

色情報と形状情報を用いた映像からのロバストな顔抽出

5 E-1

孔 秋雁
東京大学佐藤 真一
学術情報センター浜田 喬
学術情報センター

1 はじめに

顔の認識の前処理として、画像に存在する人物の顔を自動的に検出し、正確に位置決めすることが必要である。顔の向きによらず、洩れなく顔を検出するため、色情報と形状情報を統一的に利用する手法を提案している。

2 色情報の利用

本研究の最終的目的地は、映像から実時間の顔追跡、認識であるため、計算量の比較的小さい色情報の分析から着手することにした。また、色情報を利用することにより、横向きの顔も洩れなく検出できるという利点もある。

2.1 肌色の学習

肌色画素の推定をする時、

$$r = \frac{R}{R+G+B}$$

$$g = \frac{G}{R+G+B}$$

というモデルを採用している。理由は以下の二つの点にある。(1)低次元;(2)照度の影響をなくすることができる。

実際、CNN ニュースの任意の 40 枚の画像から 17526 個の肌色画素を元に、平均 (\bar{r}, \bar{g}) と共分散行列 Σ をシステムに学習させた。 \bar{r}, \bar{g} と Σ の定義を以下に示す。 r_i, g_i は学習する肌色画素 i の r, g 値、 N は肌色画素数。

$$\bar{r} = \frac{1}{N} \sum_{i=1}^N r_i \quad (1)$$

$$\bar{g} = \frac{1}{N} \sum_{i=1}^N g_i \quad (2)$$

$$\Sigma = \frac{1}{N} \sum \begin{bmatrix} r_i - \bar{r} \\ g_i - \bar{g} \end{bmatrix} \begin{bmatrix} r_i - \bar{r} \\ g_i - \bar{g} \end{bmatrix}^t \quad (3)$$

学習するデータ(図1)を、 (r, g) の確率密度分布にプロットした結果が図2に示してある。

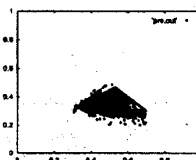


図 1:

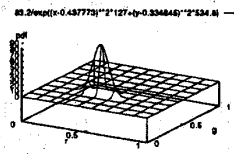


図 2:

2.2 肌色画素の推定

ある画素が肌色かどうかを判断する時に、Mahalanobis 距離式(4)を利用している。Mahalanobis 距離が小さければ小さいほど、その点が平均値に近い。

$$Dis(r, g) = \begin{bmatrix} r_i - \bar{r} \\ g_i - \bar{g} \end{bmatrix}^t \Sigma^{-1} \begin{bmatrix} r_i - \bar{r} \\ g_i - \bar{g} \end{bmatrix} \quad (4)$$

ここで、Mahalanobis 距離が 2 より小さい画素は肌色画素と認識する。

2.3 実験結果

SGI デジタルカメラで撮影した画像 13 枚と CNN ニュースから任意に選んだ画像 47 枚を対象に、実験した結果を表 1 に示す。



図 3: 原画像



図 4: 肌色画像

表 1: 肌色画素抽出実験結果

正検出	60
検出洩れ(光線が強(弱)過ぎる)	5
検出洩れ(肌色が特殊)	3
誤検出(背景が大範囲に取られた)	2
正解率	85.71%

3 形状情報の利用

肌色画像を元に、肌色画素のかたまりを探索できれば、顔の位置が決められる。ここでは、顔の形状は楕円とな

ると仮定し、 (x, y) 平面でガウス分布を利用した楕円の抽出手法を提案している。

楕円は、長軸・短軸が x, y 軸と平行となることを仮定すると、中心 (x_0, y_0) 、長・短径 (a, b) で表される。このような楕円中にある点群は、平均点が (x_0, y_0) 、 x 軸と y 軸で分散がそれぞれ a^2, b^2 の二次元正規分布に従うことになる。計算を容易にするため、 (a, b) として、元の画像の大きさ (352×240) に対して、一定値 $(30, 50), (40, 60), (50, 70)$ の三組を想定している。以下に、与えられた肌色画像から、それぞれに対応する楕円の中心 (x_0, y_0) を算出するアルゴリズムについて述べる。

3.1 評価関数

評価関数を次のように定義している。

$$J(x_0, y_0) = \int \int_{xy} I(x, y) g(x_0 - x, y_0 - y) dx dy$$

ただし、

$$g(x, y) = \frac{1}{2\pi\sqrt{|C|}} e^{-\frac{1}{2} \begin{bmatrix} x \\ y \end{bmatrix}^T C^{-1} \begin{bmatrix} x \\ y \end{bmatrix}}$$

$$I(x, y) = \begin{cases} 1 & \text{ただし、}(x, y) \text{ に肌色素} \\ 0 & \text{それ以外} \end{cases}$$

$g(x, y)$ は (x, y) 平面での正規分布関数。C は共分散行列。長・短径 (a, b) の楕円を探索する時に、C は次のようになる。

$$C = \begin{bmatrix} a^2 & 0 \\ 0 & b^2 \end{bmatrix} \quad (5)$$

(x_0, y_0) が画像の左上から右下へと順に走査していく時に、評価関数 J が一番大きい値となる (x_0, y_0) が顔の中心であると判断する。

一つ以上の顔がある場合は、評価関数の最大値を取るかわりに、極大値を求めることにより、洩れなくすべての顔を探索できることも実験により検証されている。ある点の J が極大値かどうかを判断するには、その点のまわりにある他の八つの点の評価関数 J と比べれば良い。この点の J より大きい値がなければ、該当点が極大値であると判断する。

図4に対応する各点 $J(x, y)$ の値を、図6に示す。



図 5: 原画像 図 6: 顔抽出結果

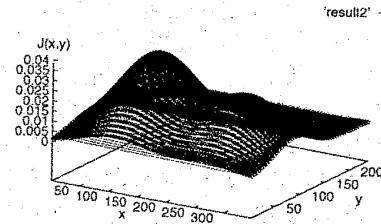


図 7: $J(x, y)$

3.2 実験結果

55 枚の画像で実験した結果を表 2 に示す。

楕円の大きさを固定したことによる誤検出や検出誤りが多く見られた。異なる大きさの顔の検出を可能とすれば、もっと良い結果が出ると思われる。

表 2: 50x70 顔抽出結果

正検出	40
誤検出 (大きさによる)	1
誤検出 (肌色素の誤り)	4
誤検出 (実際顔がない)	2
検出誤差あり (大きさによる)	7
検出誤差あり (肌色素の誤り)	1
正解率	72.73%

4 おわりに

本稿では、ロバストな顔抽出手法について検討した。今後は、異なる大きさの顔抽出及び映像情報への適応などについて、検討していく予定である。

参考文献

- [1] Nuria Oliver, Alex P. Pentland and Francois Berard, "LAFTER: Lips and face real time tracker", IEEE computer society conference on computer vision and pattern recognition June 17-19, 1997 Puerto Rico, pp.123-128