

顔画像からの表情認識のための一手法の提案

7D-7

井澤広泰[†] 佐治 斉[†] 品川 嘉久[†] 國井利泰[‡]東京大学[†] 会津大学[‡]

1 導入

人間の表情を解析することは、近年さまざまな分野において重視されてきている。そのためには、顔画像から、顔の各特徴（目、耳、鼻、口など）を認識する必要がある。さらに、その特徴を認識するためには、その特徴を構成しているエッジを抽出する必要がある。本研究においては、最適曲線を利用することにより、特定の顔の特徴の知識を用いずにエッジを求める。

2 エッジの目立ちやすさ

人間の見方によって、画像上のエッジの数は変化する。つまり、エッジの中には、目立ちやすいものと目立ちにくいものが存在する。本研究においては、エッジの目立ちやすさを次の二つの面から考えることにした。

- エッジの空間的な広がり（長いエッジは目立つ）
- エッジのコントラスト（エッジの両側の輝度値の差が激しいエッジは目立つ）

3 アルゴリズム

3.1 最適曲線の定義

画像は、 $N \times M$ の正方形のピクセルから構成されているとする。そして、そのピクセルの座標 p を (i, j) ($1 \leq i \leq N, 1 \leq j \leq M$) で表現することにする。そして、そのピクセルの輝度値を $I(p)$ で表現する。曲線は n 個の微小有向線分からなる折れ線に近似する。微小有向線分は、いずれかのピクセルの辺に向きをつけたものとなっており、その線分を辺として持つ二つのピクセルの座標の組 (a, b) で表現される。 a をその向きに向かって左側のピクセルの座標、 b を向かって右側のピクセルの座標とする。二つのピクセルは隣り合っていないから、 $a = (a_i, a_j), b = (b_i, b_j)$ とすると $|a_i - b_i| + |a_j - b_j| = 1$ が成り立つ。

曲線が連続であるためには、 k 番目の微小有向線分を q_k とすると、 $2 \leq k \leq n$ において、 q_k の始点と q_{k-1} の終点は一致しなくてはならない。

The Presentation of a Method for Recognition of Human Facial Expressions from an Image of a Face
Hiroyasu Izawa[†], Hitoshi Saji[†], Yoshihisa Shinagawa[†], and Toshiyasu L. Kunii[‡]
The University of Tokyo[†], The University of Aizu[‡]

さらに $q_k = (a_k, b_k)$ と表現することによって、 k 番目の微小有向線分の評価値 $E_k(q_k)$ は次の式で算出される。

$$E_k(q_k) = |I(a_k) - I(b_k)|$$

曲線全体の評価値 $E(q_1, q_2, \dots, q_n)$ は、微小有向線分の評価値の総和である。

$$E(q_1, q_2, \dots, q_n) = \sum_{k=1}^n E_k(q_k)$$

この最大値を与える曲線を最適曲線と定義する。この最適曲線は、ダイナミックプログラミングにより求められる。[1, 2, 3]

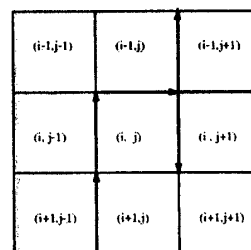


図 1: 最適曲線における微小有向線分

3.2 最適曲線による領域分割閾値の算出

s を、空間的な広がり小さいエッジを除去するためのパラメータ、 c を、コントラストが小さいエッジを除去するためのパラメータとする。

原画像 P_{org} は $N \times M$ に分割され、正方形ピクセルによって構成されていることにする。ピクセルの存在する領域を $D_{org} = \{(i, j) \mid 0 \leq i < N, 0 \leq j < M\}$ と表現し、 $(i, j) \in D_{org}$ で指定されるピクセルの輝度値を $I_{org}(i, j)$ で表現する。この画像を $s \times s$ ピクセルからなる正方形ブロックに分割し、その平均値を用いて、低解像度画像 P_{block} を構成する。すなわち、そのピクセル存在領域 $D_{block} = \{(i, j) \mid 0 \leq i < \lfloor \frac{N}{s} \rfloor, 0 \leq j < \lfloor \frac{M}{s} \rfloor\}$ であり、ピクセル $(i, j) \in D_{block}$ の輝度値 $I_{block}(i, j)$ は、

$$I_{block}(i, j) = \frac{1}{s^2} \sum_{k=0}^{s-1} \sum_{l=0}^{s-1} I_{org}(si + k, sj + l)$$

で表現される。 P_{block} においては、 s に対して空間的な広がり小さいエッジが平滑化により、除去されると考えられる。

D_{org} の任意の部分領域 D_1 に対応する D_{block} の部分領域 $COR_{ob}(D_1)$ を次のように定義する。

$$COR_{ob}(D_1) = \bigcup_{k=0}^{s-1} \bigcup_{l=0}^{s-1} ((i, j) \mid (si+k, sj+l) \in D_1)$$

一方、 D_{block} の任意の部分領域 D_2 に対応する D_{org} の部分領域 $COR_{bo}(D_2)$ を次のように定義する。

$$COR_{bo}(D_2) = \bigcup_{k=0}^{s-1} \bigcup_{l=0}^{s-1} ((si+k, sj+l) \mid (i, j) \in D_2)$$

1. D_{block} において最適曲線を求める。求まらなければ手続き終了。
2. 最適曲線を含む領域を $D_{block,oc}(\subset D_{block})$ として、 $D_{org,oc} = COR_{bo}(D_{block,oc})$ とすると、 $D_{org,oc}$ にエッジが存在すると予想し、 $D_{org,oc}$ における平均輝度値 $threshold_1$ を求める。
3. $D_{org,oc}$ を $threshold_1$ よりも高輝度の領域と低輝度の領域に分割し、それぞれの連結成分の面積をピクセル単位で計算する。 $s \times s$ ピクセルよりも小さい面積の連結成分しかないのならば、 $D_{org,oc}$ にはエッジが存在しないと見なし、 D_{block} から、 $D_{block,oc}$ を除いて1に戻る。
4. 3において得られた連結成分のうち、 C_{max} が最大面積を持つとする。 $D_{block,max} = COR_{ob}(C_{max})$, $D_{org,max} = COR_{bo}(D_{block,max})$ とし、 $D_{org,max}$ を $threshold_1$ よりも高輝度の領域と低輝度の領域に分割し、それぞれの領域の平均輝度値 $average_h, average_l$ を計算する。
5. $average_h - average_l < c$ である場合はコントラストが小さいエッジとみなして、その領域からはエッジを検出せず、 D_{block} から、 $D_{block,max}$ を除いて1に戻る。
6. $threshold_2 = \frac{(average_h + average_l)}{2}$ とし、周辺領域を $threshold_2$ よりも高輝度の領域と低輝度の領域に分割する。その中で最適曲線の軌跡を含む領域の輪郭をエッジとする。 D_{block} から、エッジを検出した領域を除いて1に戻る。

4 検出結果

163×87ピクセル、256階調モノクロの口の画像(図2)を検出対象として用いた。図3,4,5にそれぞれ $(s, c) = (8, 0), (4, 0), (4, 20)$ でエッジ抽出した結果を示す。

5 結論

本稿において、特定の顔特徴の知識を用いないエッジの検出手法について述べた。この手法を、低解像度から高解像度へ適用し、顔特徴の知識と結び付けて、表情の自動認識に役立てることが、今後の我々の研究における課題である。

参考文献

- [1] Ramesh C. Jain Amir A. Amini, Terry E. Weymouth. Using dynamic programming for solving variational problems in vision. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 12(9):855-867, sep 1990.
- [2] Alberto Martelli. An application of heuristic search methods to edge and contour detection. *Communications of the ACM*, 19(2):73-83, feb 1976.
- [3] Ugo Montanari. On the optimal detection of curves in noisy pictures. *Communications of the ACM*, 14(5):335-345, may 1971.

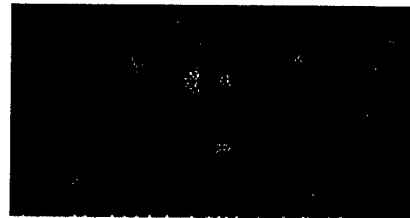


図 2: 原画像



図 3: 8×8ブロック、コントラスト制限なし



図 4: 4×4ブロック、コントラスト制限なし



図 5: 4×4ブロック、コントラスト制限つき