

卒業論文要旨  
テレビ電話を用いた表情分析の精度向上に関する研究

(情報伝達システム学) 嶋田 亮一

## 1. 緒言

当研究室の表情分析の研究[1-4]では、動画像から OpenCV[5]を用いて人物の口領域画像の抽出を行ってきた。そして、口領域画像を用いて表情強度[1]を測定し、表情分析に利用してきた。被験者が無発声で無表情である際に取得されたフレーム(以下、「基準フレーム」と表記)から抽出された口領域画像の特徴ベクトルと、表情分析の対象となるフレームから抽出された口領域画像の特徴ベクトルとの差分ベクトルのノルムを表情強度と定義している[1]。ところが、OpenCVを用いて抽出された口領域画像における口中心位置が一定となる保証はなく、実際に変動していることが観察されている。以下では、この変動を、「口領域画像抽出変動」と呼ぶ。

そこで、本研究では、抽出された口領域画像における口中心位置が表情強度に及ぼす影響を調査し、口領域画像抽出変動が表情強度に与える影響を小さくする方策を検討した。

## 2. 開発環境

OS : Windows 7 Professional

PC : DELL OPTIPLEX780 (CPU : Intel Core2 Duo CPU E8400 3.00GHz, RAM : 4.00GB)

使用プログラミング言語 : Visual C++ 2008 Express Edition

## 3. 表情分析法における口領域抽出と表情強度測定

### 3.1 口領域抽出

まず、動画ファイルから各フレームを抽出し、OpenCVの正面顔の分類器を用いて顔検出を行う。この分類器では、Haar-like 特徴量と Adaboost 学習アルゴリズムを用いている[6]。顔領域の検出を終えると、更にその中で口領域の検出を行う[4]。

### 3.2 表情強度測定

検出された口領域の各  $8 \times 8$  画素領域で 2D-DCT 処理を施し、直流成分を除く低周波の帯域における 2D-DCT 係数だけを後述する表情特徴量を求める計算に用いる[1]。各周波数帯域における 2D-DCT 係数の絶対値の口領域における平均値を 1 つの表情特徴量(DCT 特徴量)[1]とする。直流成分を除く計 15 個の DCT 特徴量を求め、これらの特徴ベクトルの要素とする。そして、表情強度[1]を測定する。

## 4. 基準フレームの決定方法

既報の方法[4]で動画像の開始から発声と判定される直前までの各フレーム(以下、「無発声フレーム」と表記)を、基準フレーム候補とする。そして、各基準フレーム候補を基準フレームとして用いて、他の全無発声フレームの表情強度の総和を計算し、その総和が最小となる基準フレーム候補を基準フレーム(図 1)とする。

## 5. 口領域画像における口中心位置の変動の測定

口領域を抽出する際、口領域画像における口中心位置に若干の変動が生じることがある。そこで、基準フレームと無発声フレームの口中心位置のズレをペイントソフトの目盛を利用し目視で測定した。

## 6. 実験

### 6.1 方法

5章で求めたズレを用いて、「閉口」「小開口」「大開口」時の3フレーム(図2)に対して口の位置を補正し表情強度を計算した。図3に「小開口」のフレームを、画素単位で、x軸方向に2、y方向に-1の補正を行った例を示す。



図1 基準フレーム



図2 閉口(左)、小開口(中)、大開口(右)



図3 小開口の補正例(x:2, y:-1)

### 6.2 結果と考察

口の位置が縦方向にずれたとき、少し口が開いたフレームの表情強度が大きく変化することが分かった(表1)。縦方向の口領域画像抽出変動の表情強度への影響を少なくするためには、口領域画像抽出後に、位置補正なしの場合と、縦方向に位置補正(例えば、画素単位で、-2, -1, 1, 2)を行った場合の表情強度を測定し、それらの最小値を採用する方策が考えられる。

表1. 口の位置補正を行った際の表情強度の変化

| 口位置補正(単位:画素) |    | 閉口   | 小開口   | 大開口   |
|--------------|----|------|-------|-------|
| x            | y  |      |       |       |
| -1           | -1 | 6.28 | 8.37  | 24.80 |
| -1           | 0  | 4.32 | 14.02 | 23.73 |
| 0            | -1 | 6.09 | 8.15  | 25.28 |
| 0            | 0  | 4.33 | 14.00 | 24.20 |
| 1            | -1 | 5.67 | 7.78  | 25.84 |
| 1            | 0  | 3.79 | 13.60 | 25.02 |
| 2            | -1 | 5.54 | 8.41  | 26.28 |
| 2            | 0  | 3.67 | 13.82 | 26.01 |

## 7. 結言

口領域画像抽出変動の表情強度への影響を調査し、この影響を小さくする方策を提案した。今後は、この方策の有効性を検証していく。

### 【参考文献】

- [1] 浅田太郎, 吉富康成, 辻愛里, 加藤亮太, 田伏正佳, 桑原教彰, 成本迅, 「テレビ電話中の表情分析手法の検討」, ヒューマンインタフェースシンポジウム 2013 論文集, pp.493-496, 2013.
- [2] T.Asada, Y.Yoshitomi, A.Tsuji, R.Kato, M.Tabuse, N.Kuwahara, and J.Narumoto, "Facial Expression Analysis While Using Video Phone", Proc. of Int. Conf. on Artificial Life and Robotics, pp.230-234, 2014.
- [3] 浅田太郎, 吉富康成, 加藤亮太, 田伏正佳, 成本迅, 「テレビ電話中の表情分析と顔動作解析」, ヒューマンインタフェースシンポジウム 2014 論文集, pp.793-796, 2014.
- [4] T.Asada, Y.Yoshitomi, R.Kato, M.Tabuse, and J.Narumoto, "Quantitative Evaluation of Facial Expressions and Movements of Persons While Using Video Phone", Proc. of Int. Conf. on Artificial Life and Robotics, pp.294-297, 2015.
- [5] OpenCV Web ページ, <http://opencv.willowgarage.com/>
- [6] P.Viola, P., Jones, M., J., "Rapid Object Detection Using a Boosted Cascade of Simple Features", Proc. of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol.1, pp.511-518, 2001.