

RGB-D カメラを使用した頭部姿勢にロバストな表情認識手法

井上 雄太[†] 任 福継[‡] 西出 俊[‡]

[†]徳島大学 大学院先端技術科学教育部 [‡]徳島大学 大学院ソシオテクノサイエンス研究部

1. はじめに

人間とロボットの柔軟なコミュニケーションを実現するためには、感情情報を利用することが不可欠である。顔表情から感情認識をする一般的な手法では正面顔が用いられるが[1], ロボットとの対話においては必ずしも正面顔が取得できるとは限らない。本研究では頭部姿勢によらない表情認識手法を提案する。

2. RGB-D カメラ

本研究では頭部姿勢の影響を補正するため、RGB-D カメラを用いる。RGB-D カメラは専用のライブラリを用いることで、顔特徴点と頭部姿勢角度を自動で取得することが可能である(図1)。これらの情報を利用することで、頭部姿勢によらない表情認識を実装する。

実験では Intel RealSense 3D Camera を用いた。これは比較的安価で購入でき、搭載されているセンサーは比較的小さいため、幅広い場面での応用が期待できる。

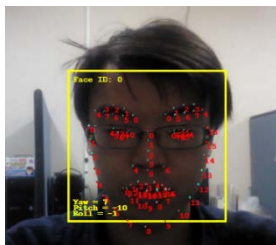


図1. RGB-D カメラで取得した特徴点

3. システム概要

本システムの流れを図2に示す。RGB-D カメラで取得したセンシングデータをもとに円柱モデルを作成し、円柱モデルに特徴点を投影する。円柱モデルを正面に向くように回転させ、各特徴点の動きベクトルを算出し、特徴点に基づく表情認識を行う。

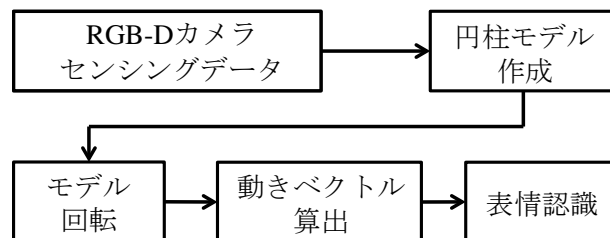


図2. 本システムの流れ

3-1. 円柱モデル

本手法では顔形状モデルとして、熊野らと同様の円柱モデル(図3)を用いる[2]。円柱モデルは、取得したカメラデータのうち、頬の両端を直径、頬の両端の中点を回転の中心として作成し、頭部姿勢と同じ角度に円柱モデルを傾ける。そこに特徴点を投影することで奥行き方向のZ座標を推定する。また、直径が一定になるよう3次元アフィン変換をし、中心の座標が原点になるよう正規化を行う。

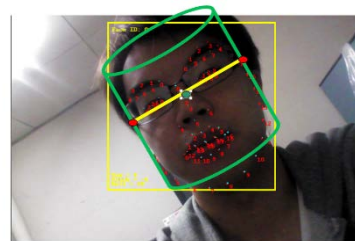


図3. 円柱モデル

3-2. モデル回転

RGB-D カメラから取得される頭部姿勢は傾き角(ピッチ)、首振り角(ヨー)、頭部傾け角(ロール)の三つである。各特徴点を上記円柱モデルの中心に対して、三種の角度の逆方向に回転させることで正面顔の特徴点の位置座標を計算する。傾いた顔画像から正面顔の特徴点を計算した結果の一例を図4に示す。

3-3. 表情認識

本研究では表情の合成法である FACS(Facial Action Coding System)に基づいて表情認識を行う[3]。FACS は表情による顔筋肉の動きを Action Unit (AU) と呼ばれる単位に分割する。AU をもと

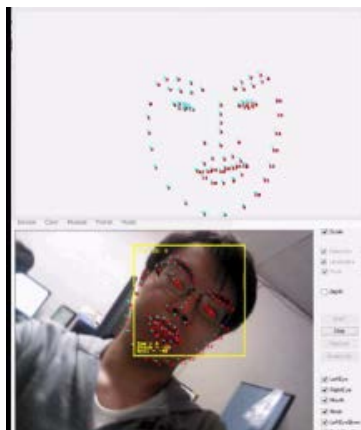


図 4. 上：回転後の特徴点
下：回転前の特徴点

に回転後の顔特徴点から動きベクトル推定に適している点を選出し、その特徴点から動きベクトルを以下のように計算する。

$$\theta = \tan^{-1} \left(\frac{\Delta x}{\Delta y} \right) \quad (-\pi < \theta < \pi) \quad (1)$$

$$\sqrt{\Delta x^2 + \Delta y^2} = 1 \quad (2)$$

ここで、 θ は頭部の回転角度、 $(\Delta x, \Delta y)$ は動きベクトルである。求めた動きベクトルと AU の一致度から表情を推定する。

4. 検証実験

提案手法の有効性を検証するため、頭部回転前後の動きベクトルの算出、表情認識の成功率を調べた。

4-1. 頭部動作による動きベクトルの除去

円柱モデルの回転により頭部姿勢の変動による影響が除去できているかを確認するため、回転前の動きベクトルと回転後の動きベクトルの大きさを比較した。結果を表 1 に示す。実験は無表情な顔で行い、表情の変化による動きベクトルの変化が発生しないようにしている。

表 1. 全特徴点の動きベクトルの大きさの平均

	回転なし	回転あり
平常時	1.082091	1.307603
頭部平行移動時	13.29419	3.259615
頭部前後移動時	20.96948	8.502889
頭部回転時	17.66412	11.45768

表 1 より、モデルを回転させた場合、回転させなかった場合と比べ、動きベクトルが減少している。これにより、頭部の姿勢変動による動き

ベクトルの影響が除去できていることが確認できた。

4-2. 表情認識

表情認識の精度を確認するために様々な頭部姿勢、表情に対し、表情認識実験をおこなった。認識の対象とする表情は無表情、怒り、悲しみ、喜び、驚きの 5 種類とした。また、カメラとの距離、照明などの影響は考慮せず自由な位置で実験を行った。実験結果を表 2 に示す。

表 2. 表情認識精度(%)

	$-\pi/6$	0	$\pi/6$
ヨー	33.3	86.7	33.3
ピッチ	53.3	86.7	40.0
ロール	60.0	86.7	75.0

正面顔である 0° での認識精度と比較してロール角に傾いた場合は比較的高い認識精度を維持している。一方で、奥行情報が必要となるピッチ角やヨー角に対しては認識精度が若干下がる結果となった。これは、頭部動作の影響が完全には除去できていないことに起因しており、今後はモデルの改善が必要だと考えられる。

5. おわりに

一台の RGB-D カメラによって得られる頭部情報を用いて頭部姿勢に頑健な表情認識手法を提案した。実験を行った結果、提案手法は頭部の動きの影響を除去することには成功したが、表情認識精度は十分とはいえない。今後は円柱モデルの改良や、より正確な動きベクトルの算出方法に改善することでロボットの対話に応用していく予定である。

謝辞

本研究は JSPS 科研費 15H01712 の助成を受けたものである。

参考文献

- [1] トーンカムチャイヤブルック, 内田理: 「Kinect を利用した感情推定手法」映像情報メディア学会技術報告, pp. 41-44, 2014.
- [2] 熊野史郎, 大塚和弘, 大和敦司, 前田英作, 佐藤洋一: 「表情認識のための変動輝度テンプレートとその頭部姿勢変動に対する頑健性の一検討」信学技報, pp. 145-152, 2007.
- [3] 中村薫, 前本知志, 藤祐佑, 谷口直嗣, 初音玲: 「Intel RealSense SDK センサープログラミング」翔泳社, 2015. ISBN:9784798139630