多視点動画像解析による舞踊動作のデジタル化

○ 向川 康博　　藤田 武史　　尺長 健
岡山大学工学部情報工学科

Digitization of Dance Motion Based on Analysis of Multiple Image Sequences

○ Yasuhiro MUKAIGAWA　　Takeshi FUJITA　　Takeshi SHAKUNAGA
Department of Information Technology , Faculty of Engineering, Okayama University

Abstract:
In this paper, we describe a method for the digitization of dance motion by analysis of multiple image sequences. We constructed a system which consists of eight cameras for capturing 3D human motion. In order to analyze the human motion, the human body is segmented into 11 parts by a combination process of bottom-up and top-down processes. Experimental results show that the human motion can be correctly captured and that the body can be segmented for each part.

1　はじめに

近年，失われつつある文化遺産の保護を目的としたデジタルコンテンツ化の重要性が増している．我々は，舞踊動作を対象とし，コンピュータビジョン技術によって動作情報を獲得・解析することを目指している．

磁気式センサを用いることで，精度良く舞踊動作を獲得できることが知られているが [1)]，コンテンツ化された舞踊動作を映像としても再利用するためには，動作情報だけではなく，衣装や装飾品を含めた見え情報も獲得する必要がある．そのため，人物動作を多数のカメラで撮影し，動作情報と見え情報を同時に獲得する研究が近年活発になされている [2)3)4)]．

舞踊動作のデジタルコンテンツ化において，動作の意味づけや編集を行う場合，単に動作情報を獲得するだけではなく，人物の動きとしての解析が重要となる．本稿では，多視点動画像から人物を対象とした 3 次元形状復元システムについて述べ，ボトムアップ処理とトップダウン処理により人物形状を体節に分割する方法について述べる．

2　システム構成

まず，多視点動画像を獲得するために構築したシステムの概要を述べる．部屋の中心付近を取り囲むように 8 台のカメラ (SONY DXC-200A) を設置した．撮影された映像は，各カメラと接続されている 8 台の PC により取り込まれ，同期のとれた 640×480 サイズの動画像としてビデオレート (30fps) で獲得できる．

このシステムを用いて多視点動画像を撮影し，背景差分法によって人物領域に対応するシルエットを抽出する．得られたシルエットから視体積交差法により，Fig.1 に示すような 3 次元形状をボクセルデータとして獲得する．観測範囲は $2m \times 2m \times 2m$ の空間とし，各ボクセルは一辺が $1cm$ の立方体とする．

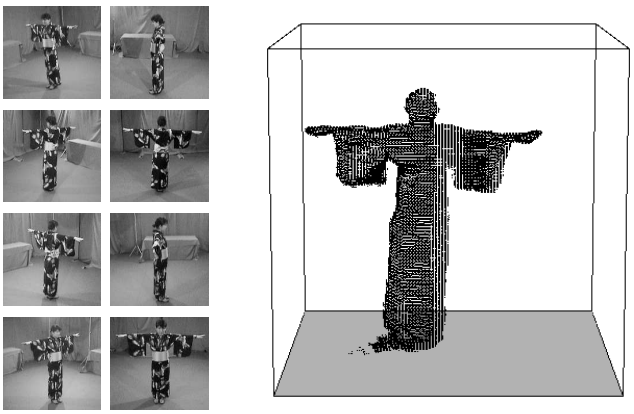


図 1:　Input images and 3D shape.

3　舞踊動作の解析

本研究では，獲得されたボクセルデータを解析し，各ボクセルに体節ラベルを付ける (以下ラベリングと呼ぶ) ことを主目的とする．以下では，ラベリングに用いる人体モデルと，ラベルの更新方法について順に述べる．

3.1　人体モデルとキャリブレーション

本研究では人体を 11 の体節に分割し，各体節が Fig.2 に示す接続関係を持つ人体モデルを用いる．なお，舞踊動作を対象とする場合，振り袖や帯などの和服に特有の形状を考慮する必要があり，これらを円筒などの一般的な形状モデルを用いて記述することは容易ではない．

そこで，初期フレームにおいて，Fig.1 に示すような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

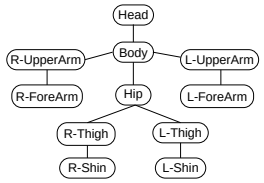


Fig. 2 Connection of parts

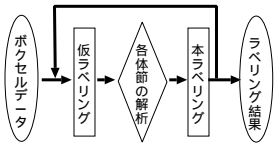


Fig. 3 Flow of the process.

状を，手入力により 11 の体節に分割する．これにより，初期フレームについては各ボクセルに対して初期ラベルを与えることができ，各体節に関する形状や体積に関するキャリブレーションも可能となる．

3.2 ラベルの更新

次に，手入力によって与えられた初期ラベルに基づいて，各時刻におけるラベルを順に更新する．このとき，ラベリングの手がかりとして，各フレーム間の動作変化は微小であること，各体節毎の体積変化は微小であることが利用できる．そこで本研究では，Fig.3 に示すように，ボトムアップ処理とトップダウン処理を組み合わせた処理によってラベルの更新を行う．以下，それぞれの処理について述べる．

(1) 仮ラベリング (ボトムアップ処理):
各フレーム間の動作変化が微小であれば，前フレームのラベリング結果が利用できる．すなわち，現フレームの各ボクセルに対して，nearest-neighbor 法に基づき，前フレームに存在したボクセルの中で最も距離が小さいものを探索し，それと同じラベルとする．

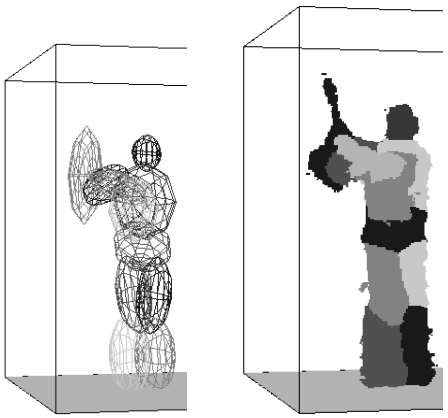
(2) 各体節の解析:
仮ラベリングによって得られた結果から，各体節の空間的分布を解析するために，各体節毎にボクセル座標の主成分分析を行う．これにより，各体節の重心位置・軸の向き・各軸方向の分散を算出する．

(3) 本ラベリング (トップダウン処理):
主成分分析の結果を利用し，各体節とのマハラノビス距離が最も小さくなる体節にラベリングすることで，各体節の大きさを考慮した本ラベリングを行う．ただし，各体節の体積の比率が初期ラベルと等しくなるように拘束条件をつける．仮ラベリングと比較すると，各体節の向きや形状変化を考慮した，より正確なラベリングが可能となる．

4 実験結果

和服を着用した人物の舞踊動作を本システムを用いて撮影し，得られた 3 次元形状に対して各体節への分割を行った．Fig.4(a) に 70 フレーム目の人物の姿勢を示す．この姿勢に対する各体節の解析と本ラベリングの結果を Fig.4(b),(c) に示す．この結果より各体節への分割が比較的安定に行われていることが確認できる．

(a)input image



(b)analysis of parts (c)labeling

Fig. 4 Experimental result.

5 まとめ

本稿では，舞踊動作をデジタルコンテンツ化し映像として再利用するために，多視点カメラを用いて獲得された人体形状を体節ごとに分割する手法について述べた．ボトムアップ処理とトップダウン処理を組み合わせることにより，形状のモデル化が難しい和服の場合にも，比較的安定に分割が可能であることを示した．今後の課題として，体節ごとに分割された結果に基づいた舞踊動作の編集方法の検討などが挙げられる．

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) 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)

2) 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).

3) 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).

4) 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)