

平石研究室

Kinectセンサによる ジェスチャー認識の研究

学籍番号 315158 鶴見 慶史郎

指導教員 平石 広典



目次

1. はじめに
 2. モーションキャプチャとは
 - 2-1. Kinectについて
 3. 卒業研究Aの研究内容
 4. Random Forestとは
 5. 使用機器, ソフトウェア
 6. 実験方法
 - 6-1. ジェスチャーの設定
 - 6-2. データの収集
 - 6-3. データの分析
 7. 実験結果
 - 7-1. 実験結果の比較
 8. まとめ
 9. おわりに
- 参考文献

1. はじめに

KinectやLeapMotionなどのモーションデバイスが登場.

→直感的な動作でゲームやコンピュータの操作ができるようになった.

(Natural User Interfaceの発達)

モーションキャプチャデバイスを利用して特定の動作を認識させるためには、各動作の特徴を個別に入力する必要がある.

→決定木を利用することで、様々な行動を自動的に認識させる手法が提案されており、ロボットの操作を対象として、利用者のポーズを認識する方法が提案されている¹⁾.

卒業研究Bでは、Kinectセンサを利用して、より動きのあるジェスチャーの認識に対して、Random Forestを利用し、認識精度の向上を目的とし、実験的に検証した.

2. モーションキャプチャとは

モーションキャプチャは、現実の人物や物体の動きをデジタル的に記録する技術である。

多くの場合は、物体に装着するマーカーとこれを検出するトラッカーを組み合わせてキャプチャする。同じ方式であればマーカーの数が精度の一因となり、トラッカーは1秒間にトラックできるフレーム数が精度に関係する。

○社会の中で活用されている例.

・医療の分野.

心臓疾患がある人に向けたカロリーメータの開発.

・スポーツの分野.

テニスプレイヤーを対象とした機械学習を用いたトレーニング支援システムの開発²⁾.

2-1. Kinectについて

Microsoftが販売する家庭用ゲーム機「Xbox360」、及びPCにおいて、コントローラやキーボードなどの入力機器を用いずに、身体の動き、ジェスチャー、音声などによる操作を可能とする周辺機器を指す。

カメラに被写体を映す事で、プレイヤーからKinectまでの距離を計測し、プレイヤーの骨格の動きを検出して、ゲーム内のキャラクターの動きに反映させることが可能になる。

Kinectの特徴

- モーションキャプチャ時に着用するマーカー付きスーツ、マーカー検出時に使用するトラッカーを必要としない。



図1. Kinectの外観

3. 卒業研究Aの研究内容

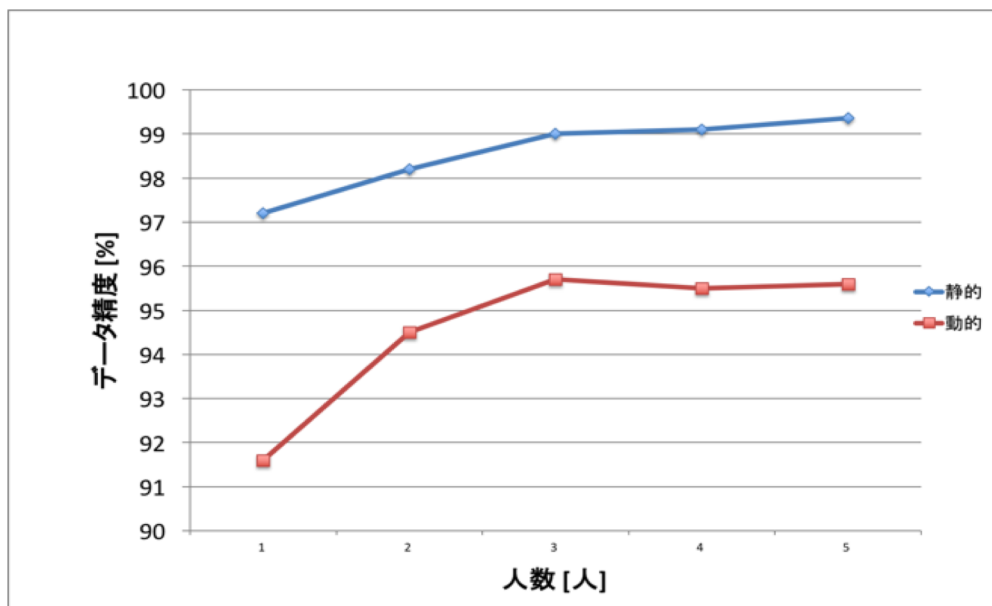


図2.決定木の手法での分類結果

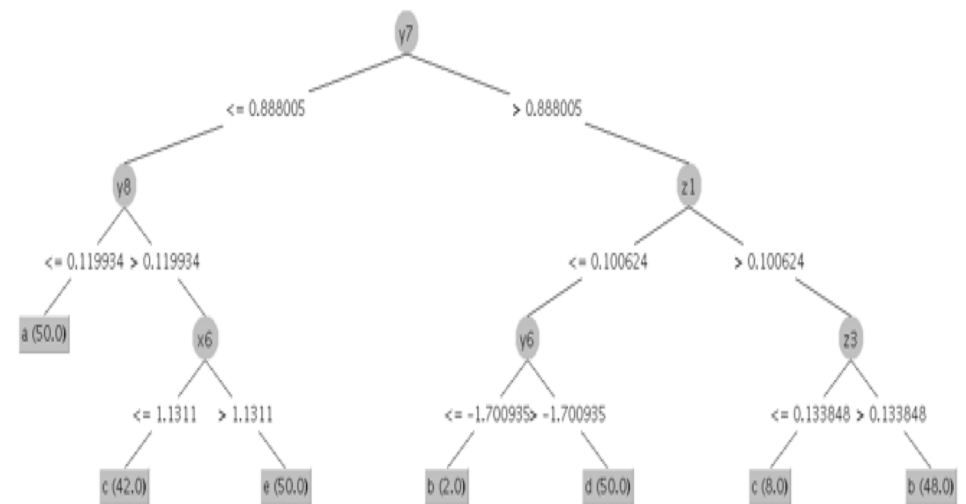


図3.分類結果の決定木

4. Random Forestとは

Random Forestは、決定木を複数組み合わせ、各決定木の予測結果を多数決することによって結果を得る。

○アルゴリズム

1. ランダムにデータを抽出し、複数の決定木を作成する。
2. 決定木を成長させる。
3. 1, 2のステップを指定回数繰り返す。
4. 予測結果を多数決することで分類閾値(ラベル)を決定する。

Random Forestはパラメータが非常に簡単になるという利点がある。

主要パラメータは、サンプリング数と決定木を成長させる際に使用する特徴量の数だけある。推定結果ごとに重要さを計算し、その平均値を変数の重要さと判定することが多い³⁾。

5. 使用機材, ソフトウェア

○ハードウェア

- ・PC ...1台 Mac mini (Mid 2011)
 - プロセッサ 2.3 GHz Intel Core i5
 - メモリ 8 GB 1333 MHz DDR3
 - グラフィックス intel HD Graphics 3000 512MB
- ・Kinect ...1台(XBOX360版) v1

○ソフトウェア

- ・Weka ver3.8.2
- ・OpenNI ver2.0

○言語

- ・Java

6. 実験方法

○Kinectを用いて座標データ収集.

→以下の2つの状態を設定した.

静的ジェスチャー …… a,b,c,d,e

動的ジェスチャー …… a,b,c,d,e

○収集したデータを用いて, Wekaによる分析を行う.

→Random Forestの手法を利用.

6-1. ジェスチャーの設定

5人の被験者に対して、動きのないジェスチャー(以下静的ジェスチャー)を5種類、動きのあるジェスチャー(以下動的ジェスチャー)を5種類を設定した。

静的ジェスチャー、動的ジェスチャーのそれぞれの状態において、a, b, c, d, eの5つの動きを、表1のように設定した。

また、それぞれ5つの状態について、各50セットずつデータを収集した。

表1. ジェスチャーの設定

	静的ジェスチャー	動的ジェスチャー
a	椅子に座っている	左右の腕でパンチし続ける
b	右手を上げる	右手を縦に動かし続ける
c	左手を上げる	左手を縦に動かし続ける
d	右手を水平に上げる	右手を水平に動かし続ける
e	左手を水平に上げる	左手を水平に動かし続ける

6-2. データの収集

図4は、OpenNIライブラリ³⁾を利用して、Kinectセンサによって利用者の骨格を認識させた様子を示したものである。

収集するデータは、各部の xyz 座標であり、上半身9箇所の座標(計27の座標)を測定した。腹部の座標を基準座標(0,0,0)として、腹部から首までの長さを1とする相対座標を用いた。また、1秒間で5個のデータを収集するように設定した。



図4. 骨格データを取得した状態

収集したデータは, 図5のようにCSV形式で保存される.

A	x0	y0	z0	x1	y1	z1	x2	y2	z2
a	0.018359861	-1.963808685	0.413359021	0.012979141	-0.99643238	0.083390965	0	0	0
a	0.041566543	-1.958523797	0.447019897	0.022106307	-0.993105845	0.115117731	0	0	0
a	0.04362961	-1.969338782	0.371636963	0.024147124	-0.996506041	0.079953907	0	0	0
a	0.045605839	-1.969117054	0.373796299	0.026848754	-0.997696261	0.062300209	0	0	0
a	0.04539139	-1.932855883	0.613919015	0.02631967	-0.984056613	0.175897293	0	0	0
a	0.019514748	-1.965446057	0.385665417	0.011406233	-0.996844421	0.078556342	0	0	0
a	0.022768181	-1.95815872	0.46639683	0.013105797	-0.993098072	0.116552384	0	0	0
a	0.0331247	-1.966078878	0.420984613	0.018332488	-0.995652866	0.091319712	0	0	0
a	0.034708597	-1.952048521	0.491198494	0.021321505	-0.991575032	0.127766775	0	0	0
a	0.037544461	-1.950214746	0.511053785	0.022100739	-0.989376285	0.143687594	0	0	0

図5. 保存されたCSVデータ

このデータを用いて, Wekaによる分析を行う.

このとき, 分類に用いるデータ数を, 被験者1人分, 2人分, ..., 5人分と増やし, 静的ジェスチャーと動的ジェスチャーについて, 分類精度を比較する.

6-3. データの分類

Wekaを用いてデータ分類を行う。

このとき、分類に用いるデータ数を、被験者1人分、2人分、..., 5人分と増やし、静的ジェスチャーと動的ジェスチャーについて、分類精度を比較する。

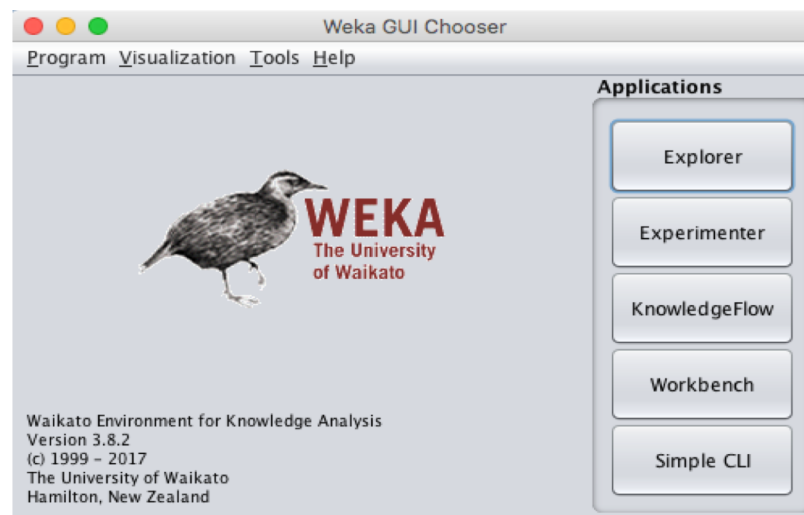


図6. Wekaの起動画面

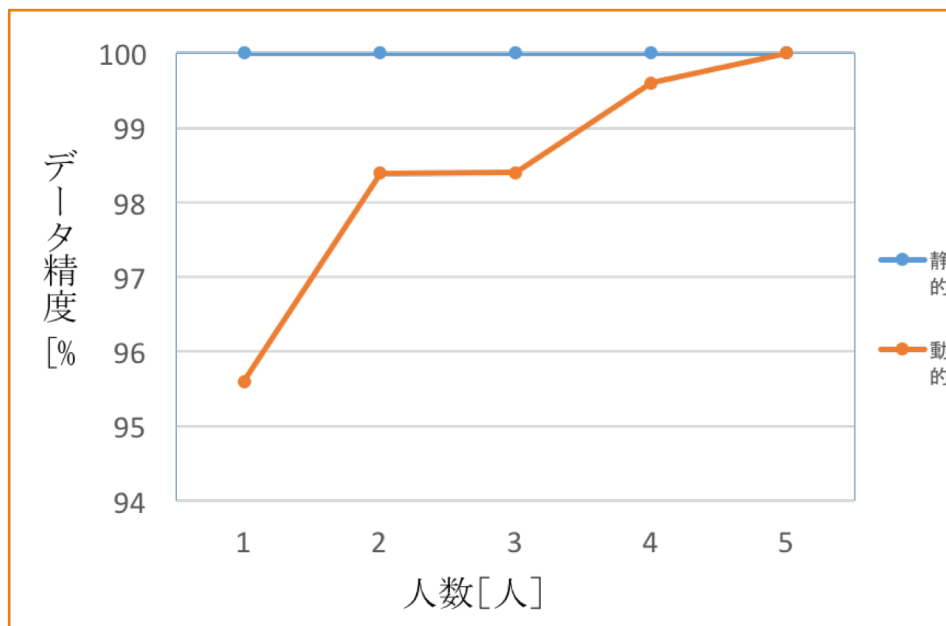
Test options

→Cross-validation(K分割交差検証)・・・ 10 Folds

教師	教師	教師	教師	教師	教師	教師	教師	教師	テスト	・・・ 1回目
教師	教師	教師	教師	教師	教師	教師	教師	テスト	教師	・・・ 2回目
教師	教師	教師	教師	教師	教師	教師	テスト	教師	教師	・・・ 3回目
.
テスト	教師	教師	教師	教師	教師	教師	教師	教師	教師	・・・ 10回目

図7. Cross-validationsの仕組み

7. 実験結果



青・・・静的ジェスチャー
赤・・・動的ジェスチャー

表2. Random Forestの分類結果

状態/人数	1	2	3	4	5
静的	100	100	100	100	100
動的	95.60	98.39	98.40	99.60	100.00

図8. Random Forestの分類結果

○静的ジェスチャー

1人目のデータ分析で100[%]の精度になった.

○動的ジェスチャー

1人目のデータ分析で96.5[%]の精度,
5人目のデータで精度が100[%]に向上した.

モデルを学習する前に, 教師データにランダム性を入れて部分集合を作り, 各部分集合のノイズを差し引く処理を繰り返し行う.

これにより重要な特徴量を判断し, ノイズが少なくなり, より精度の高い結果を得ることができたと考えられる.

7-1. 実験結果の比較

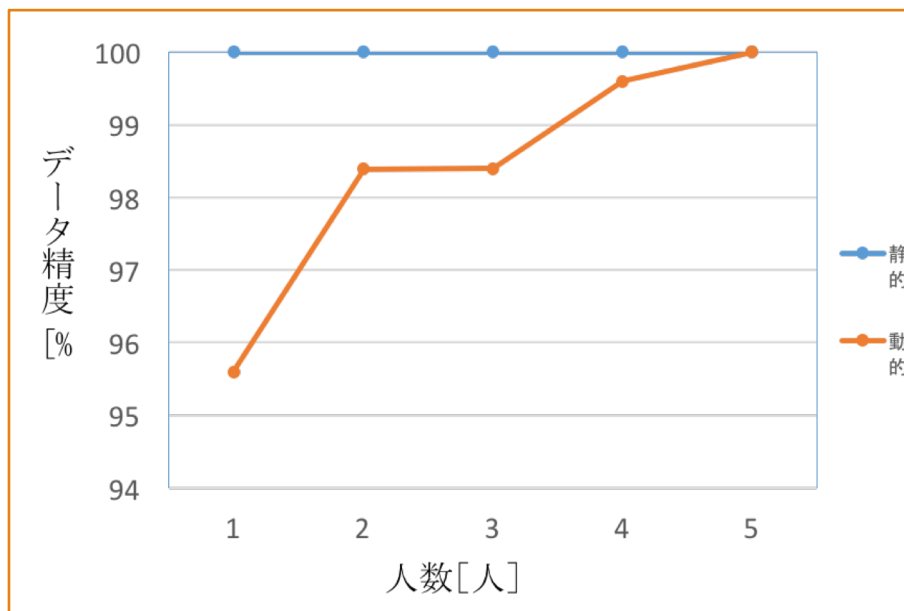


図8. Random Forestの分析結果

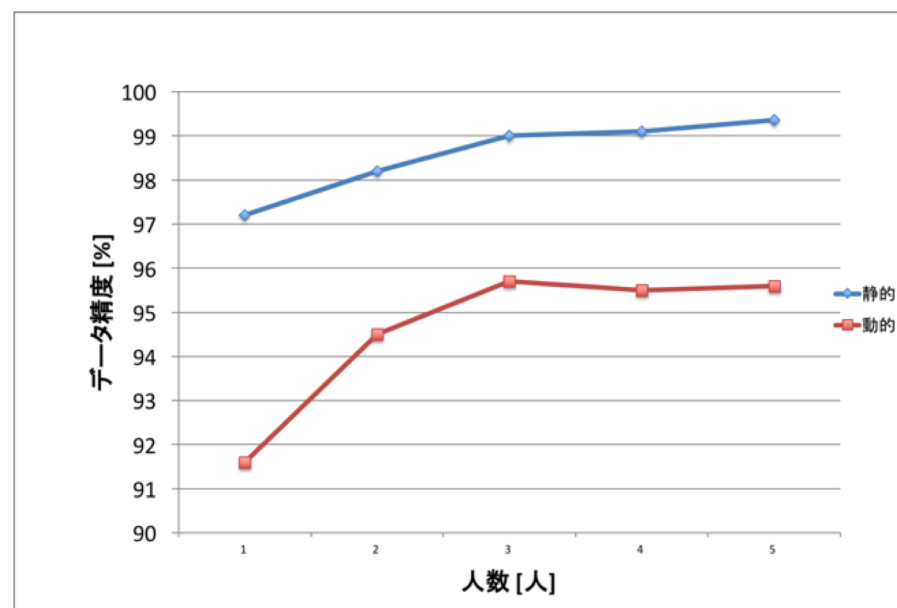


図9. 決定木の分析結果

8. まとめ

ジェスチャーの認識において,
静的ジェスチャー, 動的ジェスチャーを設定し,
決定木, またRandom Forestの手法を用いて分類精度を求めた.

○決定木の手法

静的ジェスチャー ... 99.4[%]の精度. 100[%]に近い精度を実現することができた.
動的ジェスチャー ... 95.6[%]の精度. これ以上の精度の向上は難しい.

○Random Forestの手法

静的ジェスチャー ... 100[%]の精度まで向上させることができた.
動的ジェスチャー ... 100[%]の精度まで向上させることができた.

9. おわりに

Random Forestの手法を用いて、データ分類を行い、100[%]の精度を求めることができた。

本研究で用いたKinectは生産終了となったが、この技術は現在でも様々な分野において利用されており、さらに新しいモーションキャプチャデバイスが開発、販売されている。

本研究は、Kinect以外のモーションキャプチャデバイスを用いる場合でも応用することができる。

参考文献

- 1) 宮城 諒, 平石 広典, “機械学習を利用した人間動作検出センサのための行動検出”, 情報処理学会第77回全国大会, Vol.4, pp.313-314, 2015.3.
- 2) 富岡 亮太, 長尾 確, “機械学習を用いたテニストレーニング支援システム”, 情報処理学会第80回全国大会, Vol.1, pp79-80, 2018.3.
- 3) 久野 遼平, 木脇 太一, “大学4年間のデータサイエンスが10時間でざっと学べる”, 株式会社 KADOKAWA, 2018.3.
- 4) 中村 薫, 新里 祐教, 鷺尾 友人, “OpenNI 3Dセンサープログラミング”, 株式会社 秀和システム, 2013.9.

○取得した座標

- | | |
|-------|--------|
| 0. 頭部 | 5. 左肘 |
| 1. 首 | 6. 右肘 |
| 2. 腹部 | 7. 左手首 |
| 3. 左肩 | 8. 右手首 |
| 4. 右肩 | |