

モーションキャプチャと音楽情報を用いた舞踊動作解析手法

白鳥 貴亮[†] 中澤 篤志^{††} 池内 克史[†]

The Structure Analysis of Dance Motions using Motion Capture
and Musical Information

Takaaki SHIRATORI[†], Atsushi NAKAZAWA^{††}, and Katsushi IKEUCHI[†]

あらまし 舞踊等を代表とする無形文化財(人の動き情報)のデジタル保存および動きの解析手法について述べる。日本の伝統舞踊や匠の技といった無形文化財は人から人へ継承することで保存、伝承されてきた。しかし近年では、後継者の減少などに伴い散逸が起りつつあるため、それらをアーカイブ化する技術の開発が急務となっている。そこで、我々はモーションキャプチャシステムを用いて人の動きをデジタル化し、舞踊動作を直接解析することにより有限個の動作プリミティブを抽出することで、この問題に対処することを考えた。動作プリミティブとは動きを構成する最小単位を意味し、本論文では従来の手足の速度や関節角の角速度のみを用いた手法に加え、音楽を考慮した動作解析手法を提案する。これは、舞踊が音楽に合わせて身体を動かすものであるという性質を持つことに由来する。本稿では音響信号からビートトラッキングを行い、そのリズムに基づいて舞踊動作中からキーポーズを抽出し、セグメンテーションによって動作プリミティブを求める手法を説明し、最後に実験結果を示しその有効性について論じる。

キーワード モーションキャプチャ, 動作プリミティブ, 音楽情報, ビートトラッキング

1. ま え が き

人の動きを計測する研究は近年の画像分野のトピックであり、画像による人物のトラッキングや姿勢推定、ジェスチャー認識等で多くの研究が行われている[1]。これらの研究は主に人物位置や姿勢など、人がとりうる数値としての値(metrics)を計測するものであり、近年のアルゴリズムや計算能力の向上により、ほぼ実用的な段階に近づいていると言える。一方で、得られたmetricsとしての人の動きを解析し利用することを想定した研究は未だ多いとは言えず、今後は計測で得られた数値的な動き情報を高度に利用する手法の開発が必要である。

我々の研究グループでは、視覚技術を用いて文化遺産をデジタル保存するプロジェクトを推進しており、レーザセンサを用いた建築物や仏像の三次元形状保存や、カラーカメラ等を用いた表面反射特性や環境光学

情報の記録等を行ってきた[2]。これら「静的」なオブジェクトによる文化遺産と同時に、伝統舞踊や匠の技等の「人の動きの文化財(無形文化財)」もデジタル保存の対象となることが考えられる。特に日本の伝統的な舞踊の保存や伝承は、主に人から人への伝承に頼ってきたのが事実であり、様々な変化や後継者不足による散逸が起こってきた。これを解決するために、人の動きをデジタル技術により保存し、高度利用することを可能にする技術が求められている。本稿では、我々が開発している無形文化財(舞踊)のデジタル保存技術のうち、人の動きをプリミティブとして分割・構造化して記述する手法について述べる。人の動き情報の獲得はモーションキャプチャシステムによって行われ、これにより人の関節位置の動きが得られる。次に得られた人の関節の動き情報を解析し、動作プリミティブを検出する。動作プリミティブとは、人間の動きは有限個の基本動作によって構成される、という仮定に由来するものであり、この基本動作のことを示す[3]。これにより、舞踊等の時系列動作を有限の動作プリミティブセットで記述することが可能になる。具体的には動作構造の把握や動作の検索、異種の舞踊間の比較、舞

[†] 東京大学 生産技術研究所
Institute of Industrial Science, The University of Tokyo

^{††} 大阪大学 サイバーメディアセンター
Cybermedia Center, Osaka University

踊情報の圧縮等が実現できる。また、動作プリミティブは舞踊中で頻出する動作であり、これを連結すれば元の舞踊を構成することが可能となる。

同様の試みは、ロボット、人工知能、コンピュータグラフィックス等の分野でいくつか見られる。ロボットの行動学習を目的とした研究としては、複数の関節角度を Self Motion Element とよばれる単位に分割し HMM を用いて基本動作を抽出した研究 [4]、人の手先の動きを平面に投影し直線・円弧等の基本パターンに分類し再現した研究 [5]、手先の動きを速度に基づいて切り出しその三次元空間中の動きを Dynamic Time Warping による距離を用いてクラスタリングした研究 [6] などが挙げられる。舞踊と対象した研究としては、八村らによるモーションキャプチャデータからの Labanotation の抽出および動作の再生成が挙げられる [7]。この研究では、身体の動き情報を Labanotation と呼ばれる舞踏譜に記述することで表現する。舞踏譜として一般的に用いられている表記方法を用いることで汎用性・有用性は高いと考えられる。しかし、Labanotation はバレエを代表とする西洋舞踊の教育現場での使用を目的としたものであり、人の動きをかなり粗い空間座標系によって記号化するため、舞踊動作の再生成に使うには本質的に適しておらず、再生成された動きも元の動きとはかなり異なったものとなった。本稿ではこういった問題点を解決した舞踊動作の解析手法を提案する。

2. 提案手法の概要

本章では、我々が提案する音楽情報を考慮した舞踊動作解析手法の概要について説明する。

我々の過去の研究 [8] では動きデータから動作プリミティブを抽出することを目的としている。しかしこの手法を舞踊動作に適用した場合、非常に細かい動作プリミティブが多く抽出されてしまい、舞踊動作のアーカイブ化として場合によっては不十分であると考えられる。またこれは両手両足の速度のみを解析対象としているため、舞踊における重要な要素のひとつであるリズムが考慮されておらず、音楽と人の動きによる複合的な活動である舞踊の解析手法としては、不完全であると考えられる。

動きのリズムを考慮した動作解析手法として、Kim らは動きデータから関節角速度を算出し、その時系列データから周波数解析などを通して動きのリズムを推定する手法を提案している [9]。この手法ではすべての

動作プリミティブの時間が一定であるという強い拘束を与えているが、実際には多くの踊りのプリミティブの時間は様々であり、その点をきちんと扱えることは不可欠である。

これらの問題を解決するために、本研究では舞踊動作における動作プリミティブとして、以下のような条件を仮定した。

仮定 1 動作プリミティブと動作プリミティブの間には、とめ動作 (キーポーズ) と呼ばれる全身がほぼ静止した状態が存在する

仮定 2 キーポーズは音楽のリズムに応じたリズムパターンに基づいて現れる

仮定 1 は、Flash らが提案する動作プリミティブの定義に由来するものであり [3]、さまざまな研究でその有効性が示されている。実際に舞踊動作においても、踊り手がとめ動作は重要であると聞き取り調査によって明らかになっており、本研究においても有効であると判断した。また仮定 2 は「人は音楽に合わせて舞踊を披露する」という知見から由来するものであり、音楽情報を利用する有効性はこの仮定 2 に由来するものである。これらの仮定を考慮してキーポーズを抽出し、抽出したキーポーズが現れる瞬間で舞踊動作をセグメンテーションすることによって動作プリミティブを抽出するアルゴリズムの概略を図 1 に示す。

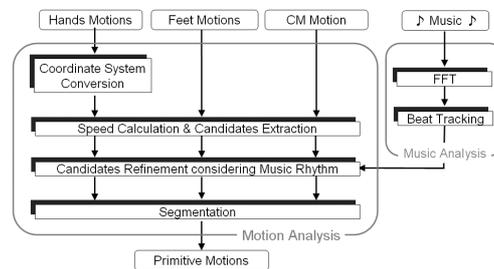


図 1 提案手法の概略
Fig.1 Algorithm overview

以降では、まず 3 章で音楽情報からのリズム推定について述べる。次に第 4 章では動き情報からセグメンテーションの候補点を求め、推定した音楽のリズムを元に動作プリミティブを抽出する。そして 5 章で実験結果を提示して本手法の有効性を示し、6 章で本稿のまとめを述べる。

3. 音楽情報のリズム解析

本章では、舞踊動作の解析に必要なリズムを音

楽情報から推定する手法について述べる．本研究では後藤の提案するビートトラッキング手法 [10] を改良し，日本の舞踊音楽への適用を可能にした．

3.1 リズムを構成する音楽的要素

後藤によると，西洋音楽の楽典に基づいた楽曲では「発音時刻」，「コード（調）の変化」，「ドラムの発音パターン」がリズムを構成している．そこで，まずこれらのリズムに関する音楽的要素が日本の舞踊音楽に適用できるかどうかについて考察する．

発音時刻は東洋音楽，西洋音楽といったジャンルに関係なくリズムのタイミングに合わせて楽器を発音させる場面が非常に多いことから，日本の舞踊音楽に対しても有効であると考えられる．コードの変化に関しては，従来手法では西洋音楽のコードに関する周波数特性を利用している．しかし，東洋音楽では西洋音楽とは異なるコードの周波数特性が異なるため適切ではないと判断し，本手法ではコード変化を利用しなかった．最後に打楽器の発音パターンであるが，従来手法ではポピュラー音楽でよく用いられるスネアドラムとバスドラムの周波数成分をテンプレートとして保持し，そのマッチングによって打楽器の発音を検出している．しかしこれらの打楽器は日本の舞踊音楽においては使用されないため，この要素も適切ではないと判断した．

以上から，本手法におけるビートトラッキングは発音時刻に基づく手法で行った．

3.2 発音時刻の検出

ある音が発せられたとき，その音に対応した周波数のスペクトルパワーが増加する．そのスペクトルパワーの増分を求めることによって発音時刻の検出を行う．ある時刻 t ，周波数 f において，スペクトルパワーの増加分を算出する関数 $d(t, f)$ を式 1 として定義する．

$$d(t, f) = \begin{cases} \max(p(t, f), p(t+1, f)) - p' \\ (\min(p(t, f), p(t+1, f)) \geq p') \\ 0 \quad (\text{otherwise}) \end{cases} \quad (1)$$

$$p' \equiv \max(p(t-1, f), p(t-1, f \pm 1)) \quad (2)$$

ここで $p(t, f)$ は時刻 t ，周波数 f におけるスペクトルパワーを表す．図 2 に $d(t, f)$ 算出における概要図を示す．

式 1 では時刻 t と時刻 $t+1$ における 2 つのスペクトルパワーを考慮したものになっている．これは発生した音が時間軸方向に連続的に続くという特性を考慮

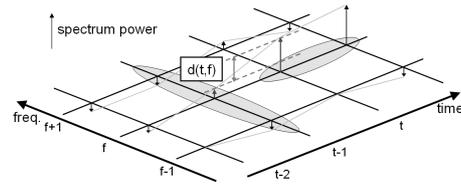


図 2 発音時刻の検出
Fig.2 Onset component extraction

したためである．また，式 2 の算出には対象となっている周波数の上下を考慮している．これはある音が発せられている間，周波数が微小ながらも変動してしまうことを考慮していることからくるものである．

次に，周波数領域全体に対して式 1 を計算し， $D(t) = \sum_f d(t, f)$ を算出する．これにより，時刻 t に発せられた音の強さを把握することができる．

3.3 リズムの開始時刻と平均間隔の算出

3.2 で算出した $D(t)$ からリズムの開始時刻とリズムの平均間隔を推定する．

まずリズムの平均間隔を推定するにあたって， $D(t)$ の自己相関関数 R_{DD} を算出する (式 3)．

$$R_{DD}(\tau) = \frac{1}{T} \sum_{t=1}^T D(t)D(t+\tau) \quad (3)$$

自己相関関数 $R_{DD}(\tau)$ は， $D(t)$ と τ 遅れた $D(t+\tau)$ の 2 つの時系列波形がどの程度類似しているか，すなわち $D(t)$ の周期性を示すものであり， $D(t)$ が最大値となる τ が平均リズム間隔 (t_{rhythm}) となる．

こうして推定された平均リズム間隔からパルス列 $P(t)$ を作成し， $D(t)$ と $P(t)$ との相互相関関数 R_{DP} を算出する (式 4)．

$$R_{DP}(\tau) = \frac{1}{T} \sum_{t=1}^T D(t)P(t+\tau) \quad (4)$$

相互相関関数は $D(t)$ と $P(t)$ の類似性を示す．そのため，相互相関関数の最大値が現れる時刻を t_{max} とすると， $D(t)$ と $P(t+t_{\text{max}})$ が類似している，すなわち t_{max} がリズム開始時刻 (t_{st}) であることを示している．

3.4 ビートトラッキング

3.3 ではリズムの開始時刻と平均リズム間隔を推定した．しかし実際にはリズム間隔にわずかな誤差があるため，その誤差が累積することによって，ビート時刻のずれが生じる可能性がある．この問題に対処するために，3.3 で求めた各リズム時刻の近傍で， $D(t)$

が最大となる時刻を新たなリズム時刻として更新する処理を適用した(図3)。これによりビートの間隔がわずかながら変化するような状況においても、従来の手法より精度よくリズム時刻を求めることが可能となった。

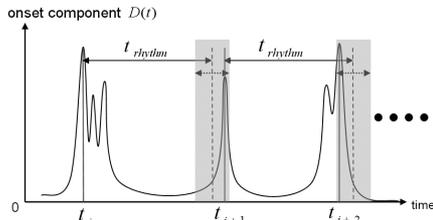


図3 リズムのひずみに対応したビートトラッキング
Fig.3 Beat tracking considering slight changes of rhythm

4. 舞踊動作の構造解析

提案する手法で用いる動きの要素は、両手、両足、重心のそれぞれの速さである。ほとんどの踊りにおいて手や足は主要な表現部位であり、様々な振り付けが考えられている。日本の伝統的な舞踊においても例外ではなく、手や足を静止させることで「とめ」を表現しており、舞踊中のキーポーズとなっている。そのため、手や足の速さはプリミティブの境界であるため動作を検出するのに非常に有効だと考えられる。

しかしながら、踊り手がリズムにのりおくれたりする場合や個人差による違いなどから、手や足だけではとめ動作を検出することが難しいことがある。そこで、本手法では重心の速さも考慮する。重心は、身体全体の大まかな動きを表しているため、この問題を解決する手がかりになると考えられる。すなわち、本手法では手足の動きだけでなく、身体全体の動き(重心の動き)もとめ動作の解析に用い、それらの候補点の中からリズムによる判定を入れることで、より正確なキーポーズの抽出が可能になった。

4.1 動きデータの抽出

手の動きの解析においては、各フレームごとに体中心座標系と呼ばれる座標系に変換し、この座標系における速度を求めて解析を行なった。体中心座標系とは、人間が直立している状態で、腰の位置を原点とし、右肩から左肩にかけての方向を x 軸、人間の正面の方向を y 軸、 x 軸と y 軸に垂直で腰から頭にかけての方向を z 軸と定義するものである(図4)。

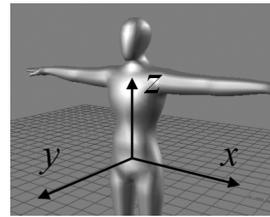


図4 体中心座標系
Fig.4 Body center coordinate system

一方、足および重心の解析においては、データが記録されているままの値(世界座標系)を用いた。これは足の場合は地面に接地していることが多く、接地している間は世界座標系における速度がほぼ0になっている、すなわち足が動いているときと動いていないときの認識が行いやすいという利点によるものである。また重心の場合も同じように、とめ動作のときは身体全体が静止している、すなわち世界座標系における重心の速度がほぼ0になっており、身体が動いているときと動いていないときの認識が行いやすいという利点から、世界座標系の値を用いて重心の速度算出を行った。

4.2 セグメンテーション候補点の算出

次に4.1で算出した体の各部位の速さを用いてセグメンテーションの候補点を求める。セグメンテーションの候補点の算出方法は手・重心と足とで異なる。まず最初に手・重心の場合のセグメンテーション候補点の算出方法について述べる。

舞踊動作の中で、手や重心の速さはとめ動作において瞬間的に0になるが、一方で速さ0の状態が継続することはほとんどない。この特徴の利用することで手・重心の速さのシーケンスからセグメンテーションの候補点を抽出する。まず速さの極小点を求め、かつその極小点がある閾値以下となる点を求める。この条件を満たす極小点の間での速さの最大値がある一定の閾値以上になっている場合、その極小点をセグメンテーションの候補点とする(図5)。

一方足に関しては、4.1で述べたように接地している間は速さが0となる。この特徴を生かし、まず足が動いているときの移動距離を求める。この移動距離に対してノイズを除去するためにしきい値処理を行い、しきい値を満たす場合に足の動き始めの時刻と静止した時刻を候補点とする(図6)。

4.3 音楽のリズムを用いたセグメンテーション候補点の絞込み

次に3章で推定した音楽のリズムを用いて、身体各

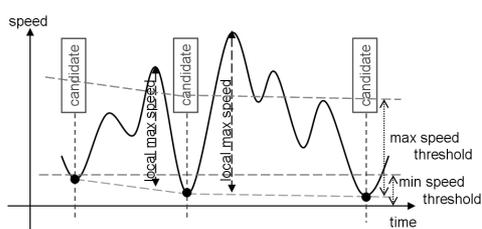


図 5 手・重心のセグメンテーション候補点の算出方法
Fig. 5 Segmentation candidates extraction - Hands & CM

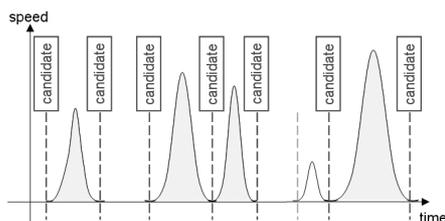


図 6 足のセグメンテーション候補点の算出方法
Fig. 6 Segmentation candidates extraction - Feet

部位のセグメンテーション候補点の絞込みを行う。その方法は、音楽のリズムの前後一定時間内 ($t_{\text{rhythm}}/10$) にセグメンテーション候補点が存在するかを認識し、存在する場合はそのリズム時刻をセグメンテーション候補点とし、存在しない場合は候補から外す処理を行う。

図 7 に身体の各部位のセグメンテーションの例を示す。1 番目、および 3 番目のリズムの前後にはセグメンテーションの候補点が存在しないため、この時刻ではセグメンテーションを行わない。一方 2 番目、および 4 番目のリズムの前後にはセグメンテーションの候補点が存在している。よってこのリズム時刻が新たなセグメンテーション候補点となる。

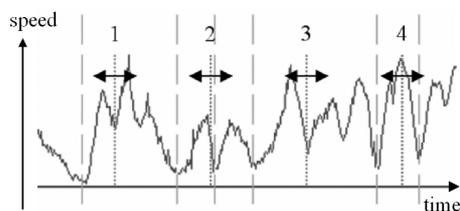


図 7 身体の各部位のセグメンテーション候補点の絞込み
破線：セグメンテーション候補点
点線：音楽のリズム

Fig. 7 Candidates refinement with musical rhythm
Broken line : segmentation candidate
Dotted line : estimated music rhythm

4.4 動作プリミティブの抽出

4.3 節の結果を踏まえて、セグメンテーションを行い動作プリミティブを抽出する。ここでのセグメンテーションは、4.3 節の結果を以下に示す 2 つの条件を同時に満たす瞬間をキーポーズとし、動作シーケンスを分割する。

条件 1 左手, 右手, 両足の候補点のうち 2 つ以上が存在する

条件 2 重心における候補点が存在する

条件 1 では「左手, 右手, 両足の結果のうち 2 つ以上の候補点が存在」としており, 両手両足がほぼ静止しているところ, すなわちとめ動作である可能性が高い瞬間を抽出することができる。これは, 両手両足の静止している時刻がずれていることがあり, 従来の「両手両足が静止している状態」を抽出する手法ではキーポーズと認識することが難しいような場合にも対処できるようにしたためである。しかしこの条件のみではとめ動作ではないところも抽出してしまうことがある。たとえば歩行動作に見られるように, 両足が接地している瞬間では両手が前, または後ろに伸びていて静止しているが, 胴体は前進している, というような状態がこれにあたる。舞踊動作においてもこのような状態が存在しており, 両手両足が静止しているかどうかを認識するだけではとめ動作であるかどうかを認識できない。そこで重心の動きを考慮する (条件 2) ことによって身体全体の動きを把握し, この問題を解決する。

5. 実験

本章では, 本稿で提案した手法を有効性を示すために行った実験の詳細とその結果, 考察について述べる。

5.1 実験データ

今回実験の対象としたデータは女性および男性による会津磐梯山踊りと女性によるじょんがら節の 3 つである。会津磐梯山踊りの特徴は手足をゆったりと動かす穏やかな舞踊であるのに対し, じょんがら節の特徴はメリハリのある素早い舞踊である。

音楽データは, 動きデータの計測時に使われたテープから USB オーディオデバイスを経由して PC に取り込み, wav ファイル形式に変換 (32000Hz, 16bit/sample) したものをを用いた。

また動きデータは Vicon Motion Systems(光学式モーションキャプチャシステム) を用いて取得した。会津磐梯山踊り, じょんがら節ともに 2 回計測を行い, 2

回目のデータを実験に用いた．取得したデータのうち，はじめの 20 秒間（会津磐梯山踊りのデータは 120fps で記録されているため 2400 フレーム，じょんがら節のデータは 200fps で記録されているため，4000 フレーム）に提案手法を適用し，その結果を比較した．

また本稿では提案手法の正当性を示すために，舞踊の師範が描いたマニュアル上のキーポーズを基に比較検討を行った．このマニュアルは師範による舞踊の認識が描かれているものであり，すなわちほぼすべての踊り手に共通な認識であると見なすことができる．本論文ではここで得られるキーポーズを真値として採用する．

これによると，はじめの 20 秒間で現れるキーポーズは，会津磐梯山踊りが 9 箇所，じょんがら節が 12 箇所であった（図 9）．

5.2 ビートトラックの結果

音楽データを周波数領域に変換する際の FFT のパラメータは，窓関数の大きさが 1024 サンプル，窓関数のシフト数が 256 サンプルであった．また，窓関数には音楽情報処理で一般的に用いられるガウス関数を用いた．

推定された平均リズム間隔は，会津磐梯山踊りが 0.704 秒 85M.M.，じょんがら節が 0.576 秒 104M.M. であった（M.M. は Mälzel's Metronome の略で，1 分間あたりの拍数を示す）．またそれぞれのビートトラックの結果を図 8(a) および図 8(b) に示す．音が発音された瞬間とリズムを示す線がほぼ一致していることから，提案したビートトラック手法が精度よくリズムを推定できていることがわかる．

5.3 動作プリミティブの抽出結果

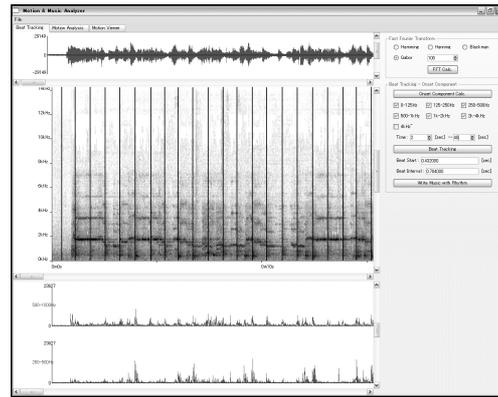
本手法の有効性を示すために，提案手法による抽出結果と動きデータのみを対象とした従来の解析手法 [8] の結果と，マニュアルでのキーポーズとの照合により認識率および誤検出数を算出し，比較を行った．

5.3.1 会津磐梯山踊りにおける結果

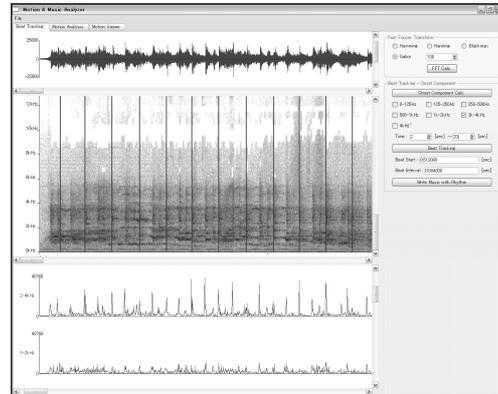
提案手法による会津磐梯山踊りからのキーポーズ抽出結果を図 10 に示す．また検出率等の比較を表 1, 2 に示す．

会津磐梯山踊りでは，基本的な振り付けは同じであるが，男性と女性とでは細部でわずかな違いがあった．

女性によって実演された会津磐梯山踊りは，男性の踊り方に比べて緩やかに手を動かすという特徴を持っていた．その特徴からリズムにしっかりと合わせて手足を動かさず場面が多かったため，従来手法では過検出



(a) Aizu-bandaisan dance



(b) Jongara-bushi dance

図 8 ビートトラックの結果
上段：音楽の振幅，中段：スペクトラム，下段：3.2
で求めた $D(t)$
スペクトラム中の直線は推定されたリズム
Fig. 8 Results of beat tracking
Top : amplitude, Middle : spectrum,
Bottom : $D(t)$ as described in 3.2
Straight line in spectrum : music rhythm

エラーが多い結果となった．具体的には，4.4 節に挙げたような，両足が接地し，両手はそのタイミングで止まっているが胴体が動いているような状態をとめ動作として抽出していることが非常に多かった．一方提案手法では，重心の動きを考慮していたためにこういった状態をエラーだと認識することができ，ほぼマニュアルに近いキーポーズ抽出が可能になった．

また，男性によって実演された会津磐梯山踊りは，女性の踊り方に比べて手を広い範囲にわたって動かすという特徴を持っていた．そのため，踊り手にとってはとめ動作と認識している姿勢であっても，手の速度が十分に遅くなっていない場面がしばしば見受けられ，その結果従来手法では非常に多くの未検出エラーが生

表 1 女性の踊り手による会津磐梯山踊りでの結果
Table 1 Result of Aizu-bandaisan dance - a Female Dancer

	総検出数	検出率 (正解数/真値)	過検出数
従来手法	13	89% (8/9)	5
提案手法	9	100% (9/9)	0

表 2 男性の踊り手による会津磐梯山踊りでの結果
Table 2 Result of Aizu-bandaisan dance - a Male Dancer

	総検出数	検出率 (正解数/真値)	過検出数
従来手法	3	33% (3/9)	2
提案手法	9	100% (9/9)	0

じた。一方提案手法では、4.4 に述べた条件 1 によりキーポーズ候補が多く残されているため、結果的にはすべてのキーポーズを抽出することが可能であった。

以上から、従来手法では動作プリミティブ抽出の精度が個人差によって大きく影響を受けるのに対し、本手法は個人差などに影響を受けずに、高い精度でキーポーズの抽出、動作プリミティブの抽出が可能であると考えられる。

5.3.2 じょんがら節における結果

提案手法によるじょんがら節からのキーポーズ抽出結果を図 11 に示す。キーポーズの検出率等の比較を表 3 に示す。

表 3 じょんがら節での結果
Table 3 Result of Jongara-bushi dance

	総検出数	検出率 (正解数/真値)	過検出数
従来手法	6	50% (6/12)	0
提案手法	9	75% (9/12)	0

じょんがら節は手足を非常に素早く動かす、という特徴があるため、手足の静止時刻と音楽のリズムのタイミングが微小ながらもずれている場面が多かった。そのため従来手法では、両手両足が同時刻に静止していると認識できず、未検出エラーが多い結果となった。一方で提案手法では、4.4 で示したように、条件 1 をゆるく設定したことによって、一部未検出エラーを克服することができたが、依然 3 点の未検出エラーが生じた。これらのエラーは、じょんがら節が非常に速い舞踊であり、とめ動作となるところで踊り手が手の速度を十分に落とすことができていないために候補点として抽出できず、生じたものである。

6. おわりに

本研究では舞踊は音楽と密接な関係があることに

注目し、音楽情報からリズムを推定し、その結果を考慮して動作解析を行ない、キーポーズを抽出、セグメンテーションにより舞踊動作を構成している動作プリミティブを抽出する手法について提案した。具体的には、微小なリズムの揺らぎに対してロバストなビートトラッキングの手法を提案し、実験を行なってその精度の良さを示した。そして、舞踊動作を構成する動作プリミティブを抽出する手法について提案した。この手法では、個人差が顕著に現れたり、人間の動きがリズムに合わない場面も多く存在することを考慮して、両手両足だけでなく重心の速度も考慮に入れることによって精度の向上を図り、実験によってその成果を示した。本手法によって抽出されたキーポーズおよび動作プリミティブの結果は、実際に踊り手が意識しているキーポーズとほぼ一致しており、新たな舞踊の記述方法への手がかりになると言える。また、今回実験対象とした会津磐梯山踊りとじょんがら節は異なる特徴を持つ舞踊であり、実験によりこれらの舞踊に対する有効性を示したことから、他の舞踊にも対応できる汎用的な手法だと考えられる。

しかし、提案したビートトラッキング手法はリズムがほぼ一定でない楽曲に対して対応できていない。そのためリズムが変わる楽曲に対応できるようにビートトラッキング手法を拡張することにより、さらに多くの種類の舞踊に本手法を適用することが可能になると考えられ、今後の課題とする。

さらに本研究の発展として、ロボットが音楽を聴いて舞踊を踊る、すなわち「サウンドフィードバック制御」が可能になると考えられる。流れてくる音楽からリアルタイムでリズムを認識し、そのリズムに応じて可能な限り人間に近いロボットの動きを生成し、ロボットに踊らせることでサウンドフィードバック機能を実現していきたい。

謝辞 本研究の一部は、科学技術振興機構戦略的基礎研究事業 (CREST) 高度メディア社会の生活情報技術、又他の一部は、科学研究補助金 (特定領域研究) 「人間行動の観察とロボットのタスク・スキル獲得に基づく人間の作業熟練過程の解明」の支援を受けて行われました。

文 献

- [1] D. M. Gavira: "The visual analysis of human movement", *Journal of Computer Vision and Image Understanding*, **73**, 1, pp. 82-98.
- [2] D. Miyazaki, T. Oishi, T. Nishikawa, R. Sagawa, K. Nishino, T. Tomomatsu, Y. Takase and K. Ikeuchi:



図 9 マニュアルから得られたキーポーズ
Fig.9 Keyposes extracted from manuals

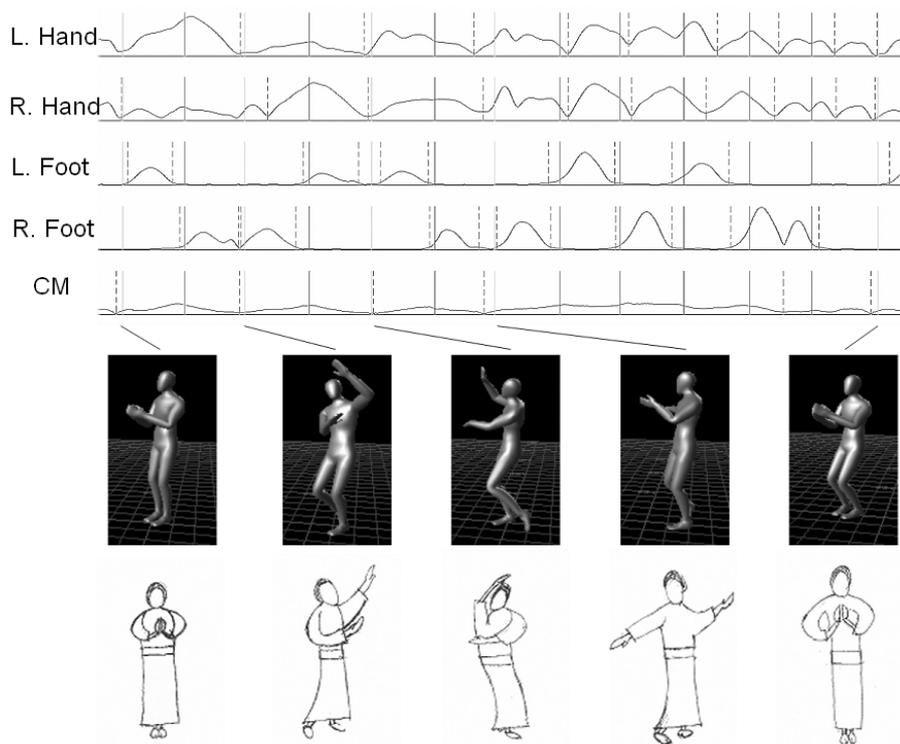


図 10 セグメンテーションの結果 - 女性による会津磐梯山踊り
上から速度グラフ, 抽出されたキーポーズ, 踊り手が認識しているキーポーズ
破線はセグメンテーション候補点, 直線は音楽のリズム

Fig.10 Results of key poses extraction - *Aizu-bandaisan* of a female dancer
Top : speed sequence, Middle : key poses extracted by our method,
Bottom : key poses extracted by dancers
Broken line : segmentation candidate, Straight line : music rhythm

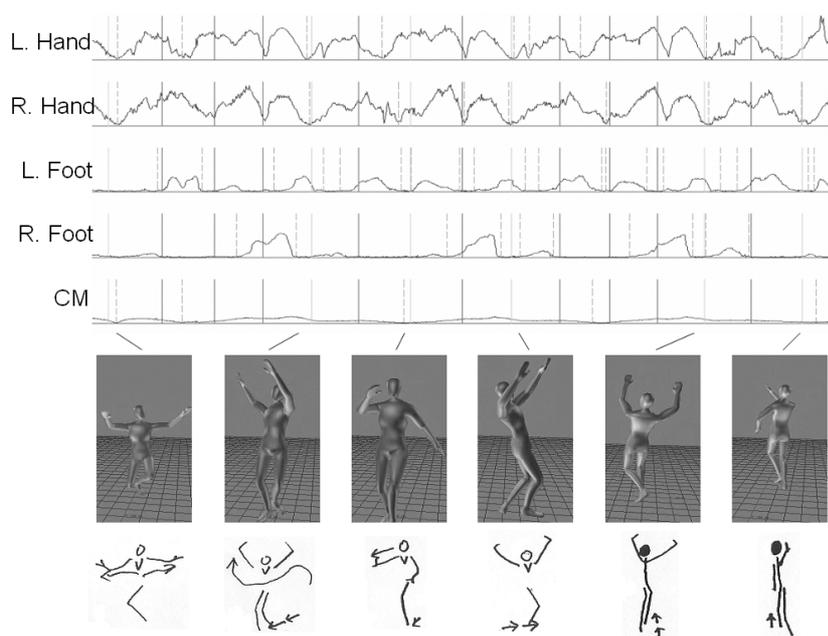


図 11 セグメンテーションの結果 - じょんがら節
上から速度グラフ, 抽出されたキーポーズ, 踊り手が認識しているキーポーズ
破線はセグメンテーション候補点, 直線は音楽のリズム

Fig. 11 Results of key poses extraction - *Jongara-bushi* dance
Top : speed sequence, Middle : key poses extracted by our method,
Bottom : key poses extracted by dancers
Broken line : segmentation candidate, Straight line : music rhythm

“Great buddha project: Modeling cultural heritage through observation”, In Proc. of the Sixth International Conference on Virtual Systems and MultiMedia (VSMM 2000), pp. 138-145 (2000).

[3] T. Flash and H. Hogan: “The coordination of arm movements”, *J. Neuroscience*, pp. 1688-1703 (1985).

[4] T. Inamura, Y. Nakamura, H. Ezaki and I. Toshima: “Imitation and primitive symbol acquisition of humanoids by the integrated mimesis loop”, In Proc. of International Conference on Robotics and Automation (ICRA2001), pp. 4208-4213 (2001).

[5] O. C. Jenkins, M. J. Metaric and S. Weber: “Primitive based movement classification for humanoid imitation”, In Proc. of First IEEE-RAS International Conference on Humanoid Robotics (Humanoids 2000) (2000).

[6] 大崎, 嶋田, 上原: “速度に基づく切り出しとクラスタリングによる基本動作の抽出”, *人工知能学会誌*, 15 巻 8 号, pp. 878-885 (2000).

[7] 吉田, 松岡, 八村: “舞踊譜 labanotation に基づく身体運動の処理-舞踊譜読取り labanreader と譜面エディタ labaneditor-”, *情報処理学会人文科学とコンピュータ研究報告*, 38-6, pp. 61-68 (1998).

[8] A. Nakazawa, S. Nakaoka, K. Ikeuchi and K. Yokoi: “Imitating human dance motions through motion

structure analysis”, In Proc. of International Conference on Intelligent Robots and Systems(IROS 2002), pp. 2539-2544 (2002).

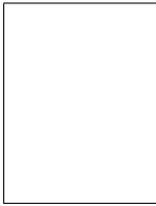
[9] T. Kim, S. I. Park and S. Y. Shin: “Rhythmic-motion synthesis based on motion-beat analysis”, In Proc. of ACM SIGGRAPH 2003 (2003).

[10] M. Goto: “An audio-based real-time beat tracking system for music with or without drum-sounds”, *Journal of New Music Research*, 30, 2, pp. 159-171 (2001).

(平成 xx 年 xx 月 xx 日受付)

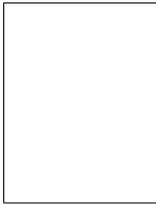
白鳥 貴亮

東京大学大学院情報理工学系研究科博士課程在学中。2002 年東京大学工学部電子情報工学科卒業, 2004 年東京大学大学院情報理工学系研究科電子情報学専攻修士課程修了。動作解析の研究に従事。



中澤 篤志

1997 年大阪大学基礎工学部システム工学科卒業，2001 年同大学院基礎工学研究科博士課程修了，博士（工学），同年科学技術振興事業団研究員（東京大学生産技術研究所）．2003 年大阪大学サイバーメディアセンター講師．現在に至る．画像計測，分散視覚システム，動作解析の研究に従事．



池内 克史（正員）

1973 年京都大学工学部機械工学科卒業．1978 年東京大学大学院工学系研究科情報工学専攻博士課程修了．MIT 人工知能研究所，電総研，CMU 計算機科学科を経て，1996 年より東京大学生産技術研究所教授．2000 年より東京大学大学院情報学環教授兼担．人間の視覚機能，明るさ解析，物体認識，人間行動観察学習ロボット，高度交通システムなどの研究に従事．工学博士．D.Marr 賞 (ICCV:1990 年)，IEEE 優秀論文賞 (CVPR:1991 年)，最多引用論文賞 (AI Journal:1992 年)，Fu 記念優秀論文賞 (IEEE Trans. R&A, 1998 年) 等受賞．IEEE Distinguished Lecturer (SPS 2000 - 2001, CS 2004 - 2006)，IEEE Fellow ．

Abstract This paper describes a dance motion analysis method considering music rhythm. In these days, many important intangible cultural properties are being lost because of the lack of successive performers, and we have started our digital archiving project of cultural intangible heritages. Our goal for archiving dance motions is to detect “primitive motions”, which consists of dance motions. In addition, we believe these primitives must be synchronized to the musical rhythm, so our proposed method automatically segments the original motion considering with estimated musical rhythm to detect the primitive motions.

Key words Motion Capture, Primitive Motions, Music, Beat Tracking