

Kinect による骨格と表情を用いた人物特定

大谷 紀子 研究室

1032103 後藤佑一郎

1. 研究の背景・目的

現在、室内に存在している人間の確認手段として在室管理システムが用いられている。在室管理システムには、非接触 IC カードやスマートフォンのアプリケーションなどを利用し、入退室を検知するものがある。しかし、システムの導入にはコストがかかるため、研究室などの小規模なグループが導入するのは難しい。また、認証の際カードを取り出してカードリーダーにかざすなどの動作が必要になる。2013 年に横尾[1]は、安価な Kinect センサの利用により取得した骨格情報と行動履歴を用いて、人物を特定する手法を提案した。しかし、すべての人物を正確に特定することはできなかった。また、あらかじめ特徴を学習された人物との類似度を求めて特定する方法であり、部外者は特定の対象ではなかった。

本研究では、人間の骨格情報と顔の情報を利用した、人物特定手法を提案するとともに、人物特定の精度の向上、および部外者の特定を可能にすることによるセキュリティ面の強化を目的とする。室内にいる人間を特定するシステムを構築し、評価実験により有用性を示す。

2. システム概要

本システムでは、人間の表情と骨格に基づいて人物を判別するシステムを構築する。RGB カメラ、深度センサを搭載した、Microsoft 社の安価なセンサデバイス Kinect for Windows(以下 Kinect)を 2 台使用する。研究室内の 2 箇所に配置し、定点観測することで、1 秒ごとに人間の骨格と顔の情報を取得する。人物の特定には、全身の部位骨格 20 箇所のうち 10 箇所の 3 次元座標、眉や目などの顔のパーツ 100 箇所の 2 次元座標を利用する。下半身が椅子などに隠れることを想定して上半身のみの骨格情報を取得する。本システムの処理手順を以下に示す。

- ① Kinect により入室した人間の骨格と顔の情報を取得する。
- ② 取得した骨格と顔の情報をある研究室メンバーを判別する多数決機械に入力する。
- ③ 多数決機械の出力した結果から、入室した人物と研究室メンバーとの適合率を算出する。
- ④ ②, ③をすべての研究室メンバーについて実行し、算出された適合率を用いて人物を特定する。

本研究では、骨格と表情に基づく人物の判別には多数決機械を利用する。多数決機械とは線形パターンを認識可能な奇数個の線形機械から構成されるもので、多数決機械が持つ各線形機械の出力の多数決により非線形パターンを認識する。本研究では、骨格情報と顔情報を別の多数決機械で処理する。研究室メンバー $P_1 \sim P_N$ の骨格情報と顔情報を学習データにして、 $P_1 \sim P_N$ を判別するための骨格情報に基づく多数決機械 $C_1 \sim C_N$ 、顔情報に基づく多数決機械 $D_1 \sim D_N$ をあらかじめ作成し、入室した人物の判別に用いる。適合率とは、入室した人物がある人物と適合した割合を表す値である。人物 P_i を判別するための骨格情報に基づく多数決機械において、 M セットの骨格情報のうち p_i セットが人物 P_i と判別されたとき、骨格情報に基づく適合率 F_i^s は p_i/M となる。同様に、顔情報に基づく多数決機械において O セ

ットの顔情報のうち q_i セットが人物 P_i と判別されたとき、顔情報に基づく適合率 F_i^q は q_i/O となる。適合率 F_i^1, F_i^2 を正規化し、式(1), (2)からエントロピー H_1, H_2 を算出して、数値が小さい方の適合率 $F_i^1 \sim F_i^N$ を人物特定に利用する。 $F_i^1 \sim F_i^N$ の中で F_i^m が最大であるとき、観測データが表す人物を P_m と特定する。また、 F_i^1, F_i^2 の平均値が 0.4 以下で、 H_1, H_2 が 2 より大きい場合、部外者と判別する。

$$H_1 = - \sum_{k=1}^N F_k^1 \log F_k^1 \quad \dots(1)$$

$$H_2 = - \sum_{k=1}^N F_k^2 \log F_k^2 \quad \dots(2)$$

3. 評価実験

本研究室のメンバー5名を被験者とし、1人ずつ90分自由に行動させ、骨格と顔の情報を5400セットずつ取得した。各被験者のデータのうち、15分の14を訓練データ、15分の1をテストデータとする15分割クロスバリデーションにより、提案した人物特定の手法を評価する。各被験者に対する適合率を表1に示す。

表 1 評価結果

		被験者					
		A	B	C	D	E	部外者
適合率	A	<u>0.8556</u>	0.0706	0.3031	0.3344	0.0022	0.1594
	B	0.0652	<u>0.8003</u>	0.3255	0.3259	0.0227	0.1908
	C	0.0296	0.0638	<u>0.7543</u>	0.4103	0.0490	0.1847
	D	0.0026	0.0405	0.3533	<u>0.7801</u>	0.0007	0.1689
	E	0.0135	0.0027	0.3324	0.3670	<u>0.8789</u>	0.1746

下線が引かれている適合率が正しい人物と判別している場合の適合率である。いずれも下線部が最も高く、他の人物と判別した場合の適合率と比べて低いので、正しく人物特定されていることがわかる。また、被験者 A, B, E においては適合率が 0.8 以上で被験者 C, D の適合率と比べて高い結果である。部外者データとの適合率がいずれも 0.2 以下と低いので、正しく部外者を判別できていることがわかる。

4. 考察

評価実験の結果、いずれの人物も正しく判別され、被験者 A, B, E については特に良好な結果を得た。しかし、0.90 以上の値は得られず、先行研究と比べて適合率が全体的に低かった。また、被験者 C, D では適合率が 0.8 以下と低く、他の被験者との適合率が 0.3~0.4 と高いため、精度の高い特定結果が得られたとはいえない。部外者については、いずれの適合率も低いことから、提案手法における部外者判別の精度は高いことがわかった。骨格情報による人物判別では、正しい人物と判別された場合の適合率で、高い数値と低い数値が確認された。被験者の頻繁な姿勢の変化により、データの取得に影響を与えたことが考えられる。顔のパーツ情報による人物判別では、顔の向きにより座標が変化し、人物の判別に悪影響を与えると考えられる。

参考文献

- [1] 横尾亮平, “Kinect センサによる骨格情報と行動履歴を用いた人物特定”, 東京都市大学情報メディア学科卒業論文, 2013.