

# 多次元時系列データ間マッチングによる 動き類似検索手法

菅野 滉介<sup>†</sup> 方 宇<sup>†</sup> 黄 宏軒<sup>†</sup> 川越 恭二<sup>†</sup><sup>†</sup>立命館大学 情報理工学部

## 1 はじめに

人の動きをデジタルデータとして計算機に取得できるモーションキャプチャが、映画、CD、ゲームをはじめ、医療やスポーツなどの分野で活用されている [1]。

また、一度取得したモーションキャプチャデータを検索再利用する研究 [2] が行われている。その方法には、映像を用いた方法や動作モデル/骨格モデルを用いた方法がある。後者の動作モデル/骨格モデルを用いた典型的な方法は、モーションキャプチャから得られる各節の3次元時系列データを用いて、時系列データのマッチングを行う方法である。しかし、既存の時系列データにおける類似度算出方法では、正確に骨格モデルが作成できれば高精度に算出可能であるものの、各節の情報である複数の3次元時系列データのマッチング処理に時間がかかるという問題点がある。

本研究では、モーションキャプチャデータ検索システムのための、動き類似検索手法の提案を行う。多くの処理時間を必要とする骨格モデル上でのマッチング処理を行うのではなく、骨格モデルを用いることなく3次元データのままにマッチングを行う方法を基本とする。このとき、処理時間の短縮は期待できるものの精度の低下を引き起こすことが予想できる。そこで、本研究では、複数の3次元時系列データを統合した多次元時系列データから特徴点となる時系列のみを抽出し、多次元データ間類似度を算出することで、精度の維持とともに処理時間のさらなる軽減を可能とする。

## 2 モーションキャプチャデータ検索に関するこれまでの研究

Bastian Demuth ら [2] は、骨格モデルを用いて、新しい特徴の構築を行いモーションキャプチャデータを検索する手法を提案した。彼らは、モーションキャプチャデータの足や腰の位置から幾何学的特徴を定義し、モーションキャプチャデータの次元数の削減を行うことで、DTWを用いた際の問題点である処理時間の削減を可能とした。また、入力にGUIを用いることでユーザの支援も可能とした。

しかし、DTWは、ユークリッド距離での座標値を用いるため、たとえデータ間の軌跡が類似していても、

データ間の座標値が大きく異なる場合、類似度が適切に求めることができないという問題がある。

本研究では、ベクトルのコサイン類似度を用いているため、座標値が大きく異なっても、ベクトルが類似していれば正しく類似度を求めることが可能である。

## 3 モーションキャプチャデータ検索手法

本研究では、モーションキャプチャデータを節ごとの複数の3次元時系列データとして扱うのではなく、各節の3次元時系列データを統合した多次元時系列データとして扱う。

### 3.1 A-LTK

A-LTK(Multi-dimensional time series Approximation with use of Local features at Thinned-out Keypoints)[3] は、処理時間を軽減可能なように考案された時系列データの類似度算出方法である。A-LTKの特徴は、2つある。まず、時系列データのキーとなるポイントのみを抽出し、時系列データを間引くことで処理時間を抑える。次に、その間引いたデータを用いて、局所特徴ベクトルを生成し、新たな時系列データを構築する。局所特徴ベクトルの生成方法には、3つの方法がある。

キーポイントの選別方法は、2つの方法の組み合わせで構成されている。以下に記述する方法では、時系列データを  $V = v_1, v_2, \dots, v_n$  と定義する。

[キーポイント選別方法 1]

点  $v_i$  の前後の点との平均を  $u_i$  とする。次に、点  $v_i$  とその平均である  $u_i$  との差分  $d_i$  を求める。先ほど求めた  $d_i$  の絶対値が閾値より大きい時、その点をキーポイントとする。

$$u_n = \frac{v_{i-1} + v_i + v_{i+1}}{3} \quad (1)$$

$$d_i = u_i - v_i \quad (2)$$

$$\text{keypoint} = \{v_i \mid \|d_i\| > \quad \} \quad (3)$$

[キーポイント選別方法 2]

点  $v_i$  と  $v_{i-1}$  の差分を  $v_i$  とする。次に、 $v_{i+1}$  と  $v_i$  の差分を  ${}^2v_i$  とする。先ほど求めた  ${}^2v_n$  の絶対値が閾値より小さい時、その点をキーポイントと

Motion similarity search using the multidimensional time-series data matching

Kousuke SUGANO<sup>†</sup>, Yu FANG<sup>†</sup>, Hung-Hsuan HUANG<sup>†</sup>, Kyoji KAWAGOE<sup>†</sup>

<sup>†</sup>Ritsumeikan University

する．

$$v_i = v_i - v_{i-1} \quad (4)$$

$${}^2v_i = v_{i+1} - v_i \quad (5)$$

$$\text{keypoint} = \{v_i \mid \| {}^2v_i \| < \epsilon\} \quad (6)$$

2つの方法から抽出されたキーポイントを元に，時系列データ  $V' = v'_1, v'_2, \dots, v'_m$  を生成する．

次に，局所特徴ベクトル生成方法を記述する．局所特徴ベクトル構築には，パラメータがある．そのパラメータによって，局所特徴ベクトルの次元数が変化する．以下に記述する方法では，そのパラメータを  $t$  とする．

[局所特徴ベクトル構築方法]

一つ目である式 (7) は，ある点  $v'_i$  の前後の点を結合し，新たなベクトル  $f'_i$  を構築する．二つ目の方法である式 (8) は，一つ前の点との差分を取り，局所特徴ベクトルを構築する．三つ目の手法である式 (9) は，先述した手法の組み合わせである．

$$f'_i = (v_{i-t}, \dots, v_i, \dots, v_{i+t}) \quad (7)$$

$$f''_i = (v_{i-t+1} - v_{i-t}, \dots, v_i - v_{i-1}, \dots, v_{i+1} - v_i, \dots, v_{i+t} - v_{i+t-1}) \quad (8)$$

$$f'''_i = (f'_i, f''_i) \quad (9)$$

A-LTK では，類似度算出には，コサイン類似度を用いる．先ほどの方法で構築された特徴ベクトルを  $A = (a_1, \dots, a_n)$  と  $B = (b_1, \dots, b_m)$  と仮定する． $A$  と  $B$  の特徴ベクトル間の類似度を算出するための関数  $SIM(A, B)$  は，以下のように定義される．

$$SIM(A, B) = \cos(\text{head}(A), \text{head}(B)) \quad (n = 0 \text{ and } m = 0)$$

$$SIM(A, B) = 0 \quad (n = 0 \text{ or } m = 0) \quad (10)$$

$$SIM(A, B) = \cos(\text{head}(A), \text{head}(B)) + \text{MAX}(SIM(A, \text{rest}(B)), SIM(\text{rest}(A), \text{rest}(B))) + SIM(\text{rest}(A), \text{rest}(B))$$

### 3.2 A-LTK を用いた部分マッチング

3.1 節で述べた A-LTK を用いるだけでは，全体の類似度しか算出できない．そこで，部分的に類似しているデータを考慮した類似度算出手法を提案する．

本提案手法では，先述した A-LTK にスライディングウィンドウ方式を用いて，多次元時系列データの部分的な類似度の算出を可能とする．スライディングウィンドウ方式では，スライド幅が類似度算出に大きな影響を与える．この時のスライド幅を  $s$  とする．

## 4 評価実験

### 4.1 評価実験の内容

今回は，2つの実験を行う．用いるモーションキャプチャデータは，ダンスを動きごとに分割したモーションキャプチャデータである．実験における A-LTK のパラメータは， $\alpha = 1.0$ ， $\beta = 0.1$ ， $t = 0$  とした．

まず，類似したモーションキャプチャデータが検出可能かを確認する実験 1 を行った．実験 1 に用いたデー

タの数は 25 で 9 つのカテゴリに属する．評価対象として，DTW と AMSS との比較を行った．

次に，スライディングウィンドウを用いることで，部分一致にも適用できるかの実験 2 を行った．9 つのモーションキャプチャデータを用意し，その内 3 つのデータを統合したモーションキャプチャデータを 3 種類作成した．入力として，同一カテゴリの 9 つのデータを用意した．評価対象として，DTW と AMSS にもスライディングウィンドウ方式を適用した．全ての手法で  $s = 20$  のスライド幅で実験を行った．

評価は，最近傍法を用いて行う．入力  $I_k$  に対して，類似度が最近接の  $T_k$  が同一カテゴリ及び同一カテゴリの動きを含んだ複合データの場合に正解とする．

### 4.2 評価実験の結果

表 1, 表 2 に実験 1, 実験 2 の結果を示す．表 1, 表 2 における平均処理時間とは，1 件の入力に対して 1 件のファイルとの類似度を算出する平均時間であり，類似データ検出精度とは，入力データの中から正解データを検出できた割合である．

表 1, 表 2 に示す実験結果から，本提案手法は，精度を維持しながら処理時間が軽減可能であることを示した．

表 1: 類似モーションキャプチャデータ検出

	平均処理時間 (s)	類似データ検出精度 (%)
A-LTK	16.1	76.0
DTW	35.0	4.00
AMSS	32.7	60.0

表 2: スライディングウィンドウを用いた部分一致検出

	平均処理時間 (s)	類似データ検出精度 (%)
A-LTK	7512.7	55.6
DTW	14084	22.2
AMSS	12109	55.6

## 5 おわりに

本研究では，新たな多次元時系列データの類似度算出手法を用いたモーションキャプチャデータ検索における処理時間の軽減を目的とした手法を提案した．今後は，データ量を増加して実験を行う．その結果を元に更なる精度向上と処理時間の軽減を行う予定である．

### 参考文献

- [1] 糟谷他. サッカー選手を対象とした一人称視点映像のための仮想カメラ運動モデル. 通信学会論文誌, Vol.J97-D, No.9, pp. 1385–1393, 2014.
- [2] Demuth et al. An information retrieval system for motion capture data. Vol. 3936 of *LNCS*, pp. 373–384. 2006.
- [3] Yu Fang et al. Multi-dimensional time series approximation using local features at thinned-out keypoints. In *ICCSIT 2014*.