

固有空間を用いた舞踊動作の定量的評価に関する検討

山根 亮 尺長 健

岡山大学大学院 自然科学研究科
〒700-8530 岡山市津島中 3-1-1

E-mail: {ryo,shaku}@chino.it.okayama-u.ac.jp

あらまし: 本稿では舞踊動作を定量的に評価する手法について検討を行う。まず、手本動作と入力動作の間に含まれる差異を、全体振りミス、部分振りミス、不完全動作および時間ズレという4つのミスに分類する。このうち、従来の手法では混在して取り扱っていた後者2つのミスが、手本動作によって作成される動作部分空間を用いることで分離できることを示す。最後に、分離したそれぞれの要因を個別の基準で評価できることを実データを用いて示す。

Quantitative evaluation of Dance Motion by Eigenspace Analysis

Ryo Yamane, Takeshi Shakunaga

Department of Computer Science, Okayama University
3-1- 1 Tsushima-naka, Okayama, 700-8530, JAPAN

Abstract: Quantitative evaluation of dance motion is discussed using eigenspaces that are constructed by motion sequences. Dance errors are classified into 4 types; global motion errors, partial motion errors, incomplete motion and time shift errors. These errors are discriminated to 2 groups by eigenspace analysis. That is, dance errors are automatically extracted by two distance measures defined for a given dance eigenspace. Experimental results show the effectiveness of the proposed measure.

1 はじめに

現在、人物の動作を測定、解析する技術は人間工学の大きな課題である。なかでも、コンピュータビジョン技術を利用した動作解析は、被験者の身体的負担も少なく、医療やスポーツ、アミューズメント利用など応用範囲も広く、様々な研究が行われている。また、舞踊解析への応用は、後継者不足や模倣を通じた伝承 [2] の難しさから近年盛んに研究されている。

動作情報は、近年のモーションキャプチャ技術の進歩によって、姿勢の時系列情報として容易に獲得することができる。これによって得られた動作データを用いて、様々な舞踊解析が行なわれている。舞踊のコンテンツ化の観点から、湯川ら [3] は舞踊を細かな単位動作に分割し、舞踊符と呼ばれる身体動作記述言語によって表現する手法を提案

している。舞踊を分割する方法としては、白鳥ら [4] が動作情報の他に音楽情報を利用し有限個動作プリミティブへ分割する手法を提案している。また、中岡ら [5] によって解析された舞踊データを元に、ロボットに舞踊を再現させる研究も行なわれている。

一方、コンテンツ化だけではなく、舞踊伝承をサポートすることも重要な課題である。舞踊動作の学習は模倣による学習を基本としているため、手本となる動作と比較して再現性を定量的に評価することで伝統舞踊の継承に役立てることができる。舞踊動作の解析や応用の分野において、動作データの比較は重要な課題となっている。

2つの動作データの比較はそれらを構成する姿勢列の比較に帰着できる。この比較として、我々は、時系列相関行列を用いることで姿勢列の比較を行なう手法を提案した [1]。これにより、2つの

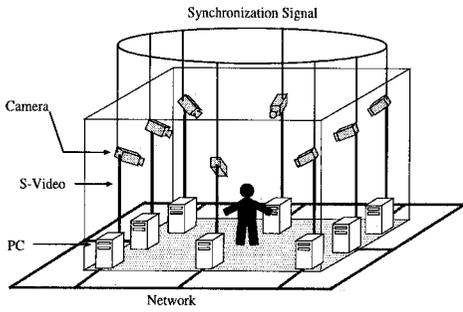


図 1: 多視点カメラシステム

動作の時間的対応と、動作に含まれる差異を抽出し、視覚的に表す方法を示した。しかし、差異には踊り手の違いや、時間的なズレや速度差、その他単純なミスなども含まれており、そのままでは舞踊の再現性の指標にはならない。したがって、これらの要因を分離し、それぞれを評価する必要がある。本稿においてはこの混在した要素を手本舞踊データによって張られる動作固有空間を用いて分類し、それぞれに対して適切な基準を用いて評価する手法を提案する。

以下 2 章では動作データ間の差異の抽出法について述べ、3 章では動作固有空間を用いた差異の分離について述べる。最後に実際の舞踊動作を用いた分類例を示す。

2 舞踊動作の定量的評価

2.1 舞踊データの撮影

本稿では舞踊動作の評価を、あらかじめ撮影しておいた手本となる舞踊動作データと、入力となる舞踊動作データの比較によって行なう。これら舞踊動作データの撮影は、図 1 に示す多視点カメラシステム [1] を用いて行なった。各姿勢データの獲得は、演者が図 2 のようなマーカ付きスーツを装着し、各マーカの 3 次元位置をステレオ法により復元することで行なった。計測に用いる座標系は床面の垂直方向を Z 軸とした世界座標系とする。ここで、マーカは体節の動きを計測できるように、関節付近に取り付けている。

本稿では演者の姿勢を図 3 に示す 9 つの体節を用いて表現する。図 3 に示す関節位置にマーカを

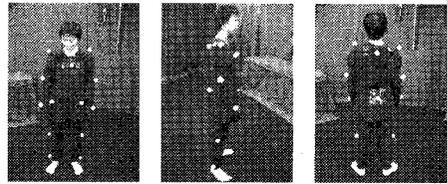


図 2: マーカ付スーツを装着した人物

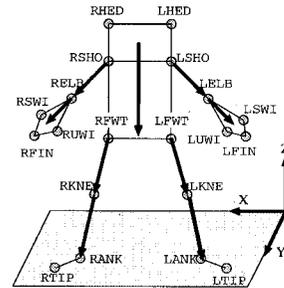


図 3: 体節ベクトル

配置し、体節ベクトル情報に変換することで、演者の立ち位置、体格に依存しない姿勢比較を行なうことができる。本稿では人物の姿勢データ \mathbf{X} を、体節 p ($1 \leq p \leq P$) に対応する体節ベクトル \mathbf{x}_p を用いて、以下のように定義する。

$$\mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2^T \ \cdots \ \mathbf{x}_p^T \ \cdots \ \mathbf{x}_n^T]^T \quad (1)$$

ここで、 \mathbf{x}_p は 3 次元ベクトルである。

$$\mathbf{x}_p = [x_{px} \ x_{py} \ x_{pz}]^T \quad (2)$$

この姿勢ベクトル集合 \mathbf{X} のとりうる範囲を動作空間と呼び、また、 \mathbf{X} の時系列集合を動作データと呼ぶ。

2.2 時系列相関行列を用いた対応付け

2 つの動作データを比較する場合、それぞれを構成する姿勢列の時系列の対応付けが必要となる。この対応付けを行う方法として、我々は時系列相関行列 [1] を用いたマッチング手法を提案している。時系列相関行列とは、2 つの姿勢列の要素のすべての組合せに対して姿勢類似度を行列表記したものである。図 4 に時系列相関行列を視覚的に表した例を示す。各軸はそれぞれの動作の時刻を表し、画素はそれぞれの時刻間での姿勢の類似度

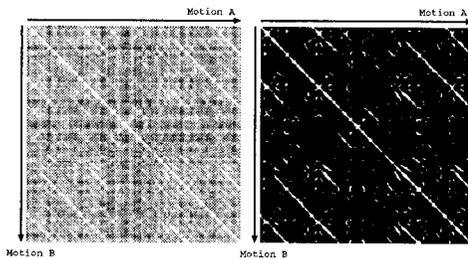


図 4: 時系列自己相関行列と高相関な領域

を表す。2つの動作が類似していた場合、時系列相関行列上では図4に示すような直線状の高相関な領域が現れる。この高相関な直線状領域に対してフィッティングを行なうことで、2つの動作の対応を得ることができる。

2.3 姿勢の比較

時間的な対応付けを行なった2つの姿勢を比較することで、その動作の差異を抽出することができる。しかし、この差異には踊り手の違いや小さな時間的なズレ、振りミスなどが含まれている。これらのうち、体格の差異は体節の長さを正規化することである程度吸収できる。しかし、舞踊に含まれるミスは混在したままである。本研究ではこれらの舞踊に関するミスを、全体振りミス、部分振りミス、不完全動作および時間ズレという4つに分類する。

(1) 全体振りミス

舞踊中のある区間の動作全体が間違っている場合を指す。この場合、対応する姿勢間の類似性が低く、時系列相関行列による対応付けに失敗する。これを利用して、類似性が低い区間をミスとして検出できる。

(2) 部分振りミス

舞踊の振り自体はほぼ一致しているが、一部の体節の動きが一致していない状態である。この場合、他の体節は正しく舞踊を再現しており、姿勢間の類似性は極端に低くなるわけではない。このため、このようなミスは、吸収しきれなかった体

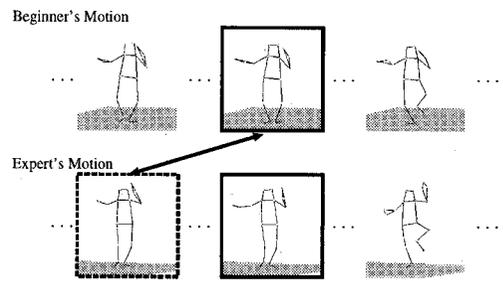


図 5: 初心者の振り切り姿勢 (上段) と熟練者の振り切り姿勢 (下段)

格差や観測ノイズに埋もれてしまう。

(3) 不完全動作

熟練者の舞踊と初心者の舞踊を比較した場合、踊りはほぼ再現しているものの、各体節の振り伸ばしが省略されることが多い。図5に振り伸ばしが省略された例を示す。上段が初心者の舞踊、下段が熟練者の舞踊、黒枠で囲んだ姿勢が右腕を振り切った時刻の姿勢を表す。熟練者に比べ、初心者の右手が伸び切っていないのがわかる。この場合、その時刻における姿勢の類似度はあまり高くない。

また、初心者の伸びの足りない姿勢が、熟練者の別の姿勢と類似度が高くなる場合がある。図5の例では、初心者の手足の伸びが中途半端なため、熟練者の振り切る手前の姿勢(破線枠)と類似度が高くなっている。このような姿勢を含む動作の場合、時系列相関行列を用いた場合、誤対応が発生する。この誤対応となる熟練者の姿勢を不完全姿勢と呼び、これによって発生するミスを不完全動作と定義する。

(4) 時間ズレ

振りの速度が異なっていたり、微妙にタイミングが遅れているなどの時間的なゆらぎを表す。このような時間ズレは時系列相関行列上で曲線状の高相関領域として現れる。この曲線を抽出することで、時間的ゆらぎを含んだ対応付けを行なうことができる。また、この対応付け結果をズレ要素として抽出が可能である。

全体振りミスと時間ズレの2つのミスは、時系

列相関行列を用いることで抽出、分類が可能である。しかし、部分振りミスと不完全動作は外的要因に埋もれやすく、それぞれの分類も容易ではなかった。本研究では、これらミスを動作固有空間を用いて分離し、評価を行なうことを目的とする。

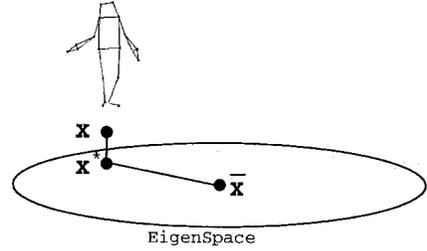


図 6: 動作固有空間

3 固有空間を用いた分類

3.1 動作固有空間

舞踊動作を構成する姿勢 \mathbf{X} は n 個の体節ベクトルに対し、それぞれ 3 つの方向情報を持つ $3n$ 次元空間で表される。しかし、人体構造を考慮した場合、とりうる範囲はこれのごく一部であり、また、動作が短時間の場合、その空間はさらに小さい多様体として局所化できる。この局所的な多様体をその動作に対応する動作空間と呼び、次元圧縮を行なったものを動作固有空間と呼ぶ。

この動作固有空間を、部分動作を表す姿勢集合を用いて作成する。動作データ中の人物はさまざまな方向を向いており、方向に関する正規化を行う必要がある。本研究では、人間の行う動作は重力に支配され、地面に拘束されることを利用し、 Z 軸まわりの回転のみ正規化を行う。この正規化は、 \mathbf{X} から Z 成分を除いた 2 次元ベクトル集合の回転角推定問題に帰着できる。

2 次元ベクトル集合の回転推定には複素相関 [6] を用いる。これにより、2 つの 2 次元ベクトル集合の Z 軸周りの位相差と、平面における類似度を同時に求めることができる。 j を虚数単位とし、 \mathbf{x}_p の Z 軸以外の 2 軸を用いた複素表現 x_p を次式で定義する。

$$x_p = x_{px} + jx_{py} \quad (3)$$

また、 x_p の複素共役を x_p^* で表す。

$$x_p^* = x_{px} - jx_{py} \quad (4)$$

このとき、姿勢 \mathbf{X} と \mathbf{X}' の 2 次元複素相関 $\mathbf{C}_{\mathbf{X}\mathbf{X}'}$ は以下のように与えられる。

$$\mathbf{C}_{\mathbf{X}\mathbf{X}'} = \sum_{p=1}^P \frac{x_p x_p'^*}{\|x_p\| \|x_p'\|} \quad (5)$$

$$= \|\mathbf{C}_{\mathbf{X}\mathbf{X}'}\| e^{j\theta} \quad (6)$$

また、 θ は姿勢データ間の Z 軸周りの位相差であ

り次式で表される。

$$\theta = \arg \mathbf{C}_{\mathbf{X}\mathbf{X}'} \quad (7)$$

この位相差を用いて方向の正規化を行なうことができる。また、人物間の体格の差異を吸収するため、各体節ベクトル \mathbf{x}_p を大きさを 1 に正規化する。これらの正規化により、環境に依存しない動作部分空間が作成できる。

正規化済み姿勢集合が与えられたとき、この集合より計算される平均を $\bar{\mathbf{X}}$ 、共分散行列を Σ で示す。ここで、 Σ を固有値分解することにより得られる固有値のうち上位 m 個に対応する固有ベクトルから構成される行列を Φ_m で表す。このとき、動作固有空間への射影 \mathbf{X}^* と残差 $\mathbf{X}^\#$ (図 6) は次式で計算される。

$$\mathbf{X}^* = \Phi_m^T (\mathbf{X} - \bar{\mathbf{X}}) \quad (8)$$

$$\mathbf{X}^\# = \mathbf{X} - \bar{\mathbf{X}} - \Phi_m \mathbf{X}^* \quad (9)$$

3.2 動作固有空間を用いた分類

手本動作を表す姿勢列 $\mathbf{S} = [\mathbf{X}_1, \dots, \mathbf{X}_t, \dots, \mathbf{X}_T]$ と入力姿勢列 $\mathbf{T} = [\mathbf{Y}_1, \dots, \mathbf{Y}_t, \dots, \mathbf{Y}_T]$ との比較を考える。2 つの動作が同じテンポの同じ舞踊である場合、同時刻の 2 つの姿勢 $\mathbf{X}_t, \mathbf{Y}_t$ には時間的な対応関係があり、またこれらの類似性も高い。

しかし、入力動作に前述の舞踊ミスが含まれている場合、この対応関係にズレが現れる。また、同時刻における類似性は必ずしも高くはならない。この同時刻におけるズレが、手本動作 \mathbf{S} によって張られる動作固有空間でどのように現れるのかを示す。

3.2.1 固有空間からの距離による分類

部分振りミスを含む場合、同時刻 2 つの姿勢 $\mathbf{X}_t, \mathbf{Y}_t$ は類似はしているものの、一部の体節でミスをしているため高くはない。また、振り全体振りミスを含む場合、同時刻の 2 つの姿勢は全く別のものであり、類似性は低い。これらのミスは、振りを構成する姿勢列が間違っているため、 \mathbf{S} の中には \mathbf{Y}_t と類似する姿勢は含まれていない。そのため、 \mathbf{Y}_t は動作固有空間からの距離が大きい (図 7)。

動作固有空間と \mathbf{Y}_t との距離 $d^\#$ は以下のように表すことができる。

$$d^\# = \|\mathbf{Y}_t - \bar{\mathbf{X}} - \Phi_m \mathbf{Y}_t^\#\| \quad (10)$$

$$= \|\mathbf{Y}_t^\#\| \quad (11)$$

ある動作中に距離 $d^\#$ が大きい区間が存在する場合、その区間で全体振りミスもしくは部分振りミスが発生していると判断できる。

3.2.2 固有空間内距離による分類

時間ズレを含む場合、同時刻の 2 つの姿勢 $\mathbf{X}_t, \mathbf{Y}_t$ の類似性は低い。しかし、時間ズレを Δt とすると、 $\mathbf{X}_{t+\Delta t}, \mathbf{Y}_t$ には対応関係があり、2 つの姿勢の類似性は高い。したがって、時間ズレを含む \mathbf{Y}_t と動作固有空間からの距離は小さい。また、不完全動作を含む場合も $\mathbf{X}_t, \mathbf{Y}_t$ の類似性は低い。しかし、手本動作 \mathbf{S} の中に \mathbf{Y}_t と類似した不完全姿勢 $\mathbf{X}_t^\#$ が含まれるため、 \mathbf{Y}_t と動作固有空間からの距離は小さい。したがって、これらのミスが含まれる場合、入力姿勢は固有空間上に存在し、姿勢間の差異もこの固有空間上に現れる。

同時刻の 2 つの姿勢 $\mathbf{X}_t, \mathbf{Y}_t$ の固有空間上での距離を d^* とすると、それぞれの射影 $\mathbf{X}_t^*, \mathbf{Y}_t^*$ の距離として以下のように表すことができる。

$$d^* = \|\mathbf{X}_t^* - \mathbf{Y}_t^*\| \quad (12)$$

$$= \|\Phi_m^T(\mathbf{X}_t - \bar{\mathbf{X}}) - \Phi_m^T(\mathbf{Y}_t - \bar{\mathbf{X}})\| \quad (13)$$

$$= \|\Phi_m^T(\mathbf{X}_t - \mathbf{Y}_t)\| \quad (14)$$

ある動作の中に距離 d^* が大きい区間が存在する場合、その区間で不完全動作もしくは時間ズレが発生していると判断できる。

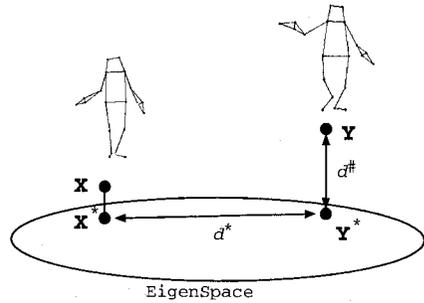


図 7: 動作固有空間を用いた評価

3.3 部分動作による動作固有空間の作成

手本動作が長い場合、その動作を表現する多様体は複雑になり、空間を表すためには次元数を大きくする必要がある。したがって、動作を適切な部分動作に分割し、それぞれに対して固有空間を作成する。また、この部分動作は舞踊にとって意味のある分割である必要がある。

舞踊を自動的に分割する方法として、辞書データとマッチングする方法や速度変化などの動作データの特性から分割する方法 [4] などが提案されている。前者の場合、あらかじめ辞書を用意する必要がある。また、後者の場合は舞踊の特徴的な部分で切り分けることが可能だが、舞踊のきめとなる姿勢を基準としているため、本稿の用途では扱い辛い。

本研究ではこの舞踊分割のために、時系列自己相関行列を用いた振り抽出法 [1] を利用する。これにより、舞踊本来の区分に近い分割が可能になり、同じ動作をまとめることもできる。

4 実験結果

4.1 動作固有空間の作成例

動作固有空間を用いて舞踊ミスが分類できることを確認するために実験を行った。対象は伝統舞踊である牛深ハイヤ節で、これを 3 人の被験者に同一のテンポで踊ってもらい、そのうち 1 人の舞踊を手本動作、残りの 2 人を入力動作として利用した。舞踊データのフレーム数は 150 フレームであり、手振り動作 1(1-70 フレーム)、回転(71-100 フ

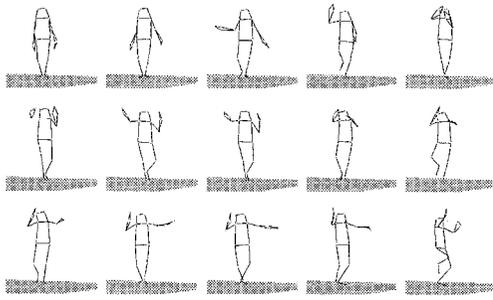


図 8: 手振り動作 1

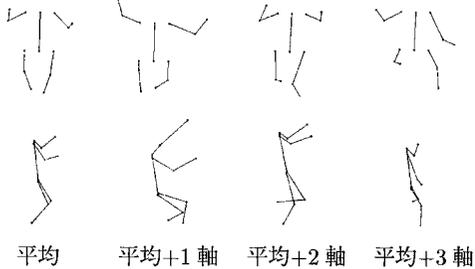


図 9: 手振り動作 1 の平均と固有軸 (上段: 正面, 下段: 横向き)

フレーム), 手振り動作 2(101-150 フレーム) の 3 つの振りによって構成される。それぞれの振りに対し個別の固有空間を作成する。固有空間の次元数はすべての振りで累積寄与率が 0.99 以上となる 7 次元とした。

手振り動作 1, 回転および手振り動作 2 を図 8,10,12 に示す。また, それぞれの動作の平均姿勢と第 1-3 固有軸要素を逆射影したものを図 9,11,13 に示す。手振り動作 1 において, 第 1 固有軸に左手の上下に関する成分, 第 3 固有軸に右手の上下に関する成分が現れている。回転動作においては, 第 1-2 固有軸に歩行動作を構成する両足運動成分が現れている。また, 手振り動作 2 においては, 右足を大きく曲げる動作が第 2 固有軸に現れている。

4.2 分類結果

図 14 に被験者 1 の d^* と $d^\#$ を示す。いくつかの区間において $d^\#$ が大きな値をとっている。 $d^\#$ が極大となる時刻の被験者 1 の姿勢を図 15 に示す。こ

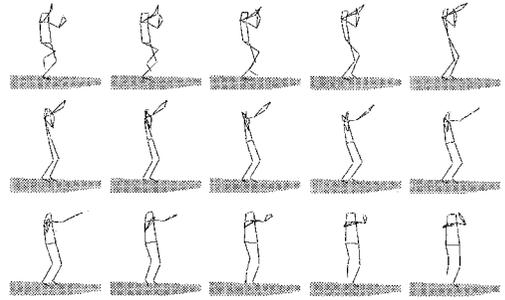


図 10: 回転動作

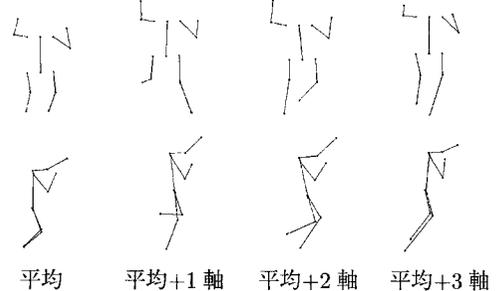


図 11: 回転動作の平均と固有軸 (上段: 正面, 下段: 横向き)

の $d^\#$ が大きな区間においては被験者 1 の動作が遅れており, また, 極大となる時刻においてズレも最大となっていることが動作データより確認できる。

また, $d^\#$ に関しては特に大きな変化は現れていない。これは被験者 1 の動作は全体舞踊ミスも部分舞踊ミス含んでおらず, 舞踊をほぼ再現していたと考えることができる。

次に, 図 16 に被験者 2 の d^* と $d^\#$, 図 17 に $d^\#$ が極大の時刻の姿勢を示す。被験者 1 の例と同様に, 極大となる第 10,71,82 フレーム前後で手本動作に比べてやや遅れていることが確認できる。

また, 第 71 フレームは時間ズレは発生していない。このフレーム付近では, 左腕を内側に曲げる動作を行なっている。しかし, 被験者 2 の左腕はあまり曲がっておらず, 手本動作に比べて不完全であった。このため, 不完全動作として $d^\#$ が大きくなったと考えられる。

また, 第 55-69 フレーム付近において d^* が大きな値をとっている。このときの姿勢例を図 18 に示す。この区間では, 被験者の動作は手本動作に

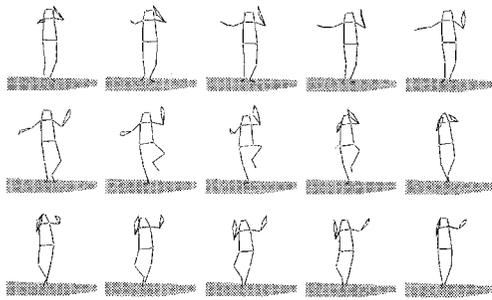


図 12: 手振り動作 2

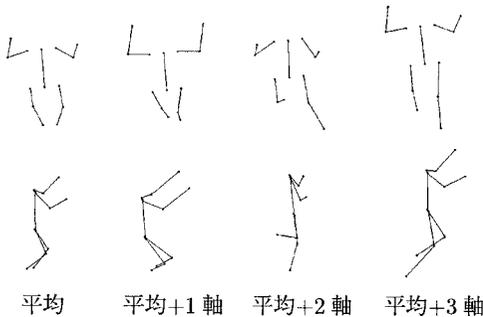


図 13: 手振り動作 2 の平均と固有軸 (上段: 正面, 下段: 横向き)

らべ右足が大きく開いており, 部分振りミスを抽出できていると考えられる。

5 まとめ

本稿では, 舞踊動作に含まれるミスを自動的に 4 つの要素に分類する方法を示した。また, この 4 つの要素の関係を, 動作固有空間を用いて説明した。最後に, 部分振りミスが固有空間との残差として, 時間ズレと不完全動作が固有空間内距離として現れる傾向にあることを実験により示した。

謝辞

実験データの撮影に協力頂いた岡山大学詠う会みみんに感謝する。本研究の一部は科学研究費補助金基盤研究 (B)(課題番号 15300062) により行われた。

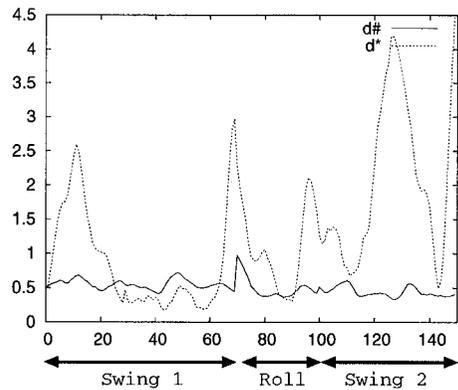


図 14: 被験者 1 の d^* および $d^\#$

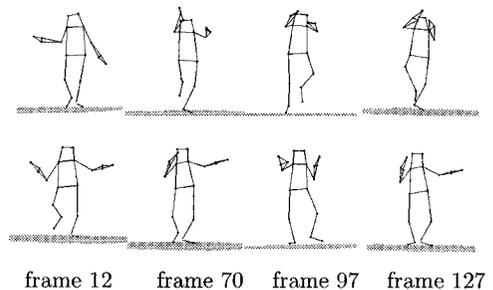


図 15: 図 14 の $d^\#$ 極大値に対応する手本姿勢 (上段) と被験者 1 の姿勢 (下段)

参考文献

- [1] 山根 亮, 戸高 千智, 川嶋 幸治, 尺長 健 “動作データの時系列相関行列による舞踊動作解析”, 電子情報通信学会論文誌 D-II, vol. J88-D-II, no. 8, pp.1652-1661, 2005.
- [2] 生田久美子, “「わざ」から知る”, 東京大学出版会, 1987.
- [3] 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”, 情報処理学会論文誌, vol. 41, no. 10, pp.2873-2880, 2000.
- [4] 白鳥 貴亮, 中澤 篤志, 池内 克史, “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”, 電子情報通信学会論文誌 D-II, vol. J88-D-II, no. 8, pp.1662-1671, 2005.

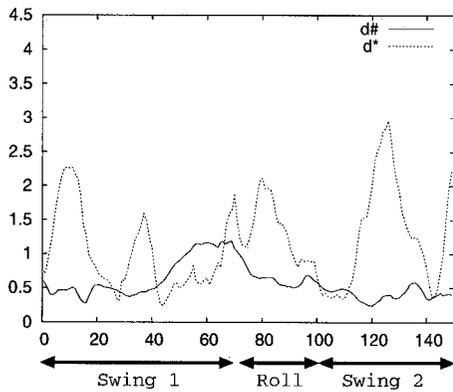


図 16: 被験者 2 に対する d^* および $d^\#$

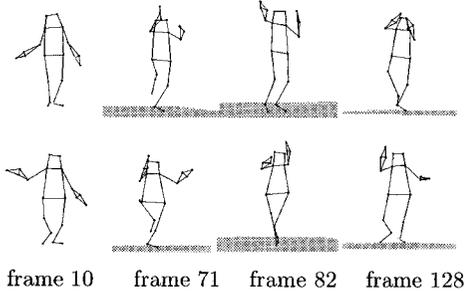


図 17: 図 16 の $d^\#$ 極大値に対応する手本姿勢 (上段) と被験者 2 の姿勢 (下段)

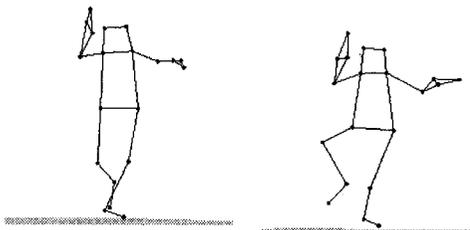


図 18: 図 16 の $d^\#$ の極大値に対応する手本姿勢 (左) と被験者 2 の姿勢 (右)

- [5] 中岡慎一郎, 中澤篤志, 横井一仁, 池内克史, “人の前進動作をロボットで模倣するための動作プリミティブ”, MIRU 2004, vol. 1, pp.177-182, 2004.
- [6] 尺長健, 金子博, “複素相関による 2 次元形状解析”, 電子情報通信学会論文誌 D, vol. J71-D, no. 2, pp.350-361, 1988.