

変化パターンの区間発生ヒストグラム学習に基づく顔表情変化認識に関する研究

情報科学研究科 情報科学専攻 H11224M 松久ひとみ 指導教員：橋本学 教授

1. はじめに

近年、うつ病や躁うつ病を中心とした気分障害を発症する人が増加傾向にあり、患者の早期発見が重要である。笑いの回数はうつ病の病状と密接に関係していることから、笑い(喜び)の顔表情変化の認識が望まれている[1]。表情変化時刻は1枚の静止顔画像では検出できないが、連続した時間における複数枚の静止顔画像を用いれば検出できると考えられる。また、一般に問診中では大仰な(表情の出方が顕著な)表情を表出することは少ないため、微妙な(表情の出方があまり顕著でない)表情を捉える必要がある。

2. 従来研究の問題点

表情認識の従来研究として、局所的な領域の方向と空間周波数を Gabor 特徴量によって捉えて、表情を識別している手法[2]がある。これは8方向のフィルタしか用いていないため、微妙な表情変化時に生じるしわや口角などの局所的な領域内の微細な特徴を捉えるのは困難と考えられる。

表情変化時刻検出の従来研究として、入力された顔画像から無表情または表情表出を識別し、一定時間における表情の出現頻度を用いる手法[3]がある。これは、表情の誤識別が起因して表情変化時刻を誤検出する問題点がある。

3. 表情変化認識

本研究では、表情変化認識を表情識別と表情変化時刻検出の2つのモジュールで構築した。表情変化認識とは、時系列にならぶ顔画像に対して、表情の種類(「喜び」または「悲しみ」など)、顕著さ(「微妙な表情」または「大仰な表情」など)の変化を認識することである。本研究では、表情変化認識の中でも特に表情の種類の変化に着目し、「無表情から喜び」または「喜びから無表情」に変化した時刻を検出する。なお、本研究では表情変化時刻を人間が目視で表情変化したと判断した2フレーム間の時刻と定義している。

3.1 微妙な表情識別

微妙な表情には、無表情と比べて顔表面のしわの深さや口角の角度、目の幅、口の開閉に見えるの違いが現われ、形状変化が生じる。これがさらに画像上では比較的安定した濃淡変化として現われる。以上のことから、本研究では、事前に設定した特定部位における濃淡変化と局所的な形状変化に着目した。微妙な特徴を捉えるためには、連続的なフィルタによって、方向と空間周波数を捉えることが必要であるため、本研究では Gabor フィルタ[4]を用いた。これは式(1)、(2)に示されるような指数関数型のフィルタである。なお、 $\sigma = \frac{1}{\pi} \cdot \frac{2^b + 1}{2^b - 1} \cdot \lambda$, $b = 1$, $\varphi = 0$, $\gamma = 0.5$ を用いる。本研究では、方向角 θ や広がり具合を規定する σ を微妙な表情表出時のしわや口角の角度やほりに合わせ、それぞれ N_θ 種類、 N_σ 種類変化させて生成される複数のフィルタを利用する。ウィンドウの候補数を N_w とすると、1つの顔画像に対して $G(N_w \times N_\theta \times N_\sigma)$ 個の Gabor 特徴を抽出する。

$$g_{\lambda, \theta}(x, y) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{(2\sigma)^2}\right) \cos\left(\frac{2\pi}{\lambda} x' + \varphi\right) \quad (1)$$

$$\begin{cases} x' = x \cos \theta + y \sin \theta \\ y' = -x \sin \theta + y \cos \theta \end{cases} \quad (2)$$

学習時には、学習顔画像の Gabor 特徴量を AdaBoost アルゴリズム[5]に適用することによって、表情識別に有効なウィンドウを自動的に設定し、表情識別器を構築する。識別時には、テスト顔画像をこの表情識別器に入力して出力を得る。

3.2 表情変化時刻検出

表情変化時刻検出では、識別された表情(無表情(0)、微妙な喜び(1)、大仰な喜び(2))の時系列のならばを用いる。このならばを変化パターンと呼ぶ。学習時には、図1に示すように、変化パターンの表情変化する確率を発生ヒストグラムとして生成する。なお、人間が目視で判断した表情変化時刻に個人差があるため、教師信号としての表情変化時刻の情報を確率的に拡張する。表情変化の発生ヒストグラムを3種類の区間長($\pm 1, \pm 2, \pm 3$ フレーム)に対して生成する。そして、3種類の区間発生ヒストグラムに対してそれぞれ重み $\omega_1, \omega_2, \omega_3$ を重み付け加算し、3種類の区間発生ヒストグラムを1つに統合する。なお、確率値の最大値が100%になるように正規化しておく。

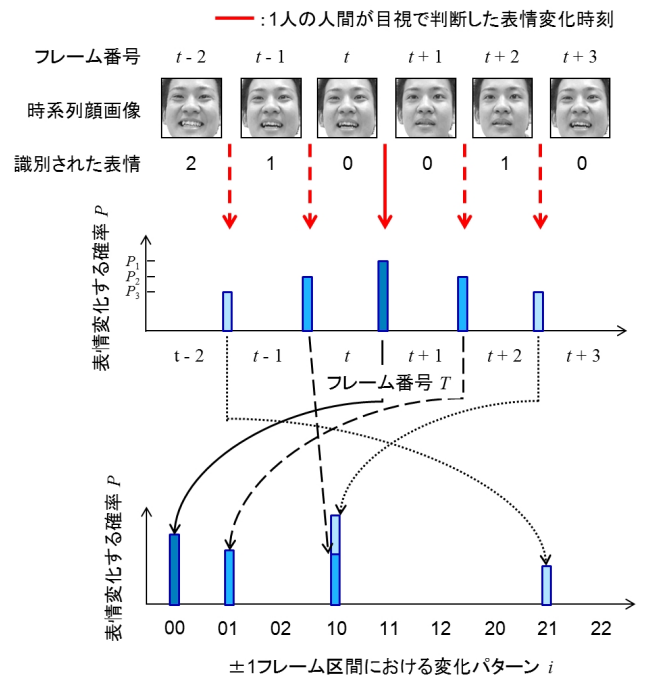


図1 表情変化パターンの発生ヒストグラムの学習の例

検出時には、時系列にならぶ顔画像を表情識別し、変化パターンと学習した発生ヒストグラムを照合する。この値が、設定したしきい値 P_{th} 以上の値であれば、その時刻で表情変化したとして検出する。

4. 実験結果と考察

男性6名、女性2名の8名の被験者(A~H)について、喜びを表出している顔画像を用いた。8名を30fpsで撮影したところ、少なくとも2秒間は表情が継続することを確認した。したがって、1fps以下であれば表情を過不足なく取得できるため、動画のフレームレートを1fpsとした。また、精神科医が患者におこなう問いかけのやり取りが1つあたり3~4分であるため、撮影時間は200secとした。

4.1 静止画像に関する表情認識の性能評価実験

被験者(A, B, C)の顔画像を用いて、無表情と微妙な表情、または無表情と大仰な表情の識別性能を評価した。識別性能の評価は、式(3)の識別成功率 P_r を用いた。なお、実験では顔領域を 119×119 画素、両目の距離を51画素に正規化した。また、方向 θ を11.25度ごとに $N_\theta = 16$ 種類、空間周波数 σ を2.8, 3.4, 3.9, 4.5[pixel]の $N_\sigma = 4$ 種類とし、 $G = 64$ 種類の Gabor フィルタを用いた。個人ごとに学習させたときの平均識別成功率 P_r を表1に示した。提案手

法は、大仰な表情の識別器では 97.0%、微妙な表情の識別器では 84.7%の識別成功率を確認した。

$$P_r = \frac{\text{正しく識別された画像数}}{\text{テスト画像に用いた画像数}} \times 100 \quad (3)$$

表 1 提案手法を用いた表情の識別成功率 P_r [%]

	Exaggerated	Subtle
Proposed	97.0	84.7

4.2 提案手法と比較手法による表情の識別成功率の比較実験

比較手法は Hong-Bo らの手法 [2] である。 N_σ 種類の広がり度合いと N_θ 種類の方向角をもつ集合を global Gabor filter bank ($G(N_\sigma \times N_\theta)$) と表わしている。さらに、global Gabor filter bank の一部をもつ集合を local Gabor filter bank ($LG(N_\sigma \times N_\theta)$) と表わしている。LG1 は、 σ の値を最小値から最大値まで増加させ、 θ の値を最小値から最大値まで増加させたフィルタを選択する方法である。LG2 は、 σ の値を最大値から最小値まで減少させ、 θ の値を最小値から最大値まで増加させたフィルタを選択する方法である。LG3 は、 σ の値を増加させる間隔を 1 つ増やしてフィルタを選択する方法である。本実験では 6 種類の local Gabor filter (LG1, LG2, LG3 ($N_\sigma = 4, 3, N_\theta = 16$)) を用い、4.1 と同様に被験者 A, B, C を用いた。local Gabor filter を用いて被験者を個別に学習させ、100 次元における平均識別成功率 P_r を求めた。比較手法の識別成功率を表 2 に示す。表 1 と表 2 より、比較手法の識別成功率の最大値が、大仰な表情では LG3 (4×16) のときに 94.4%、微妙な表情では LG1 (4×16) のときに 80.7%であることを確認した。また、提案手法の識別成功率は、比較手法の識別成功率の最大値よりも大仰な表情では 2.6%、微妙な表情では 4.0%高いことを確認した。

表 2 比較手法 [2] を用いた表情の識別成功率 P_r [%]

Local Gabor filter	Exaggerated	Subtle
LG1 (4×16)	94.3	80.7
LG2 (4×16)	93.5	80.0
LG3 (4×16)	94.4	79.5
LG1 (3×16)	92.8	76.4
LG2 (3×16)	92.3	76.4
LG3 (3×16)	90.7	76.6

4.3 表情識別に有効な特徴量

表情識別に有効な特徴量を図 2 に示す。これは弱識別器の信頼度が上位 15 位以内のときの特徴量の位置である。実際には、弱識別器の信頼度は 100 位程度までであるが、順位が低いことを省略した。微妙な表情の識別器では口の輪郭付近、大仰な表情の識別器ではほうれい線付近の信頼度が高かった。2 つの識別器で選択した特徴量は鼻孔付近や口付近であり、大まかな顔部分領域としては共通していた。信頼度が 6 位以降になると、目付近の特徴量が選択されることがわかる。また、3 名において喜びとニュートラルの表情の識別に有効な顔部分領域は、特に鼻孔付近であった。この顔部分領域の特徴量は、喜びの表情の表出時に、個人によらず現われることがわかった。

4.4 動画像に関する表情変化時刻検出の性能評価実験

被験者 (D~H) の顔画像を用いて、表情変化時刻の検出性能を評価した。なお、表情変化する確率値 P_1 を 80、 P_2 を 50、 P_3 を 30、区間 1, 2, 3 に対応する重み ω_1 を 0.5、 ω_2 を 0.25、 ω_3 を 0.25、さらに、しきい値 P_{th} を 75 とした。表情変化時刻検出の性能評価には、式 (4)(5) に示される適合率 R_p と再現率 R_r を用いた。また、提案手法と、野宮らの手法 [3] による表情変化時刻の検出性能を比較した。表情変化時刻の検出性能を表 3 に示す。提案手法は比較手法と比べ、適合率・再現率ともに高いことを確認した。提案手法は平

均適合率 87.8%、平均再現率 84.5%を確認し、比較手法 [3] と比べて適合率 29.3%、再現率 39.3%の性能が向上した。なお、表情変化の発生ヒストグラムを 4 種類の区間長 (± 1 , ± 2 , ± 3 , ± 4 フレーム) に対して生成した。適合率と再現率が最も高くなる組み合わせは、区間 1, 2, 3 であることを確認した。

$$R_p = \frac{\text{正しく検出された表情変化数}}{\text{検出された表情変化数}} \times 100 \quad (4)$$

$$R_r = \frac{\text{正しく検出された表情変化数}}{\text{テスト画像に含まれる表情変化数}} \times 100 \quad (5)$$

Height of reliability : ○1~5位, △6~10位, □11~15位

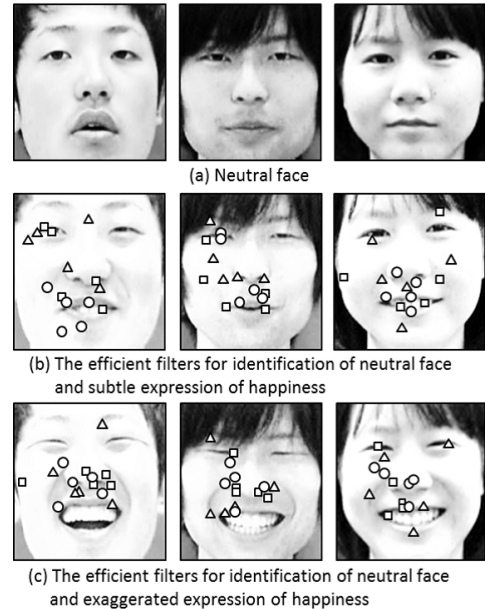


図 2 表情識別に有効な特徴量

表 3 表情変化の認識性能 [%]

Method	Precision rate	Recall rate
Comparative [3]	58.5	45.2
Proposed	87.8	84.5

5. おわりに

表情識別に有効なウィンドウ、Gabor フィルタの種類を最適化することによって、無表情と微妙な表情を識別する手法を提案した。さらに、識別された表情の時系列のならばを変化パターンと定義し、これを考慮することにより、安定して表情変化時刻を検出する手法を提案をした。提案手法は、微妙な表情の識別では、識別成功率 84.7%となり、比較手法よりも識別成功率が 4.0%高いことを確認した。さらに、提案手法は、表情変化時刻の検出では適合率 87.8%、再現率 84.5%を達成し、比較手法よりも適合率 29.3%、再現率 39.3%の性能向上を確認した。

参考文献

- [1] 角辻豊, “精神科医からみた顔画像認識への期待,” 日本医用画像工学会誌, vol.12, no.6, pp.700-709, 1994.
- [2] Hong-Bo Deng, Lian-Wen Jin, Li-Xin Zhen, and Jian-Cheng Huang, “A New Facial Expression Recognition Method Based on Local Gabor Filter Bank and PCA plus LDA,” International Journal of Information Technology, vol.11, no.11, pp.86-96, 2005.
- [3] 野宮浩揮, 宝珍輝尚, “アンサンブル学習を用いた効率的な映像からの表情表出シーン検出,” 信学論 (D), vol.J95-D, no.2, pp.193-205, 2012.
- [4] S.E.Grigorescu, “Comparison of texture features based on Gabor filters,” IEEE Trans. Image Process, vol.11, no.10, pp.1160-1167, 2002.
- [5] P.Viola and M.Jones, “Rapid object detection using a boosted cascade of simple features,” Proc. Of CVPR, pp.511-518, Dec. 2001.