

コミュニケーションを円滑に進めるための 表情変化検出方式の検討

森 博章[†] 宮脇 健三郎[‡] 佐野 睦夫[‡] 西口 敏司[‡]

^{† ‡}大阪工業大学情報科学部 〒573-0196 大阪府枚方市北山 1-79-1

E-mail: [†] e1c05117@info.oit.ac.jp, [‡] {miyawaki, sano, nishigu}@is.oit.ac.jp

あらまし 人とロボットの円滑なコミュニケーションの実現において、顔の表情は重要な役割を果たしており、表情認識はコミュニケーションの状態理解において極めて重要である。本研究では、横顔や後顔も含めた微妙な表情変化を認識することを目標として、第1段階として、検討例がほとんどない横顔の表情認識を扱う。具体的には、Active Appearance Model (AAM)を用いて顔の時系列画像に対して顔モデルの追跡を行い、顔モデルを構成する特徴点から、FACSモデルに対応したAU (Action Unit) 特徴を算出し、AU特徴の時系列データの変動パターンを詳細に分析することにより、微妙な表情の変化に着目した認識方式について検討を行った。

キーワード 表情認識, Active Appearance Model

A study of expression change detection method for smooth communication

Hiroaki MORI[†] Kenzaburo MIYAWAKI[‡] Mutsuo SANO[‡] and Satoshi NISHIGUCHI[‡]

^{† ‡} Faculty of Information Science and Technology, Osaka Institute of Technology

1-79-1 Kitayama, Oosaka-shi, Oosaka, 573-0196 Japan

E-mail: [†] e1c05117@info.oit.ac.jp, [‡] {miyawaki, sano, nishigu}@is.oit.ac.jp

Abstract In order to realize the smooth communication between human and robots, it is important to recognize facial expression. The aim of this study is to recognize a subtle expression change reliably in the variation of face directions. As the first stage, we treat with recognizing profile images or back profile images that haven't been studied sufficiently. In the proposed method, the facial feature points of time-series images are tracked using Active Appearance Model (AAM) and the Action Unit (AU) features of FACS model are calculated. Then we investigated the recognition methods based on a subtle expression change by analyzing the variation of feature patterns,

Keyword Facial expression recognition, Active Appearance Model

1. はじめに

近年、ロボットが生活に入り込み、支援を行う機会が増えてきている。ロボットが円滑な支援を行うためには、人間の状態に応じて適応的に振る舞うコミュニケーション機能が不可欠である。このようなロボットによる自然なコミュニケーション機能を実現する上で、表情は豊富な情報量を含んでおり、表情認識はコミュニケーションの状態理解において極めて重要である。

本研究では、コミュニケーションロボットにおける表情認識に着目し、表情変化を検出するための方式について検討を行う。表情認識の研究では種々の手法([1]~[4])が提案されているが、個人差がある中、表情の微妙な変化を感情モデルの全空間(種類や強弱で分類される空間)に渡り、顔の見え方によらず、信頼性よく検出することはなかなか困難である。時系列画像から本当の笑いと作り笑いなど微妙な表情の変化を検出する研究例[5][6]もあるが、実験例として正面の顔

への適用に留まっていたり、強弱を含めた種々の感情に対して研究している例は少ない。さらに、複数話者のコミュニケーションにおいては、図1に示すように、認識対象とする顔の向きは±90度以上に広がり、横向き顔や後向き顔の表情認識も必要となる。

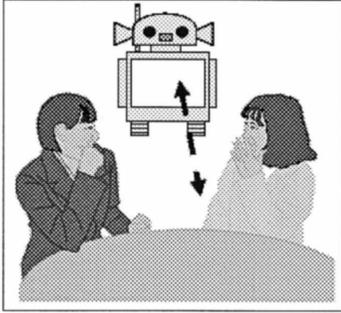


図1 コミュニケーションロボットによる表情認識シーン

本研究では、横顔や後顔も含めた微妙な表情変化を認識することを目標として、第1段階として、検討例がほとんどない横顔の表情認識を扱う。

2. 表情変化検出の概要

表情の獲得には、Active Appearance Model(AAM)[7]を使用し、動画像より特徴点位置の抽出および追跡を行う。さらに、Facial Action Coding System (FACS)[8]に基づいて表情ごとの各AUを認識するために必要な特徴点間の距離を特徴量とし認識を行う。

2.1. AAMによる特徴点抽出

特徴点位置の抽出および追跡にはAAMを用いる。AAMは顔を含む変形を伴う物体の形状と輪郭の内部のテクスチャを同時にモデル化する統計的なモデルであり、安定したマッチングを行うことができる。

AAMではあらかじめ特徴点の座標がすべて与えられた学習データより、形状とテクスチャをパラメータに持つappearanceモデルを作成する。各点の座標を並べたベクトルを x で表すと、これをすべての学習データについて平均して得られる平均ベクトルは \bar{x} となり、これを平均形状とする。 x は \bar{x} からの偏差を主成分分析して得られる固有ベクトル P_s を用いると、

$$x = \bar{x} + P_s b_s \quad (1)$$

と表すことができる。ここで、 b_s は形状のパラメータを表している。

次に、すべての学習データを共通の形状に変換すると、 \bar{x} により正規化された形状についての明度分布である g を得る。 g についても、それらベクトルを平均して \bar{g} とし、これを平均明度分布とする。さらに g は \bar{g} からの偏差を主成分分析して得られる固有ベクトル P_g を用いて、

$$g = \bar{g} + P_g b_g \quad (2)$$

と表すことができる。ここで、 b_g は明度分布のパラメータを表している。

最後に、形状とテクスチャ間の単位を正規化するための行列を W_s とし、式(1)と(2)より、

$$b = \begin{pmatrix} W_s b_s \\ b_g \end{pmatrix} = \begin{pmatrix} W_s P_s^T (x - \bar{x}) \\ P_g^T (g - \bar{g}) \end{pmatrix} \quad (3)$$

なるベクトル b を定義する。これを主成分分析すると、

$$b = Qc \quad (4)$$

と表すことができる。ここで、 Q は固有ベクトルであり、 c は、物体の画像上の形状と明度分布の両方を表すappearanceパラメータである。このappearanceパラメータに加え、対象物体の位置や画像のスケールなどを考慮し、appearanceモデルを作成することで、未知画像の入力に対して安定したマッチングを行うことができる。

本研究では、横顔に対応したモデルを作成し、横顔の時系列画像に対してマッチングを行い、特徴点の追跡および座標値を取得した。特徴点は図2に示す点とし、眉に4点、目に8点、口に8点の計20点とし、主に目と口の追跡を行った。

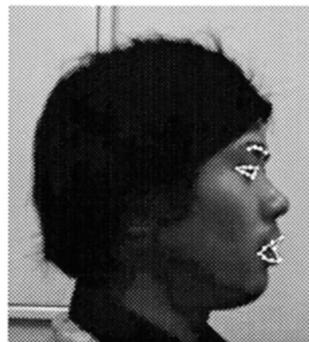


図2 特徴点の定義

2.2. FACS モデルに基づく表情動作認識

FACS は Ekman らによって提案された表情記述法の 1 つで、Action Unit(AU)と呼ばれる表情動作の最小単位の組み合わせにより表情を記述する方法である。また AU は、基本 6 感情[9](幸福、驚き、恐怖、怒り、嫌悪、悲しみ)と対応している。今回は基本 6 感情の中から Positive(幸福、驚き)と Negative(恐怖、怒り、嫌悪、悲しみ)に分け、Positive に属する感情に対して判別を行う。また、幸福の感情は笑いによって判断した。

動作認識は AU の構成に基づいて、AAM から得られた目、眉、口の各特徴点の位置より特徴量を算出した。今回は横顔の表情が認識対象なので、図 3 に示すように、目尻の点と目の中央の点を結んだ線と口の端の点、同線と眉の端の点、さらに同線と口の上の点、口の上下の点、目の上下の点、口の端と口の中央の点間の 6 つの各距離を特徴量 F1~F6 として定義した。

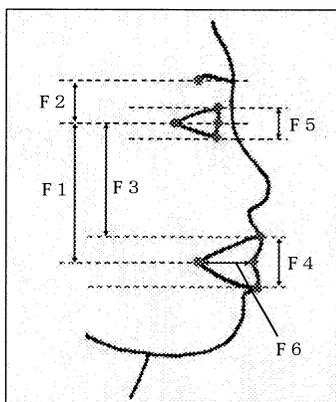


図 3 抽出する特徴量

3. 実験

3.1. 実験環境

動画像取得には IEEE1394 カメラ(Point Gray Research 社 Flea)を被験者の真横に設置し、解像度 320 × 240pixel, フレームレートは微妙な表情の変化に対応するために 60fps で撮影した。被験者は 10 人とし、無表情、笑い、驚きの計 3 種類の表情を各 4 回ずつ行ってもらった。笑いは被験者に漫才の映像など自分が笑いやすいと思うものを視聴してもらい、その時に表出された表情を撮影した。驚きに関しては、笑いと同じく何かを見て驚いてもらうことが困難であったため、各自のタイミングで意図的に驚いてもらい、表出された表情を撮影した。動画像の処理を行う環境については CPU が Intel Xeon 3 GHz プロセッサ、3 GB メモリ

の PC で行った。

3.2. 結果

得られた顔画像系列に対して、AAM より図 2 に示す 20 個の特徴点を追跡した結果を図 4 の(a), (b)に示す。これらの図より、表情の変化に伴う特徴点、また目、口の形状においても精度良く追跡できていることがわかる。



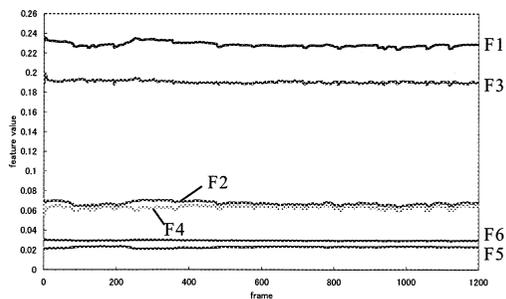
(a) 笑いの表情変化



(b) 驚きの表情変化

図 4 表情の変化に伴う特徴点の追跡

次に、得られた特徴点に対して 2.2 節で述べた方法を用いて無表情、笑い、驚きの 3 種類について特徴量の算出を行った。その結果を図 5 の(a), (b), (c)に示す。図中の F1~F6 は図 3 で示した各特徴量を表している。図の縦軸は特徴量を表し、横軸は時間(frame)軸である。



(a) 無表情における特徴量の変化

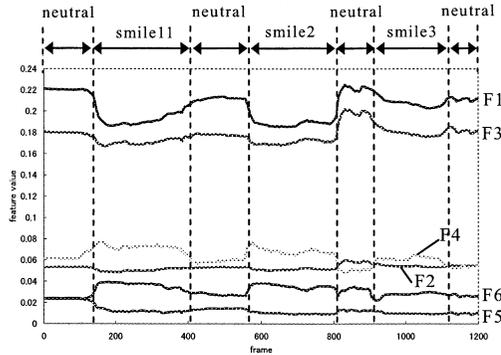
がどのように有効に働くかの検討も加えるつもりである。

4. おわりに

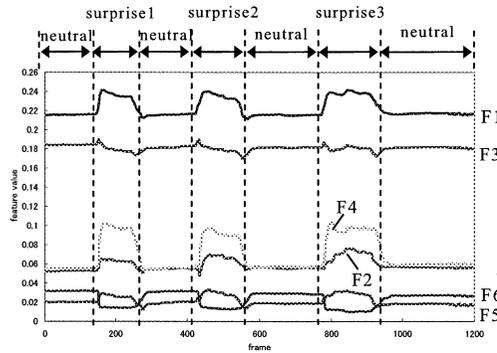
今回の報告では、AAMに基づき顔モデルを追跡し、横顔に着目した表情の時系列データの比較検討を行い、表情認識の実現可能性を示した。現在、学習識別系を組んでおり、個人差も含めて、認識率の評価を行っている。どの顔向き角度から表情認識が可能となるかについても検討したい。また、横顔や後顔を含めた任意の顔方向に対してシームレスな表情認識の検討も行う予定である。

文 献

- [1] Y. Tian, T. Kanade and T. F. Cohn: "Recognizing action units for facial expression analysis", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.23, No.2, pp.97-115, Feb. 2001.
- [2] I. Cohen, N. Sebe, A. Garg, L. S. Chen, T. S. Huang: "Facial Expression Recognition from Video Sequences: Temporal and Static Modeling", Computer Vision and Image Understanding, Vol. 91, No. 1-2, pp.160-187, 2003
- [3] K. Schmidt, L. Cohn and Y. Tain: "Signal Characteristics of Spontaneous Facial Expressions: Automatic Movement in Solitary and Social Smiles", Biological Psychology, Vol. 65, pp.49-66, Oct. 2003.
- [4] M. Bartlett, G. Littlewort, M. Erank, C. Lainscsek, I. Fasel and J. Movellan: "Automatic recognition of facial actions in spontaneous expression", JMM, 1, 6, pp.22-35, 2006.
- [5] Krumhuber, E., Kapps, A.: "Moving Smiles: The Role of Dynamic Components for the Perception of the Genuineness of Smiles", Journal of Nonverbal Behavior, Vol. 29, No. 1, pp.3-24, 2005.
- [6] 平山高嗣, 川嶋宏彰, 西山正紘, 松山隆司, "表情譜: 顔パーツ間のタイミング構造に基づく表情の記述", ヒューマンインタフェース学会, Vol.9, No.2, pp.201-211, 2007.
- [7] T. F. Cootes, G. J. Edwards and C. J. Taylor: "Active Appearance models", Proceedings of the 5th European Conference on Computer Vision, Vol. 2, pp.484-498, 1998.
- [8] P. Ekman, W. V. Friesen "The Facial Action Coding Systems", Consulting Psychologists Press, 1978.
- [9] P. Ekman, W. V. Friesen, 工藤力訳: "表情分析入門", 誠信書房, 1987.



(b) 笑いにおける特徴量の変化



(c) 驚きにおける特徴量の変化

図5 無表情, 笑い, 驚きにおける特徴量の算出結果と変化

3.3. 考察

図5は、横顔における無表情, 笑い, 驚きの3表情の典型的な時系列データを示している。特徴量(F1~F6)の変化を比較すると3表情に関しては、適切な特徴選択や直交変換を行えば、ほぼ類別可能であると言える。図5(b)のsmile3はsmile1およびsmile2と比べて応答変化が小さい。smile3は含み笑いとなっており、普通の笑い(smile1, smile2)とは類別できるものの、どこまでの含み笑いが検出できるかについては現在検討中である。現在、他の表情も含め、微妙な表情の変化についてデータを採取中であり、今後報告する予定である。また、横顔でしか得られない特徴量F6