

## グループダンス練習支援のための時空間的なダンスモーション可視化 Spatio-temporal Visualization of Dance Motion Data for Group Dance Practice Support

大越 瑠衣子<sup>1)</sup> 藤田 秀之<sup>1)</sup> 大森 匡<sup>1)</sup> 新谷 隆彦<sup>1)</sup>  
Ruiko Ohkoshi Hideyuki Fujita Tadashi Ohmori Takahiko Shintani

### 1 背景と目的

グループで踊るダンスにおいて、ダンスに統一感を与え、完成度を高めるためには、共通の振り付けに対して、全体のモーションのずれをできるだけ小さくする必要があります。モーションのずれを改善するために、グループで踊っている様子を撮影した動画がよく利用される。ダンス動画を利用したずれを改善する過程は、ずれを発見する、ずれの原因を特定する、ずれの情報共有をする、各自がずれを修正する、という4つの作業の繰り返しである。

例えば、以下の図1の画像は、グループダンス練習時に撮影した動画から、ダンスのずれを指摘するために作成した再生時刻「0:21」のスクリーンショットにずれの内容のコメントを記入した画像である。これは、画像内中央の対象者が周りのメンバーに対して動き出しが早いことを指摘しているものである。実際に、このずれを動画から発見し、ずれの原因が動き出しのタイミングである事を特定し、対象者に図1のような画像を共有し、画像を受け取った対象者がずれを確認し修正するまでの過程を具体的に追っていく。まず、ずれを指摘する人はダンス動画を再生したり、シークバーを手動で動かしたりすることで対象者の身体の形が周りのメンバーに対して異なる再生時刻を探し、再生時刻0:21にずれが生じていることを見つける。次に、このずれが生じた原因を特定するために、見つけた該当時刻の前後を繰り返し確認することで、該当時刻よりも少し前に、対象者のみが動き出すタイミング、該当時刻よりも少し後に対象者以外のメンバーが動き出すタイミングが確認でき、ずれの原因が動き出しのタイミングのずれであることが分かる。ずれの原因を特定したら、その情報を対象者に伝えるために、再生時刻の映ったスクリーンショット画像を撮り、更に、ずれの原因が伝わるように画像に文字を入れる編集を行い、SNSを用いて情報共有をする。ずれの情報を受け取った対象者は、自身の動き出しのタイミングを修正するために、共有された画像のみではなく、元のダンス動画を確認し、該当時刻付近の自身の動きと、他のメンバーの動きを比較しながら、正しいタイミングを確認し練習する。

この改善方法は、グループ練習時の動画を利用して、グループ練習後に各自で練習できる点で有用だが、一連の作業に非常に時間がかかるという課題がある。ずれの原因を特定するためには、再生速度を変えながら、繰り返し動画を再生して観察する必要がある。ずれの情報共有のために、動画の再生時刻を指定するほか、該当箇所をスクリーンショットで切り取った静止画を作成する場合もあるためだ。加えて、ずれの修正にも、ずれている箇所とその原因を理解するために、指定された再生時刻前後の動画を繰り返し再生し、確認する必要がある。

1) 電気通信大学 大学院情報理工学研究所  
〒182-8585 東京都調布市調布ヶ丘 1-5-1



図1 ダンスのずれの情報共有の例

近年では、身体の動きをモーションデータとして取得するモーションキャプチャが身近になった。個人用のデバイスも普及し、幅広く活用されている。ダンスの情報処理に関する研究はHCI分野で盛んであり[1]、ダンスモーションデータの分析・活用法が多数提案されている。そこで本研究では、グループダンスの練習において、モーションのずれを改善する過程を支援するために、モーションキャプチャを用いてダンスモーションデータを可視化する手法を提案する。

### 2 関連研究

ダンスモーション間のずれを対象とする可視化分析システムとして、フレーム単位での身体の形である、ポーズに着目した手法が複数提案されている。ダンスモーションの可視化分析システムの提案の多くは、映像と、モーションキャプチャデータによる身体各部位の位置を結んだ棒人間を用い、さまざまな属性値のフレーム単位の遷移を可視化する。集団的傾向の可視化手法[2]を、多数の身体部位の移動軌跡に対して適用する手法[3]や、棒人間を時系列で高密度に並べることで、一覧性を高めた可視化手法[4]が提案されており、複数のダンスモーションの比較にも応用されている。

各時刻における複数のダンサー間のポーズに関する差異を、大規模な主観評価により調整した指標により定量化し、身体各部位のずれの大きさとして、映像上で可視化する手法[5]が提案されている。

個人のダンスの反復練習の支援を目的に、同じ振り付けのダンスを複数回踊ったデータを用い、異なる練習回をまたいだ身体各部位のクラスタリングに基づき、練習回ごとの差異や変化を可視化する手法 [6] が提案されている。

モーションの分類や比較のために、ポーズの類似度に基づき、ポーズをノードとするグラフやツリー構造で可視化する手法 [7] や、平面上に配置して可視化する手法 [8] が提案されているが、いずれも同一のモーション間のずれの分析にフォーカスした手法ではない。

本研究では、モーションキャプチャデータを用いて、グループダンスにおける同一のモーションのタイミングのずれを、その他のずれと区別して扱う可視化分析システムを検討する。

### 3 予備調査

本研究で提案するシステムの要件整理を目的とした予備調査を実施した。この調査では、著者が過去にグループでダンスを練習した際に作成した 2 種類の資料、資料 A(図 2)、資料 B(図 3) を用いて、ダンスのずれの扱いや情報共有の方法について考察することで、実際の作業内容やダンスの改善の過程を整理した。資料 A は、グループ練習時に撮影したダンス動画を利用し、練習後にダンスのずれについて情報共有するために、動画の該当箇所をスクリーンショットに取めた静止画上に、ずれの原因や解決策のコメントを書き込んだ画像である。2018 年 9 月 13 日に撮影された動画からグループのメンバー 1 名のずれを発見、原因を特定し、前述の静止画を作成し、翌 9 月 14 日に SNS を用いて共有した。ここでは、1 分 37 秒の動画の中で 8 件のずれを指摘している。資料 B は、別のグループダンスの動画に関して、ダンスのずれの情報を共有するために著者が作成した、手書きのイラストと文章によるメモの画像である。この手書きメモは、2017 年 9 月 5 日に、グループのメンバー 2 名それぞれに SNS で共有した。2 分 25 秒の動画に対してそれぞれ 4 件ずつ計 8 件の指摘をしている。



図 2 資料 A : ダンス静止画、動画スクリーンショットにコメントを記入

資料 A,B を見比べると、伝達方法は異なるものの、共有している情報は同じである。まず資料 A の共有している内容を整理すると、1 件のずれに対して、画像下部の動画の再生時刻でずれの発生時刻を、動画のスクリーンショット画像でずれ発生時のポーズを、ピンク色のコ

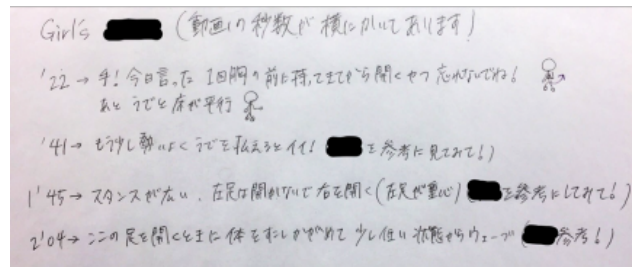
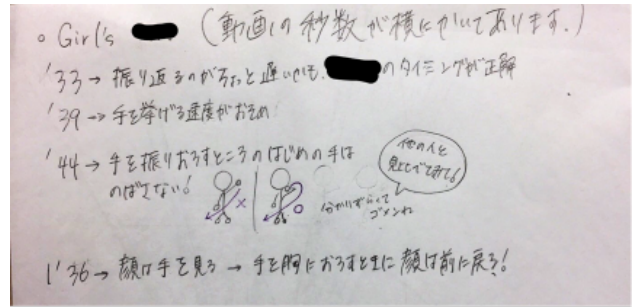


図 3 資料 B : 手書きメモの画像

メントでずれの内容や改善方法をテキストで共有している。次に資料 B の指摘している内容を整理すると、画像左端の動画の再生時刻でずれの発生時刻を、棒人間のイラストでずれ発生時のポーズを、テキストでずれの内容や改善方法を共有している。このように伝達方法は異なるが、どちらの資料でも、ずれの再生時刻、ポーズ、問題点と改善方法の情報を共有していることが分かる。

2 つの資料の計 16 件の具体的なずれの指摘を表にまとめると以下の表 1 のようになる。

表 1 各資料の指摘の内容

資料 A	資料 B
●足はあげない	◇振り返るのがちょっと遅い
●下の手はまっすぐ出す	◇手を挙げる速度がおかしい
●もう少し上体を下げる	●手を振りおろすところのはじめの手はのぼさない
●首を傾ける	●顔は手を見る
●もう少し首を傾ける	●うでと床が平行
★肩の助走は抑える	◇もう少し勢いよくうでを払う
◇動き出しが早い	●スタンスが広い
●両足は対象	●足を開くときに体を少しかがめる

指摘の内容に着目すると、「首をもう少し傾ける」や「足を対称に開く」といった、表内で「●」を付けた指摘である、動画の中の一瞬を切り取ったときの身体の形のずれである空間的なずれを指摘しているものと、「動き出しが早い」や「振り返るのが遅い」というように、表内で「◇」を付けた指摘である、動きのタイミングのずれである時間的なずれを指摘しているもの、更には表内で「★」を付けたどちらにも属さないずれがあることが分かる。空間的なずれと時間的なずれの 2 種類のずれについて比較すると、空間的なずれは、一枚の画像から相手にずれが直感的に伝わりやすく、ずれを修正する人はその一枚の画像から他者と自分の違いを比較しやすいが、時間的なずれは、画像からずれそのものは伝わる

が、他の人とどのくらいタイミングがずれているのか、ずれ始めはいつなのかを理解し修正するために、元の動画をモーション単位で何度も確認する必要があることが分かった。

また、この資料の画像を作成するにあたり、ずれの発見とその原因の特定という作業を行った。ずれの発見は動画をポーズ単位で確認しながら発見したが、そのずれが空間的なずれなのか、時間的なずれなのかを判別するためには、前後のフレームも確認する必要があるため、モーション単位での確認を行った。つまり、ずれの発見や、空間的なずれの情報共有や修正はポーズ単位で行っているが、ずれの原因の特定や時間的なずれの情報共有、修正はポーズ単位ではなくモーション単位で行うことが自然であると考えられる。この考えを表にまとめたものが以下である。

表 2 改善過程とずれを扱う単位

過程	空間的なずれ	時間的なずれ
発見	ポーズ	ポーズ
原因特定	モーション	モーション
情報共有	ポーズ	モーション
修正	ポーズ	モーション

以上の予備調査より、ダンスモーションのずれの情報共有や修正では、空間的なずれと時間的なずれを区別して扱う事が有用であることが確認できた。

## 4 提案手法

### 4.1 利用データと定義

#### ダンスモーションデータ

本研究では、ダンスのモーションキャプチャに Microsoft 社の AzureKinect を利用する。同機器は、赤外光を用いた Timeof-Flight 方式の深度カメラを搭載し、同時に複数の人体のモーションキャプチャをリアルタイムで行う。人体に特別なマーカーを装着する必要はない。出力されるモーションキャプチャデータは、一定の時間間隔 (最大 30fps) の時系列データであり、時系列データの各フレームを持つデータは、取得時刻と、同時刻における、各人体の各関節についての mm 単位の 3 次元絶対座標値と、四元数による回転角である。Azure kinect で検出される関節 (目、鼻、つま先等の特徴点を含む) は 32 箇所であり、関節をノードとし、うち、骨盤を根とする木構造として、スケルトンが定義されている。本研究の前処理で、以降に定義するモーションの形式に変換する。

#### ポーズ

ポーズとは各時刻における瞬間的な身体の形である。時刻  $t$  におけるダンサー  $i$  のポーズ  $p_{i,t}$  を、骨盤を除く関節  $j$  の単位方向ベクトル  $v_{i,j,t}$  の組として定義する。

$$v_{i,j,t} = (x_{i,j,t} - x_{i,parent(j),t}) / |x_{i,j,t} - x_{i,parent(j),t}| \quad (1)$$

ここで、 $x_{i,j,t}$  は時刻  $t$  におけるダンサー  $i$  の関節  $j$  の座標値、 $parent(j)$  は、スケルトンの木構造における関節  $j$  の親の関節を示す。目や鼻を除く 27 の関節をよる 26 の単位方向ベクトルによるポーズを用いる。各関節の位置ではなく、単位方向ベクトルを用いるのは、ポ

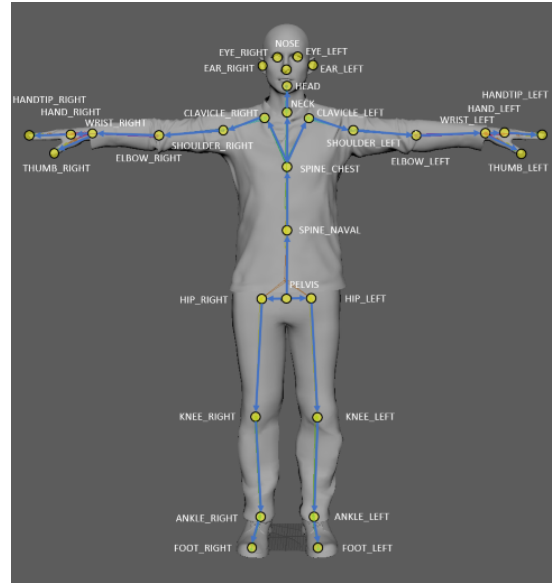


図 4 AzureKinect によって取得できるスケルトンデータ。黄色い点で示す関節をノード、青い矢印をエッジとする木構造。(出典:Microsoft <https://learn.microsoft.com/ja-jp/azure/kinect-dk/body-joints>)

ズの比較の際、人の身長やカメラからの距離が異なることによるスケールの問題を回避するためである。

#### ポーズ非類似度

ポーズ非類似度は、ふたつのポーズの非類似性の指標である。本研究では、ふたつのポーズ間の非類似度  $d(p_{i_1,t_1}, p_{i_2,t_2})$  を以下で定義する。

$$d(p_{i_1,t_1}, p_{i_2,t_2}) = \sum_j w_j |v_{i_1,j,t_1} - v_{i_2,j,t_2}| \quad (2)$$

ここで、 $w_j (0 \leq w_j \leq 1)$  は、関節  $j$  の重みであり、所与の実数値である。

#### モーション

モーションとは、一定時間をかけた身体の連続的な動きである。ダンスの振り付けには、ターンやステップといった、典型的なモーションが含まれることが多い。本研究では、時刻  $s$  から時刻  $e < s$  までのダンサー  $i$  のモーション  $m_{i,s,e}$  を、ダンサー  $i$  のポーズの時系列として以下で定義する。

$$m_{i,s,e} = \langle p_{i,s}, \dots, p_{i,t}, \dots, p_{i,e} \rangle \quad (3)$$

$t (s \leq t < e)$  は時刻である。ダンス開始時刻から終了時刻までのダンス全体も、ひとつのモーションである。

#### 基準モーションと比較モーション

本研究では、同一の振り付けの 2 つのダンスモーションを比較する。2 つのダンスモーションのうち、基準とするモーションを、基準モーション、比較対象とするモーションを比較モーションと呼ぶ。

#### 時間的なずれ

本研究において、図 5 のように、2 つのダンスモーションを比較した際に、基準モーションに対して、比較モーションに動き出しのタイミングの違いや、動きの速度の違いが生じた場合、比較モーションに時間的なずれが生じているという。2 つのダンスモーション間に時間

的なずれが生じているとき、基準モーションの該当時刻のポーズと、同時刻の比較モーションのポーズは異なるものの、ポーズ非類似度が相対的に小さい、同じポーズをしていると思われる別の時刻が存在する。

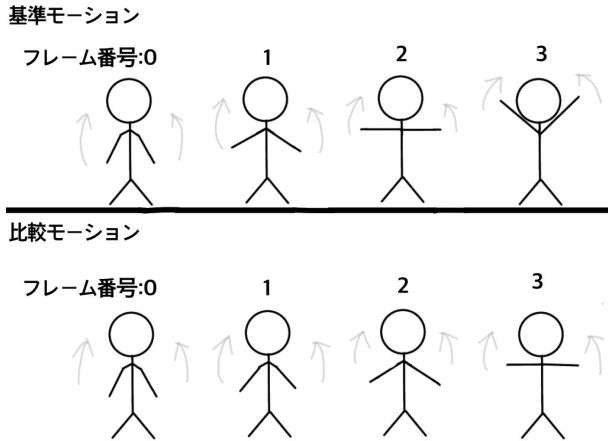


図5 時間的なずれの例。腕を挙げるモーションをしていて、基準モーションに対して、比較モーションに時間的な遅れが生じている。

空間的なずれ

本研究において、図6のように、2つのダンスモーションを比較した際に、基準モーションに対して、比較モーションの動きそのものが異なり、モーションの一瞬を切り取ったときのポーズが異なるとき、比較モーションに空間的なずれが生じているという。2つのダンスモーションに空間的なずれが生じているとき、比較モーションのどの時刻のポーズも、基準モーションの該当時刻のポーズとポーズ非類似度が相対的に大きい。

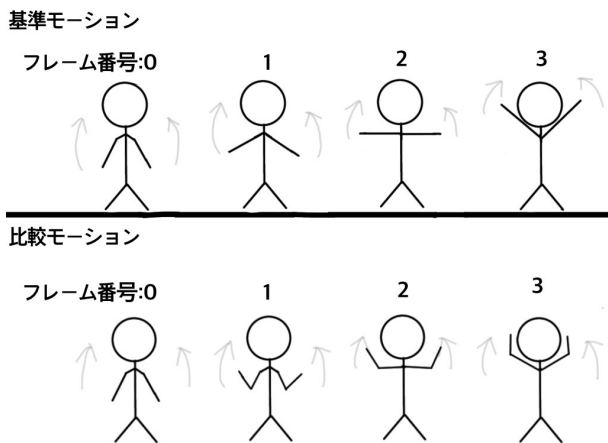


図6 空間的なずれの例。腕を挙げるモーションをしていて、基準モーションに対して、比較モーションの腕の形が異なり、空間的なずれが生じている。

4.2 ダンスモーションの遅れの算出手法

Dynamic Time Warping (DTW) は、2つの時系列データ間で、境界条件、連続性、単調性を満たし、フレームの要素間の非類似度の合計が最小となるよう、各フレームを対応付ける手法である。このとき、ある時系列データの一点の時刻に対して、別の時系列

データの複数点の時刻を対応付けることができるため、時間方向に非線形な対応付けが可能である。そのため、長さの異なる2つの時系列データの対応付けも行うことができる。得られた対応付けにおける非類似度の合計をDTW距離という。長さ  $n$  の時系列データ  $X = (x_0, x_1, \dots, x_{n-1})$  と、長さ  $m$  の時系列データ  $Y = (y_0, y_1, \dots, y_{m-1})$  について  $X$  の時刻  $i_k$  と  $Y$  の時刻  $j_k$  対応関係を  $w_k = (i_k, j_k)$  ( $0 \leq i_k \leq n-1, 0 \leq j_k \leq m-1$ ) と表し、その非類似度を  $\delta(w_k)$  とする。同様に全ての点の総当たりの対応関係の集合  $W = (w_0, w_1, \dots, w_K)$  とするとき、DTW距離は以下の式で与えられる。

$$DTW(X, Y) = \min_w \sum_{k=0}^K \delta(w_k) \tag{4}$$

本研究では、DTWの結果、系列  $X$  の要素  $x_i$  に対応付けられた、長さ  $n_i$  の  $Y$  の部分系列を  $\langle y_{j_i}, y_{j_i+1}, y_{j_i+2}, \dots, y_{j_i+n_i-1} \rangle$  として、 $X$  のフレーム  $i$  に対する、 $Y$  のDTW対応フレーム  $j_{i,X,Y}$  を以下で定める。

$$j_{i,X,Y} = \operatorname{argmax}_{j_i \leq j < j_i+n_i} |i - j| \tag{5}$$

これは、比較モーションの時刻  $j$  のポーズ  $y_j$  に対して複数の時刻の基準モーションのポーズが対応付けされた場合、遅れ時間の絶対値が最大になる一点のみを対応点として採用し、比較モーションの各時刻に対して対応する基準モーションの時刻が一点になるようにする処理である。このとき、遅れ時間の絶対値が最大になる一点を採用するのは、時間的なずれの評価において、取りうる値の中で最も厳しい評価をするためである。

また、 $x_i$  の系列  $Y$  に対する遅れ時間(フレーム数)を以下で定める。遅れは正、進みは負となる。

$$\operatorname{lag}_{i,X,Y} = i - j_{i,X,Y} \tag{6}$$

図7に提案手法の概略図を示す。基準モーション、比較モーションそれぞれ  $n = m = 10$  について、DTWによる対応付けが行われている。赤い線で示す、比較モーションの4フレーム目には、基準モーションの2~7フレーム目までの5フレームが対応付けられている。この場合、遅れ時間の絶対値が最大になる一点のみを対応点とするため、比較モーションの4フレーム目の対応点は、基準モーションの7フレーム目となり、遅れ時間は-3である。

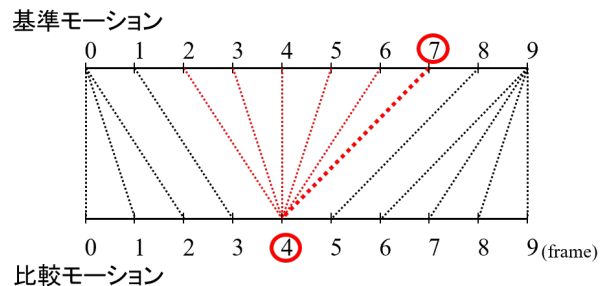


図7 提案手法の概略図

また,  $X$  の  $Y$  に対する時間的なずれの総計  $total\ lag_{X,Y}$  を以下で定める.

$$total\ lag_{X,Y} = \sum_{0 \leq i < |X|} |lag_{i,X,Y}| \quad (7)$$

上記の各値を,  $X$  に比較モーション,  $Y$  に基準モーション, 非類似度にポーズ類似度を適用して求める.

## 5 実験

提案手法をダンスモーションキャプチャデータ可視化アプリケーションとして実装した(図 8). 画面左の各折れ線グラフは, x 軸をフレーム番号とする各指標の遷移を示し, 対象とする指標は (a) 遅れ時間, (b)DTW 対応フレーム間と同時刻フレーム間の各非類似度, (c) 基準モーションおよび (d) 比較モーションの各関節の速度 (m/s) である. 画面下のシークバーでフレーム番号を指定可能であり, 画面右の棒人間 (e1-4) は, 指定フレームおよび関連フレームにおける, 基準および比較モーションのポーズを示す.

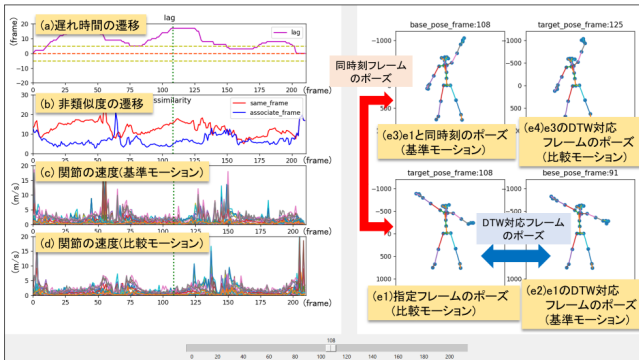


図 8 可視化アプリケーションのスクリーンショット画像

提案手法とシステムの有用性を確認する実験を行った. グループでの実験を見送り, 著者が, 時空間的なずれが生じるよう意識して, 4 種類の振り付けのダンスをそれぞれ複数回踊り, モーションデータを取得した.

まず, 基本的な単一の動きの短い振り付け (30fps で約 200 フレーム) について, 比較モーションの時間的なずれが大きくなるよう踊った結果を図 9 に, 空間的なずれが大きくなるよう踊った結果を図 10 に示す. それぞれ, (a) は遅れ時間の遷移, (b) は同時刻フレーム間と DTW 対応フレーム間の非類似度の遷移である. 図 9 と図 10 を比較すると, 時間的なずれが相対的に大きいこと, 遅れ時間の絶対値が相対的に大きいこと, 同一フレーム間の非類似度に対し, DTW 対応フレーム間の非類似度が小さいことが分かる. 以上より, 空間的なずれと時間的なずれのある各区分は, 区別して抽出可能と考えられる.

続いて, 複数の基本的なモーションを組み合わせた, 相対的に長いダンスモーション (570 フレーム) に提案手法を適用した. 図 11 のように, 遅れ時間の遷移のグラフに基本モーションの区切りを与えると, ずれ始めとずれ終わりが一つの基本モーションの区切りと一致することが分かる. また, 遅れ時間が正の基本モーションと負の基本モーションが混在することが分かる. これより, 時間的なずれは基本モーション単位で発生すると考えられる.

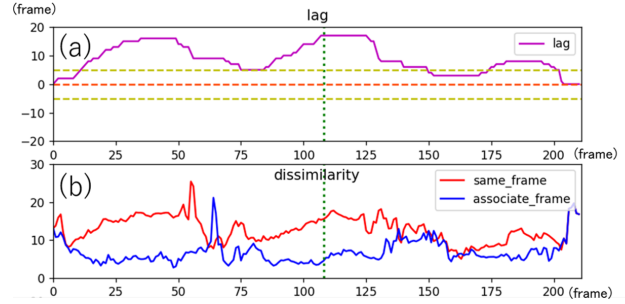


図 9 時間的なずれが大きいモーションデータ間の (a) 遅れ時間と (b) 対応フレーム間 (青線), 同時刻フレーム間 (赤線) のポーズ比類似度の遷移

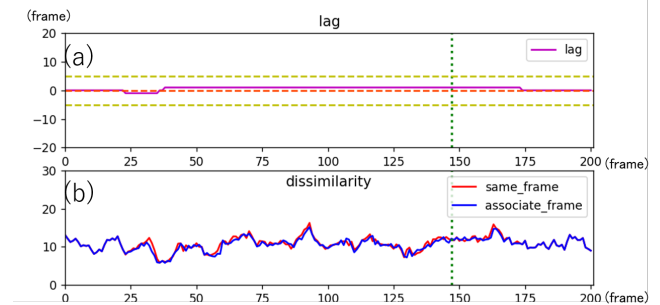


図 10 空間的なずれが大きいモーションデータ間の (a) 遅れ時間と (b) 対応フレーム間 (青線), 同時刻フレーム間 (赤線) のポーズ比類似度の遷移

さらに, 同様の長尺のダンスモーション全体における, 時間的なずれの大きさの指標として, (a)DTW 距離, (b) 同時刻フレーム間のポーズ非類似度の総計, (c) ずれ区間の個数, (d) 時間的なずれの総計を比較した結果を, 図 12 に示す. 各指標とも, 同一の基準モーションに対する 5 つの比較モーション (D1-D5) の結果であり, D3 のみ時間的なずれが大きくなるよう意識して踊っている. この結果より, 上記の指標として, DTW 距離は不適切であり, 時間的なずれの総計 (提案手法) がもっとも適切であると考えられる.

## 6 結論と今後の課題

本研究では, グループダンス練習支援のために, モーションキャプチャデータを用いて, グループメンバーのダンスモーション間のずれを可視化することを目的として, ポーズの非類似度を用いた DTW による各フレームの対応付けに基づき, モーション間の時間的なずれを算出する手法を提案した. 提案手法を可視化アプリケーションとして実装し, 多数のダンスモーションデータに適用して, 時間的なずれの可視化を行った. また, ダンスモーション間のずれに関する指標の比較を行い, 提案手法により算出される, 各時刻の遅れの総計が時間的なずれを, 反映することを確認した.

今後の課題として, グループ実験による提案手法の有用性の分析, 各関節の速度, 加速度, 移動距離, 関節の階層構造等の新たな指標の活用, 提案した各指標と, 主観的印象との関係の分析等が挙げられる.

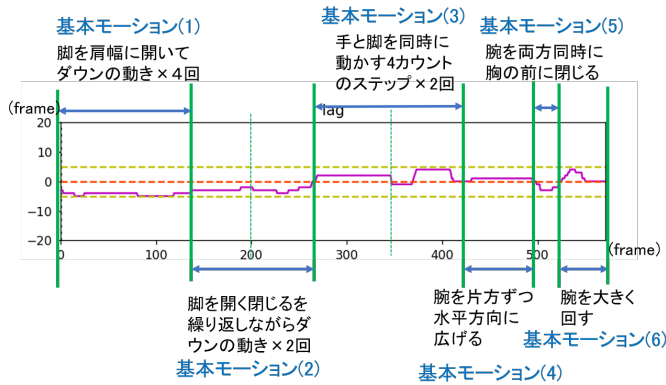


図11 基本モーションを組み合わせた長尺のダンス  
モーション間の遅れ時間の遷移

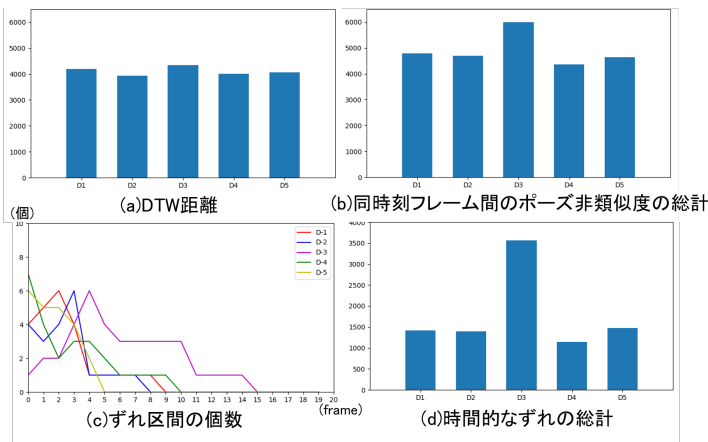


図12 時間的なずれに関する指標の評価

#### 参考文献

- [1] Q. Zhou, CC. Chua, J. Knibbe, J. Goncalves, E. Velloso, Dance and Choreography in HCI: A Two-Decade Retrospective, Proc. ACM CHI 2021, 1-14, 2021
- [2] J. Buchmüller, D. Jäckle, E. Cakmak, U. Brandes, D. A. Keim, MotionRugs: Visualizing Collective Trends in Space and Time, Transactions on Visualization and Computer Graphics, 25, 1, 76-86, 2019
- [3] V. Arpatzoglou, A. Kardara, A. Diehl, B. Flueckiger, S. Helmer, R. Pajarola, DanceMoves: A Visual Analytics Tool for Dance Movement Analysis, Proc. EuroVis 2021, 1-5, 2021
- [4] J. Buchmüller, D. Jäckle, E. Cakmak, U. Brandes, D. A. Keim, MotionRugs: Visualizing Collective Trends in Space and Time, IEEE Transactions on Visualization and Computer Graphics, 25, 1, 76-86, 2019
- [5] Z. Zhou, A. Xu, K. Yatani, SyncUp: Vision-based PracticeSupport for Synchronized Dancing. Proc. ACMIMWUT, 5-3, 1-25, 202
- [6] 川西真美, 土田修平, 伊藤貴之: ダンスモーションの反復練習の可視化とその有効性の検証, 情報処理学会インタラクティブシンポジウム2022論文集, 649-654, 2022
- [7] S. Jang, N. Elmqvist, K. Ramani, MotionFlow: Visual Abstraction and Aggregation of Sequential Patterns in Human Motion Tracking Data, Transactions on Visualization and Computer Graphics, 22, 1, 21-30, 2016
- [8] Y. Sakamoto, S. Kuriyama, T. Kaneko, Motion map: Image-based retrieval and segmentation of motion data. Proc. ACM SIGGRAPH/Eurographics Symposium on Computer Animation, 259-266. 2004