

## 顔画像を用いた気持ちに寄り添う対話のための感情状態推定手法

加藤 孝浩<sup>†</sup> 石田 和生<sup>‡</sup>NEC ソリューションイノベータ株式会社<sup>†</sup>

## 1. はじめに

従来の対話システムは、一問一答形式の命令実行型を前提に設計されており、決まった受け答えしかできない「単純な音声入力的手段」となっている。これは、AI スピーカーといった「誰でも手軽に対話する」ことを目的としたシステムでも同様である。

今後の対話システムは、相手の状態に配慮し、気持ちに寄り添う対話によって、否定的な感情の抑制・解消、肯定的な感情の促進・継続することで、高齢化を迎える日本社会で、ICT の先進的な知識を持たない人でも、様々なサービスを楽しむことができることを目的とした、高度な対話システムとなることが望まれる。

このような高度な対話システムに必要な技術要素として、非言語情報である顔の画像から表情を判定し、その組み合わせから感情状態を推定する手法を紹介する。

## 2. 気持ちに寄り添う対話システム

対話における非言語情報としては、表情、身振り手振り、視線、声の大きさや高さ、話す速度やイントネーションといった様々なものが考えられる。

本稿では、対話する相手に特殊な機器を装着することなく、非言語情報を取得できることから、カメラによる画像を入力情報とし、表情の判定、感情状態を推定する。

感情状態推定に必要な表情は、Paul Ekman[1]が人間の基本的な感情で、全人類に普遍的なものとして結論付けている「幸福、驚き、怒り、恐怖、悲しみ、嫌悪」の6つを判定し、組み合わせで感情状態を推定する。

判定した表情と感情状態の組み合わせとして「幸福、驚き」から「快」、「怒り、恐怖、悲

しみ」から「不快」、「嫌悪」から「不同意」を推定し、対話に反映することで、否定的な感情の抑制・解消、肯定的な感情の促進・継続させ、気持ちに寄り添う対話システムを構築する。

## 3. 表情と感情状態の判定・推定

手順①～⑤で表情判定・感情状態推定する。

## ①顔・顔特徴点の検出

対話の相手をカメラで撮影して、顔および顔特徴点を検出する。(図 1)  
検出には Dlib[2]および既存の学習済みモデルを用いる。

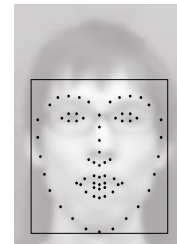


図 1 顔・顔特徴点の検出

## ②顔器官の定義・状態測定

まず検出した顔特徴点をつないだ線や領域を、表情判定に必要な顔器官として定義する。線や領域の定義は、表情筋の動く方向と範囲を考慮する。例えば、頬は上下の動きを測定できる領域定義とする。(図 2)

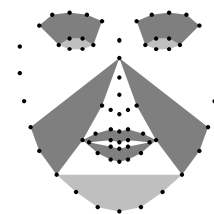


図 2 顔器官の定義例

次に定義した顔器官の状態を測定する。測定する状態は「面積・重心・明暗・傾き・皺状態」の組み合わせである。例えば、面積は目や口の開き具合といった状態を測定し、各表情との対応付けの対象とする。

An emotional state estimation method based on facial images for dialogue.

<sup>†</sup> Takahiro Katoh

<sup>‡</sup> Kazuo Ishida

<sup>†</sup> NEC Solution Innovators, Ltd.

1-18-7 Shinkiba, Koto-ku, Tokyo, 136-8627 Japan

顔器官により測定項目の組み合わせは異なる。例えば「額は皺状態を測定」するが、「頬は皺状態と重心を測定」する。

### ③表情と顔器官状態の対応付け

表情と測定した顔器官の状態を対応付ける。(図3)

その表情で特徴が表れる顔器官と対応付ける。例えば「驚きの表情」は Paul Ekman の表情分析より「額、眉、上下まぶた、口、顎」に特徴が表れるため、これらの顔器官と対応付ける。同様に「幸福、怒り、恐怖、悲しみ、嫌悪」についても、それぞれ特徴が表れる顔器官と対応付けて、表情スコアの算出項目とする。

	幸福	驚き	怒り	恐怖	悲しみ	嫌悪
額 全体 皺		対応				
額 中央 皺				対応		
眉 重心 上がる		対応				
眉 重心 下がる			対応	対応		
眉 下側 皺		対応				
眉 寄る			対応	対応		
眉 内側 上がる					対応	
眉間 皺			対応			
目 面積			対応			
目尻 皺	対応					
上まぶた 重心 上がる		対応		対応		
上まぶた 内側 上がる					対応	
上まぶた 重心 下がる			対応			
下まぶた 重心 上がる						対応

図3 対応付けの例

### ④表情スコアの算出

まず表情に対応付けた顔器官の状態に特徴が「表れる」「表れない」という組み合わせで配点(加点する/減点する/加点・減点しない)する。表情分析より各表情に特徴的な状態である場合(幸福で頬が上がるなど)に、スコアが高くなるよう配点する。逆に特徴的でない状態である場合(幸福で眉間に皺ができるなど)は、スコアが低くなるよう配点する。

対応付けた表情によって、同じ顔器官の状態であっても配点が異なる。例えば「下まぶたに皺あり」は「幸福は加点」するが、「悲しみは減点」する。

次に算出した表情スコアから、表情を判定する。表情は一つに決まるとは限らない。例えば「驚き」と「怒り」を同時に判定することもあり得る。

### ⑤表情から感情状態を推定

判定した表情「幸福・驚き」から「快」、  
「怒り・恐怖・悲しみ」から「不快」、

「嫌悪」から「不同意」の感情を推定する。例えば「快」の感情を推定する場合、「幸福・驚き」の表情は、Russell の円環モデル[3]から「快と覚醒」の2軸で示す角度で表すことができる。幸福は約7.8°、驚きは約69.8°の角度である。(X軸を快、Y軸を覚醒とする)

この時のX軸方向の長さから「快」感情の大きさを推定する。(図4)

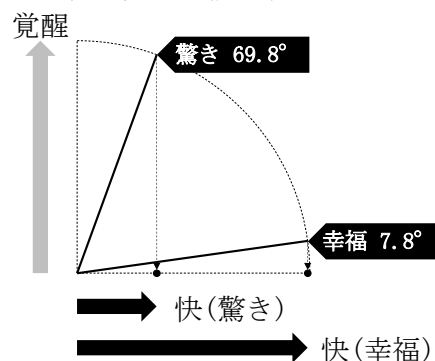


図4 表情の角度と感情「快」の関係

## 4. まとめ

本稿では、非言語情報である顔の画像から表情を検出し、その組み合わせから感情状態を推定する手法を紹介した。

推定した「快不快・不同意」を対話に反映し、否定的な感情の抑制・解消、肯定的な感情の促進・継続につなげることは、気持ちに寄り添う対話に必要な技術要素である。

今後は、他の非言語情報(例:ジェスチャー)と組み合わせ、推定する感情の種類を増やしたり、特定の状況(例:自動運転中の車内)を考慮したりして、より高度な寄り添い対話システムの構築を目指していきたいと考える。

本稿に記載した内容の一部は、総務省 平成30年度 情報通信技術の研究開発に係る提案「高度対話エージェント技術の研究開発・実証」の研究成果によるものである。

## 5. 参考文献

- [1] <http://www.paulekman.com/> - Paul Ekman, 表情の分類
- [2] <http://dlib.net/> - Dlib C++ Library
- [3] James A. Russell, "A Circumplex Model of Affect", Journal of Personality and Society Psychology 1980, Vol. 39, No. 6, 1161-1178