

一般化学習ベクトル量子化に基づく顔検出 Face detection based on Generalized LVQ

細井 利憲[†]
Toshinori Hosoi

鈴木 哲明[†]
Tetsuaki Suzuki

佐藤 敦[†]
Atsushi Sato

1. まえがき

近年、バイオメトリクス認証や監視において、顔照合技術への期待が高まってきている。画像からの顔検出は、顔照合の前処理として重要であり、そのアプローチは、顔に関する先験知識を利用する Feature-based 型と、パターン認識技術に基づく Image-based 型に分けられる [1]。一般に、Feature-based 型は処理が高速で位置精度が高いが、検出性能は背景や撮影環境に強く依存する。Image-based 型は複雑な背景からの顔検出性能が高い反面、検出位置精度が低い問題がある。本稿では、顔・非顔の判別器に一般化学習ベクトル量子化 [2] を用いた Image-based 型と、目検出を行う Feature-based 型を組み合わせた顔検出手法を提案する。

2. 一般化学習ベクトル量子化 (GLVQ)

一般化学習ベクトル量子化 (GLVQ) とは、最小分類誤り (MCE) に基づく参照ベクトルの学習方法であり、確率的降下法によって経験損失を最小化する。MCE に基づく学習方法としては、Generalized Probabilistic Descent (GPD) が知られるが、参照ベクトルが発散する問題があった。GLVQ は、新たな誤分類尺度の導入によってこれを解決し、さらに、決定論的アニーリングによって、連続関数で定義される損失を漸近的に経験損失に一致させる点に特長がある。最近傍分類法に対する GLVQ の学習式を以下に示す。

$$m_r(t+1) = m_r(t) + \alpha w(\rho_k) \frac{d_s(\mathbf{x})}{d_r(\mathbf{x}) + d_s(\mathbf{x})} (\mathbf{x} - m_r) \quad (1)$$

$$m_s(t+1) = m_s(t) - \alpha w(\rho_k) \frac{d_r(\mathbf{x})}{d_r(\mathbf{x}) + d_s(\mathbf{x})} (\mathbf{x} - m_s) \quad (2)$$

ここで、 α は微小な正の実数、 m_r は正解クラスに属する参照ベクトルで入力 \mathbf{x} に最も近いもの、 m_s は競合クラスに属する参照ベクトルで \mathbf{x} に最も近いものであり、 $d_r = \|\mathbf{x} - m_r\|^2$ 、 $d_s = \|\mathbf{x} - m_s\|^2$ 、 $w(\rho_k) = 4\ell(\rho_k)\{1 - \ell(\rho_k)\}$ 、 $\ell(\rho_k) = 1/\{1 + e^{-\xi(t)\rho_k}\}$ である。ここで $\xi(t) > 0$ は時間 t とともに単調増加する。 ρ_k は誤分類尺度と呼ばれ

$$\rho_k = \frac{d_r(\mathbf{x}) - d_s(\mathbf{x})}{d_r(\mathbf{x}) + d_s(\mathbf{x})} \quad (3)$$

と定義される。 $-1 \leq \rho_k \leq 1$ であり、 \mathbf{x} が正しく分類されている場合には $\rho_k < 0$ となる。

3. 提案手法

3.1 顔検出処理フロー

提案手法の処理フローについて説明する (図 1)。処理時間を短縮するために、階層的な探索を行っている。粗探索では、まず粗い多重解像度画像を作成し、各解像度について画像を粗く回転させながら、GLVQ による顔信

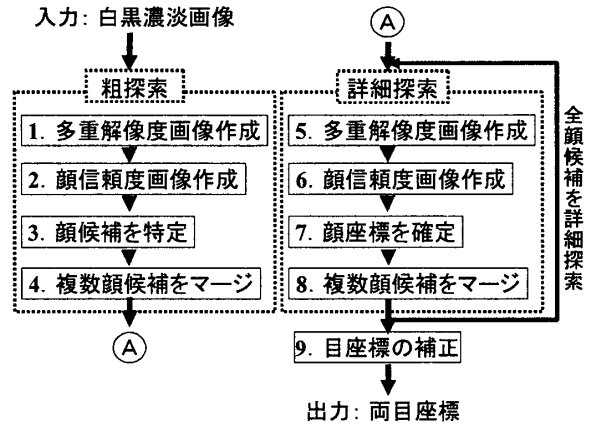


図 1: 検出処理フロー

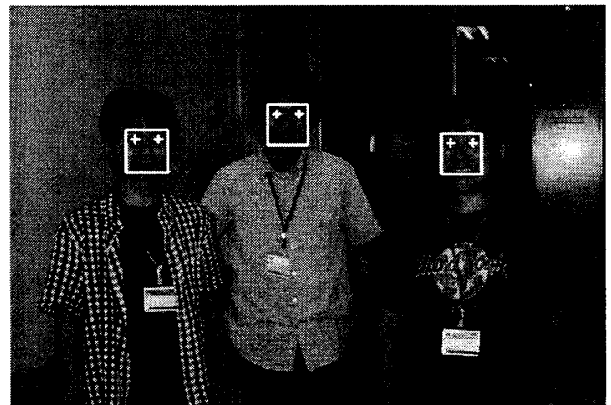


図 2: GLVQ による顔検出結果例

頼度画像を生成する (顔信頼度については後述)。信頼度画像をしきい値処理し、ラベリング後の各重心座標を顔候補位置とする。複数の多重解像度の結果をマージし、最終的な顔候補位置を求める。次に、得られた全顔候補位置に対し詳細探索を行う。処理内容は粗探索と同様であるが、解像度、画像の回転角の刻み幅を細かくとり、GLVQ による顔信頼度を求め直す。最後に、得られた顔候補位置から推定される両目位置について、リングフィルタを用いた目位置補正を行う [3]。

本処理フローに沿い、GLVQ による顔検出を実行した例を図 2 に示す。

3.2 顔・非顔判別

顔信頼度画像は、大きさ固定の部分画像を切り出し、GLVQ による顔・非顔判別を行うことによって作成する。参照ベクトルは、学習用の顔画像、非顔画像を用いて予め学習されている。入力された部分画像を特徴ベクトル

[†]日本電気 (株) マルチメディア研究所, Multimedia Research Labs., NEC Corp.



図 3: 顔・非顔判別実験に用いた画像の例

\mathbf{x} に変換し、参照ベクトルとの距離計算を行い、顔のクラスへの距離 $d_s(\mathbf{x})$ と非顔のクラスへの距離 $d_r(\mathbf{x})$ を求め、式 (3) で得られる値を信頼度とする。信頼度が正の時は顔と判定されるので、しきい値を 0 として信頼度画像をラベリング処理し、顔候補位置を求める。

特徴ベクトルへの変換は、以下のように行う。まず、ヒストグラム均等化によって、コントラストを補正し照明変動の影響を抑える。カラー画像の場合は、予め白黒濃淡画像に変換されているものとする。得られた濃淡画像から、文字認識などで有効とされる輝度勾配方向に基づく特徴を抽出する。具体的には、Sobel フィルタにより輝度勾配方向を D 方向に量子化し、各方向について微分強度を求める。画像を $N \times N$ に分割した領域それぞれについて微分強度の総和を求めることで、 DN^2 次元のベクトルを得る。今回の実験では、部分画像の大きさを 30×30 画素とし、 $D = 8$, $N = 8$ と設定した。即ち、512 次元の特徴ベクトルとなる。

4. 実験

4.1 顔・非顔判別実験

GLVQ による顔・非顔判別性能を評価するため、勾配特徴を用いた顔画像と非顔画像の判別実験を行った。実験に用いた画像の例を図 3 に示す。画像の大きさは 30×30 画素であり、顔画像と非顔画像を 7,855 枚ずつ計 15,710 枚用意し、2 分して学習と評価に用いた。顔画像は室内環境において撮影したほぼ正面の顔で、両目位置を基準に顔の大きさを揃えてある。非顔画像は、室内・風景・動植物等の画像からランダムに切り出した。GLVQ の参照ベクトルは、顔・非顔のクラスそれぞれ 20 個とし、初期値はクラスごとに K 平均法で求めた。比較のために、サポートベクトルマシン (SVM) による同様の実験も行った。

評価画像に対する判別結果を表 1 に示す。M は参照ベクトル数 (SVM の場合はサポートベクトル数) を表す。SVM は SVM^{light}[4] による結果で、濃淡画像をそのまま用いた場合の結果を併せて示した。カーネル関数には Radial Basis Function: $\exp(-\gamma\|\mathbf{x} - \mathbf{x}_i\|^2)$ を使い、 γ は 10^{-7} から 10^7 の範囲で評価し、最も誤りの少ない $\gamma = 10^{-5}$ を選んだ。本実験により、GLVQ の顔・非顔判別性能が最も高く、K 平均法や濃淡特徴を用いた SVM よりも優れていることがわかった。勾配特徴を使えば SVM の性能が改善されることが確認できたが、GLVQ には及ばなかった。また、必要な参照ベクトル数を比べると GLVQ は SVM の 1/20 以下と少なく、処理速度の点からも有利であることが示された。

4.2 顔検出実験

室内で撮影された静止画 9,720 枚を対象に、提案する顔検出手法の性能評価実験を行った。画像は逆光や低露光を含む様々な照明条件で撮像されたもので、ほぼ正面

表 1: 顔・非顔判別結果

手法	総数	誤り数	誤り率%	M
GLVQ	7854	21	0.27	40
K 平均法	7854	127	1.62	40
SVM(濃淡特徴)	7854	105	1.34	2466
SVM(勾配特徴)	7854	36	0.46	803

表 2: 顔検出実験結果

判別法	顔総数	正検出数	誤検出数	検出率%
GLVQ	9720	9493	227	97.7
K 平均法	9720	9092	628	93.5

を向いた顔が画像中に必ず 1 個存在する。顔・非顔判定部に用いた GLVQ は、前述の学習画像をさらに拡充した顔画像、非顔画像それぞれ 5,150 枚の画像を用いて学習した。検出結果の正誤判定は、正解両目位置と検出両目位置を基に行い、次の 3 つの条件を満たす場合に正解と判定した。1) 正解両目間距離 L に対する検出両目間距離の比が 80% から 120% の範囲、2) L に対する、両目間中点位置のずれの比が 20% 以下、3) 正解両目位置を結ぶ直線と出力両目位置を結ぶ直線の傾きが 10 度以下。検出結果を表 2 に示す。顔・非顔判別処理を K 平均法に置き換えた場合と比較して、GLVQ は検出率が高いことが確認された。誤検出は顔判別と目位置補正で発生しており、前者では照明が暗くコントラストが極端に低い場合や極端な逆光の場合に信頼度が低下する、後者では眼鏡フレームや眉毛を目の位置に誤ることが主な原因であった。

5. まとめ

GLVQ による顔・非顔判定を用いた顔検出手法を提案した。階層的探索で処理時間を短縮し、目補正を適用して位置精度を高めた。実験を通して、GLVQ が顔・非顔判別において高い性能を実現することを確認するとともに、画像からの顔検出実験によって検出性能を評価し、問題点を調査した。今後の展望としては、照明条件の悪い場合の顔・非顔判定の改善、眼鏡フレームや眉毛に対する誤りの改善、正面以外の顔への拡張があげられる。

参考文献

- [1] E. Hjelm and B. K. Low, "Face Detection: A Survey," *Computer Vision and Image Understanding*, Vol. 83 No. 3, pp. 236-274, 2001.
- [2] A. Sato and K. Yamada, "Generalized Learning Vector Quantization," In *Advances in Neural Information Processing Systems*, Vol. 8, pp. 423-429, MIT Press, Cambridge, MA, 1996.
- [3] 櫻井, 井上, 佐藤 "顔検出のための濃淡特徴を用いた高速な両目ペアの検出", 第 8 回画像センシングシンポジウム講演論文集, pp. 557-562, 2002.
- [4] T. Joachims, "Making large-Scale SVM Learning Practical", In *Advances in Kernel Methods Support Vector Learning*, B. Shölkopf and C. Burges and A. Smola (ed.), MIT-Press, 1999.