

# 時系列ボリュームデータの領域追跡を用いた 関節物体の体節への分割手法

飯山将晃

亀田能成

美濃導彦

京都大学

京都大学

情報学研究科

総合情報メディアセンター

## 1 はじめに

本研究では、人体に代表される関節物体の動作モデルを複数方向から撮影された映像を用いて獲得する手法を提案する。ここでいう動作モデルとは、それぞれの体節の形状とその運動を表現するモデルを指す。

関節物体の動作モデルを獲得する手法はいろいろなアプリケーションに応用することができる。例えば、動作モデルを用いてCGやVR上で物体の動作を表現すること、人間工学やロボティクスなどの分野において、実データから得られた人間の動作モデルを解析し、それを人間が取り扱う機械器具の設計に活かしたり、ロボットの制御に用いたりすることなどが挙げられる。

近年、レンジセンサやカメラを用いた形状計測方法が提案され、物体の形状をボリュームデータとして時系列で獲得することが可能となった [1]。しかしながら、このとき得られる情報は関節物体全体の形状であり、物体の形状を計測したからといって、その物体がどのように運動するのかを知ったことにはならない。そのため、物体の動作モデル獲得するためには計測された形状からそれぞれの体節（例えば手や足）などがどのような形状をし、またそれらがどのような運動をしているかについて知る必要がある。

計測された時系列ボリュームデータには、ボリュームデータをどのように各体節に分割すべきかという情報は含まれていない。そこで、関節物体の構造について何らかのモデルを定め、そのモデルに基づいてボリュームデータを体節毎に切り分けを行う。

これまで関節物体、特に人体について行われてきた研究では、体節の幾何形状に対して何らかのモデルを与え、体節の切り分けを行うものが多かった [2, 3, 4, 5]。

与える幾何形状のモデルが対象と大きく異なる場合、これらの手法ではうまく体節の切り分けを行うことができず、取り扱う対象に応じて逐一幾何形状のモデルを修正する必要があった。

そこで、本研究では幾何形状のモデルを必要とせずにボリュームデータを体節毎に分割する方法を提案する。関節物体の構造についてのモデルとして、各体節が剛体であること、そして体節の形状が空間的に連続した領域で表されることの2つを与える。そして、各映像に予め与えておいた対応点とボリュームデータの局所的な形状特徴を用いてボリュームデータの運動を追跡し、その結果を基にボリュームデータを体節毎に切り分ける。

複数方向から同時に撮影された映像を用いて各体節の形状を獲得し、動作モデルを獲得する方法として、Ioannis[6]は人体を対象とし、画像から人体の輪郭線を抽出し、輪郭線の形状の変化をもとに体節の形状とその位置を同時に推定する手法を提案した。しかしながら、この手法ではある定められたシナリオに沿った運動を被験者にとってもらう必要があり、被験者の任意の運動には対応できない。

これに対して提案手法では、各映像に対して予め対応点を与えておく必要があるものの被験者の任意の運動に対応できるアルゴリズムである。

以下2節では、複数方向から撮影された映像から時系列ボリュームデータを作成する手法について述べ、3節では、予め与えられた対応点とボリュームデータの局所的な形状を用いて、関節物体の運動を推定し、関節物体の運動をもとにボリュームデータを体節に分割することについて述べる。4節ではシミュレーション実験によって提案手法の有効性を検証する。5節はま

とめである。

## 2 映像からの時系列ボリュームデータの獲得

本研究では、複数方向から撮影された画像からボリュームデータを獲得する手法として視体積交差法 [1, 7] を利用する。視体積交差法は再構成できる形状に一部制限はあるものの、データとして物体が投影されたシルエット画像を与えるだけでよく、画像で生じるノイズ等に対するロバスト性に優れている。

**撮影環境** まず、観測に用いるカメラのモデルと、各映像に予め与える対応点について述べる。

観測に利用するカメラ台数を  $n$  台とし、各カメラを  $C_1, \dots, C_n$  で表す。ここで、各カメラ  $C_i$  はピンホールカメラモデルで表現可能な光学系を持つとし、その三次元空間座標から画像平面座標への投影変換行列は既知であるとする。

また、各映像に対していくつかの対応点を以下の条件を満たすように与える。

- 関節物体の表面にある点 (観測点) に対応する映像上の点を対応点とし、少なくとも 2 台のカメラが 3 次元空間上の同一の観測点を観測する。つまり、対応点からステレオ計測の原理を用いて観測点の 3 次元座標を計算できる環境にある。
- 映像において常に対応点が観測される状態にある。つまり、映像において対応点が消失することはないとする。
- 各体節に対して少なくとも 1 つ以上の観測点が存在する。

**視体積交差法** 次に、関節物体の 3 次元形状を復元するために用いる視体積交差法について述べる。

関節物体はカメラ  $C_i$  で撮影された画像上に投影され、投影像は二次元画像平面上の領域として得られる。この領域を物体投影領域と呼ぶ。このとき、物体は必ず、カメラ  $C_i$  のレンズ中心を端点とし物体投影領域上の任意の点を通る半直線の集合からなる錐体状の開空間に内接して存在する。この開空間をカメラ  $C_i$  における物体存在可能空間と呼ぶ。

従って、対象空間内の物体が複数のカメラ  $C_1, \dots, C_n$  によって観測されるとき、物体は必ず各カメラにおけ

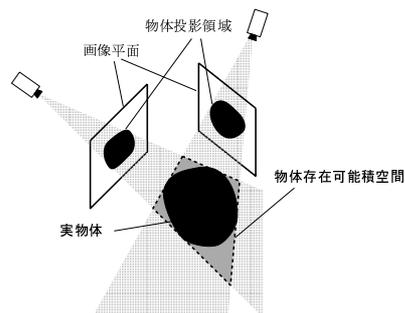


図 1: 物体存在可能積空間と物体形状

る物体存在可能空間の積空間に内接する (図 1)。この積空間を、物体存在可能積空間と呼ぶ。

物体存在可能積空間の表面は、一般には物体表面に一致しない。しかし実際的には、十分な数のカメラを用いれば、物体形状を十分近似するに足る物体存在可能積空間を得ることが可能である。よって、本稿では物体存在可能積空間を物体の再構成形状とし、これを体節毎に分割することによって体節の形状を得る。

## 3 提案手法の概要

本節では、本研究で用いる関節物体の構造モデルについて述べ、次にその構造モデルに基づいてボリュームデータを体節毎に分割する方法について述べる。

### 3.1 関節物体の構造モデル

関節物体の構造モデルとして以下の知識を与える。

- 各体節は剛体である。つまり体節の運動は 6 自由度の剛体運動パラメータで表現される。
- 体節の形状は空間的に連続している。

この構造モデルに基づいて、視体積交差法によって得られた時系列ボリュームデータを体節毎に分割し、動作モデルを獲得する。構造モデルとして体節が剛体で、かつ空間的に連続していると仮定しているので、その仮定に基づき、ボリュームデータの中で同一の剛体運動に従っているものをクラスタリングすればよい。

しかしながら、連続する 2 つのボリュームデータからだけでは関節物体がもつすべての体節の形状を得ることができるとは限らない。連続する 2 つのボリュームデータ間で、2 つの体節があたかも同一の体節であ

るかのように運動した場合、その2つの体節を別々の体節としてクラスタリングすることは出来ない。

そこで、この問題を解決するために逐次的にポリウムデータを用いることを考える。逐次的にポリウムデータを用いることにより、与えられたポリウムデータにおいて2つの体節が「常に」同一の体節であるかのように運動しない限り、関節物体が持つすべての体節毎にクラスタリングすることが出来る。

また、2時刻間で2つの体節の相対的な運動が小さい場合にも、その2つの体節を別々の体節としてクラスタリングすることは難しいが、この場合でも、逐次的にポリウムデータを用いることにより別々の体節としてクラスタリングすることが可能となる。

以降、時刻  $t_0$  から  $t_n$  におけるポリウムデータと対応点を用い、時刻  $t_0$  のポリウムデータを体節毎にクラスタリングすることを考える。

ポリウムデータの体節へのクラスタリングの処理は以下に挙げる3つのプロセスで構成される。

最初のプロセスでは、体節の運動が剛体であるという仮定のもとに、観測点間の3次元距離の時間変化を観察し、同一の体節上に存在すると考えられる観測点同士をクラスタリングする。クラスタリングによって得られた観測点の集合を点クラスタと呼ぶ。

次のプロセスでは、2時刻間における点クラスタの剛体運動を、点クラスタに含まれる観測点の位置変化と、観測点の周辺に存在するポリウムを用いて推定する。

そして最後のプロセスでは、体節の形状が空間的に連続しているとの仮定のもとに、点クラスタの剛体運動を用いてポリウムデータを剛体とみなせる領域毎にクラスタリングする。クラスタリングによって得られたポリウムをポリウムクラスタと呼ぶ。

上で述べた3つのプロセスをまず、時刻  $t_0, t_1$  におけるポリウムデータに対して適用し、点クラスタとポリウムクラスタを得る。さらに、時刻  $t_0, t_2$  におけるデータ、時刻  $t_0, t_3$  におけるデータというように逐次的にデータを用いて上の3つのプロセスを繰り返し、点クラスタやポリウムクラスタを再構築していくことによって、最終的に得られたポリウムクラスタを各体節の形状とする。また、各体節の運動は各時刻における点クラスタの剛体運動によって表される。

### 3.2 観測点の体節へのクラスタリング

以降、時刻  $t_0$  におけるポリウムデータと時刻  $t_k$  におけるポリウムデータを用いると仮定して議論を進める。

まず、同一の体節上に存在すると考えられる観測点同士をクラスタリングして、点クラスタを得る。

体節は剛体であるという仮定から、同一体節に属する任意の2つの観測点間の3次元空間上での距離は時間的な変化がないと仮定できる。そこで、任意の2つの観測点間の距離を計算し、その距離の時間変化を調べることによって、2つの観測点が同一の体節に属しているかどうかを決定する。

各時刻における観測点の3次元座標は、カメラの投影変換行列を用いることによりステレオ計測の原理を用いて計算することができる。

なお、与えられた対応点の位置には誤差が含まれ、また、体節の形状も厳密には剛体では表現されないことが予測されるため、2つの観測点間の距離の時間変化が一定の閾値以内の場合、その2つの観測点は同一の体節に属していると判定する。

### 3.3 点クラスタの剛体運動の推定

次に、点クラスタそれぞれに対して、剛体運動を推定する。

剛体運動は6つの独立したパラメータ(剛体運動パラメータ)で表現される。点クラスタに含まれる観測点の数が3つ以上の場合、一意に剛体運動を推定することができるが、各々の体節に対して常に3つ以上の観測点を与えられている保証はなく、2つ以下の観測点しか与えられていない場合は、観測点のみからでは原理的に剛体運動パラメータを計算することはできない。

そこで、観測点が3つ未満の場合でも、観測点だけでなく観測点の周辺に存在するポリウムを利用することにより、点クラスタの剛体運動パラメータを推定する。

点クラスタに含まれる観測点の数が3つ以上の場合観測点の数が3以上の場合、原理的には剛体運動パラメータが得られるが、実際には観測点の3次元位置の計算に誤差が生じたり、体節の運動が厳密な剛体運動に従わなかったりするため、剛体運動パラメータを得ることができない。

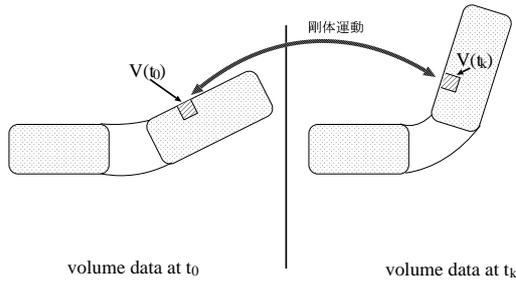


図 2: 体節上に存在する領域

そこで、エラー関数  $E$  を次のように定め、エラー関数を最小にする剛体運動パラメータを最急降下法を用いて計算する。

$$E = \sum_{i=0}^n \|R(\mathbf{p}_i(t_0)) - \mathbf{p}_i(t_k)\|^2 \quad (1)$$

ここで、 $\mathbf{p}_i(t_k)$  は時刻  $t_k$  における各観測点の 3 次元位置、 $R$  は剛体運動を表すパラメータ数 6 の関数である。

点クラスタに含まれる観測点の数が 2 つ以下の場合観測点の数が 2 以下の場合、原理的に剛体運動パラメータを計算することができない。そこで、体節の形状がほぼ剛体であるという性質を用い、観測点の近辺のボリュームを用いてこの点クラスタの剛体運動パラメータを推定する。

時刻  $t_0$  におけるボリュームデータ上のある領域  $V(t_0)$  が時刻  $t_k$  のボリュームデータにおいて占有する領域（以降これを  $V(t_0)$  の「時刻  $t_k$  における占有領域」と呼ぶ）を  $V(t_k)$  とした場合、もし  $V(t_0)$  が剛体領域であるならば、 $V(t_0)$  と  $V(t_k)$  との間の運動は剛体運動で表現することができる（図 2）。

そこで、点クラスタの近傍に存在する  $V(t_0)$  に対して、その占有領域  $V(t_k)$  を探索することによって点クラスタの剛体運動パラメータを計算する。 $V(t_0)$  に対して剛体運動パラメータのうち 6 つを変化させることで  $V(t_0)$  と形状が一致する  $V(t_k)$  を探索し、 $V(t_0)$  と形状が一致する  $V(t_k)$  が検出されたときの剛体運動パラメータを点クラスタの剛体運動パラメータとする。

なお、形状が一致するかどうかの評価は、剛体運動パラメータを用いた剛体変換を施して  $V(t_k)$  を計算し、 $V(t_k)$  と時刻  $t_k$  のボリュームデータとの領域の重なりを計算することで行う。

点クラスタに含まれる観測点の数が 2 つの場合、6

自由度の剛体運動パラメータのうち、自由度 5 の制約を与えるため、1 次元空間の探索となる。

また、点クラスタに含まれる観測点の数が 1 つの場合、与えられる制約は自由度 3 となるため、3 次元空間の探索となる。

以上のようにして得られた点クラスタの剛体運動パラメータを各体節の運動とする。

### 3.4 観測点の移動に基づくボリュームデータの体節への分割

最後にボリュームデータを剛体毎にクラスタリングする。

ボリュームデータ中のある領域  $V(t_0)$  が点クラスタに属するための必要条件は以下の通りである。

1.  $V(t_0)$  の運動が点クラスタの剛体運動パラメータで表現される。
2.  $V(t_0)$  が点クラスタと空間的に連続した領域に存在する。

そこで、時刻  $t_0$  のボリュームデータを均等な大きさの立方体領域に分割し、それぞれの立方体領域に対して、どの点クラスタに属するかを判定する。

例として、時刻  $t_0$  のボリュームデータ中のある領域  $V(t_0)$  が点クラスタに属するかどうかを判定することを考える。

最初に、 $V(t_0)$  が点クラスタに属するための 1 の条件を満たすかどうかを判定する。点クラスタの剛体運動パラメータを用いた剛体変換を  $V(t_0)$  に対して施し  $V(t_k)$  を計算する。この計算によって得られた  $V(t_k)$  が時刻  $t_k$  のボリュームデータ上に存在する場合、この領域  $V(t_0)$  は 1 の条件を満たし得ると考えられる。

次に、2 の条件を満足するかどうかを判定するために、1 の条件を満たす領域の集合に対してラベリング処理を行い、点クラスタと空間的に連続している領域を抽出する。抽出された領域がボリュームクラスタとなる。

分離すべき 2 つの体節の相対的な運動が小さい場合、上に述べたような判定法では、領域が 2 つ以上の点クラスタに属し得るという状況が発生する。しかしながら、逐次的にボリュームデータを用いていく過程で、2 つの体節の相対的な運動が大きくなることにより 2 つ以上の点クラスタに属し得る領域は減少していくことが期待できる。

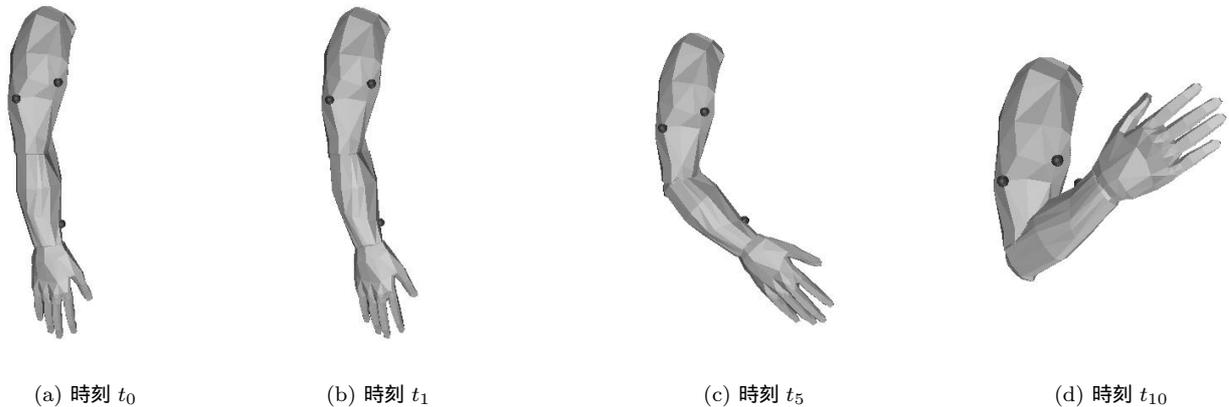


図 3: 実験データ

## 4 実験と考察

提案手法の有効性を評価するため、シミュレーション実験を行った。

実験データとして、図 3 に示される右腕の 3 次元形状の動作系列を作成した。図 3 で黒い点で示されているのが観測点である。8 台の仮想カメラをこの 3 次元形状を取り囲むように円周上に配置し、それらを用いてこの 3 次元形状を撮影することで映像データを作成した。

まず、時刻  $t_0$  における画像に対して視体積交差法を適用して得られたポリウムデータを図 4(a) に示す。ひとつのボクセルの大きさを  $1\text{cm}$  とした。

次に、時刻  $t_0$  と時刻  $t_1$  のポリウムデータを用いて時刻  $t_0$  のポリウムデータを分割した結果を図 4(b) に示す。分割の結果上腕と下腕の 2 つのポリウムクラスタが生じているが、時刻  $t_0$  と時刻  $t_1$  とのポリウムの相対的な変位が小さいため、両方のポリウムクラスタに属する部分が肘付近に大量に生じている。

次に、時刻  $t_0$  から時刻  $t_5$  までのポリウムデータを逐次的に用いて時刻  $t_0$  のポリウムデータを分割した結果を図 4(c) に、時刻  $t_0$  から時刻  $t_{10}$  までのポリウムデータを逐次的に用いて時刻  $t_0$  のポリウムデータを分割した結果を図 4(d) に示す。逐次的にポリウムデータを適用していくことによって、ポリウムの相対的な変位が大きくなるため、両方のポリウムクラスタに属する部分が少なくなり、最終的に上腕部分と下腕部分の形状を得ることができている。

## 5 まとめ

本稿では、複数方向から撮影された映像より視体積交差法を用いて時系列ポリウムデータを獲得し、ポリウムデータから関節物体の体節の形状を抽出することによって動作モデルを獲得する手法を提案した。

まず、観測点間の 3 次元距離の時間変化を観測することにより、同一の体節に属すると考えられる観測点同士を点クラスタとして獲得した。次に、点クラスタの剛体運動を、観測点の位置変化と観測点の周辺に存在するポリウムを用いて推定した。最後に、点クラスタの剛体運動をもとにポリウムデータを剛体とみなせる領域にクラスタリングし、得られたポリウムクラスタを各体節の形状とした。

提案手法の問題点として、予め与えておく対応点が満たすべき条件が厳しいという点が挙げられる。今回提案した手法では、観測点の 3 次元座標を常に得ることができなければならない、また、それぞれの体節について少なくとも 1 つ以上の観測点が存在しなければならない。

ある体節について観測点が全く存在しない場合、その体節が存在する領域はどのポリウムクラスタにも属さないことが予想される。このようなもののポリウムクラスタにも属さない領域をポリウムクラスタの生成の過程で検出することができれば、それを独立した体節として認識することで、観測点を与えられていない体節についても形状と運動の復元が可能となると考えられる。

また、この他にも現実的には異なるカメラ間での点

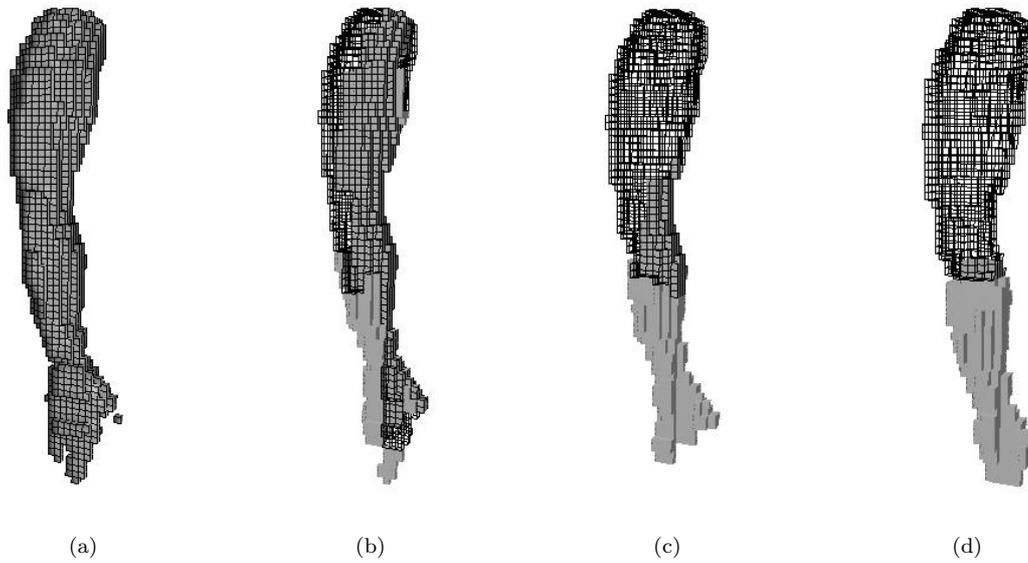


図 4: 実験結果

の対応付けが得られない状況や、特徴点がある時刻では隠蔽されてしまう状況などが考えられる。これらの状況に対応することのできる方法についても今後考察していく必要がある。

## 参考文献

- [1] 亀田 能成, 太尾田 健男, 角所 考, 美濃 導彦, “時空間の分割とビデオ画像のパイプライン処理による高速三次元再構成,” 情処論, Vol.40, No.1, pp.13-22, 1999
- [2] Daniel D. Morris, James M. Rehg “Singularity Analysis for Articulated Object Tracking” Proc. Computer Vision and Pattern Recognition, pp.289-296, 1998
- [3] James M. Rehg, Takeo Kanade, “Model-Based Tracking of Self-Occluding Articulated Objects,” Proc. Fifth International Conference on Computer Vision pp.612-17, 1995
- [4] D.M.Gavrila and L.S.Davis, “3-D model-based tracking of humans in action: a multi-view approach,” Proc. Computer Vision and Pattern Recognition, pp.73-80, 1996
- [5] 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,” 信学論, D-II, Vol.J79-D-II, No.1, pp71-83, 1996
- [6] Ioannis A. Kakadiaris and Dimitri Metaxas, “3D Human Body Model Acquisition from Multiple Views,” Fifth International Conference on Computer Vision, pp.618-623, 1995
- [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, 1995