

自然写真と赤外線写真からの顔表情認識 Recognition of Facial Expression by Natural Color and Infrared Images

潘 昕[†] 青野 雅樹[†]
Xin Pan Masaki Aono

1. はじめに

人の表情には人間の豊富な心理及び感情の情報が含まれている。人間と人間がコミュニケーションを行う場合、言語だけでなく非言語も有効に使われている。非言語、特にその中でも表情は、異文化間でも普遍的な特徴を持つ。そのため、人間が機械とコミュニケーションを行う場合も非言語に対する期待が高まっている。近年のマルチメディアの進展に伴い、顔表情を表す画像が得られたとしても、その品質や明るさなどが異なり、表情特徴量を抽出することは容易でない。一方、赤外線画像が利用できれば人体の体温など色画像とは別の情報を利用することができる。本研究では、顔表情の静止色画像と赤外線画像から、特徴量を抽出して、併合した特徴量を用いて、表情認識正解の向上を目指して実験を行なったので、これを報告する。

2. 関連研究

表情などによる非言語のコミュニケーションに関する研究者である Paul Ekman らによって 1978 年に FACS (FACS: Facial Action Coding System) が提案された[1]。FACS では人の感情表現において、文化や教育の影響を受けず、人種にかかわらず顔を通じて普遍的に認知する、または表現することができる感情として次の 6 つをあげている。喜び (Happiness)、悲しみ (Sadness)、怒り (Anger)、嫌悪 (Disgust)、驚き (Surprise)、恐れ (Fear) の 6 種の表情と無表情 (Neutral) を含め、7 個の表情が基本表情となっている。

2002 年には Joseph Hager らによって、顔表情を表現するユニットとして顔の表情動作の最小単位である AU (Action Unit) を修正した新しいバージョンがインターネットで公開された[2]。

3. Gabor Filter

Gabor Filter は、2次元ガウス曲面に2次元平面上を一方に伝わる平面波を畳み込んだフィルタであり、線エッジに対する反応が強く、かつ鋭い方向選択性を持っている[3]。この性質は視覚系における方向選択性に対応するものと捕らえることができる。

本報告では、Gabor Filter による特徴検出における能力に注目し、Gabor Filter を顔の自然写真と赤外線写真に適用し、得られたデータを表情の特徴量として使用する。

4. 次元圧縮処理

抽出された特徴はベクトルの形で表すが、データのサンプル数に対してベクトルの次元数が膨大になると、処理が困難になる。線形次元圧縮処理は特徴空間の変換と考えることができるので、ある n 次元のベクトル x を m ($n \gg m$) 次元のベクトル y に変換する処理、すなわち、 $y = A^T x$ となる変換行列 A を求める問題に帰着する。

PCA (Principal Component Analysis: 主成分分析) と LDA (Linear Discriminant Analysis: 線形判別法) が有名な手法である。PCA はパターン全体の分布を最良近似する部分空間を求める方法で、LDA は各クラスのパターン分布の距離度を最大にする部分空間を求める方法である。実験には PCA で削減したデータから、LDA で処理する。

5. 実験

5.1 実験データ

今回実験で使用したデータでは、中国科学技術大学が作り、2010年9月に発表した USTC-NVIE Database である[4]。その USTC-NVIE Database は Posed Database と Spontaneous Database を含む。Posed Database は通常のデータベースのように参加者が指示された表情を、大げさに表情して作成されたデータベースである。一方で、Spontaneous Database はビデオを使って、参加者が自然な表情を作ったデータベースである。且つ、USTC-NVIE Database の特徴は自然画像と赤外線画像がセットになっている点である。図1は各6種類表情(左から:喜び、嫌悪、驚き、恐れ、悲しみ、怒り)の自然画像と赤外線画像の例である。

今回の実験は Spontaneous Database から嫌悪 (79 枚)、恐れ (58 枚)、喜び (86 枚)、合計で 223 枚の自然画像と同じ 3 種類の赤外線画像 223 枚を使った。

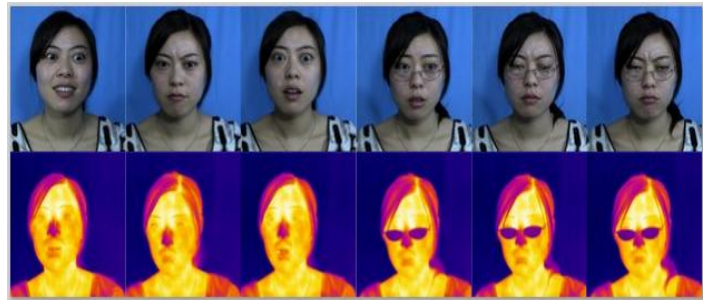


図1 6つの表情例 (上: 自然画像、下: 赤外線画像)

5.2 実験の流れ

Spontaneous Database に含まれる自然画像 (704×480) と赤外線画像 (304×230) のサイズは異なる。また、図1の例を見ると、自然画像や赤外線画像には表情認識に不要な情報の背景や髪などが大量に含まれている。そのため、前処理が必要となってくる。自然画像と赤外線画像を同じサイズ (120×120) にさせ、画像平滑化して元画像を変化させる。次に、Gabor Filter に関する周波数と傾きを変えた事前実験より、自然画像では 40 個フィルタを使い、赤外線画像では 10 個フィルタを使う場合が最も有効であり、この設定で各特徴量を抽出する。その後、抽出された特徴量は PCA と LDA で順次次元圧縮する。最

[†] 豊橋技術科学大学
Toyohashi University of Technology

後に、自然画像の特徴量と赤外線画像の特徴量を併合して、テスト特徴量と学習特徴量を分け、SVM (Support Vector Machine) 分類器を用いてテスト特徴量に対して、分類実験を行う。全体的な実験流れは図2のようになっている。

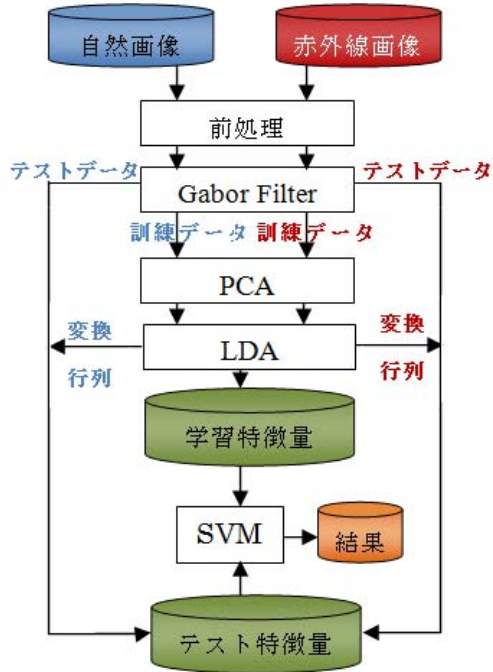


図2 実験の流れ

5.3 実験の結果

図3は各特徴量による表情の認識率結果である。自然画像特徴量を用いた正解では60.9%、赤外線画像特徴量を用いた正解では28.6%、両者を併合した特徴量を用いた正解では、58.2%という結果を得た。

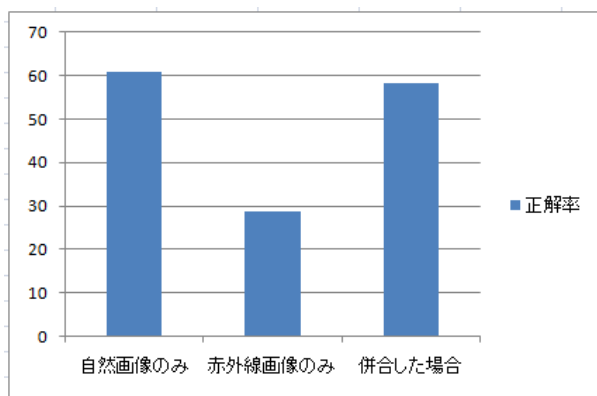


図3 各特徴量の認識率

図4は各3種類の表情の認識率結果を示す。嫌悪と恐れを併合した特徴量を用いた場合、自然画像だけの特徴量を用いるより認識の正解率は低くなっている。しかし、喜びの正解率は自然画像特徴量を用いた場合75.5%、赤外線画像特徴量を用いた場合43%、両者を併合した特徴量を用いた場合76.7%と正解率の向上が見られた。

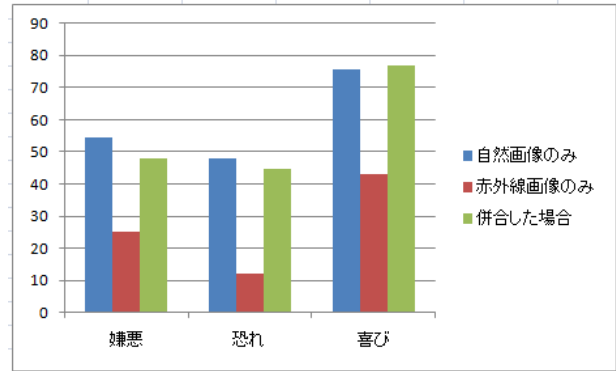


図4 各表情の認識率

6. 考察

自然画像と赤外線画像の特徴量を組み合わせることで認識率の向上を試みたが、3つの感情表現の平均では、下がる結果になってしまった。その理由として、赤外線画像には体温の影響をある程度反映し、感情の変化に伴う、赤外線画像の変化が併合により大きく寄与すると思っていたが、期待していたほど大きくなかったことが考えられる。

今回の実験では表情の認識において、自然画像と赤外線画像から各特徴量を抽出する Gabor Filter を制御するパラメータが最終的な認識率に影響を与えると考えられる。具体的には Gabor Filter のパラメータ周波数と傾きの2種類を変化させて実験を行った。実験結果より、表情の種類を問わず、総正解数が一番良い40個(自然画像用)のフィルタと10個(赤外線画像用)のフィルタを決めた。図4の実験結果を見ると、この10個を用いた赤外線用フィルタは嫌悪と恐れに対して、有効な特徴量をあまり抽出していなかったと考えられる。各表情に対して、有効な特徴量を抽出する点で、まだ改善の余地がある。今後は、Gabor Filter 以外の特徴量の考慮や、別な次元削減の方法などを試す予定である。

謝辞

USTC-NVIE Databaseは中国科学技術大学に提供いただいた。本研究の一部は科学研究費(基盤研究(C)No.23500119)による。

参考文献

- [1] Ekman, Paul, Freison, Wallace V., Anli, Sonia "Facial signs of emotional experience" Journal of Personality and social Psychology, Vol 39(6), pp.1125-1134, Dec 1980
- [2] Donato, G., Bartlett, M.s, Hager, j.c, Ekman, P., Sejnowski, T.j "Classifying facial action" Pattern Analysis and Machine Intelligence, IEEE Transactions on , Issue.10 pp.974-989 Date: Oct 1999
- [3] Y. Hamamoto, S. Uchimura, M. Watanabe, T. Yasuda, Y. Mitani and S. Tomita, "A Gabor Filter-based method for recognizing handwritten numerals" Pattern Recognition, vol31, no.4, pp.395-400, 1998
- [4] Pattern Recognition, vol31, no.4, pp.395-400, 1998 Shangfei Wang, Zhilei Liu, Siliang Lv, Yanpeng Lv, Guobing Wu, Peng Peng, Fei Chen, Xufa Wang, A Natural Visible and Infrared Facial Expression Database for Expression Recognition and Emotion Inference, IEEE Transactions on Multimedia, VOL. 12, No. 7, Page(s): 682-691, NOVEMBER 2010.