

相補的分類器の組み合わせによる顔検出システムの高度化 Improvement of Face Detection System by Combining A Complementary Classifier

徳永 浩成† 黄 琳琳† 清水 昭伸† 小畑 秀文†
Hironari Tokunaga Lin-Lin Huang Akinobu Shimizu Hidefumi Kobatake

1. はじめに

近年、個人認証システムとして、顔画像を使ったシステムへの注目が高まっている。このようなシステムは、ATMやコンピュータへのアクセス認証、立ち入りが制限されたエリアへの立ち入り許可などへの幅広い応用が考えられている[1]。ここで、画像から顔を自動認識するためには、顔の自動検出が必須であるが、そこでは画像からいかに顔を安定かつ正確に得ることができるかが重要となる。

これまでに著者らが開発した顔検出システム[2]は、従来提案された代表的なシステムである多層パーセプトロン(MLP)を用いたシステム[3]と同程度の性能でありながら、より実装が簡単かつ高速であるという特徴があった。

本論文では、更なる高度な顔検出システムの構築を目的とし、新しい特徴量を用いたもう一つの判別処理を組み合わせ、システムの性能を高度化する。具体的には、従来は、PCA (Principal Component Analysis) を適用して抽出した主成分特徴を用いて判別を行っていたが、これに DPC (Difference Principal Component, 差分主成分) [4]特徴に基づく判別処理を並列に組み合わせた。

2. 従来の顔検出システム[2]

従来の顔検出システム[2]の処理の流れを Fig.1 に示す。

まず、検出対象となる原画像のスケールを 10 段階に変えながら、それぞれをシステムに入力し、入力された画像内を 20×20 pixel のウィンドウを走査させる。次に、ウィンドウ内の画像から特徴量を抽出する。具体的には、ウィンドウの 4 隅を除いた画像を作成し、それに対して sobel フィルタを用いて求めた水平・垂直の濃度勾配を、顔の特徴を考慮して 3 方向に分解する[2]。今回は、上記の 4 隅を除いた画像 (gray) と、3 方向の濃度勾配画像 (grad) とをそれぞれ半分圧縮した画像の濃度値を特徴量とした。それぞれの特徴の例を Fig.2 に示す。次に、各特徴を軸とした特徴空間に対し PCA を行い、主成分を特徴として抽出する。さらに、抽出された特徴を 2 次の PNN (Polynomial Neural Network) に入力することで、ウィンドウ内の画像が顔であるかどうかを判別する。なお、顔と判定されたウィンドウが原画像上でオーバーラップした場合は、PNN の出力が最も高いものを顔とする。また、PNN の学習は、顔 29900 個、非顔 90994 個 (Bootstrap 法[3]によって収集) の学習データに対して、最急降下法を用いて行った。

3. DPC を用いた顔検出システムの高度化

まず、DPC[4]について説明する。PCA で抽出された従来の特徴空間において、顔と非顔が Fig.3 に示すように分布し、PNN の決定境界が図中の実線で示される位置にあるとすると、誤判別である FP (False Positive, 拾いすぎ) と FN (False Negative, 見落とし) は決定境界付近に分布している。ここで、FP クラスと顔クラス、FN クラスと非顔ク

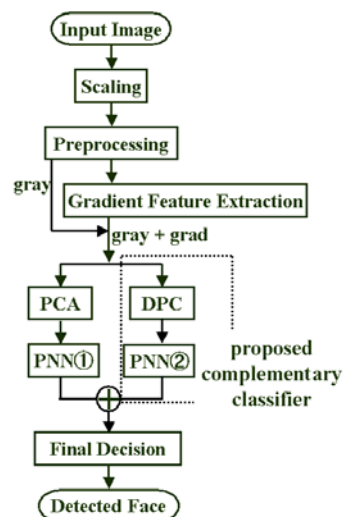


Fig.1 顔検出システム[2]と今回追加した相補的分類器 (点線枠内)

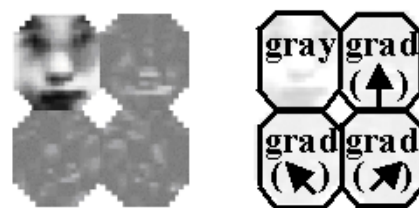


Fig.2 濃度値特徴量 (gray+grad)
(右図中括弧内の矢印は濃度勾配を求める方向)

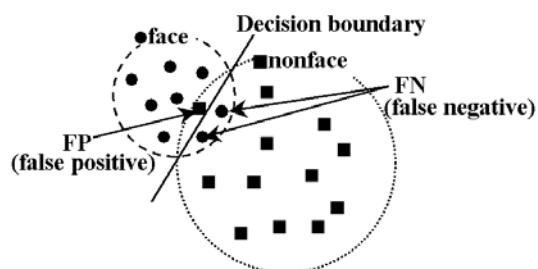


Fig.3 特徴空間における顔と非顔の分布例

ラスの距離ができるだけ離れるような新しい特徴空間を求めることにより、FP と FN は判別し易くなると考えられる。DPC は以上のような考えに基づき、FP クラスと顔クラスとの分離度と、FN クラスと非顔クラスとの分離度の和を最大にするような軸を求めることで、従来の特徴空間では誤判別し易かったデータを判別し易い空間に射影する方法である。また今回は、その新たな特徴空間で PNN を学習させ、従来の PNN と組み合わせることで能力の向上を計る。提案する顔検出システムの流れを Fig.1 に示す。Fig.1 の点線枠で示すように、提案するシステムでは、従来の分類器に対して DPC を特徴量とする分類器を並列に付け加え

†東京農工大学大学院 生物システム応用科学研究科

ている。2つの分類器の出力は(1)式によって統合され、その値に対する閾値処理によって最終的な判定が行われる。

$$\text{output} = \frac{1}{1+\alpha} (\text{PNN}\textcircled{1} + \alpha \text{PNN}\textcircled{2}) \quad (1)$$

4. 実験

新しい顔検出システムの性能を評価するために、以下の実験を行った。DPCの計算のための2章で示した学習サンプルからFPサンプルとFNサンプルを収集した。ここで、FPやFNは、システムの出力量に対する閾値ごとに定義されるものであるが、実際にどの閾値を用いるかは決まっていなかったために、今回は便宜的に顔なのに出力が低いものをFNサンプル、非顔なのに出力が高いものをFPサンプルとする。具体的には、Table 1に示すようなFP、FNサンプルを学習セットから採取した。Sample①は、従来のシステム[2]が苦手な順（FPなら出力の高い順、FNなら出力の低い順）に1000個採取したものであり、括弧内はPNNの出力（0~1）に対する閾値を示す。また、Sample②は、やや高い値を閾値として収集したものであり、FPなら括弧内に示した閾値以上、FNなら閾値以下のものを採取した。

以上の手順で求めたFP、FNサンプルを用いて、まずはFPクラスと顔クラス間の分離度を最大にするDPCをSample①と②の場合で2通り、FNクラスと非顔クラス間の分離度を最大にするDPCをSample①と②の場合で3通り求めた。ここで、(1)式 α の値は0.25~5.0まで9通りに変化させた。その他のパラメータは2章に示した従来のシステム[2]と同様である。

次に、求められたDPCを用いて提案したシステムによって顔検出を行った。テスト用の試料は、従来のシステム[2]やMITのシステム[3]でも用いられた、CMUのテストセット（151個の顔を含む23枚の画像）[6]である。また最後に、上記の5通りの中から誤りの少ない結果を示したサンプルを用いて、FPクラスと顔クラス間の分離度と、FNクラスと非顔クラス間の分離度の和を最大にするDPCを求めた。

上記の実験では、6通りのDPCと、9通りの α により、合計54通りの場合について検討を行ったわけだが、それぞれの結果をFROC曲線を用いて評価した。その中で、MITのシステムや従来のシステムと比べた場合に、よい結果を示したFROC曲線をFig.4に示す。横軸はテストセットに対するFP数、縦軸は顔検出率であり、分類器の出力に対する閾値を様々に変化させることにより、曲線を描いた。Fig.4のNF-FN1000 $\alpha=1.5$ の意味は、DPCの計算においてFNサンプル1000個と非顔の学習データとの分離度を用い、式(1)の α を1.5にしたということである。

ここで、このFROC曲線は、FNサンプル1000個と非顔クラス間の分離度を最大にするDPCを用いた場合であったが、従来と比べてFROC曲線は全体的に左上にあり、性能が向上したことが分かる。また、具体的にはMITのシステム[3]とFP数（13個）に注目すると、検出率は3.4%上回る。また、このFROC曲線を、従来のシステム[2]と比較すると、検出率91.3%の時（Fig.4破線）、従来のシステムのFP数が57個であるのに対して、提案するシステムのFP数は21個であり、従来と比べてFP数を63.2%削減することに成功した。この時の検出例Fig.5に示す。

Table 1 FP, FN サンプル数
(括弧内は従来システム[2]の出力に対する閾値)

	FP	FN
Sample①	1000(≥ 0.001)	1000(≤ 0.00001)
Sample②	428 (≥ 0.01)	3018 (≤ 0.01) 4302 (≤ 0.1)

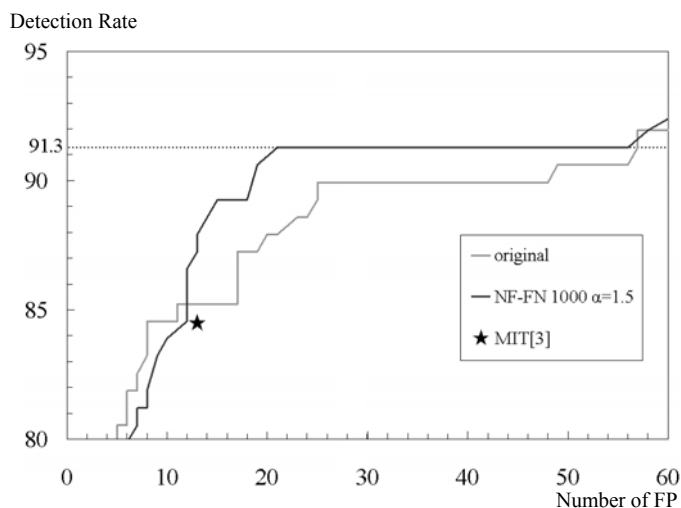


Fig.4 提案するシステムと他のシステムのFROC曲線



Fig.5 検出例（左：提案システム 右：従来システム[2]）

5. まとめ

差分主成分という従来の主成分空間に対して相補的な特徴空間を求め、それに対して学習したPNNを並列に組み合わせることによって、従来のシステムと比べ、FPを約1/3に抑えることが出来た。今後は、従来の主成分空間の次元数を縮小し、検出能力を落とすことなく高速化を図る。また、特徴量を追加し、判別器を並列に増やすことで検出能力の向上を図る。

参考文献

- [1] A. V. Nefian et al., "Face detection and recognition using Hidden Markov Models", IICIP, vol.1, pp. 141-145 1998.
- [2] L.-L. Huang et al., "Gradient feature extraction for classification-based face detection", Pattern Recognition, in printing.
- [3] K.K. Sung, T. Poggio, "Example-based learning for view-based human face detection", IEEE Trans. PAMI, Vol.20, No.1, pp.39-50, 1998.
- [4] 川谷隆彦, "相補的認識系に差分主成分を用いる手書き漢字認識", 信学技報, PRMU98-39 1998.
- [5] 徳永浩成 他, "新しい特徴を用いた顔検出システムの高度化", 電子情報通信学会総合大会講演論文集, pp.215 2003.
- [6] http://vasc.ri.cmu.edu/idb/html/face/frontal_images/index.html