多視点動画像を用いた舞踊動作のデジタル化

藤田 武史　　向川 康博　　尺長 健

岡山大学 工学部 情報工学科

〒 700-8530 岡山市津島中 3-1-1

{mukaigaw, shaku}@chino.it.okayama-u.ac.jp

あらまし

本稿では，多視点動画像を用いて人物の動作情報と見え情報を獲得することで，舞踊動作をデジタルコンテンツ化する手法について述べる．複数カメラで撮影された映像を基に，各時刻の人物形状をボクセルデータとして得ることで動作情報を獲得する．また，獲得した人物形状の表面ボクセルの色を決定することにより見え情報を獲得する．更に，獲得した舞踊動作を解析するために，人体モデルをフィッティングすることで人物形状を各体節に分割する．実際に，8 台のカメラからなる多視点カメラシステムを構築し，獲得した舞踊動作を各体節へ分割し，本手法の有効性を確認した．

Digitization of Dance Motion from Multiple Image Sequences

Takeshi FUJITA, Yasuhiro MUKAIGAWA, Takeshi SHAKUNAGA

Department of Information Technology, Faculty of Engineering, Okayama University
3-1-1 Tsushima-naka, Okayama, 700-8530, JAPAN
{mukaigaw, shaku}@chino.it.okayama-u.ac.jp

Abstract

This paper discusses a method for digitization of dance motion. The motion information is recovered from multiple image sequences, and represented as a sequence of voxel data. The appearance information is also recovered from the image sequences by selecting the most feasible surface color. In order to analyze the dance motion, the human body is segmented into parts by fitting a human model. Experimental results show that our eight-camera system works for the digitization of dance motion.

1　はじめに

文化財の保護を目的としたデジタルコンテンツ化の研究が，近年盛んに行われている [1]．その中でも，伝統舞踊や能といった無形文化財は，後継者不足などの理由から失われつつあるのが現状である．そのため，舞踊動作を単に映像として保存するだけではなく，動作情報をも保存しておく重要性が増している．一方，従来から人物の動作を獲得する手段として，磁気式センサを利用したモーションキャプチャシステムが広く用いられており，精度良く人物動作を獲得できることが知られている [2]．しかし，舞踊動作のコンテンツ化を考えた場合，人物の動作情報のみならず，身に付けている衣装や装飾品などの見え情報も伝統舞踊の一部として重要な情報と考えられ，動作情報と見え情報を同時に獲得する必要がある．

この目的を達成するために，多視点カメラシステムを用いた研究が活発になされている [3][4][5] ．これらの手法は，時系列の人物形状と表面テクスチャを獲得することで，対象人物の動作情報と見え情報を同時に獲得することができるため，舞踊動作のコンテンツ化に適していると考えられる．しかし，獲得した舞踊動作をコンテンツ化する際には，舞踊動作の意味づけや加工・編集等のさらなる応用に利用可能な形で保存することが望ましいと考えられ，これらに関する検討を同時に行う必要がある．

そこで，本研究では，多視点カメラシステムを入力手段とし，舞踊動作の動作情報と見え情報を同時に獲得し，動作解析に利用することを目指す．このため本稿では，人体モデルを用いて，舞踊動作を解析することで，対象人物の姿勢を推定し，得られた人物形状を各体節に分割する方法についてもあわせて論じる．以下，2 節では多視点カメラシステムを用いて舞踊動作の動作情報と見え情報を獲得する方法について述べる．3 節では，人体モデルを用いた動作情報の解析方法について述べ，4 節では実験結果を示す．

2　舞踊動作のデジタル化

2.1　多視点カメラシステム

本研究で用いる多視点カメラシステムを図 1 に，このシステムによって撮影された画像の一例を図 2 に示す．対象を洩れなく撮影するために，観測範囲を取り囲むように複数の校正済みカメラを設置する．この時，カメラからの映像を取り込む PC を各カメラに対して用意し，各 PC 間をネットワークで接続することで獲得した情報を共有する．また，あらかじめ時刻を合わせた各 PC に，撮影開始トリガとして同期信号を送信し，各カメラ間での撮影時刻の同期を実現する．

2.2　動作情報の獲得

獲得した複数の映像から各時刻での対象人物の 3 次元形状を復元することで，動作情報を獲得する．まず，カメラから得られた映像の各時刻で，人物領域を抽出してシルエット画像を獲得する．人物領域は背景画像と入力画像との差分をとることで得られる．この時，照明変動に対応するために，YCrCb 表色系に変換し，Cr，Cb 成分のみの差分を用いる．次に，得られた複数視点からのシルエット画像から視体積交差法を用いて人物形状を復元する．視体積交差法は，それぞれのシルエット画像から得られる多角錐の共通部分を求めることで対象物体が存在する空間を求める手法である．この空間を物体存在空間と呼び，この空間は十分な数のカメラを用いることで実際の物体形状に近付けることが可能である．

物体存在空間は多面体であり，3 次元空間中の複数の平面を用いて定義できるが，実際に物体形状を獲得する際には，取り扱いが複雑になる．そのため，本研究では，計測空間をボクセル空間として表現し，人物形状をボクセルの集合体として獲得する．ボクセル空間とは，3 次元空間を格子状に配置した小立方体 (ボクセル) の集合で構成される空間である．ボクセルをシルエット画像上に射影し，全てのシルエット画像でシルエットの内部にあれば占有状態，それ以外の場合は空状態とする．全ボクセルについてこの処理を行い，占有状態のボクセルの集合体を求めることにより人物形状を獲得する．以上の処理を各時刻で行

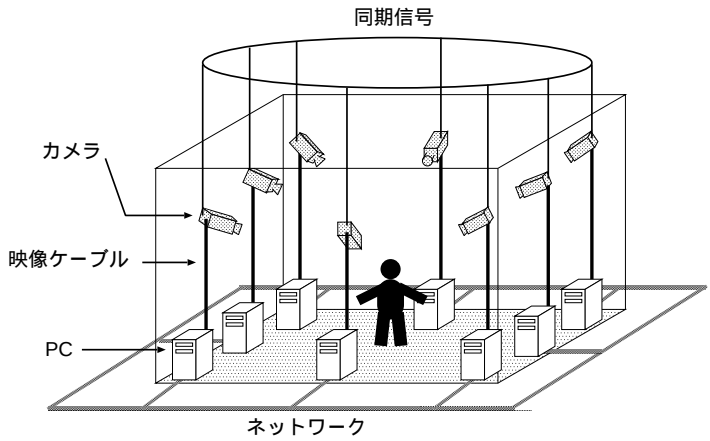


図 1: 多視点カメラシステム

図 2: 撮影画像の例

い動作情報を獲得する．

2.3 見え情報の獲得

前節で述べた視体積交差法では，人物の形状しか得られない．そこで，人物形状として獲得したボクセルの色を決定することで，舞踊動作の見え情報を獲得する．ただし，人物の内部に存在するボクセルは，画像中に現れないので，表面ボクセルの色のみを決定すればよい．

まず，各画像上で観測されている表面ボクセルを求め，画像上に投影された位置にある画素の色を候補色として保持する．大部分の表面ボクセルは複数の画像上で観測されているため，候補色を複数持つ表面ボクセルが多数存在すると考えられる．本来ならば，全ての候補色は同じ色になるが，獲得した形状や射影行列の誤差等の影響により全ての候補色が同じになるとは限らない．この時，単純に候補色の平均色をとってしまうと，実際には存在しない色となる場合があるので，平均色に最も近い候補色を表面ボクセルの色とする．これにより，舞踊動作の見え情報を獲得する．

3 舞踊動作の解析

3.1 体節への分割

前章で述べた方法を用いて舞踊動作の動作情報と見え情報の獲得を行えるが，獲得した舞踊動作の意味づけや加工・編集等のさらなる応用を可能にするためには，舞踊動作の解析が必要である．そこで本研究では，ボクセルデータとして獲得した各時刻での人物形状に対し，定義した人体モデルのフィッティングを行うことで，各体節の関節位置，関節角度を算出し，人物の姿勢を推定する．更に，推定した姿勢に基づいて各ボクセルに体節ラベルを割り当てる (以下ラベリングと呼ぶ) ことで，獲得した人物形状を各体節へ分割する．これにより，各体節の動作を独立の動作情報として獲得することが可能となる．

3.2 人体モデル

動作情報の解析を行うために人体モデルを定義する．これまでに，人体を対象として各体節を円筒等の単純な幾何形状で近似し，モデルフィッティングを行うことで，姿勢を推定する研究が行われている [6][7]．これらの手法は，近似した形状モデルが対象の形状と大きく異なる場合には安定にフィッティングを行うことが難しい．一方，各体節を剛体と仮定し，各映像にあらかじめ与えておいた対応点と局所的な形状情

表 1: ノード ω の持つパラメータ

$\omega \in \mathbf{\Omega}$	ノード番号．
$\nu(\omega) \in (\mathbf{\Omega} - \{\omega\})$	ノード ω の親ノード．
$\mathbf{\Lambda}^{\omega} \subseteq (\mathbf{\Omega} - \{\omega\})$	ノード ω の子ノード $\lambda(\omega)$ の集合.
$\boldsymbol{J}^{\omega} = [t_x^{\omega}, t_y^{\omega}, t_z^{\omega}]^T$	関節位置．親ノード $\nu(\omega)$ の局所座標系に対する値.
$\boldsymbol{r}^{\omega} = [r_x^{\omega}, r_y^{\omega}, r_z^{\omega}]^T$	関節角度．親ノード $\nu(\omega)$ の局所座標系に対する値.
$\boldsymbol{g}^{\omega} = [g_x^{\omega}, g_y^{\omega}, g_z^{\omega}]^T$	関節の重心位置．ノード ω の局所座標系に対する値.
$\boldsymbol{\tau}^{\omega} = [\tau_x^{\omega}, \tau_y^{\omega}, \tau_z^{\omega}]^T$	関節の空間的分布を表す軸方向．ノード ω の局所座標系に対する値.
$\boldsymbol{V}^{\omega} = [V_x^{\omega}, V_y^{\omega}, V_z^{\omega}]^T$	関節の空間分布の分散の組．

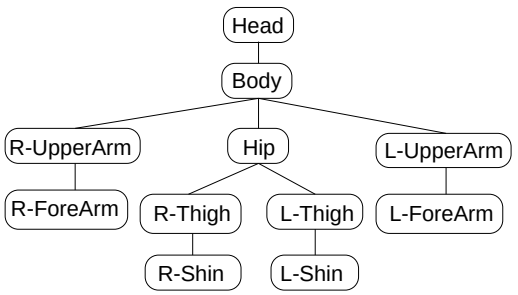


図 3: 各体節の接続関係

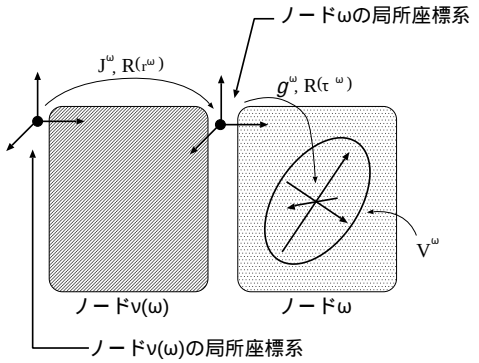


図 4: ノード ω の姿勢パラメータと形状パラメータ

報を用いて各体節へ分割する手法も提案されている [8]．しかし，本研究では和服を着用した人物の舞踊動作を対象としていることから，各体節を剛体と仮定することが難しく，単純な幾何形状で近似ができないため，これらの手法を用いることはできない．

一方，各体節の形状の変化は一定の範囲であることから，明確な形状情報ではなく，体節の空間中での重心位置，分布を表す軸方向，分散等の情報を持つ人体モデルを定義することで，体節の形状の変化に対応できると考えられる．そこで，本研究では各体節の空間的分布を形状情報として持つ人体モデルを定義する．

まず，人体を 11 個の体節からなるものとし，図 3 に示すように，根ノードが Head 部となる木構造で表現する．根ノードを 0 とするノードの集合を $\mathbf{\Omega}$ とし，ノード $\omega \in \Omega$ の持つパラメータを表 1 に示す．各ノード ω は姿勢パラメータとして，関節位置 $\boldsymbol{J}^{\omega}$，関節角度 $\boldsymbol{R}(\boldsymbol{r}^{\omega})$ を持ち，それによって決定される局所座標系を持っている．ただし，局所座標系は常に X 軸上に子ノード $\lambda(\omega)$ の関節位置を持つものとし，$\boldsymbol{J}^{\omega}$ $\boldsymbol{R}(\boldsymbol{r}^{\omega})$ は親ノード $\nu(\omega)$ の局所座標系に対する値で，根ノード以外の関節位置は親ノードの局所座標系で固定された位置にあるものとする．また，形状パラメータとして空間的分布の重心位置 $\boldsymbol{g}^{\omega}$，軸方向 $\boldsymbol{R}(\boldsymbol{\tau}^{\omega})$，各軸方向の分散 $\boldsymbol{V}^{\omega}$ を持つ．ただし，$\boldsymbol{g}^{\omega}$ $\boldsymbol{R}(\boldsymbol{\tau}^{\omega})$ はノード ω の局所座標系に対する値である．ノード ω の持つパラメータを図 4 に示す．

3.3 形状パラメータと初期姿勢パラメータの決定

人体モデルの各ノードが持つパラメータのうち，形状パラメータは，対象人物によって異なる値を持つため，分割処理を行う前に求める必要がある．また，形状パラメータの重心位置，軸方向は姿勢パラメータに依存する．このことから，本研究では，初期時刻 $t=0$ において，図 5 に示すように両手を水平に伸ばした姿勢で 3 次元形状を獲得する．この形状を手入力により各体節に分割を行うことで得られる各ボクセルの初期ラベルと，手動により得た各体節の関節位置 $\boldsymbol{J}_W^\omega$ から，各ノードの形状パラメータと時刻 $t=0$ の各ノードの姿勢パラメータを決定する．

まず，与えられた関節位置から姿勢パラメータを求める．原点が $\boldsymbol{J}_W^\omega$，X 軸の方向が $\boldsymbol{J}_W^\omega - \boldsymbol{J}_W^{\lambda(\omega)}$ となる座標系がノード ω の局所座標系になることから，この座標系からワールド座標系への変換行列を $\boldsymbol{A}_W^\omega$ とすると，以下の式が成り立つ．

$$\begin{bmatrix} \boldsymbol{R}(\boldsymbol{r}^\omega) & \boldsymbol{J}^\omega \\ \boldsymbol{0}^T & 1 \end{bmatrix} = (\boldsymbol{A}_W^{\nu(\omega)})^{-1}\boldsymbol{A}_W^\omega \tag{1}$$

ただし，$\boldsymbol{A}_W^{\nu(0)} = \boldsymbol{I}$ とする．これにより，ノード ω の姿勢パラメータを決定する．

次に，ノード ω の形状パラメータを求める．各体節ごとにその体節ラベルが与えられたボクセルの座標について平均と分散共分散行列を求めて主成分分析を行うことにより体節の空間的分布を得る．これにより得られた重心位置 $\boldsymbol{g}_W^\omega$，軸方向 $\boldsymbol{R}(\boldsymbol{\tau}_W^\omega)$ はワールド座標系に対する値であるから，以下の式を用いて，局所座標系に対する値を求めることで，ノードの形状パラメータを決定する．

$$\begin{bmatrix} \boldsymbol{R}(\boldsymbol{\tau}^\omega) & \boldsymbol{g}^\omega \\ \boldsymbol{0}^T & 1 \end{bmatrix} = (\boldsymbol{A}_W^\omega)^{-1} \begin{bmatrix} \boldsymbol{R}(\boldsymbol{\tau}_W^\omega) & \boldsymbol{g}_W^\omega \\ \boldsymbol{0}^T & 1 \end{bmatrix} \tag{2}$$

以上の処理により，各ノードの形状パラメータと時刻 $t=0$ の姿勢パラメータを決定する．

3.4 体節ラベルの更新

時刻 $t=1$ 以降は，定義した人体モデル及び利用可能な拘束条件を用いて，各ボクセルのラベルを順次更新することで各体節への分割を行う．分割に用いる拘束条件を以下に列挙する．

(1). 各体節は常に一つのみ

(2). 各体節の体積の比率は一定

(3). 抽出した顔・手領域位置

(4). 各時刻間での動作変化は微小

(1)(2) は各体節は二つ以上に分かれることはなく，形状が変化しても体積の変化は少ないことを示す．また，(3) は 2.3 節で獲得した色情報から容易に顔，手領域を検出することが可能であることを示す．(4) はカメラの撮影間隔は対象人物の動作変化に対して十分小さいことを示す．

しかし，列挙した全ての拘束条件を同時に満たす解を求めるためには，計算コストが高くなるという問題はもちろん，解が見つからないおそれもある．そこで，本研究では，図 6 に示すようにラベリング処理を仮ラベリング，姿勢推定，本ラベリングの 3 段階に分けて行い，人体モデルの構造と拘束条件を一つずつ満たしていくことで，人体モデルと全ての拘束条件を考慮したラベリングを行う．以下，各処理について詳しく述べる．

図 5: 初期時刻における姿勢

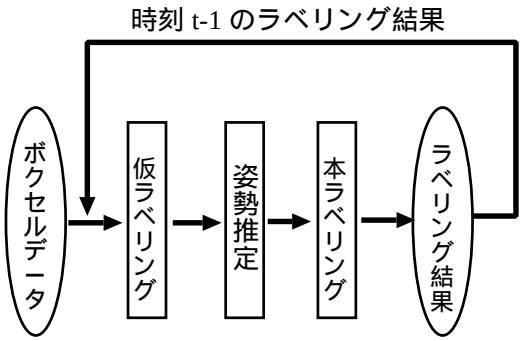


図 6: 処理の流れ

3.4.1 仮ラベリング

仮ラベリングでは，時刻 t の人物姿勢が未知であるため，拘束条件のみを用いてラベリングを行う．時刻 $t-1$ でのラベリング結果を利用してラベリングを行い，その後，体積の不変性を考慮した調整を行う．これにより，時刻 t の各ボクセルに仮の体節ラベルを割り当てる．

まず，拘束条件 (4) に基づき，時刻 $t-1$ のラベリング結果を利用して nearest-neighbor 法によりラベリングを行う．現時刻 t に存在している全ボクセルに対し，時刻 $t-1$ に存在したボクセルの中で距離が最も小さくなるボクセルの体節ラベルを割り当てる．ただし，拘束条件 (3) を満たすために，顔，右手，左手領域部分にそれぞれ，Head 部，R-ForeArm 部，L-ForeArm 部の体節ラベルを割り当てる．顔，手領域は，時刻 $t=0$ において手入力で与え，時刻 $t=1$ 以降は $t-1$ の顔，手領域位置に最も近い位置にある肌色領域を追跡することで検出する．

次に，体積の不変性を考慮したラベリングの調整を行う．まず，各体節において体積が最大の領域を求め，この領域以外に属する全てのボクセルの体節ラベルを未確定状態に戻す．未確定状態のボクセルには空間的に連続でかつ最も近い体節ラベルを割り当てる．これにより，拘束条件 (1) を満たすようにする．次に，拘束条件 (4) に基づき，体積比率が初期ラベル時の比率より最も下回っている体節を求め，この体節ラベルを割り当てられた領域に隣接するボクセルのラベルを，この体節のラベルに変更することで体積を増加させる．以上の膨張処理を繰り返すことで，各体節の体積比率を初期ラベル時の比率に収束させる．

3.4.2 姿勢推定

仮ラベリングによって得られた結果を用いて人体モデルをフィッティングする．これにより時刻 t の各体節の位置・角度を推定する．本研究で用いる人体モデルは多数の姿勢パラメータを持つため，全てのパラメータを同時に変化させてフィッティングを行うと，計算コストが非常に高くなる．そこで，和服の影響を最も受けにくく，形状の変化が最も小さいと考えられる頭部からフィッティングを行う．

3.2 節で定義した人体モデルは明確な体節の形状を持っておらず，体節の空間的分布情報のみを持っている．そこで本研究では，各ノードは対応する体節の空間的分布として重心位置，分布を表す軸方向，各軸方向の分散値から決定される楕円体を定義する．仮ラベリングで同じ体節ラベルを割り当てられたボクセルのうち，楕円体の内部に含まれたボクセルの数が最大となるようにフィッティングを行うことで，人体モデルの各ノードの持つ体節の空間的分布と仮ラベリングの結果から得られる各体節の空間的分布が一致するようにする．

この時，ノード ω の体節ラベルを割り当てられたボクセル $\boldsymbol{M}=[X,Y,Z]^T$ がノード ω の持つ空間的分布を表す楕円体に含まれるかどうかの判定式 $E_\omega(\boldsymbol{M})$ を以下の式で定義する．ただし，$\sigma_x^\omega, \sigma_y^\omega, \sigma_z^\omega$ を空間的分布の各軸方向の標準偏差とし，各ノードの持つ空間的分布で定義される楕円体と体節の体積が同程度になるように，楕円体の各軸方向の半径を，$2\sigma_x^\omega, 2\sigma_y^\omega, 2\sigma_z^\omega$ とする．

$$E_\omega(\boldsymbol{M}) = \begin{cases} 1 : (\boldsymbol{M} - \boldsymbol{g}_W^\omega)^T (4\boldsymbol{K}_\omega)^{-1} (\boldsymbol{M} - \boldsymbol{g}_W^\omega) \leq 1 \\ 0 : otherwise \end{cases} \tag{3}$$

$\boldsymbol{K}_\omega$ は以下の式で定義される．

$$\boldsymbol{K}_\omega = \boldsymbol{R}(\boldsymbol{\tau}_W^\omega) \begin{bmatrix} (\sigma_x^\omega)^2 & 0 & 0 \\ 0 & (\sigma_y^\omega)^2 & 0 \\ 0 & 0 & (\sigma_z^\omega)^2 \end{bmatrix} \boldsymbol{R}^T(\boldsymbol{\tau}_W^\omega) \tag{4}$$

頭部は関節位置と関節角度の 6 パラメータ，それ以外の体節は関節角度の 3 パラメータを変化させて，フィッティングを行う．また，拘束条件 (4) を利用し，各姿勢パラメータの探索範囲を時刻 $t-1$ の値から一定の範囲内にする．

3.4.3 本ラベリング

本ラベリングでは，前節の姿勢推定で得られた人体モデルの姿勢から得られる時刻 t の各体節の空間的分布情報に基づくラベリングを行う．このため，各ボクセルから各体節の空間的分布の重心までのマハラノビス距離をラベリング時の評価値として用いる．$\boldsymbol{M} = [X, Y, Z]^T$ にあるボクセルと体節 ω とのマハラノビス距離 $D(\boldsymbol{M}, \omega)$ は以下の式で表される．

$$D(\boldsymbol{M}, \omega) = [\boldsymbol{M} - \boldsymbol{g}_W^\omega]^T \boldsymbol{K}_\omega^{-1} [\boldsymbol{M} - \boldsymbol{g}_W^\omega] \tag{5}$$

各ボクセルに対して，マハラノビス距離が最小となる体節のラベルを割り当てる．最後に，拘束条件 (1) を満たすように，各体節において体積が最大の領域を求め，この領域以外に属する全てのボクセルの体節ラベルを未確定状態に戻し，未確定状態のボクセルには空間的に連続でかつ最も近い体節ラベルを割り当てる．これにより，最終的なラベリング結果を得る．

ここで，仮ラベリングと本ラベリングでは nearest-neighbor 法の実装が異なる点に注意されたい．本ラベリング時には現時刻の各体節の空間的分布が既知であるため，マハラノビス距離を用いる．一方，仮ラベリング時には，現時刻の各体節の空間的分布が未知であるため，前時刻のラベリング結果をそのまま用いる nearest-neighbor 法 を用いるのが妥当である．

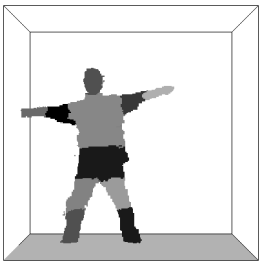
4 実験

4.1 実験環境

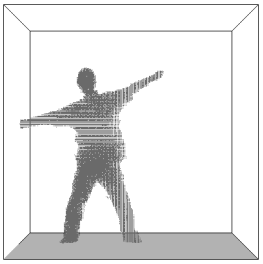
舞踊動作を獲得・解析するために，8 台のカメラ (SONY 製 DXC-200A) と，それぞれのカメラからの映像を取り込むための 8 台の PC (PentiumIII 750MHz Dual 搭載) で構成される多視点カメラシステムを構築した．本システムの観測範囲を部屋の中心付近の 2m×2m×2m の空間とし，その周りにカメラを設置した．各カメラで得られる画像は 320×240(pixel)RGB 各 8 ビットのカラー画像で，画像中の 1 画素が対応する空間中での大きさが，観測範囲の中心付近で約 1cm となることから，ボクセルの一辺の大きさを 1cm とした．なお，人物抽出を容易に行うために部屋の背景は緑色に統一した．

4.2 分割処理の検証

提案手法により，人物形状を各体節に安定に分割できるかどうかを確認する実験を行った．入力データとして，多視点動画像から取り出した連続する 2 フレームを用い，それぞれのフレームの時刻を $t = 0, t = 1$ とした．また，本節の実験では，人体の各体節の位置・形状を確認しやすくするために和服を着用してい

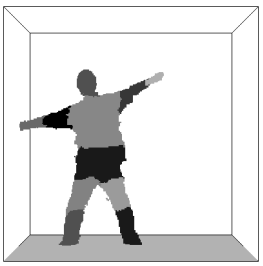


(a) 時刻 $t = 0$ (分割済み)

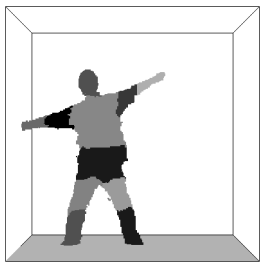


(b) 時刻 $t = 1$

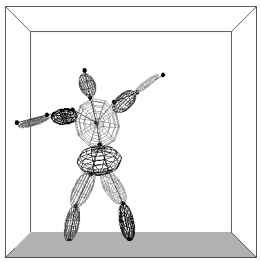
図 7: 3 次元形状復元結果



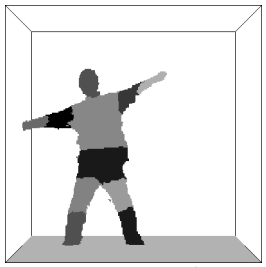
(a) nearest-neighbor 法



(b) 仮ラベリング結果



(c) 姿勢推定結果



(d) 本ラベリング結果

図 8: ラベリング結果

ない人物を対象とした．入力データから各時刻の 3 次元形状の復元を行い，時刻 $t = 0$ の形状に対して手動で行った分割結果をもとに時刻 $t = 1$ の形状に対して分割処理を行った．

図 7(a),(b) に各時刻の形状復元結果を示す．また，図 7(a) の各色は各体節を表している．図 8(a) は nearest-neighbor 法を用いた結果である．これより，大まかなラベリングしか行えておらず，各体節の接合部が乱れていることが分かる．図 8(b) は，図 8(a) の結果に対して体積の不変性を考慮した調整を行った仮ラベリングの結果であり，図 8(a) で小さくなっていた L-ForeArm 部が適切な大きさを確保できている．しかし，依然として Head 部と Body 部との接合部などが乱れている．図 8(c) は図 8(b) の結果に対してモデルフィッティングを行った結果で，適切に姿勢推定ができていることが確認できる．最後に，図 8(d) は姿勢推定の結果に基づいて本ラベリングを行った結果である．この結果より，Head 部と Body 部の分割が良好に行えており，各体節の体積も時刻 $t = 0$ と同程度の大きさが確保できていることが分かる．以上の結果より，本手法により人物形状を各体節に分割できることが確認された．

4.3 舞踊動作への適用

提案手法の有効性を確認するために，実際に，一連の舞踊動作を獲得し，各体節への分割を行った．入力として，本システムを用いて得られた 150 フレームの動画像を用いた．獲得した人物形状に対して表面ボクセルの色を付加した結果を，時刻 $t = 0$ から 15 おきに示したものが図 9 である．

この結果より，主に足元付近で，誤った色を持つ表面ボクセルの存在が確認される．これは人物形状の復元誤差やカメラの射影行列の誤差の影響によって，背景色が表面ボクセルの色として推定されてしまったと考えられる．しかし，大部分の表面ボクセルについては適切な色を推定できていることが分かる．これにより本手法を用いて，舞踊動作の動作情報と見え情報を，比較的良好に獲得できることを確認した．

また，図 10 に，獲得した人物形状と見え情報を用いて生成した任意視点画像の生成例を示す．ボクセル解像度が粗いために，生成画像の品質はそれほど高くないが，カメラを配置していない視点からの舞踊動作の見え方を画像として提示可能であることを確認した．

次に，獲得した舞踊動作に対して体節への分割処理を行った．図 11 に各時刻における姿勢推定結果を示す．また，図 12 に各体節への分割結果を示す．結果から，一時的にモデルフィッティングが不安定となる場合があるが，最終フレームまで比較的良好に各体節へ分割できていることが分かる．しかし，姿勢推定

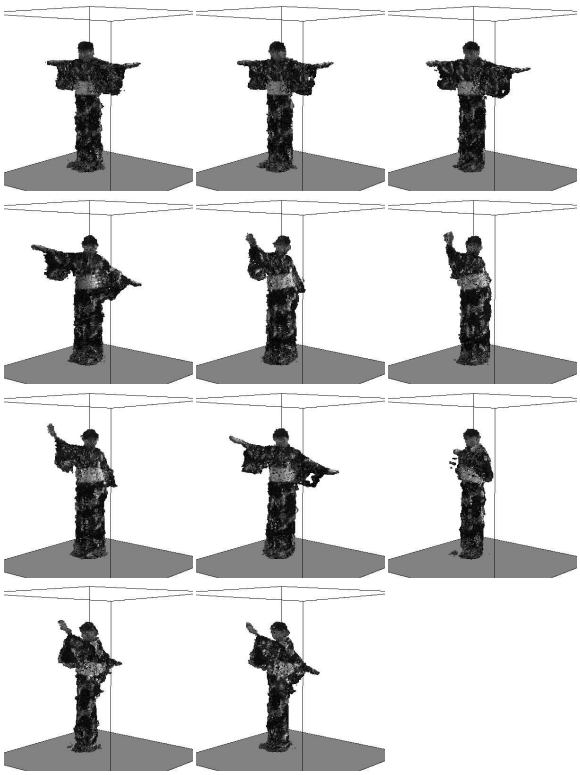


図 9: 舞踊動作獲得結果

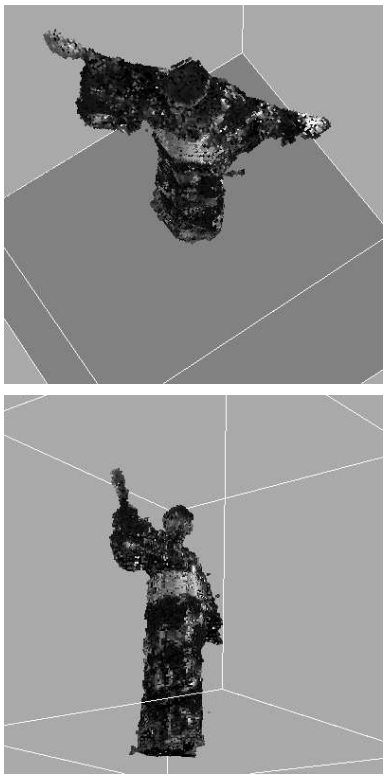


図 10: 任意視点画像の生成例

に関しては図 11 に示すように，$t = 70$ 付近で左足，$t = 100$ 付近で左腕の姿勢推定を大きく誤る場合があり，左腕に関しては $t = 100$ 以降，姿勢推定は誤ったままであった．

姿勢推定を失敗した原因の一つとして，提案手法は，根ノードの Head 部から順にフィッティングを行っているため，親ノードである根ノードから離れている体節は親ノードのフィッティングの誤差の影響を受けやすいくなっているためで，$t = 70$ はこれにより左足のフィッティングを失敗したと考えられる．また，$t = 100$ の左腕に関しては L-UpperArm 部及び体節の形状が大きく変化し，空間的分布が大きく変化したため，L-UpperArm 部のフィッティング結果が正解とずれが生じ，L-ForeArm 部の関節位置が誤った位置に推定されてまい，これが L-ForeArm 部のフィッティングにも影響を及ぼしたためと考えられる．

しかし，いずれの場合においても，各体節への分割は姿勢推定の失敗の影響をあまり受けていない結果となっている．これは，各体節を剛体と仮定せず，空間的分布で体節の形状を表現しているため，多少の姿勢推定の誤差を許容できる手法となっているためであると考えられる．以上の結果より，姿勢推定がやや不安定になる場合があったものの，一連の舞踊動作に対しても比較的良好に各体節へ分割できることを確認した．

5 まとめ

本稿では，多視点カメラシステムを用いて舞踊動作における動作情報と見え情報を同時に獲得し，舞踊動作の解析を行うことにより，人物形状を各体節に分割する手法について述べた．また，実際に，8 台のカメラを用いた多視点カメラシステムを構築し，舞踊動作のデジタル化を行った．実験結果より，動作情報・見え情報共に比較的良好に獲得ができることを確認した．また，各体節への分割は比較的安定に行えたが，和服の形状変化が大きい場合には，フィッティングが不安定となることが分かった．今後の課題として，人物のテクスチャ情報を有効活用し，より安定に姿勢推定を行うこと，及び，体節への分割結果に

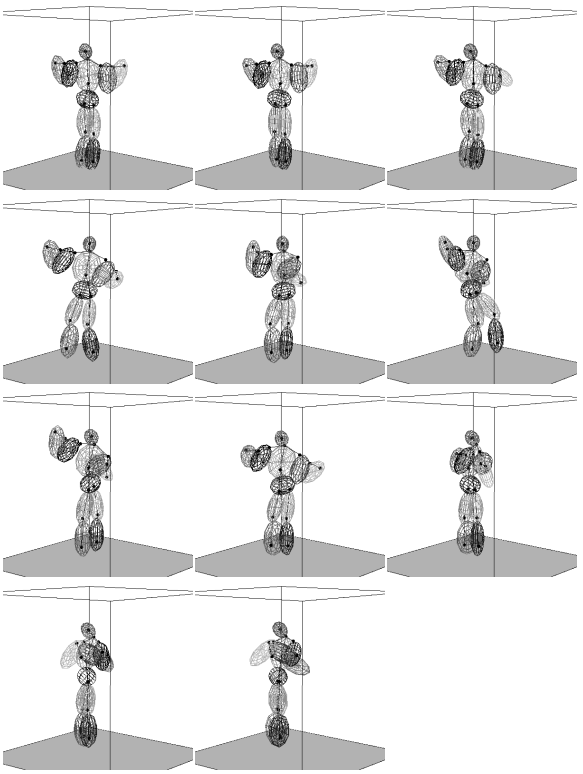


図 11: 姿勢推定結果

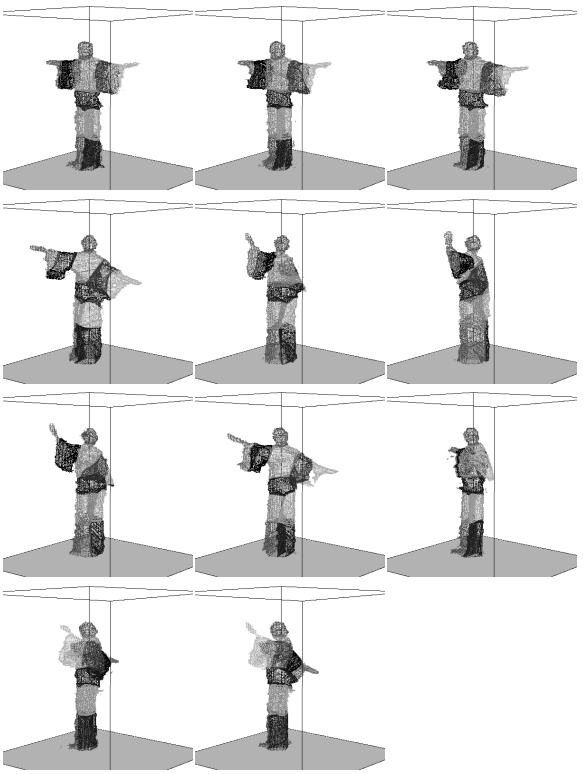


図 12: 各体節への分割結果

基づく舞踊動作の加工・編集が考えられる．

本研究は，科学技術振興事業団 CREST 池内プロジェクトの援助を受けて行った．

参考文献

[1] 池内克史，“CREST 池内プロジェクト -観察に基づく文化遺産のデジタル保存-”，計測自動学会第 2 回システムインテグレーション部門学術公演会論文誌，pp.217-218, 2001.

[2] 湯川崇，海賀孝明，長瀬一男，玉本英夫，“舞踊符による身体動作記述システム”，情処学論，vol.41, no.10, pp.2873-2880, 2000.

[3] ウ小軍，圓藤康平，和田俊和，松山隆司，“3 次元ビデオ映像の能動的実時間撮影と対話的編集・表示”，信学技報 PRMU2000-187，pp.9-16, 2001.

[4] 北原格，大田友一，“多視点映像の融合によるスポーツシーンの自由視点映像生成 -3 次元形状表現用平面の適応的配置-”，信学技報 PRMU2000-189，pp.23-30, 2001.

[5] H.Saito, T.Kanade，“Shape Reconstruction in Projective Grid Space from Large Number of Images”, IEEE Proc. CVPR’99, vol.2, pp.49-54, 1999.

[6] 佐藤明知，川田聡，大崎善彦，山本正信，“多視点動画像からの人間動作の追跡と再構成”，信学論，vol.J80-D-II, no.6, pp.1581-1589, 1997.

[7] 毛呂功，水田忍，美濃導彦，“表面作成のための点群データからの人体姿勢推定”, 信学技報 PRMU2001-128, pp.105-112, 2001.

[8] 飯山将晃，亀田能成，美濃導彦，“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”, 画像電子学会第 190 回研究会, 01-05-09, pp.57-64, 2001.