

分離度特徴を用いた顔画像解析*

目、瞳の検出

山口 修 福井 和広

(株) 東芝 関西研究所

1 はじめに

個人認証、視線検出、まばたき検出などの顔画像解析は、ヒューマンインタフェース構築のための重要な技術である。本研究では、これらの要素技術の開発を目的として、画像中から顔領域を抽出し、自動的に目、瞳の検出を行う方法について報告する。

目、瞳の検出にはエッジの不明瞭さ、エッジ強度の弱さなどの問題がある。これらの問題に対応するため、分離度に基づいたエッジ抽出 [1] を行い、形状パターンの抽出を行う「組合せ分離度法」を提案する。

2 顔領域の抽出 [2]

顔画像解析では、最初に画像中から顔を探す必要がある [5][6]。本研究では、顔の濃淡パターンを辞書として登録し、複合類似度 [3] を用いてパターンマッチングを行い、全画面を探索する。

さらに詳細に顔領域のみを抽出するためには、分離度 Snakes[2] を用いる。図 1 はその抽出結果である。



図 1: 顔領域の抽出例

3 目、瞳の検出

顔領域を定めた後、顔部品の検出を行う。この際、目領域を中心として顔領域のズームを行い、より解像度の高い画像を得ることが必要となる。本報告では、手動でズームアップして取得した画像を実験に用いた。

3.1 目領域の検出

本報告では、目を検出対象とし、まず目領域の検出を行う。眼鏡の有無、環境の変化などに対応するため、検出は次のステップで行う。

1. 色相画像の 2 値化による候補生成

目領域は「肌色とは異なる色相を持つ」ことから、

*Face Image Analysis based on Separability Feature Extraction. Eye and Pupil Detection.
Osamu Yamaguchi and Kazuhiro Fukui
TOSHIBA Kansai Research Laboratories

色相画像 (図 2 右上) を 2 値化 (図 2 左下) して領域の候補を生成する。

2. 領域の形状制約を用いた候補の制限

候補領域の形状特徴量としてモーメント不変量 [4] を求め、目領域の形状と異なるものは候補から除く (図 2 右下、矩形内のみ選択)。

3. パターン類似度による候補の検証

あらかじめ目領域の画像から辞書を作成し、複合類似度を用いて目候補を選択する (図 3 の 2 つの矩形領域が選択結果)。



図 2: 目領域の検出

3.2 瞳の検出

次に、瞳の検出を行う方法について述べる。瞳が楕円形状として撮影されていると仮定し、高精度な楕円検出を行うことを目指す。瞳を検出する際の問題点として、(1) まぶたによる瞳輪郭の一部隠蔽、(2) 蛍光灯などの映り込みによる強いエッジ点の存在、(3) 瞳輪郭のなだらかな明度変化などが挙げられる。

ここでは、エッジの方向性を考慮して形状パターンを構成し、エッジ強度の類似度を評価関数に組み込んだ「組合せ分離度法」によって検出を行う。

[1] で提案した分離度とは、エッジを輝度、色、テクスチャの異なる領域間の境界としてとらえ、領域間の特徴量からエッジ強度を計算する手法である。

分離度 $S(0.0 \leq S \leq 1.0)$ は次の式で与えられる。

$$S = \frac{\sigma_b^2}{\sigma_T^2}$$

$$\sigma_b^2 = n_1(\bar{P}_1 - \bar{P}_m)^2 + n_2(\bar{P}_2 - \bar{P}_m)^2$$

$$\sigma_T^2 = \sum_{i=0}^{N-1} (P_i - \bar{P}_m)^2$$

ここで、 N は矩形領域内の全画素数、 n_1 は領域1に含まれる画素数、 n_2 は領域2に含まれる画素数、 P_k は位置 k における画像の特徴量、 \bar{P}_m は領域全体での画像の特徴量の平均値、 \bar{P}_1 は領域1での画像の特徴量の平均値、 \bar{P}_2 は領域2での画像の特徴量の平均値を示している。図3に示すように、マスクの形状、大きさを設定し、そのマスク内で上述した式を計算し、分離度を求める。

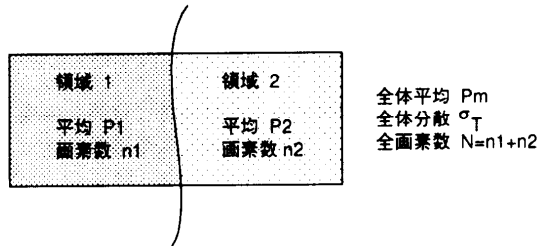


図 3: 分離度

分離度によるエッジ抽出は、従来のエッジ抽出と異なり、エッジ強度が 0.0 ~ 1.0 に正規化されており、これらを複数個組み合わせる利用することが可能である。そこで、この分離度マスクを図4のように、抽出したい形状(ここでは楕円)のエッジに沿って、垂直にマスクを複数個配置する。

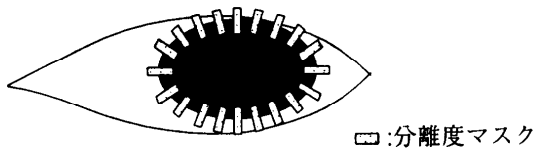


図 4: 組合せ分離度マスク

それぞれのマスクにおいて計算された分離度に基づいて、組合せ分離度 C は次式で計算される。

$$C = \frac{\alpha}{\alpha + \beta} \frac{G}{n} + \frac{\beta}{\alpha + \beta} \frac{E}{n} = \frac{\alpha G + \beta E}{n(\alpha + \beta)}$$

(n はマスクの数、 α, β は重み定数)

ここで、 G は隣接するマスク間の分離度の類似度を評価する関数で、各マスク $M_i (i = 1 \sim n)$ の分離度を $S(M_i)$ で表すとき、隣接した各マスク間の差の絶対値を、1.0 から引いた値を類似度とし、その総和として次のように定義する。

$$G = \{1 - |S(M_n) - S(M_1)|\} + \sum_{i=1}^{n-1} \{1 - |S(M_{i+1}) - S(M_i)|\}$$

類似度による評価項が入っているために、抽出対象が全体的に弱いエッジで構成されている場合でも抽出が可能になる。

E は各部分マスクで計算された分離度の総和をとり、全体のエッジ強度を表す。

$$E = \sum_{i=1}^n S(M_i)$$

抽出したい形状(楕円)に対し、平行移動、回転、拡大縮小といったアフィン変換を施し、変換された輪郭形状に対して組合せ分離度を求める。この組合せ分離度を最大にするアフィン変換パラメータを画像中から求めることにより、瞳を抽出する。図5が抽出結果である。



図 5: 瞳の抽出結果

4 実験結果

一般環境に設置されているワークステーションの前に座った日本人男女 14 名、84 枚のデータを用いて実験を行った。撮影場所は同一であるが、通常の利用状況でのデータ取得を行うことを目的として、眼鏡の有無、髪型などに制限は設けなかった。また、固定焦点で撮影したため、カメラからの距離は撮影された人により変化しており、焦点のあっていない画像も含まれている。なお、顔領域、目領域の辞書パターンを作成には、実験用の画像とは別の画像を用いた。

目の領域抽出は 75/84(89%) の画像に対し抽出が成功した。また、目領域として抽出された部分画像に対して、瞳の抽出を行った結果、78% の画像に対して抽出できた。比較実験として、瞳の輪郭のエッジ点を抽出し、最小二乗法により楕円当てはめを行った。輪郭の欠けが多い場合に当てはめを誤るため 61% の検出率であり、本手法の検出能力の高さが認められた。本手法における失敗の原因としては、眉の誤認や、まぶた部分のエッジとの誤認などがある。

5 おわりに

本報告では、分離度特徴に基づいた顔画像処理として、顔輪郭検出、目、瞳検出について報告した。今後は、より多くの画像に対しての実験を行うとともに、上記問題点に対し、アルゴリズムの改良を行っていく。

参考文献

- [1] K.Fukui: "Edge Extraction Method Based on Separability of Image Features", MVA'94, pp391-396(1994)
- [2] 福井: "領域間の分離度を用いた輪郭抽出", 情処研報 CV88-2, pp.9-16(1994)
- [3] 飯島泰蔵: "パターン認識理論", 森北出版 (1989)
- [4] M.K.Hu.: "Visual Pattern Recognition by Moment Invariants", IRE Trans. Vol.IT-8, pp179-187 (1962)
- [5] M.A.Turk, A.P. Pentland: "Face Recognition Using Eigenfaces", Proc. IEEE Conf. CVPR, pp586-591 (1991)
- [6] K.Sung and T.Poggio: "Example-based Learning for View-based Human Face Detection", Technical Report A.I.Memo No.1521, A.I. Lab., MIT (1994)