

センサフュージョンを用いた個人識別

米谷和記[†] 三好力[†]

龍谷大学理工学部[†]

1 はじめに

近年、テレビゲーム機のゲームデバイスである Kinect が様々な方面で注目され始めている。このデバイスは、人物を認識し追跡が可能であること、複数のセンサを搭載していること、USB インタフェースがあることからゲームコントローラ以外の利用も多くなされている。その一つにホームネットワークシステムへの応用があり、これは家電をネットワークに接続しリンクさせることでより便利に使えるようにするシステムのユーザインタフェースとして用いるものである。実際に、Kinect を家電リモコンとして使用した次世代住宅も存在しており、家電の操作をより直観的に行うことが可能となっているが、個人識別による使用制限機能のようなものがない。そこで本研究では Kinect が複数のセンサを搭載していることに着目し、センサフュージョン技術を用いることで高い識別精度を持った個人識別システムの構築を試みる。

2 提案手法

2.1 センサフュージョンを用いた個人識別

本研究では Kinect に搭載されている複数種のセンサを用いてセンサフュージョンを行うことで個人識別システムの識別精度の向上を目指す。そのためにはまず、複数のセンサから得られた情報をもとにそれぞれ別の特徴量を算出する。次に、複数の特徴量を用いることでどの程度識別精度が向上するかを比較、検証を行う。今回は特徴量として、深度画像からの骨格情報と RGB 画像からの顔画像を用いた。

2.2 深度画像からの特徴量（骨格特徴量）

深度情報から得られる特徴量にはジョイント間の距離を利用する。使用する部位はなるべく頭に近いものを選択する。これはカメラに上半身しか映ってなくても深度画像・RGB 画像から特徴量を得ることができ、探索範囲を限定することで処理を軽くすることが目的である。事前実験を行った結果、最も数値のブレが小さくかつ個人差の出やすかった肩幅の長さを今回利用する骨格特徴量とした。

2.3 RGB 画像からの特徴量（顔特徴量）

カメラ画像センサで得た RGB 画像から、目・鼻・口などの顔部品がどの座標にあるかを検出しそれらの大きさや配置パターンを数値化したものを特徴量とする。今回座標データを得るために無料 API サイトである「detectFace0:」を利用する。このサイトは顔を含んだ画像を送ると顔部品の位置を認識し、顔を構成する特徴的な部分(以下特徴点)の座標データを送り返す API を無料で提供している。特徴点座標から特徴量を得る手法として、瞳間距離と基点・特徴点間距離の比率を使用している。基点は瞳間距離の midpoint としている。この手法を用いることで得られる特徴量はユーザとカメラの距離および座標のずれによるノイズの影響が小さい。そのためより正確な数値を得ることができ識別精度の向上につながると考える。今回用いる特徴点は、位置が変化することが少ない目頭・目じり・鼻翼とした。算出式は以下の通りである。

顔特徴量

$$= \sum \frac{\sqrt{(\text{特徴点のX座標} - \text{基点のX座標})^2 + (\text{特徴点のY座標} - \text{基点のY座標})^2}}{\text{瞳間距離}}$$

3 実験・評価

3.1 概要

単一の特徴量を用いたシステムと複数の特徴量を用いたシステムの認識精度を比較することで、センサフュージョン技術が個人認証システムにどの程度有効かを検証した。システムの性能は情報検索における性能評価指数の1つである適合率・再現率・F 値で評価する。この3つの値は、最大値 1、最小値 0 であり F 値が 1 に近い程性能が良いことを意味している。算出式を以下に示す。

適合率

$$= \frac{\text{登録済みの人物を認識した際の正答数}}{\text{登録済み人物を識別した際の正答数} + \text{未登録人物を識別した際の誤答数}}$$

再現率

$$= \frac{\text{登録済みの人物を認識した際の正答数}}{\text{登録済み人物を識別した際の正答数} + \text{登録済み人物を識別した際の誤答数}}$$

$$F \text{ 値} = \frac{2 \times \text{適合率} \times \text{再現率}}{\text{適合率} + \text{再現率}}$$

Title: 「Identification using the sensor fusion」

[†]MAIYA Kazuki, MIYOSHI Tsutomu,

Faculty of Science and Technology, Ryukoku University

3.2 骨格特徴量を用いた個人識別

3.2.1 実験目的

センサフュージョンを用いた個人認識システムの性能を評価するための比較対象として、骨格特徴量を使用した個人認識の精度を検証した。

3.2.2 実験結果

骨格特徴量のみを用いた個人認識システムの性能評価は表1のようになった。

表1：性能評価（骨格特徴量）

適合率	再現率	F値	平均処理回数
0.41	0.50	0.450	1.0

3.2.3 考察

実験より骨格特徴量の長所と短所について考察する。まず、長所として処理回数の少なさがある。これにより Kinect が人物を認識してから結果が表示されるまでの時間を短縮することが可能である。短所として体格の似ている人物の識別が困難であり、精度を向上させるためには他の要素を取り入れる必要がある。

3.3 顔特徴量を用いた個人識別

3.3.1 実験目的

センサフュージョンを用いた個人認識システムの性能を評価するための比較対象として、顔特徴量を使用した個人認識の精度を検証した。そのために、信頼度閾値を変化させることで識別結果がどのように変化するかを計測した。

3.3.2 実験結果

顔特徴量のみを用いた個人認識システムの性能評価は表2のようになった。

表2：性能評価（顔特徴量）

信頼度 閾値	適合率	再現率	F値	平均処理 回数
0.6	0.49	0.48	0.485	7.2
0.7	0.59	0.64	0.614	11.0
0.8	0.76	0.74	0.750	25.3
0.9	測定不能	測定不能	測定不能	測定不能

3.3.3 考察

実験より顔特徴量の長所と短所について考察する。顔特徴量を用いた場合、識別精度の安定性が長所といえる。信頼度の閾値を大きくすることで確実に F 値が上昇していることがわかる。しかし処理回数の上昇はそれ以上の欠点となっており、閾値 0.7 と 0.8 の処理回数に 14.3 回の差があることは見逃せない問題である。閾値を大きくすると処理回数が増加する理由として、Kinect の RGB カメラの解像度があまり高くないため、閾値を満足する画像を得るのに時間がかかるためであると推測する。

3.4 センサフュージョンを用いた個人識別

3.4.1 センサフュージョン方法

処理の速い骨格特徴量を用いてデータベースの登録データとある程度絞り込む。そして、取得した RGB 画像より得た顔特徴量と絞り込んだデータとの比較によって人物を識別する。今回は信頼度の閾値を 0.7 とした。

3.4.2 実験目的

センサフュージョンを用いた個人認識システムの性能の比較実験を行った。単一の特徴量での識別の欠点をうまく補い性能の向上が図れているのか検証した。同じ閾値での識別率・処理回数の比較によって性能が向上したかどうかを判断した。

3.4.3 実験結果

センサフュージョンを用いた個人認識システムの性能評価は表3のようになった。

表3：性能評価比較

	適合率	再現率	F値	平均処理 回数
骨格特徴量	0.41	0.50	0.450	1.0
顔特徴量 (閾値0.7)	0.59	0.64	0.614	11.0
センサフュ ージョン	0.73	0.70	0.715	11.7
顔特徴量 (閾値0.8)	0.76	0.74	0.750	25.3

3.4.4 考察

性能評価に関して表3より、閾値 0.8 の顔特徴量識別には至らないものの近い値を得ることができた。処理回数を加味して評価すると上回る性能のシステムであるといえる。

4 まとめ

本論文では、センサフュージョンを用いた個人識別システムを提案した。提案手法のシステムの性能を評価するためにまず、単一のセンサからの特徴量による個人識別の精度について検証した。その結果、骨格特徴量のみを用いた場合、処理回数に優れるが識別率の不安定さが目立った。顔特徴量のみを用いた場合、閾値を大きくするほど識別率も上昇するが、それ以上に処理回数が増してしまう。手案手法では、骨格特徴量で登録データをあらかじめ絞り込み、顔特徴量の比較を行う回数を減らすことで、F 値 0.101 の上昇に対し処理回数の上昇を 0.7 回に抑えることができた。これによりセンサフュージョンを用いて個人識別システムを構築することは有効であることが判明した。