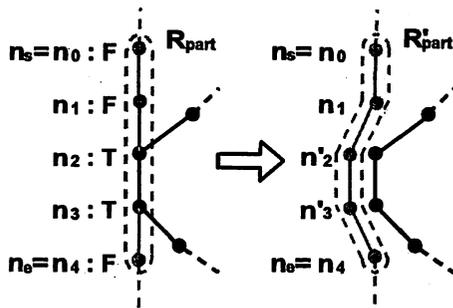


図9 フラグに基づいた部分骨格線の修正

### 3.3.2 部分骨格線に基づいた円筒軸方向の探索範囲の決定

円筒多関節モデル中の円筒  $c$  は、軸の一方の端点  $p_s$  を固定点としたとき、式2で表される評価関数  $E(c(v))$  が最大となるような軸方向へとフィッティン

グする。

三次元空間においては、円筒軸の方向  $\mathbf{v}$  は二つの角度変数  $\theta$ 、および  $\delta$  によって定義される。ところが評価関数  $E$  は各ボクセルごとに、それが円筒  $c$  に含まれるかを判定するため非常に計算コストが大きい。そのため、全軸方向 ( $0 < \theta < 2\pi, 0 < \delta < \pi$ ) について  $E$  を計算することは処理時間の観点から非現実的である。

そこで円筒  $c$  に相当する骨格線モデル上の部分骨格線  $R_{part}$  の両端ノード  $n_s, n_e$  について、 $n_s$  から  $n_e$  へ向かうベクトル

$$\mathbf{v}_t = \mathbf{n}_e - \mathbf{n}_s = (\theta_t, \delta_t) \quad (8)$$

を円筒軸の初期方向とし、軸方向  $\mathbf{v}$  に関して以下の制約を設ける。

$$\theta_t - \Delta\varepsilon < \theta < \theta_t + \Delta\varepsilon \quad (9)$$

$$\delta_t - \Delta\varepsilon < \delta < \delta_t + \Delta\varepsilon \quad (10)$$

この範囲内で  $E(c(\mathbf{v}))$  が最大となる軸方向へと円筒  $c$  をフィッティングする。

#### 4. 評価実験

第3章で述べた原理に基いて、三次元の人体形状データに対して姿勢推定を行った。また、獲得された推定結果がどの程度正確であるかについて評価し、これに対する考察を加える。

なお、ボクセルデータをリープグラフの原理に基いて骨格線モデルを取得する方法として、aMRG<sup>9)</sup> と呼ばれる手法を用いた。

##### 4.1 姿勢推定の精度評価指標

入力となる人体の三次元ボクセルデータに対して、推定された姿勢がどの程度の精度であるかを定量的に評価する指標が必要となる。ここでは、全円筒のフィッティングが完了した時点で、少なくとも一つの円筒に含まれるボクセルの数  $V_f$  と、ボクセルデータ中に含まれるボクセル総数  $V$  の割合を評価値とする。すなわち、

$$E_f = \frac{V_f}{V} \quad (11)$$

により定められる評価値  $E_f$  が大きいほど、より姿勢推定結果が高精度であるとみなす。

ただし、円筒が実際の位置から大きくずれている場合でも、円筒に含まれるボクセル数が増えることがある。また、体積の大きい円筒ほどより多くのボクセルを含む可能性が高いため、各円筒ごとに評価値  $E_f$  に対する寄与率に差が生じる。

##### 4.2 結果と考察

実環境においてダンス動作を行う人物を撮影し、それを元に視体積交差法から得られたボクセルデータを入力とした、姿勢推定アルゴリズムの精度評価実験を

行った。これは計 100 フレームからなる時系列ボクセルデータであり、図 10 はそのうちのある 2 つのフレームに対する本手法を用いた姿勢推定結果である。また、図 11 には各フレーム毎の精度評価値  $E_f$  を示した。

図 11 の×印のプロットは、骨格線モデルを利用せず、ボクセルデータのみを入力とした姿勢推定アルゴリズム<sup>1)</sup>の結果である。このアルゴリズムは一つ前のフレームでの推定姿勢を、次フレームにおけるフィッティング時の初期姿勢とするため、誤差の累積により後のフレームほど  $E_f$  の値が低下している。この手法では、10 フレーム目以降から精度評価値  $E_f$  の値が下降しているため、有効フレーム数は約 10 フレームである。すなわち、この手法を有効に利用するのであれば、10 フレーム毎に初期姿勢を手動で訂正する必要がある。

これに対して、同図の+印のプロットは骨格線モデルを利用した本手法の結果である。全フレームにおいて  $E_f$  の値は 0.8 から 0.9 の間に位置しており、ほぼ一定の精度で推定姿勢が得られていることが分かる。これは、本手法では各フレーム毎に骨格線モデルに基づいて円筒軸の初期方向を決定しており、各フレーム間における姿勢推定結果が互いに無関係なためである。したがって、仮にあるフレームで致命的な誤差が発生しても、他のフレームにおける推定姿勢は影響を受けることがないため、フレーム数の多い長時間のデータに対しても、安定した結果を得ることが可能である。

ボクセルデータのみを入力とした手法において推定姿勢の精度が低下するのは、図 10 上の姿勢のように異なる体節同士が接触している場合である。これは、本来別の体節に含まれるはずのボクセルが、式 2 で表される円筒の軸方向に対するフィッティング評価値  $E$  に寄与してしまうためである。本手法では、骨格線モデルを元に円筒のおおまかな軸方向を決定し、軸の可動角度をその方向を中心とした微小領域に限定した上でフィッティングしているので、異なる体節に含まれるボクセルの影響を緩和できる。

ただし、このような問題を完全に回避するためには、単純に円筒に含まれるボクセル数のみによってフィッティング評価値を決定するのではなく、それらのボクセルが円筒内部でどのように分布しているかを考慮に入れた評価関数の導入が必要であると考えられる。

## 5. 結論

本研究では、三次元ボクセルデータをリープグラフという手法に基いて骨格線モデルへと変換し、これを入力として円筒多関節モデルのフィッティングに利用した姿勢推定手法を提案した。また、提案した手法を元に姿勢推定プログラムを実装し、精度評価実験を行った。

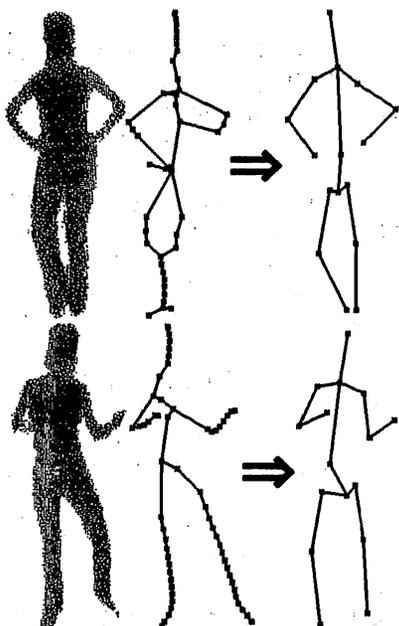


図10 ダンス動作における姿勢推定結果:左から順にボクセルデータ, 骨格線モデル, 推定姿勢を表す

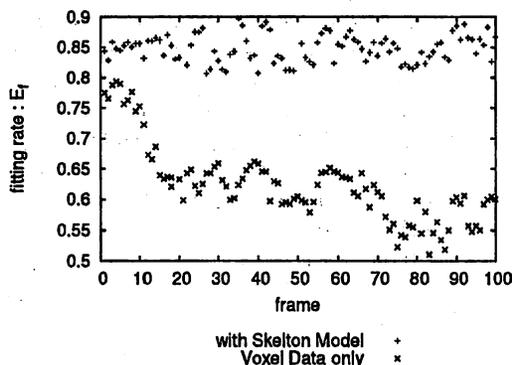


図11 各フレーム毎の評価値  $E_f$

実験結果による考察から, 姿勢推定の成功率や処理に要する時間という点で, 以下のような改善の余地があると考えている。これらは今後取り組む課題として検討する。

- ある円筒について, 式4で間違った交点  $r_i$  を  $n_e$  として選んでしまうと, そこから派生する円筒においても骨格線モデルの探索の際に致命的な失敗を引き起こす可能性がある。したがって, 複数の交点から  $n_e$  を決定するためのより信頼性の高い評価値を導入するか, あるいはある円筒においてフィッティングに失敗した場合, 既にフィッティングされている円筒まで処理を遡ることで再度探索し直すという方法が必要であると考え。

- 本手法では, 式9, 10により円筒軸方向の探索範囲に制約を与えているが, それでも円筒のフィッティング時における評価関数の計算に要するコストが大きく, 1フレームの人体形状から姿勢推定をするにはある程度の時間がかかる。また, この方法では式8の初期軸方向  $v_i$  が実際の軸方向から大きくずれている場合には, 軸方向の探索範囲の制約のため, 正確なフィッティングをすることができない。したがって, フィッティングの際に,  $v_i$  を初期値として, より効率良く円筒の軸方向を更新するような探索手法の導入を検討する。

### 参考文献

- 1) 村上 祐介, “3次元ボクセルデータに基づく人体の姿勢推定”, 京都大学大学院情報学研究科知能情報学専攻
- 2) K. Rohr, “Human Movement Analysis Based on Explicit Motion Models”, chapter 8, pp.171-198, Kluwer Academic Publishers, 1997.
- 3) C. Sminchisescu and A. Telea, “Human Pose Estimation from Silhouettes: A Consistent Approach using Distance Level Sets”, WSCG International Conference on Computer Graphics 2002, pp.413-420, Visualization and Computer Vision, 2002.
- 4) Thomas B. Moeslund and Erik Granum, “3D Human Pose Estimation using 2D-Data and an Alternative Phase Space Representation”, IEEE Workshop on Human Modeling Analysis and Synthesis (HuMANs 2000), South Carolina, June 16, 2000.
- 5) C. Christensen and S. Corneliusen, “Tracking of Articulated Objects using Model-Based Computer Vision”, Technical report, Laboratory of Image Analysis, Aalborg University, Denmark, June 1997.
- 6) 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次元動画追跡”, 電子情報通信学会論文誌, Vol.J79-D-2, No.1, pp71-83, 1996.
- 7) Aldo Laurentini, “How Far 3D Shapes Can Be Understood from 2D Silhouettes”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.17, No.2, pp.188-195, February 1995.
- 8) M. Hilaga, Y. Shinagawa, T. Kohmura, T.L. Kunii, “Topology Matching for Fully Automatic Similarity Estimation of 3D Shapes”, Proc. of SIGGRAPH 2001, pp.203-212, Dordrecht Boston, 2001.
- 9) Tony Tung et al., “The augmented multiresolution Reeb graph approach for content-based retrieval of 3d shapes”, International Journal of Shape Modeling (IJSM), Vol.1, No.1, pp.91-120, June 2005.