

特徴選択された LBP-TOP を用いた微表情による感情推定 Emotional Estimation by Micro-expression Using Selected Features of LBP-TOP

加藤 功基[†] 高野 博史[†] 西光 雅弘[‡] 平池 龍一[‡] 今岡 仁[‡]
Koki Kato Hironobu Takano Masahiro Saiko Ryuichi Hiraike Hitoshi Imaoka

1. はじめに

「悪い知らせ」は、患者の将来の見通しを根底から否定的に変えてしまうものである[1]。そのため医師と患者のコミュニケーションが重要視されており、医療の間では CST (Communication Skills Training) によって、コミュニケーションの質の向上を図っている。しかし、この CST では実際の患者を診るわけではないため、認知的共感という相手の立場で気持ちを考えることが検討されない。また、慌ただしい医療現場は、医師が患者の感情を推測することが困難な状況にある。そこで本研究では、医師が悪い知らせを患者に伝える際のコミュニケーションをサポートするために、表情の中でも自分では制御しにくいといわれている微表情を用いた感情推定法の開発を目的とする。

2. 関連研究

2.1 微表情データセット

表 1 に微表情のデータセットの一覧を示す。SMIC は、positive、negative、surprise の 3 感情に、CASME II は、happiness、surprise、disgust、repression、other の 5 感情に、SAMM は、happiness、surprise、disgust、anger、fear、sadness、contempt の 7 感情に分類されている。

表 1 微表情データセット

Dataset	Subjects	Samples	Emotion Classes
SMIC(HS) [2]	16	157	3
SMIC(VIS)	8	71	3
SMIC(NIS)	8	71	3
CASME II [3]	26	247	5
SAMM [4]	32	159	7

2.2 先行研究

Liong らは、SMIC 及び CASME 微表情データセットを用いて感情推定を行った[5]。特徴量には LBP-TOP (Local Binary Patterns Three Orthogonal Planes)、分類器には SVM (Support Vector Machine) を用い、LOVO (Leave One Video Out) 交差検証を用いて評価を行った。ROI (Region of Interest) を目、口で選択し、顔全体で特徴量を取得した場合の感情推定精度と比較した。顔全体での Accuracy は 59.15%であったのに対し、目と口を切り出したときの Accuracy は 62.74%となり約 3%精度が向上した。

3. 提案法

本研究では、部分顔に対して LBP-TOP 特徴量を算出し、特徴選択を行った後、SVM を用いて感情を推定する手法を提案する。次に、提案法の詳細について述べる。

[†] 富山県立大学 Toyama Prefectural University

[‡] 日本電気 NEC Corporation

3.1 前処理

特徴量を計算する画像は、微表情が表れる動画をフレームレート 30 fps で切り出した。Dlib の face landmark detector を用いて 68 点ランドマークを検出し、顔部分及び ROI を切り出した[6]。本研究での ROI の選択方法を図 1 に示す。ROI1 は目と口、ROI2 は眉間と口、ROI3 は眉間を含めた両目と口とした。

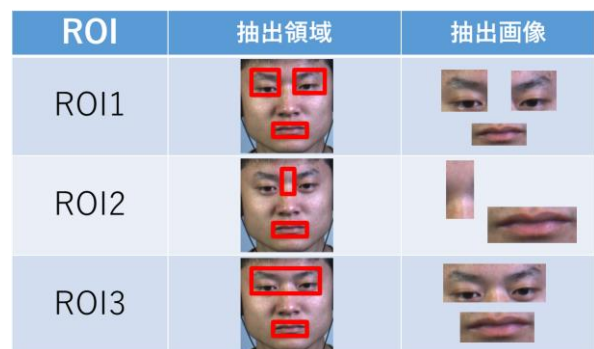


図 1 ROI の選択方法

3.2 特徴量抽出

本研究では、特徴量として LBP-TOP 特徴量を用いた。LBP-TOP 特徴量は、局所領域の輝度の差を捉える LBP 特徴量を、時空間の情報を扱えるように拡張したものである。図 2 に LBP-TOP の算出方法を示す。LBP-TOP 特徴量は、画像を示す XY 平面と画像の X 軸、Y 軸と時間の T 軸によって作られる XT 平面、YT 平面それぞれに対して LBP 特徴量の計算を行い、これらのヒストグラムを連結することにより得られる。特徴量を算出する領域を ROI に分割した際は、分割した部位それぞれで LBP-TOP 特徴量を算出し、各顔部位で得られたヒストグラムをさらに連結することで、最終的な LBP-TOP 特徴量が得られる。

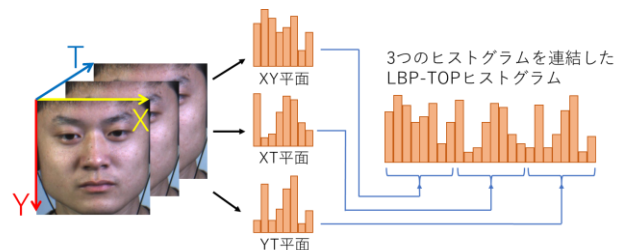


図 2 LBP-TOP の算出方法

3.3 特徴選択

LBP-TOP 特徴量は、一つの顔部位に対して 256 次元の特徴ベクトルとして表される。そのため、ROI の数が増えるたびに特徴量の次元は増える。特徴量の次元が増加すると、感情推定に有用な情報が他の情報に埋もれてしまう可能性

が考えられる。そこで、本研究では、あるクラスのまとまりを示すクラス内分散 σ_W^2 と、各クラスの広がりやクラス間分散 σ_B^2 を用いて、クラス間分離度を評価するクラス内分散・クラス間分散比 J_σ を用いて特徴選択を行った。 J_σ は、式(3)のように式(1)に示すクラス内分散と式(2)に示すクラス間分散の比によって求められる。ただし、 χ_i はクラス i に属するパターンの集合、 \mathbf{x} は χ_i に含まれる特徴ベクトル、 \mathbf{m}_i は χ_i の平均ベクトル、 \mathbf{m} は全パターンの平均ベクトル、 n は全パターンの個数、 t は転置を表している。特徴選択では、LBP-TOP 特徴量の各次元に対して J_σ を求め、顔部位ごとに 256 次元の特徴ベクトルから次元削減の割合に応じて、 J_σ の値が大きい成分から選んだ。

$$\sigma_W^2 = \frac{1}{n} \sum_{i=1}^c \sum_{\mathbf{x} \in \chi_i} (\mathbf{x} - \mathbf{m}_i)^t (\mathbf{x} - \mathbf{m}_i) \quad (1)$$

$$\sigma_B^2 = \frac{1}{n} \sum_{i=1}^c n_i (\mathbf{m}_i - \mathbf{m})^t (\mathbf{m}_i - \mathbf{m}) \quad (2)$$

$$J_\sigma = \frac{\sigma_B^2}{\sigma_W^2} \quad (3)$$

4. 実験

4.1 評価方法

本研究では、感情推定の評価に LOVO 交差検証を用いた。これは、ビデオ一つをテストデータとして用い、残りのビデオをトレーニングデータとして用いる方法である。すべてのデータがテストデータとして扱われるように、テストデータとトレーニングデータを入れ替えて学習を繰り返した。分類された結果から識別率 (Accuracy) と F 値を算出し、提案手法の評価指標とした。

4.2 実験結果

3種類 of データセットについて、顔全体と各 ROI、特徴選択ありとなしの感情推定精度を図 3～図 5 に示す。棒グラフは Accuracy を示し、実線は F 値を示す。また、特徴選択における Accuracy と F 値は、特徴ベクトルの次元削減によって得られた最大のものを示している。顔全体及び特徴選択なしと比較して、ROI の選択、特徴選択によって得られた Accuracy、F 値は増加した。

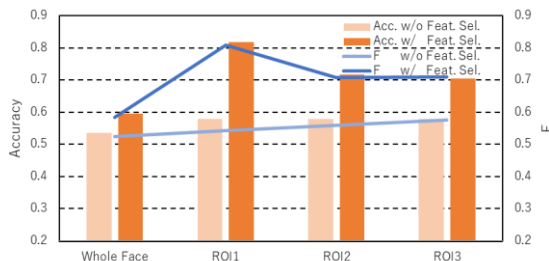


図 3 SMIC における Accuracy と F 値

4.3 考察

実験結果から ROI1 が最も良い選択方法と考えられる。しかし、両目を除いた ROI2 の分割方法でも顔全体よりも Accuracy は増加し、データセット CASME II においては最も良い Accuracy の値を示した。このことから眉間も微表情に

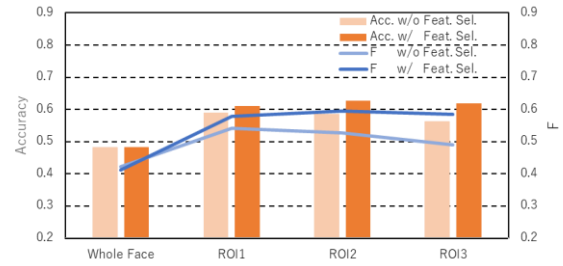


図 4 CASME II における Accuracy と F 値

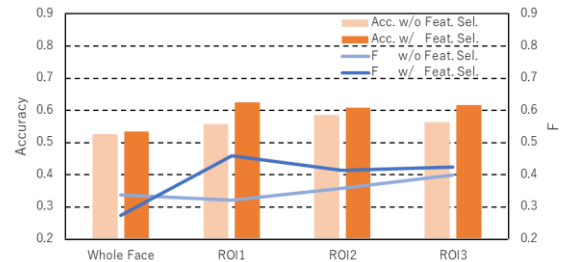


図 5 SAMM における Accuracy と F 値

よる感情推定において一つの有用な材料と考えられる。

クラス内分散・クラス間分散比を利用して特徴量の次元を減らして特徴選択を行った結果、Accuracy は増加し、特に SMIC において良い結果をもたらした。SMIC は 3 クラス分類であるので、クラス間の差がより顕著に現れたのではないかと考えられる。より多くのクラスを分類する場合、クラス内分散・クラス間分散比に大きな差が表れず、あまり精度が向上しなかったと考えられる。

F 値に関しては SAMM での値が小さくなった。このデータセットでは、7 クラスの感情があり、データに偏りがあるため、分類されにくいクラスが存在してしまったためであると考えられる。

5. おわりに

本研究では、特徴量を抽出する領域を ROI に分割することと、クラス内分散・クラス間分散比を用いて特徴選択をすることにより、微表情による感情推定の識別精度が向上するか試みた。実験結果から、ROI の分割、特徴選択により精度は向上した。今後は、メトリックラーニングを導入し、感情推定精度の評価を行う。

参考文献

- [1] R. Buckman, "Breaking Bad News: Why Is It Still So Difficult?", *British Medical Journal*, Vol.288, pp.1597-1599 (1984).
- [2] X. Li, T. Pfister, X. Huang, G. Zhao, and M. Pietikainen, "A spontaneous micro-expression database: Inducement, collection and baseline", *Proc. 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition*, pp.1-6 (2013).
- [3] W.-J. Yan, X. Li, S.-J. Wang, G. Zhao, Y.-J. Liu, Y.-H. Chen, and X. Fu, "CASME II: an improved spontaneous micro-expression database and the baseline evaluation", *Plos One*, Vol.9, No.1, pp.1-8 (2014).
- [4] A. Davison, C. Lansley, N. Costen, K. Tan, and M. H. Yap, "SAMM: a spontaneous micro-facial movement dataset", *IEEE Trans. Affect. Comput.*, Vol.9, No.1, pp.116-129 (2016).
- [5] S. Liong, J. See, R. Phan, K. Wong, and S. Tan, "Hybrid facial regions extraction for micro-expression for recognition system", *J. Sign. Process. Syst.*, Vol.90, pp.601-617 (2018).
- [6] Dlib C++ Library, <http://dlib.net>