

モーションキャプチャを用いた動作判別手法の高精度化 High precision human motion recognition method with motion capture device

○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, 高木基樹
○KOUYAMA Yuma, ADILJAN Yimit, HAGIHARA Yukari, HAGIHARA Yoshihiro, MIYOSHI
Tasuku, TAKAGI Motoki

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), リハビリテーション(Rehabilitation), キネクト(KINECT)

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

近年の高齢者の増加に伴い、怪我等により低下した機能の改善を目的としたリハビリテーションの需要は高まりつつあり、その施術を行う作業療法士などの医療従事者の負担は増加している。こうした負担の増加は医療従事者だけではなく、患者にとっても十分な施術を受けることができないといった影響が現れる。

このような背景から、現在リハビリテーションにおける医療従事者、患者双方の負担軽減を目的とした支援システムの開発が行われている。それらのシステムの中には近年、装置の小規模化を低価格化が進んでいるモーションキャプチャシステムを活用したものがある。[1]

こうしたシステムで用いられる安価なモーションキャプチャデバイスは赤外線等を用いることで距離情報を取得、解析を行うことで対象者の姿勢を判別するため、機器の設置場所が1カ所のみであることや、患者に関節の位置を示すマーカーの装着を必要としないという利点がある一方、オクルージョンによる計測点の消失が発生するなど、大規模なキャプチャデバイスと比較し、姿勢の計測精度が劣る。

また、判別に用いる特徴量はその数を単に増加させても判別精度が向上するわけではなく、計算量の増加を招く。そのため判別の結果を映像に反映させるといったリアルタイムでの処理が求められる場合、判別精度が低下しない範囲で判別に用いる特徴量を制限する必要があるため、複数の特徴量の中から判別を行う学習データ毎に最適なものを選択する手法が必要となる。

本研究では、安価なモーションキャプチャデバイスを用いた動作判別システムを作成し、複数の特徴量から前向き逐次選択法を用いて判別に用いる特徴量を選択した結果を述べ、本手法における課題を検証する。なお本研究では、動作判別の指標としてマハラノビス距離を、判別性能の評価値としてROC曲線下の面積を用いた。

2. システムの概要

本研究で使用した動作判別システムは、モーションキャプチャデバイスを用いて一定間隔で被験者の姿勢を取得し、学習データおよび標本データの素となる動作データを作成し、その後動作判別とその結果を用いて前向き逐次選択法による特徴量の選択を行った。

学習データ、標本データはそれぞれ特徴量として各コマ、各関節毎に取得したクオータニオンをすべて列挙したものを使用し、動作の判別には、学習データと標本データ間のマハラノビス距離を用い、判別精度の評価にはROC曲線下の面積を用いた。

2.1. マハラノビス距離

マハラノビス距離とは、多変量正規分布のもとで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次元ベクトルから構成されるクラス w_1, w_3 から構成されているとき、判別対象である情報 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(1)で表わされる。

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

ここで、 $P(x)$ はクラスに依存しない生起確率で、あり $P(w_i)$ は w_i の事後確率である。式(1)から共通したものを除き簡略化すると、

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

さらに、確率密度関数 $P(x|w_i)$ が式(2)で表される正規分布になる場合、式(3)のようになる。

$$P(x|w_i) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} \times \exp\left\{-\frac{1}{2}(x - m_i)^t \Sigma_i^{-1} (x - m_i)\right\} \quad (3)$$

($i=1,2$)

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 m_i, Σ_i はそれぞれ以下のように定義される。

$$m_i = \frac{1}{n} \sum_{x \in w_i} x \quad (4)$$

$$\Sigma_i = \frac{1}{n} \sum_{x \in w_i} (x - m_i)(x - m_i)^t \quad (5)$$

式(3)を式(2)へ代入すると、式(6)のようになる

$$g(x_i) = \max\left\{-\frac{1}{2}(x - m_i)^t \Sigma_i^{-1} (x - m_i) - \frac{1}{2} \log \left[\sum_i \left| -\frac{d}{2} \log 2\pi \right| \right]\right\} \quad (6)$$

$$D_M^2(x - m_i) = (x - m_i)^t \Sigma_i^{-1} (x - m_i) \quad (7)$$

式(7)のように式(6)を $D_m^2(x-m_i)$ 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。今回の手法では識別関数 $g(x_i)$ はマハラノビス距離を用いて式(8)のように定義される。

$$g(x_i) = \min\{D_M^2(x - m_i)\} \quad (8)$$

2.2. ROC 曲線

ROC 曲線(Receiver Operating Characteristic curve、受信者動作特性曲線)は横軸に偽陽性率(False Positive)を縦軸に真陽性(True Positive)を取り、分類結果の内、陽性とする閾値を変え算出したそれぞれの値をプロットすることにより得られる曲線である。

曲線下の面積は 0.5 から 1.0 の値を取り、値が 1.0 に近づくほど分類器の性能が高いことを示す。

2.3. 前向き逐次選択法

この選択法では、判別結果が最も良好な特徴量を 1 点選択する。その後、選択した点と残りの点を 1 つ組み合わせた 2 個組の中から最も判別性能が良いものを選択する。以降の点についても同様の操作を行い、追加することでその時点で最も性能が高くなる特徴量を逐次的に追加していく。

判別性能の評価では前述のとおり ROC 曲線化面積を用いた。

また、本研究で判別に使用するマハラノビス距離は 1 点では算出することができないため、本研究では、最初に最も判別性能が良好な 2 個組となる 2 点を総当たりで探索した。また、 $n+1$ 点特徴量を選択した場合よりも、 n 点特徴量を選択した場合のほうが判別性能が良好だった場合、以後の選択を打ち切った。

3. 実験

3.1. 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、Kinect を使用する。Kinect は Microsoft 社が開発したモーションキャプチャデバイスであり、RGB カメラに加え赤外線カメラ、赤外線レーザーパターン投影機、マイクを搭載し、赤外線カメラが撮影した赤外線パターンの歪みから対象までの距離を割り出し、その情報と RGB カメラから取得したデータと組み合わせ各関節の位置ベクトルからなる骨格情報を算出する。また、親関節を基準としてときの子関節の向きを回転行列やクオータニオンとして取得することができる。



Fig.1 Motion capture device Kinect

3.2. 動作データ

本研究では、試験動作としてラジオ体操第 1 内で行われる「腕を振って足を曲げ伸ばす運動」と「腕を回す運動」を行い、それぞれ動作 A,B とした。

動作データの取得ではまず、20 代男性 5 名に装置の前で各動作を 100 回程度繰り返し行ってもらい、0.4 秒間隔でその姿勢を取得した。その後、取得したデータの中から動作の開始点を目視により割り出し、各開始点から事前実験より 4 コマを連続して取得し動作データを作成した。

今回取得した特徴量は特徴量選択の際の試行回数を制限するため Kinect for Windows SDK を用いることで取得することができる 20 関節の内、両動作を比較した時、変化が小さいと考えられる 7 関節に絞り、各点の親ボーンを基準とした回転を示すクオータニオンを用いた。取得する特徴量の総数は、7 関節×クオータニオン(4 点)×4 コマ=112 点となる。

3.3. 動作判別

実験では、正解動作群と不正解動作群を用意し、各データを標本とし、学習データとのマハラノビス距離を求め、ROC 曲線を作成した。

本研究では正解動作を動作 A、不正解動作を動作 B として動作判別を行った。この時、動作 A は学習データとして用いるものと正解データとして用いるものの 2 つに分け使用した。

4. 結果・考察

実験の結果、112 点の特徴量の中から 32 点の特徴量を選択し、ROC 曲線化面積は 0.9916 となった。Fig.2 に 2 点から 32 点まで選択する間の ROC 曲線の推移を示す。

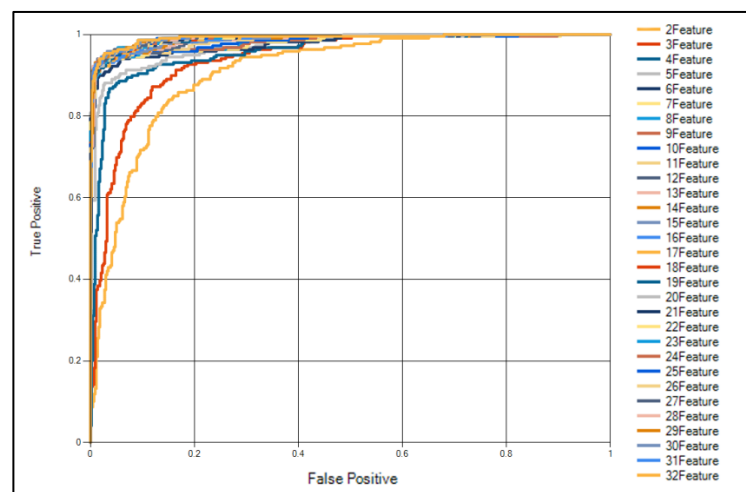


Fig.2 Trends in ROC curve

この選択に要した試行回数は、9161 回であったが、選択した特徴量毎に見ていくと、その大半は最初に行われる、2 点を選択する段階での試行であることがわかる。今後、さらに多くの特徴量の中から最適なものを選択しようとした場合、初期段階の選択はさらに増大すると考えられるため、初期の特徴量選択について少ない試行数を軽減する手法の検討が必要である。

選択した特徴量の数に注目すると、総特徴量数の 3 割ほどの特徴量を選択している。今回判別に用いたマハラノビス距離では、その計算過程で特徴量の数に比例した正方行列を用いるため、特徴量の選択により判別時の計算量の減少が見込まれる。

この実験後、112 点すべての特徴量を用いた判別を行ったところ、ROC 曲線化面積は最大値である、1.0 を示した。このことから、特徴量を 33 点選択したとき一時的に判別性能が低下するものの、さらに使用する特徴量を増加させた場合、再度判別性能が向上する可能性があるものと考えられる。

5. 結言

前向き逐次選択法を用いた特徴量選択を行った時の判別性能と、特徴量選択を用いる際の課題を検証した。

今後の課題は、選択に用いる特徴量を増やした場合更なる試行数の増大が見込まれる初期段階での特徴量選択への対処、遺伝的アルゴリズムを用いた手法など、他の特徴量選択手法を用いた場合の評価などである。

参考文献

- 1) 狐崎直文,安達栄治郎,増田卓,水澤純一,KINECT のリハビリテーション用アプリケーションの実用化,映像メディア学会技術報告 36(48),41-46,2012
- 2) 鳥脇純一郎,認識工学,pp.141-142,コロナ社,東京,1993