

モーショキャプチャを用いた動作判別手法の検討
A Study of human motion recognition method using motion capture

○幸山侑磨, アデルジャンイミティ, 萩原由香里, 萩原義裕, 三好扶
○Yuma Kouyama, Adiljan Yimit, Yukari Hagihara, Yoshihiro Hagihara, Tasuku Miyoshi

岩手大学
Iwate University

キーワード: モーションキャプチャ(Motion capture), 動作判別(Motion recognition),
マハラノビス距離(Mahalanobis distance), クォータニオン(Quaternion), リハビリテーション(Rehabilitation)

連絡先: 〒020-5881 岩手県盛岡市上田 4-3-5 国立大学法人 岩手大学 工学部 機械システム工学科 萩原義裕, TEL.:
(019) 621-6400, FAX.: (019) 621-6400, E-mail: dhag@iwate-u.ac.jp

1. 諸言

運動療法の中には理学療法士により適切な運動の指導・練習を行うものがある。これについて、近年装置の低価格化が進んでいるモーショキャプチャを用いた運動検出の活用が注目されている。運動検出の導入により、適切な運動を基準とし、実際に行われた運動と比較することにより指導された運動を再現することができているか判別を行うということが可能になる。

一般に非接触型のモーショキャプチャ装置では関節の位置情報のみを用い動作を判別する。高精度な装置ではハードウェア上で関節の位置情報を正確に捉える工夫がなされているが、3次元データのみを使い関節位置を推定する低価格な装置では、高精度な測定を期待できない。

本研究では、位置情報から関節の角度情報を示すクォータニオンを算出し、動作認識のための特徴量として用いることで認識の高精度化を図る。

判別を行う手法には様々なものがあるが、今回はマハラノビス距離を用いた判別手法について検討を行った。

2. クォータニオン

本研究では、特徴量として従来手法で用いられてきた位置情報に変わり、クォータニオン(4元数)を特徴量として用いる。クォータニオンを式(1)に示す。ここで q_0 を Q のスカラー部分、 \mathbf{q} を Q のベクトル部分と呼ぶ。

$$Q = (q_0, q_1, q_2, q_3) = (q_0, \mathbf{q}) \quad (1)$$

また、 Q は虚数単位を用いて式(2)のように

表すことができるが、 i, j, k には交換法則が成立せず、式(3)のように表わすことができる。

$$Q = s + iu + jv + kw \quad (2)$$

$$\begin{aligned} ii = hh = kk &= -1 \\ ij = k, ik &= i, ki = j \\ ji &= -k, kj = -i, ik = -j \end{aligned} \quad (3)$$

$\mathbf{v} = iu + jv + kw$ とすると、スカラー部とベクトル部に分けることができ、式(4)のようになる。また、共役クォータニオンは式(5)で表わされる。

$$Q = s + \mathbf{v} \quad (4)$$

$$\bar{Q} = s - iu - jv - kw \quad (5)$$

クォータニオンは3次元ユークリッド空間における回転を記述するという幾何学的意味を持ち、回転後の位置ベクトルクォータニオンと共役クォータニオンとの積により式(6)のように表現することができる。

$$\mathbf{x}' = \mathbf{q}\mathbf{x}\bar{\mathbf{q}} \quad (6)$$

回転前の位置ベクトルを \mathbf{x} 、回転後の位置ベクトルを \mathbf{x}' とし、単位ベクトル \mathbf{u} を回転軸として角度 θ だけ回転させた場合、クォータニオンを用いて式(7)のように表現することができる。

$$\mathbf{q} = \cos \frac{\theta}{2} + \mathbf{u} \sin \frac{\theta}{2} \quad (7)$$

$$\mathbf{u} = iu + jv + kw \quad (8)$$

3次元ユークリッド空間での回転を表現するクォータニオンは回転前の位置ベクトル \mathbf{A} と回転後の位置ベクトル \mathbf{B} を用いて求めることができる回転軸ベクトル \mathbf{u} と回転角 θ は

A, B の外積を用いて式(9)で示される.

$$\begin{aligned}
 u &= \frac{A \times B}{|A \times B|} \\
 &= \frac{1}{|A \times B|} (a_y b_z - a_z b_y, a_z b_x - a_x b_z, a_x b_y - a_y b_x) \\
 \theta &= \sin^{-1} \frac{A \times B}{|A \times B|} \\
 &= \sin^{-1} \frac{\sqrt{u_x^2 + u_y^2 + u_z^2}}{\sqrt{a_x^2 + a_y^2 + a_z^2} \sqrt{b_x^2 + b_y^2 + b_z^2}} \quad (9)
 \end{aligned}$$

3. マハラノビス距離を用いた判別法

マハラノビス距離は多変量分布において確率密度と距離を対応させたものである.

学習データが n 次元のベクトルから構成されたクラス w_1, w_2 から構成されているとき, 判別を行う情報 x がクラス w_i に所属する事後確率を $P(w_k|x)$ で表わすと, $P(w_k|x)$ が最大となる w_i を判別結果とする手法がよく用いられる. $P(w_k|x)$ はベイズの定理より式(10)で表わされる.

$$P(w_i|x) = \frac{P(x|w_i)}{P(x)} P(w_i) \quad (i = 1, 2) \quad (10)$$

ここで $P(x)$ はクラスによらない生起確率であり, $P(w_i)$ はクラス w_i の事後確率である. 式(10)の対数を取り, 共通する部分を削除すると, 識別関数は式(11)のように簡略化することができる.

$$P(x_i) = \max\{\log P(x|w_i)\} \quad (i = 1, 2) \quad (11)$$

確率密度関数 $P(x|w_i)$ が式(11)で表わされる正規分布になる場合を考えると, 式(12)のようになる.

$$\begin{aligned}
 P(x|w_i) &= \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \\
 &\times \exp\left\{-\frac{1}{2}(x - m_i)^t \sum_t^{-1} (x - m_i)\right\} \quad (12) \\
 &\quad (i = 1, 2)
 \end{aligned}$$

ここで m_i, Σ_i はそれぞれクラス w_i の平均ベクトルと分散共分散行列であり, $|\Sigma_i|$ は Σ_i の行列式である. m_i, Σ_i はそれぞれ式(13), 式(14)のように定義される.

$$m_i = \frac{1}{n} \sum_{x \in w_i} x \quad (13)$$

$$\sum_i = \frac{1}{n} \sum_{x \in w_i} (x - m_i)(x - m_i)^t \quad (14)$$

式(12)を式(11)へ代入すると式(16)のように展開することができる.

$$\begin{aligned}
 g(x_i) &= \max\left\{-\frac{1}{2}(x - m_i)^t \sum_i^{-1} (x - m_i) \right. \\
 &\quad \left. - \frac{1}{2} \log \left| \sum_i \right| - \frac{d}{2} \log 2\pi\right\} \quad (15)
 \end{aligned}$$

$$D_M^2(x - m_i) = (x - m_i)^t \sum_i^{-1} (x - m_i) \quad (16)$$

式(15)において式(16)のように $D_M^2(x - m_i)$ とおくと, $D_M(x - m_i)$ は x と m_i のマハラノビス距離となる. 本手法では識別関数 $g(x_i)$ はマハラノビス距離を用いて式(17)で定義される.

$$g(x_i) = \min\{D_M^2(x - m_i)\} \quad (17)$$

4. 実験

4.1 実験装置

本研究では, モーションキャプチャデバイスとして Microsoft 社製モーションキャプチャデバイス, Kinect を使用する. Kinect に以下のセンサが搭載されている.

- RGB カメラ
- 赤外線レーザーパターン投影部
- 赤外線レーザー撮影用カメラ
- マイク

この装置を用いてモーションキャプチャを行う際は赤外線レーザーパターンの歪みから対象物の表面形状を復元し RGB カメラから取得した画像データと組み合わせることで, 特殊な器具を計測対象となる被験者の身体に取り付けることなく骨格情報を取得することができる.

4.2 骨格情報

Kinect から得られた情報から骨格情報を算出するまでの操作は予め設定されており, 骨格情報をスケルトンデータという形で取得することができ, 取得したい関節を指定するだけでその関節の位置情報を取得することができる.

Kinect は奥行き方向についても計測可能であることから身体が傾いても身体が写っていれば骨格情報は追従する. また, 本体内部にグラフィックチップを内蔵し, 画像処理の一部を内部で行うことで PC での処理負荷を低減させており, 激しい運動のキャプチャも可能である.

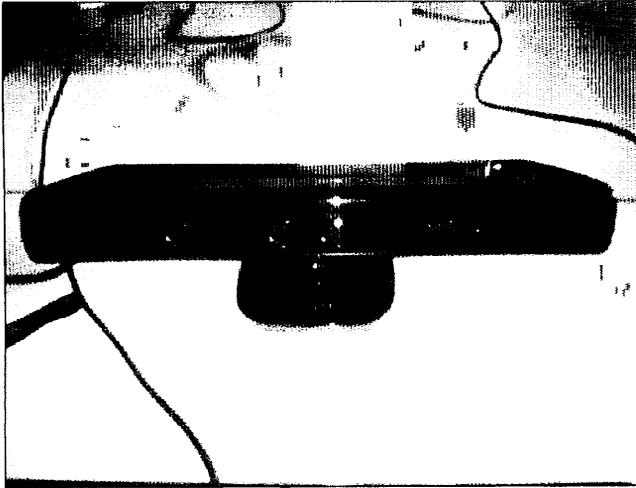


Fig.1 Motion capture device Kinect

Kinect を用いて骨格情報の取得を行うと、20ヶ所の関節の位置ベクトルを取得することができるが、本研究ではその中から

- ・左右の肩
- ・左右の肘
- ・股関節センター

の5ヶ所の位置ベクトルを取得し判別に用いた。

4.3 実験の流れ

本研究では判別を行う動作としてオーバースロー、サイドスロー、アンダースローを対象とした動作の判別を行った。

実験ではまず模範動作としてそれぞれの動作を100回撮影した。その際、1回あたりの撮影時間を3秒とし0.1秒間隔で骨格情報を格納したフレームの取得を行った。その後、取得したフレームから各コマでの各関節の位置ベクトルを抽出しそれぞれクォータニオンを算出した。これらの操作により算出した各動作100回分のクォータニオンは各関節、各コマごとに集積し、その平均値と逆行列化した分散共分散行列を動作ごとに動作判別に用いる学習データとして格納した。

判別では学習データ作成時と同様の時間、間隔で各関節の位置ベクトルを取得し、クォータニオンを算出する。次に、各動作を示す学習データを用いてマハラノビス距離を各関節、各コマごとに算出する。算出後、各コマで、関節ごとに最もマハラノビス距離が小さくなった動作を挙げ、最も多かった動作をそのコマにおける判別結果とする。その後コマごとに挙げられた判別結果についても同様に操作を行い、最も多く判別結果として挙げられたものを最終的な判別結果とした。

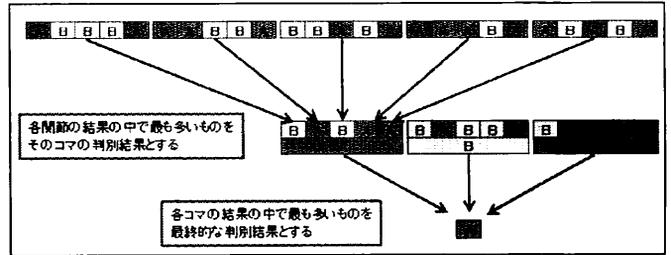


Fig.2 Recognition flow

5. 実験結果・考察

実験では3種の動作(オーバースロー・サイドスロー・アンダースロー)について、それぞれ10回ずつ動作を判別し、それぞれの動作において正答率を求めた。また比較手法として、特徴量として位置ベクトルを用いた判別を行いその結果を比較した。

以下にサイドスローの動作を行った際の判別結果、各動作での提案手法と比較手法の正答率をそれぞれ示す。

実験の結果、いずれの動作においても特徴量としてクォータニオンを用いた判別では高い確率で正しく判別することが分かった。

一方位置ベクトルを特徴量として用いた場合アンダースロー、サイドスローの判別において正答率がそれぞれ、50%、30%にとどまった。フレームごとの正答数を比較すると、クォータニオンを用いた場合では、正しく判別した回数がいずれも過半数を超えていたが、位置ベクトルを用いた場合、たとえ正しく判別したものであっても、フレームごとに検証すると正しく判別した回数が過半数を下回るものがあり、判別が不安定であることが分かった。

位置ベクトルを特徴量として用いた場合に判別の正答率が低下した理由として、モーションキャプチャデバイスと被験者との間の距離が変動することによる影響が考えられる。位置ベクトルを用いた場合、たとえ動作が模範動作に近い場合でも、被験者と装置間の距離や被験者の体格により値が大きく異なる。一方クォータニオンが保有する情報は回転軸のベクトルと回転角に関するもので、いずれも立ち位置や体格等の影響が少ないことから、判別が高精度に行われたと考える。

Table1 Recognition result of sidearm-
throwing
(a) Quaternion

	Result	Number of recognition results for each frames (Total 30 frames)		
		Over	Side	Under
1	○	2	17	11
2	○	2	19	9
3	○	1	21	8
4	○	1	22	7
5	○	0	18	2
6	○	1	18	11
7	○	5	20	5
8	○	0	23	7
9	○	1	21	8
10	○	1	23	6
Accuracy rate	100%			

(b) Position vector

	Result	Number of recognition results for each frames (Total 30 frames)		
		Over	Side	Under
1	○	8	13	9
2	○	6	14	10
3	×	4	4	22
4	×	4	4	22
5	×	3	8	19
6	×	0	11	19
7	×	11	1	8
8	○	3	15	12
9	×	3	5	22
10	×	3	9	18
Accuracy rate	30%			

Table2 Accuracy Rates

Motion	Accuracy rate	
	Quaternion	Position vector
Over	100%	100%
Side	100%	30%
Under	100%	50%

6. 結言

モーションキャプチャを用いた動作判別について、マハラノビス距離を用いた判別を行う場合、クォータニオンを特徴量とした判別が位置ベクトルを用いた判別と比べ高精度な判別が可能であることを示した。

今後の課題は、他の特徴量を用いた場合の判別精度の検証、判別に最適な特徴量の選択である。

参考文献

- 1) 榎本美香,野口広彰, Kinect センサを利用した調理場の行為認識コンポーネントの提案-高次脳機能障害者支援のための環境制御技術の実装に向けて, 言語・音声理解と対話処理研究会 62, 41-44, 2011-07-22
- 2) 竹内 俊貴, 中島 統太郎, 西村 邦裕, 谷川 智洋, 廣瀬 通孝, PRIMA-異なる時間軸上のユーザとのインタラクションを実現するシステム-, 情報処理学会研究報告. EC, エンタテインメントコンピューティング 2011-EC-20(2), 1-6, 2011-05-06
- 3) 杉本論, 訪問理学療法により ADL 能力に改善を示した 慢性期脳卒中患者の 1 例, 理学療法科学, 2003 Vol.18(2003), No.4 pp235-239