

複数舞踊動作からの個性の抽出および適用

中澤 篤志 中岡 慎一郎 池内 克史^{*,††}

本稿では、伝統舞踊に代表される人の動きをデジタル保存する研究の一環として、舞踊における「基本動作」と「個人の個性（動作スタイル）」を自動的に抽出することを目標としている。従来の研究例では、モーションキャプチャ等によって得られた動作データを何らかの基準でセグメンテーションし、そのデータをそのまま用いてアニメーションを生成したり、動作解析に用いるものが多かった。これに対し我々は、舞踊における動作は個々人に共通する基本動作と、個人の個性を示す動作スタイルに分離できると考え、アルゴリズムを設計した。たここで得られた解析結果を用いて、ユーザが設定したパラメータによって複数動作を合成し、新たな動作を合成することが可能になった。実験では、同一舞踊を男女で演じたデータを用い、本手法を適用することによって、その有効性を確認した。キーワード：無形文化財、舞踊解析、人物動作、モーションキャプチャ、基本動作、動作スタイル、アニメーション

Retrieving and applying human dance motion styles from several motion data

ATSUSHI NAKAZAWA, SHINICHIRO NAKAOKA, KATSUSHI IKEUCHI

This paper presents the method to analyze human motion for the purpose of digital archive of intangible cultural heritages, such as folk dances. In the recent studies, the whole motion sequence is segmented and directly used for computer animation or motion analysis. We propose the idea that the human dance motion consists of 'Basic Motion' and 'Motion Styles'. The Basic Motion is common motion for any dancers, and Motion Styles represents the uniqueness of the individual dancers. In the experiment, we confirmed proposed method works effectively through different motion data of male and female dancers. Keywords: folk dance, human motion, motion analysis, motion capture data, basic motion, motion style, animation

1. はじめに

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。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²⁾、デジタル保存された動き情報の獲得、解析、編集、提示に関する技術を開発してきた³⁾。ここでは、人の動きをモーションキャプチャを用いて入力し、得られた人の動き情報を解析することで基本動作(動作プリミティブ)に分割することにより、舞踊動作を有限のプリミティブセットで記述することが可能になった。また、これを連結することにより元の舞踊を構成したり、新たな動作を生成することが可能になった。また、ここで得られた動き情報は、ヒューマノイドロボットに入力されて動作、提示される。

ここで提案した手法では、与えられた動きを目標点の速度変化により分割・合成することを行ったが、いくつかの問

題点がある。すなわち、

- (1) セグメンテーションが不安定である。
- (2) 入力されるデータは特定の個人の動作であり、必ずしもその舞踊における一般的な動きであるかはわからない。
- (3) 2と同様に手、足の動きの時間的関連性に関する考慮が十分ではない。

といった点が上げられる。特に2および3の問題は、ある舞踊の基本構造を理解するためには重要な問題である。そこで本稿では、同一の舞踊を複数の個人で演じたデータを用いて時間的なセグメンテーションを行い、その時間的配置の同一性と動きの同一性を抽出することで舞踊の基本動作を抽出することを考える。また、入力されたデータと基本動作を取り除くことで、個人特有の動きを取り出すことが可能になり、任意の動きに特定の個人の特性を付加することができる。

2. 従来手法とその問題点

本章ではまず、我々が提案した動きの解析手法について述べる。次に、この手法の問題点について考察する。

2.1 従来手法

従来手法においては、以下のような解析アルゴリズムで、舞踊動作における各部位の動きがセグメント化され、有限個のラベル列として記述される(図1)。

- (1) 体中心座標系の設定

モーションキャプチャデータから得られたデータのうち、腰の方向をX軸としZ軸を鉛直上向きとした座標系を「体中心座標系」とする。

- (2) 目標点の座標変換

* 大阪大学サイバーメディアセンター

†† 東京大学生産技術研究所

Institute of Industrial Science, University of Tokyo

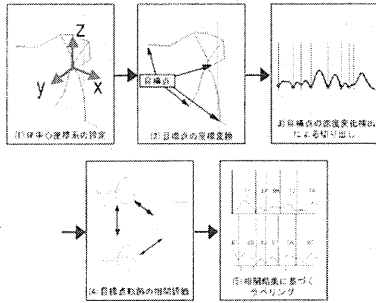


図 1 従来の動作解析アルゴリズム

目標点(両手足先)を体中心座標系に変換する。

(3) 初期セグメンテーション

体中心座標系に変換された各目標点の速度を計算し、加速度0の点をセグメント分割点とする。

(4) セグメントの相関評価

DP マッチングによりセグメント内の目標点軌跡の相関を評価する。

(5) セグメントのクラスタリング

両手間の全セグメント、両足間の全セグメントに対して相関を評価し、Nearest Neighbor アルゴリズムでクラスタリング、ラベル付けする。これにより、目標点が同様の軌跡を描くものに共通のラベルがつけられる。

2.2 従来手法の問題点

以上のアルゴリズムにより、舞踊における各目標点の動きを有限個の動作セグメントで記述することができる。ここから、舞踊の構造が把握でき、動作セグメントを連結することで動作編集等が可能になった。一方で、前述したように本手法には以下のような問題点が存在する。

(1) セグメンテーションが不安定である

動作のセグメンテーションは目標点の速度の極小値をとることによって行われるが、この検出方法は非常に不安定であり、特に局所的な速度変化に対して過剰にセグメント点検出されるという問題点がある。これに対していくつかの提案がなされているが⁽⁴⁾⁽⁵⁾、仮定に基づくヒューリスティクスが多く汎用性があるとは言いがたい。特に、複数の人物が演じた舞踊のセグメント対応を取る場合、セグメンテーションエラーは大きな問題である。

(2) 使用データは特定の個人の動作であり、舞踊における一般的な動きとはいえない

従来手法では、特定の個人の動作を分割し「基本動作」としたが、舞踊動作が個人間においても大きな差があることを考えると、この仮定は正しくないと考えられる。複数の舞踊動作を観察し共通の要素を抽出する新たな手法が必要である。

(3) 手、足の動きの時間的相関性に関する考慮が十分ではない。

提案手法では、異なる部位の動作セグメント間の時間的共起確率を元に、動作を区切る手法を提案していた。しかし、このタイミングに関しても(2)と同様、個人差に依存す

るものと考えられ、何らかのモデル化を行うべきである。

以上のような点を踏まえ従来手法の問題点を解決し、同時に複数の動作データから共通の要素・相違のある要素を抽出することで、動作プリミティブや動作スタイルが検出できるものと考えられる。

3. 複数人物動作の同時解析

本章ではまず、舞踊動作における基本動作と個性(スタイル)について述べ、その後スタイルはどのような要素から構成されているかを考察する。

3.1 舞踊動作における基本動作とスタイル

我々の従来の研究から、舞踊動作は体の各部位の動きセグメントより構成されると考えることができる。一方で、舞踊動作を観察すると、同一の舞踊であっても個人間で大きな相違があることがわかる。例(図2)は「会津磐台山」の男女の舞踊演者における動作列であるが、同じ時刻の姿勢であっても、その間にはかなりの相違がある。また同時に、目標点(手、足)の「止め」が入るタイミングにも時間的相違があり、これらの点も舞踊の個性を表現していると考えられる。一方で、舞踊全体における体の動きとしては共通要素があり、これを基本動作と考えることができる。「止め」が前述した動作セグメントの境界フレームであると考え、これらの観察により、舞踊の個性を表す要素としては、(1)止め時刻(セグメントフレーム)の相違、(2)止め姿勢(セグメント境界での姿勢)の相違、(3)セグメント内の動作の相違の3点を挙げる事ができる。模式的に表すと、以下のように記述できる。

$$\begin{aligned} [HumanMotionSequence] &= [MotionSegment\#1] \\ &\quad + [MotionSegment\#2] + \dots \\ [MotionSegment] &= [MotionBase] \oplus [MotionStyle] \end{aligned}$$

ここで、セグメントフレームは、 $MotionSegment$ の長さによって影響を受け、 $MotionBase$ は、 $MotionSegment$ の終端条件から導き出される基本動作であり、終端条件が定まれば唯一求まる動作である。 $MotionStyle$ はセグメント内の動作の相違を意味し、人物間の微妙な体の動きの違いを表す。

4. 提案アルゴリズム

上記の考え方に基づき、複数の舞踊動作から動きの個性を抽出アルゴリズム、およびその動きを適用する技術を設定する。動作解析アルゴリズムでは、(同一舞踊の)複数動作データ間のマッチングをとり、セグメンテーションを行う。ここでは、EM (Expectation Maximization) アルゴリズムを用いて、従来よりロバストなセグメンテーションを提案する。次にセグメンテーションされた動作列から基本動作と動作スタイルを抽出する。動作生成では、対応が取られた動作セットから、任意の合成パラメータにより、互いの動きを合成するアルゴリズムを提案する。ここでは、間接位置や角度等の単純なブレンドではなく、上記概念に基づいた時空間的なスタイルの合成を行う。

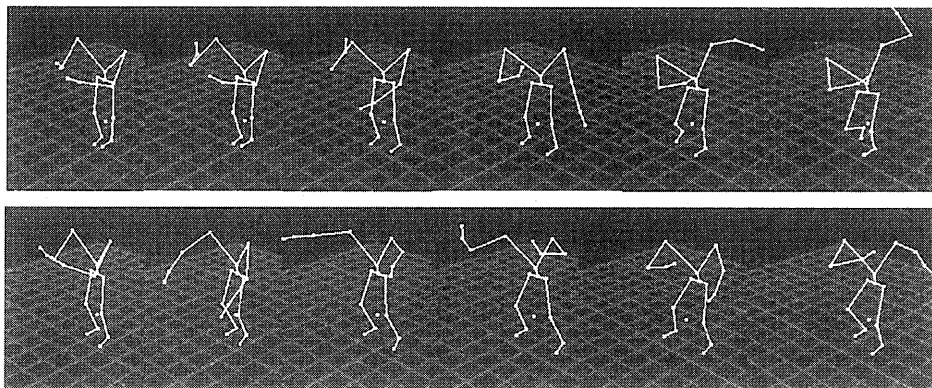


図2 「会津磐台山」における同一個所での動作の違いの例（上：女性の動作，下：男性の動作）

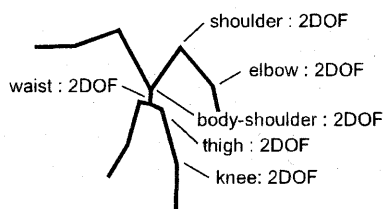


図3 体座標系モデル

4.1 体座標系の表現

ここで動作解析・生成を行う体座標系を設定する。全身は図3のように、17セグメントからなるリンク構造で表現される。主となる座標系（体中心座標系） q_0 は腰位置に設定され、各リンクの位置関係は、接続している親リンクとの相対座標系によって規定される。この位置関係は四次数パラメータ $q_n = (x_n, y_n, z_n, w_n)$, $1 \leq n \leq 17$ で表される。また、腰位置の世界座標系での位置を $p_0 = (x, y, z)$ とすると、全身姿勢 Pose は、

$$Pose = (p_0, q_0, q_1, q_2, \dots, q_{17}) \quad (1)$$

と表現できる。

4.2 動作解析アルゴリズム

解析アルゴリズムは以下になる。

(1) 体リンク長の正規化

与えられた動作データセットのリンクの長さを、与えられたリンク長で正規化する。

(2) 体中心座標系の設定と目標点の座標変換

正規化された動作データの各フレームにおいて体中心座標系を求め、目標点（両手足先）をこの座標系に変換する。

(3) 速度によるフレーム対応付け

データセット間の目標点の速度変化を対応付けする。得られた全フレームでの速度変化を規定範囲に収まるよう正規化し、DP マッチングによって対応付けする。すなわち、セグメント m, n が各々 $V_m = \{vm_1, vm_2, \dots, vm_{im}\} | vm_i \in R^3$, $V_n = \{vn_1, vn_2, \dots, vn_{in}\} | vn_i \in R^3$ で表されたとなると、これらの相関 $D(M, N)$ は以下のように表される。

$$D(M, N) = S(V_m, V_n) \quad (2)$$

$$S(k, l) = d_{kl} + \min(S_{k,l-1}, S_{k-1,l-1}, S_{k-1,l}) \quad (3)$$

$$d_{i,j} = (1 - \alpha \exp(\beta(i - j)^2)) |vm_i - vn_j| \quad (4)$$

ここで、 α および β は任意のパラメータで、両データセットにおける時間差の許容範囲を示すパラメータである。ここで、(2) 式 $\min(S_{k,l-1}, S_{k-1,l-1}, S_{k-1,l})$ 部分において最小となったフレーム番号の関係を記録しておく。

(4) 対応フレーム間速度によるセグメンテーション

(3) で得られた対応フレーム間の速度を合計し、新たな速度変化グラフを生成する。次に、各セグメントの速度変化が Bell Shape であるという知見⁶⁾を用いて、これらを Gaussian Mixture として考え、EM アルゴリズム⁷⁾によるフィッティングを導入する。すなわち、元となる速度変化グラフを $x(t)$ とすると、以下の式を満たす最適なパラメータ w_i, c_i, s_i を求める。

$$x(t) - \sum_{i=1}^N w_i \exp(s_i^2(t - c_i)^2) \rightarrow \min. \quad (5)$$

これにより、隣接する Gaussian Distribution の交点から動きのセグメントフレームを行うことができる。ここから、各動作データにおけるセグメントフレームとその間の対応が得られる。

(5) 基本動作と動作スタイルの抽出

次に、各動作セグメントから基本動作と動作スタイルを抽出する。ここで、動作セグメントのリンク n の動きを $Q = \{q_n(0), \dots, q_n(T)\}$ とし、基本動作 $Q^B = \{qb_n(0), \dots, qb_n(T)\}$ としては開始・終了フレーム ($q_n(0) \sim q_n(T)$) 間を等速に回転する運動であると考え、動作スタイル $Q^S = \{qs_n(0), \dots, qs_n(T)\}$ は実際の動きとの相違であると考え、基本動作と動作スタイルは以下のように分離することができる (図4)。

$$qb_n(t) = \frac{\sin(\delta\theta(1 - \frac{T-t}{T}))}{\sin(\delta\theta)} q_n(0) + \frac{\sin(\delta\theta(\frac{T-t}{T}))}{\sin(\delta\theta)} q_n(T). \quad (6)$$

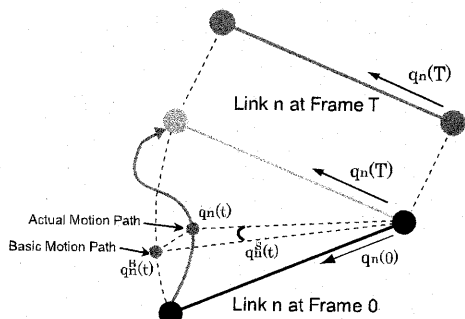


図4 基本動作と動作スタイル

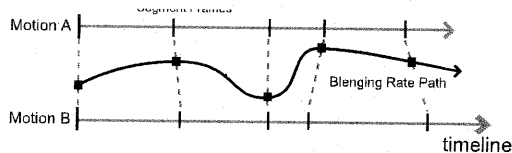


図5 動作の合成割合の設定

$$q_s(t) = \frac{q_n(t)}{q_b(t)} \quad (7)$$

ただし、 T は動作セグメントのフレーム数、 $\delta\theta$ は境界条件 $q_n(0), q_n(T)$ 間のなす角である。

4.3 基本動作と動作スタイルの性質

以上のアルゴリズムにより、動作セットにおける対応セグメントフレーム、動作セグメントにおける基本動作と動作スタイルが抽出できる。基本動作 Q^B は動作セグメントにおける開始・終了姿勢を線形に補間する動きであり、動作スタイル Q^S は実際の動きとの差異、すなわち個人間で異なる体部位のひねりや速度変化などを表すパラメータである。動作スタイルは、元の動作から基本動作を除算したものであるため、理論的には逆演算（積演算）によって元の動作を復元できる。すなわち、セグメント境界での体部位姿勢（ポーズ）と動作スタイルによって、元の動きが表現できることになる。

4.4 動作生成アルゴリズム

動作生成においては、対応付けられた2つの動作を任意のパラメータで合成して、新しい動作を生成することを考える。ユーザーが支持するパラメータとしては、解析アルゴリズムで得られた各部位のセグメント境界に対し、合成割合を示す値を設定する（図5）。動作合成は、以下の項目について行われる。

（1）止め時刻の合成

止め時刻の合成は、各部位の止めのタイミングを制御するために行われる。動作A、Bのセグメントフレームを f^A, f^B 、合成割合を r^A, r^B ($r^A + r^B = 1$) とした場合、新たなセグメントフレームはその線形補間として $f^A r^A + f^B r^B$ で表される。

（2）止め姿勢の合成

止め姿勢はセグメント境界でのポーズであり、動作セグメントの境界条件を意味する。この場合においても止め時刻と同様に線形合成を用いる。すなわち、動作A、Bのセグメントフレームにおける姿勢 $Pose^A, Pose^B$ があった場合、その合成された姿勢 $Pose^{AB}$ は以下で表される。

$$Pose^{AB} = (r^A p_0^A + r^B p_0^B, q_0', q_1', q_2', \dots, q_{17}'). \quad (8)$$

$$q_n' = \frac{\sin(r^A \delta\theta)}{\sin(\delta\theta)} q_n^A + \frac{\sin(r^B \delta\theta)}{\sin(\delta\theta)} q_n^B. \quad (9)$$

（3）動作スタイルの合成

止め姿勢の合成により動作セグメントでの境界条件が設定されるため、ここから基本動作が合成できる。一方、動作スタイルに関しても両者のスタイル合成を行い反映させる。対象とする二つの動作セグメント $C(t_c) | 1 \leq t_c \leq T^C$ および $D(t_d) | 1 \leq t_d \leq T^D$ を考える。ここで、両動作セグメントからは、基本動作 $C^B(t_c), D^B(t_d)$ 、動作スタイル $C^S(t_c), D^S(t_d)$ が得られるものとする。ここで、合成結果となる動作セグメント $E(t_e)$ のフレーム数は（1）より $T_E = T_C r^C + T_D r^D$ となる。また、 $C^S(t_c), D^S(t_d)$ の長さを T_E に変形したものをそれぞれ $C^S(t_e), D^S(t_e)$ とすると、合成結果は、

$$E(t) = E^B(t) \left(\frac{\sin(r^C \delta\theta)}{\sin(\delta\theta)} C^S(t) + \frac{\sin(r^D \delta\theta)}{\sin(\delta\theta)} D^S(t) \right). \quad (10)$$

として得られる。

5. 実験と考察

以上のアルゴリズムを実装し、実データについて適用した。使用したデータは民族舞踊のひとつである「会津磐台山」で、男性、女性1サンプルずつを用いた。データは光学式モーションキャプチャシステムである Oxford Metrics 社 Vicon8 で取得した。取得レートは 120Hz である。またデータ取得時は録音テープに記録された音楽にあわせて演じられる。本舞踊は、一連の動きを繰り返して何度も演じるものであるため、そのうちの1回を抜き出し、実験に用いた。抜き出されたデータは約 1000frame (約 8秒) である。男女の動きを 100 フレームごとに取り出したものを図6に示す。前述したように、全体としてみると同様の動きであるが、局所的に動きの大きく異なる部分があることがわかる。

本アルゴリズムは、体の各部位ごとに評価をすることが可能であるが、ここでは左手の動きを元に動作のセグメンテーションを行った。男・女の左手先の速度を示したグラフを図7に示す。ここからわかるように、全体としては同様の速度変化をしているようだが、局所的に速度変化のゆらぎがある。この速度変化に対し、本手法で Gaussian Mixture にフィッティングし、その結果をセグメンテーションに使用した結果を図8に示す。従来のアルゴリズムでは誤検出された部分が回避できることがわかり、また同時に両者の「止め」が生じるフレームは微妙に相違があることが確認できる。次に、基本動作を抽出した例を図9に示す。人の動きの滑らかさが失われ、線形補完による機械的な動

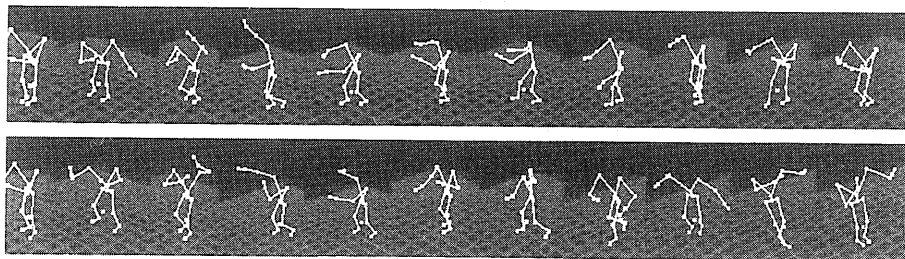


図6 「会津磐台山」の100frame 毎の動作変化（上：女性の動作，下：男性の動作）

作となることがわかる。

動作合成アルゴリズムの例として、男・女の動作の合成比を変化させ生成した動作を図10に示す。合成比を変化させる事によって、男女いずれかの動きに近づいていくことが確認できる。特に、左手を振り上げる動作を観察すると、女性は滑らかに手の振り上げ動作が生じるのに対し、男性は上腕から手先に向かって、回転させながら上昇させる動作をとっており、この違いが合成パラメータの割合によって徐々に変化していることが確認できる。

J. Neuroscience, pp.1688-1703, 1985

6. ま と め

モーションキャプチャによって得られた複数の動作を解析し、そこから動きの基本構造（基本動作）および個性をあらわす動作スタイルを抽出する手法を提案した。また、動作スタイルを用いた動作合成により、個人の個性を増減させながら動作を合成させることができるようになった。これにより、舞踊における個性が数値的に表され、応用することが可能になったと考えられる。今後は、いくつかの他の動作に対して本手法を適用し確認を行ったり、得られた動作スタイルパラメータをより深く解析することを進めたいと考えている。

7. 謝 辞

本研究は、科学技術振興事業団CRESTプロジェクト、高度メディア社会のための情報技術からの援助を受けた。

参 考 文 献

- 1) D.M.Gavrila: The Visual Analysis of Human Movement: A Survey, CVIU vol.73, no.1, pp.82-98, 1999
- 2) D. Miyazaki et.al : The Great Buddha Project: Modeling Cultural Heritage through Observation, VSMM2000, pp.138-145, 2000.
- 3) 中澤, 中岡, 原田, 工藤, 池内: 視覚による舞踊動作の保存・解析及び生成, 画像の認識・理解シンポジウム2003(MIRU2003), 2003.
- 4) 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 5 号, pp.878-885, 2000
- 5) 松本, 八村: モーションキャプチャデータからの基本身体動作の抽出, じんもんこん 2000, pp.17-24, 2000
- 6) 川人光男: 脳の計算理論, 産業図書, 1996
- 7) T.Flash, H.Hogan: The coordination of Arm Movements,

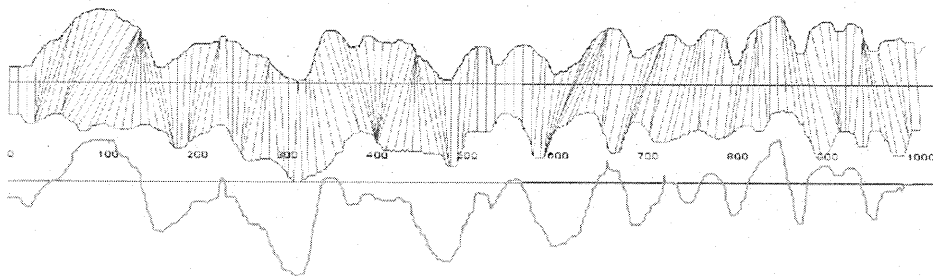


図 7 「会津磐台山」の体中心座標系における左手先の速度変化（上：女性，中：男性，下：対応付け結果）

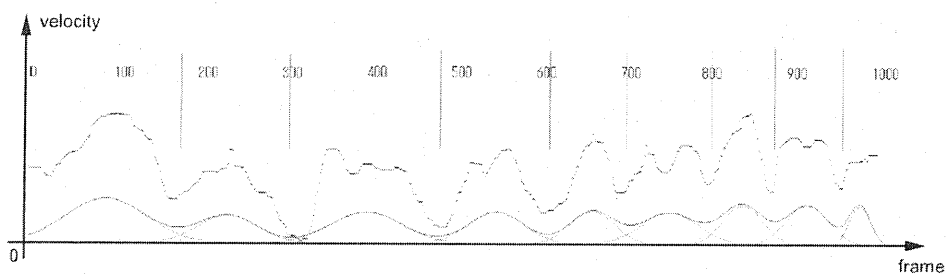


図 8 合成速度に対するEMアルゴリズムによるフィッティングとセグメンテーション結果，速度変化グラフとモデルとなったガウス分布関数，ガウス分布関数より合成されたフィッティング結果，およびセグメンテーション結果（縦線）

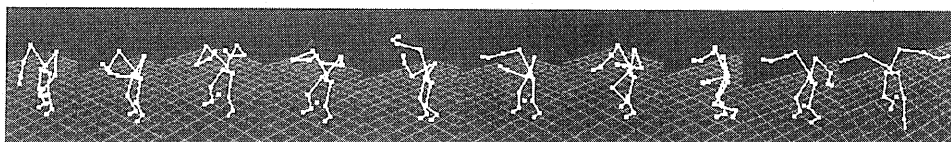


図 9 セグメンテーション結果より得られた基本動作

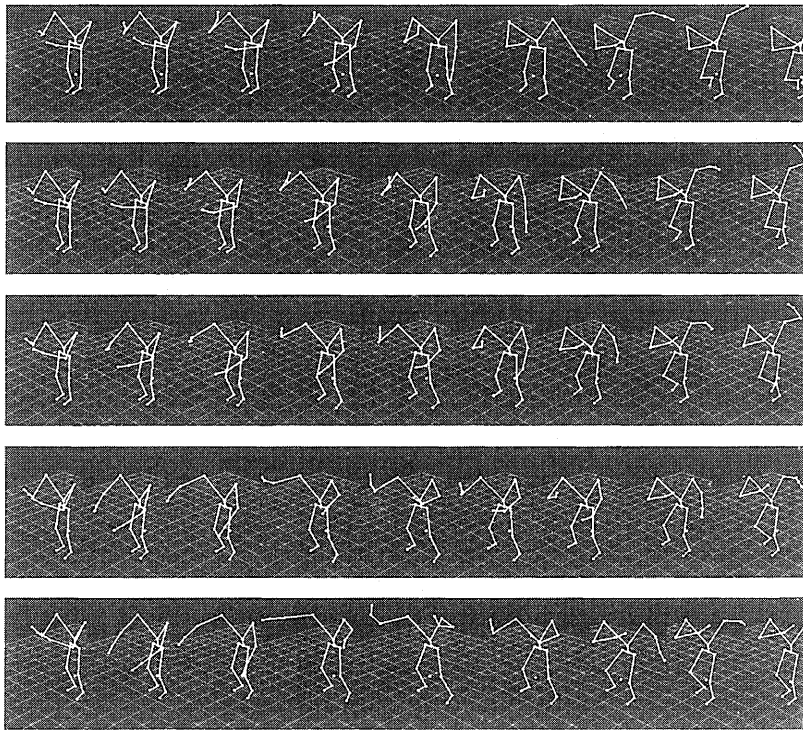


図 10 オリジナルの動きと合成結果 (会津磐台山 Frame0-200). 1 段目: オリジナルの動き (女性), 2 目: 合成結果 (女性 8, 男性 2), 3 段目: 合成結果 (女性 5, 男性 5), 4 段目: 合成結果 (女性 2, 男性 8), オリジナルの動き (男性)