

モーションキャプチャデータの類似検索に関する一手法

Similarity Retrieval for Motion Capture Data

竹林 佑介† 西尾 孝治‡ 小堀 研一‡
Yusuke Takebayashi Koji Nishio Ken-ichi Kobori

1. はじめに

近年、映画やゲームなどの映像制作分野やバーチャルリアリティ分野などの3次元CGキャラクターアニメーションにモーションキャプチャが広く利用されている。しかし、モーションキャプチャシステムは、高価な専用機器や専用の技術が必要になるため、利用できる環境が限られてしまう。そこで、一度取得したモーションキャプチャデータを再利用する試みが行われている。また、モーションキャプチャデータを販売、配布されることも多くなった。そのため、モーションキャプチャデータをデータベース化して再利用することが可能になった。しかし、膨大な数のモーションキャプチャデータから目的の動作を目視で検索することは非常に困難である。

そこで、モーションキャプチャデータを再利用するためにモーションキャプチャデータの効率的な検索が必要になる。従来法として、モーションキャプチャデータから動作の特徴を抽出し、検索を行う手法^[1]が提案されているが、必要となる動作の特徴をユーザが指定する必要がある、自動で動作の特徴抽出が行われなかったといった問題がある。

そこで本研究では、モーションキャプチャデータから関節の位置の特徴と動作全体の特徴を抽出し、これら2種類の特徴を利用することで類似検索を行う手法を提案する。

2. 提案手法

提案手法は、検索キーをモーションキャプチャデータとし、検索キーに類似した動作を検索する。また、モーションキャプチャデータから2種類の特徴を抽出し、2種類の特徴を合成することで類似検索を行う。

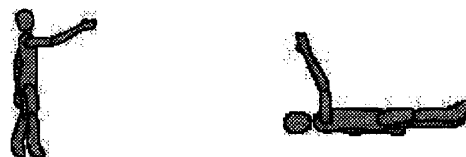
まず、モーションキャプチャデータから動作の局所的な特徴である関節の位置の特徴と、動作の大局的な特徴である動作全体の特徴を抽出する。抽出した2種類の特徴をモーションキャプチャデータの特徴量とする。そして、検索キーのモーションキャプチャデータに対して求められた特徴量と、データベース内のモーションキャプチャデータに対して求められた特徴量を比較し、類似している数種類のモーションキャプチャデータを検索する。以下に、類似検索を行う方法について述べる。

2.1 関節の位置の特徴

人間が目視で類似動作を検索する場合、手や足といった局所的な部位の動きに着目し、類似動作を検索していると考えられる。そこで提案手法では、このような局所的な部位の先端にある関節の位置を局所的な特徴として利用し、関節の位置の特徴を抽出する。以下、部位の先端にある関節を末端関節とする。また提案手法で着目する局所的な部位は、人間の動作の特徴を表していると考えられる頭、両手、両足の5箇所とする。提案手法で得られる関節の位置

の特徴は、“腕を上げる”や“脚を上げる”といった局所的な変化を表す。例えば、図1に示すような“手を挙げる”姿勢の場合、同図(a)、(b)とも同じ特徴量となる。以下で、このような特徴量を計算する方法を述べる。

まず、各部位の付け根にある関節のローカル座標系を基準とし、可動範囲を考慮した固有の空間を作成する。ここで作成する空間は、各部位の付け根から末端までの長さを半径とする部分球である。次に、作成した空間を緯度、経度を用いて量子化する。そして、量子化した各領域に対して、関節の位置の特徴量を計算する。関節の位置の特徴量は、固有の空間上にプロットしたモーションキャプチャデータの全フレームにおける各末端関節の位置を利用して求める。量子化した各領域に存在する末端関節の位置から求めた x, y, z 値についての標準偏差を関節の位置の特徴量とする。



(a) 直立状態 (b) 仰臥状態

図1 手を挙げる姿勢

2.2 動作全体の特徴

人間が目視で動作を検索する場合、2.1節で述べた局所的な部位の動きだけでなく、動作そのものの特徴も考慮していると考えられる。そこで提案手法では、動作全体の特徴をモーションキャプチャデータの大局的な特徴として抽出する。提案手法で得られる動作全体の特徴は、“前進している”や“その場に留まっている”といった大局的な変化を表す。

提案手法では、腰関節の位置を利用して動作全体の特徴量を計算する。モーションキャプチャデータの全フレームにおける腰関節の位置から求めた x, y, z 値についての標準偏差を動作全体の特徴量とする。

2.3 類似度判定

検索キーとなるモーションキャプチャデータをA、データベース内のモーションキャプチャデータをBとしたとき、2種類の特徴量を用いて類似度を判定する方法を以下に述べる。

まず、関節の位置の特徴を用いた類似度を求める。モーションA、Bの量子化した m 番目の領域において、末端関節 j の位置から求められた標準偏差をそれぞれ $ASD(i, j, m)$ 、 $BSD(i, j, m)$ ($i \in \{x, y, z\}$) とすると、関節の位置の特徴を用いた類似度LSimは式(1)で表される。

† 大阪工業大学大学院情報科学研究科

‡ 大阪工業大学情報科学部情報メディア学科

$$LSim = \frac{1}{J_n} \cdot \frac{1}{L+1} \sum_{j \in J} \sum_{k=0}^L \sum_{m=1}^{R_k} S(j, m) \quad (1)$$

$$S(j, m) = \frac{1}{3} \cdot W(j, m) \sum_{i \in \{x, y, z\}} \frac{\min(ASD(i, j, m), BSD(i, j, m))}{\max(ASD(i, j, m), BSD(i, j, m))}$$

$$W(j, m) = \frac{AP_m^j + BP_m^j}{2}$$

ここで、 J は末端関節の集合、 J_n は末端関節の数、 L は量子化レベル、 R_k は量子化レベル k における領域数、 AP_m^j 、 BP_m^j は、量子化した m 番目の領域における末端関節 j の出現確率である。また、量子化レベルとは、量子化する領域数を 2^n としたときの n のことである。

次に、動作全体の特徴を用いた類似度を求める。モーションA、Bの腰関節の位置から求められた標準偏差をそれぞれ $A\sigma_i$ 、 $B\sigma_i$ とすると、動作全体の特徴を用いた類似度CSimは式(2)で表される。

$$CSim = \frac{1}{3} \sum_{i \in \{x, y, z\}} \frac{\min(A\sigma_i, B\sigma_i)}{\max(A\sigma_i, B\sigma_i)} \quad (2)$$

最後に、2種類の類似度LSim、CSimを用いて、モーションA、B間の類似度Simを式(3)で計算する。

$$Sim = t \cdot LSim + (1-t) \cdot CSim \quad (0 \leq t \leq 1) \quad (3)$$

ここで、類似度LSim、CSim及びSimは0から1までの値をとり、1に近いほど類似度が高い。

3. 実験・考察

提案手法の有効性を検証するために実験を行った。検索キーとなるモーションキャプチャデータを図2に示し、6種類のモーションキャプチャデータをデータベースとして図3に示す。また、35人の被験者に対して主観評価実験を行った。提案手法による検索結果と主観評価による検索結果を表1に示す。主観評価における順位は、被験者それぞれが決定した順位の合計値の小さい順に並べている。主観評価における括弧内の数字はその合計値である。また、提案手法における括弧内の数字は、式(3)で求めたSimの値である。提案手法は、量子化レベルを4に設定し、実験を行った。また、式(3)における t の値を0.0、0.5、1.0と変化させて実験を行った。ここで、 t が0.0のときは、大局的な特徴量だけを用いた類似度となり、 t が1.0のときは、局所的な特徴量だけを用いた類似度となる。

表1より、主観評価における上位3モーションはC、E、Fである。 $t=1.0$ のときの局所的な特徴量を用いた場合の上位3モーションは、E、C、Fであった。主観評価と類似した結果を得ることができたが、1位と2位が主観評価と異なってしまった。これは、モーションEの両手の動きと検索キーとなるモーションの片手の動きが類似していたため、モーションCよりも類似していると判断されたと考えられる。 $t=0.0$ のときの大局的な特徴量を用いた場合の上位3モーションは、F、C、Eであった。この場合も主観評価と類似した結果を得ることができたが、同じ直立姿勢で行うモーションF、CとモーションEとの差が大きくなった。これは、同じ直立姿勢ではあるが、モーションEだけが大きく体を動かす動作であるため、類似度が低くなったと考えられる。これらのことから、2種類の特徴がモーションデータの局所的、大局的特徴を表現できていると考えられる。また、 $t=0.5$ のときの検索結果が主観評価と一致した。この

ことから、2種類の特徴を合成して類似度を求める方法が有効であることが確認できた。

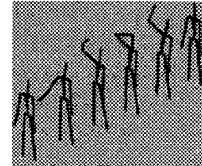
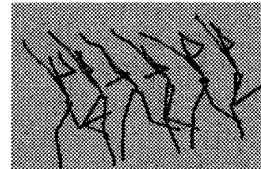
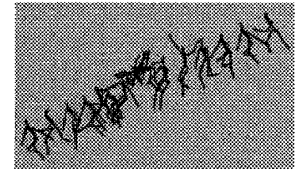


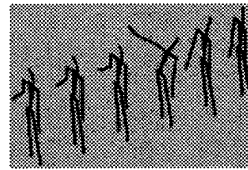
図2 検索対象(右手を振る)



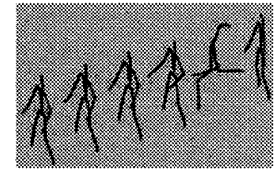
(a) A (パニック状態で走る)



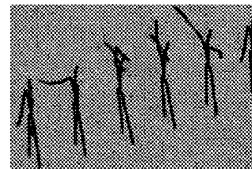
(b) B (前方宙返りをする)



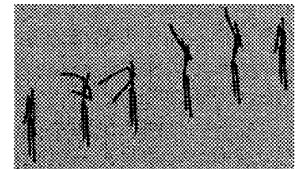
(c) C (頭上の物を取る)



(d) D (ボールを投げる)



(e) E (両手を振る)



(f) F (バンザイをする)

図3 データベース

表1 検索結果

順位	t=1.0	t=0.5	t=0.0	主観評価
1	E (0.517)	C (0.513)	F (0.559)	C (48)
2	C (0.475)	E (0.500)	C (0.553)	E (68)
3	F (0.272)	F (0.416)	E (0.482)	F (104)
4	D (0.249)	D (0.205)	A (0.212)	D (145)
5	B (0.176)	A (0.187)	D (0.162)	A (166)
6	A (0.162)	B (0.151)	B (0.126)	B (204)

4. おわりに

本研究では、モーションキャプチャデータから動作の局所的な特徴である関節の位置の特徴と、動作の大局的な特徴である動作全体の特徴を抽出し、2種類の特徴を合成することで類似検索を行う手法を提案した。実験より、2種類の特徴を合成することで主観に一致した検索を行えることを確認した。

今後の課題として、より多くのデータを用いて提案手法と主観評価との順位的相关及びSimの値を検証し、精度の高い類似検索を行うことが挙げられる。

<参考文献>

- [1] Meinard Müller, Tido Röder, Michael Clausen, "Efficient Content-Based Retrieval of Motion Capture Data", Proceedings of ACM SIGGRAPH 2005, pp.677-685, 2005