

テクニカルノート

顔らしさの評価値分布を利用した顔検出の提案

高塚 皓正[†] 田中正行[†] 奥富正敏[†]

顔検出はコンピュータビジョンにおいて、近年注目されている分野の1つであり、様々な研究が行われている。従来の枠組みでは、各サブウィンドウで独立に顔または非顔を識別するため、識別器の評価値の高い非顔画像を誤検出してしまうことがよくある。本研究では、顔と非顔における識別器の評価値分布の違いを利用した分布ベースの顔検出の枠組みについて提案する。実験ではMIT+CMUのテストデータセットおよびWWWより収集した様々な種類の実画像170枚を用いて、提案手法の有効性を確認した。その結果、従来の顔検出と比較し、同一の誤検出数のときに最大5%程度の検出率の向上を確認した。

Face Detection Based on the Distribution of Classifier Outputs

HIROMASA TAKATSUKA,[†] MASAYUKI TANAKA[†]
and MASATOSHI OKUTOMI[†]

Face detection is a useful technique in computer vision. Many face detection methods have been developed in the literature. In previous face detection process, since the classifiers also evaluate the scanned sub-windows independently, non-faces with high face likelihood are often misdetected. In this paper, we propose a novel face detection framework which explicitly uses difference of the classifier output distribution between faces and non-faces. Experiments with real-world images, from MIT+CMU dataset and WWW, show that the proposed framework improves the detection rate approximately by 5%.

1. はじめに

画像中から物体を検出する技術は画像処理における基盤技術であり、監視カメラや顔認証システムといった高度な画像処理システムにおいてなくてはならない重要な技術である^{2),5)}。なかでも顔検出の分野は近年、めざましい進歩をとげており、様々な顔識別器、および学習アルゴリズムが提案されている³⁾⁻⁵⁾。たとえば、Rowleyらはニューラルネットワークを用いた識別器を提案している³⁾。また、Violaらはカスケード構造をなす複数のAdaBoost識別器を提案している⁴⁾。従来手法の中でも、Violaらの手法⁴⁾は高速かつロバストであるため、特に広く利用されている。

識別器は与えられたサブウィンドウを顔か非顔かに識別する機能しか有していないため、識別器のみではシーン中から直接、顔を検出することはできない。シーン中から顔を検出するためには、まず、シーン中

を走査しながらサブウィンドウを切り取り、次いで、切り取られたサブウィンドウに対して識別を行う。このような手法を用いた場合、1つの真の顔付近に複数個の検出が生じることが多い。そのため、近傍の検出を1つにまとめる処理が必要となる。この処理のことをグルーピングと呼ぶことにする。

このような従来の顔検出の枠組みにおける問題点は、各サブウィンドウをそれぞれ独立に評価している点である。つまり、評価値が閾値以上の値をとる非顔画像は近傍の値をみることなく、顔として検出されてしまう。

提案手法では従来の顔識別器を用いて評価値の分布を生成する。次に、この分布の形状を評価することで、入力画像からの顔検出を行う。これにより、局所的に高い評価値をとる非顔画像に対し、より正確な分類が可能となる。

本研究では、識別器の提案または改良ではなく、顔検出の枠組みを見直している。識別器は画像中から顔を検出する際の重要な要素であり、顔検出の精度を向上させるために、多くの研究者が識別器を改良してい

[†] 東京工業大学大学院理工学研究科
Graduate School of Science and Engineering, Tokyo
Institute of Technology

る．一方で，顔検出の識別器以外の部分に関する検討はあまりなされていない．本研究では顔検出の枠組みを見直し，同じ識別器を使用した場合でも，より良い検出結果を得る方法を検討している．

以下では，まず今回の提案手法において用いた識別器の説明をする．次に，顔画像と非顔画像における識別器の評価値分布の違いについて具体的に示す．そして，この分布を用いた顔検出の枠組みについて説明し，提案手法を用いた実験を行う．実験では，入力画像から顔を検出する際の提案手法の有効性について確認する．

2. AdaBoost 識別器

AdaBoost 識別器およびカスケード型 AdaBoost 識別器^{1),4)} は従来手法の中でも実用性に優れた方法として知られており，多くの研究者が関連する研究成果を発表している．

本研究では，WWW から集めた 5,000 の顔画像と 5,000 の非顔画像の学習データを用いて，学習を行った．5,000 の顔画像は手動で切り取り，非顔画像は顔の含まれていない画像よりランダムに生成した．すべての画像の大きさは 19×19 ピクセルにそろえられ，それらを使用し，識別器を学習した．弱識別器には HaarLike 特徴を用いた⁴⁾．今回の実験では，500 個の弱識別器からなる AdaBoost 識別器を使用している．

以下では式 (1) の $H(x)$ を AdaBoost 識別器の評価値と呼ぶ．

$$H(x) = \sum_{l=1}^T \alpha_l h_l(x) \quad (1)$$

ここで T は弱識別器の総数， $H(x)$ は AdaBoost 識別器の評価値， α_l は各弱識別器の重み， $h_l(x)$ は弱識別器の識別結果である．AdaBoost 識別器では，評価値 $H(x)$ に対して閾値処理をすることで，識別を行っている．本研究では，閾値処理を行う前の評価値の分布を利用する．

次章では，この評価値の分布の違いを利用した顔検出を提案する．

3. 分布を利用した顔検出

3.1 顔画像と非顔画像における評価値の分布

我々は真の顔 (true positive) と，あるサブウィンドウで高い評価値を示す非顔 (false positive) における評価値分布の違いを実験的に調べた．

図 1 は，典型的な顔と非顔画像において，サブウィンドウの位置を変化させた際の評価値の分布を示して

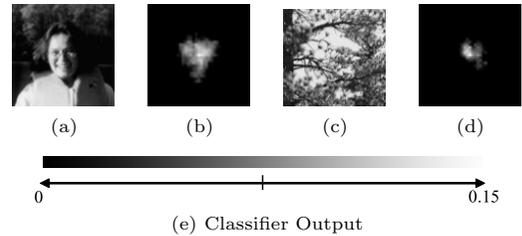


図 1 顔画像と非顔画像における評価値分布の差：(a) 顔画像例，(b) 顔画像 (a) の分布，(c) 非顔画像例，(d) 非顔画像 (c) の分布．(e) 識別器の評価値と表示色の対応

Fig. 1 Evaluation Distribution: (a) a typical face example, (b) an evaluation distribution of image (a), (c) a typical false alarmed non-face example, (d) an evaluation distribution of image (c). (e) a scale for evaluation value

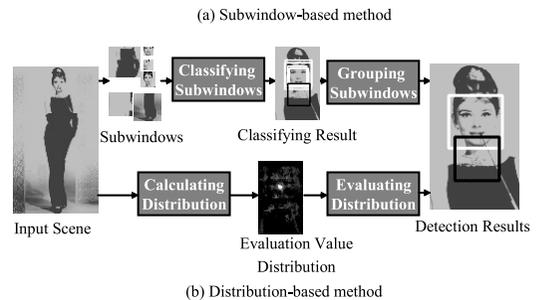


図 2 分布ベース (提案手法) およびサブウィンドウベース (従来手法) の処理の流れ

Fig. 2 The procedure of subwindow-based and distribution-based method.

いる．サブウィンドウの大きさを固定し，位置のみを変化させた際の各サブウィンドウの評価値をサブウィンドウの中心位置にプロットした．各入力画像は評価値が最大になる位置が中心になるように調整されている．図 1 の (a)，(c) は評価値の最大値がほぼ同じになる顔画像と非顔画像である．まず，顔画像の例に着目すると，顔画像の分布 (b) においては，中心から離れた領域においても評価値が高い値を持っていることが分かる．一方で，非顔画像の分布 (d) では，高い評価値を持つ領域が，(b) と比較して狭いことが確認できる．もし，中心のサブウィンドウの評価値のみを使用して画像 (c) を評価した場合には (c) の画像を顔として誤検出してしまう．しかし，近傍の評価値を含めて判断した場合には (c) の画像を正しく分類することが可能である．

3.2 分布を利用した顔検出

図 2 は従来手法であるサブウィンドウベースの顔検出と，提案手法である分布ベースの顔検出の処理の流れを示している．一方，図 2 (b) は分布ベースの顔

検出の流れを示している．まず，従来の顔識別器を用いて評価値を計算し，評価値の分布を作成する．この分布は横位置，縦位置，大きさの三次元の分布を持っている．評価値の分布を作成した後，この分布の形状を評価する．

最後に，画像全体で分布の評価の最大値を探索し，顔として検出する．2 番目，3 番目...と順に閾値を超える局所最大値がなくなるまで探索を繰り返す．

サブウィンドウベースの顔検出が，サブウィンドウを独立に評価しているのに対して，分布ベースである提案手法はシーン中から顔を探すことを積極的に考慮している．各サブウィンドウの評価値だけで識別を行うのではなく，その近傍領域の評価値まで含めた分布を評価する．これにより，局所的に高い評価値を持つ非顔画像を正しく分類することができる．さらに非顔画像の評価が相対的に低くなることで，サブウィンドウベースの顔検出では検出できなかった顔も正しく検出することが可能となる．

3.3 分布の評価法

ここでは，位置に関する顔画像の分布を評価する方法について説明する．分布を評価する方法はいくつか存在するが，本研究では簡単のため，あらかじめ推定した真の顔の評価値分布を利用し，入力画像中の注目点および近傍の分布の評価を行う．この処理はあらかじめ算出した真の顔の評価値分布との類似度を求めることに相当する．本研究では，式 (2) を分布の評価式として用いた．

$$g(u, v, s) = \sum_{\zeta} \sum_{\eta} \sum_{\theta} w(\zeta - u, \eta - v, \theta/s) d(\zeta, \eta, \theta). \quad (2)$$

ここで， $d(\xi, \eta, \theta)$ は位置 (ξ, η) ，大きさ θ における識別器の評価値， $g(u, v, s)$ は分布の評価， $w(\xi - u, \eta - v, \theta/s)$ は真の顔の分布を示している．ここで，扱う分布は位置と大きさの三次元を考慮している．真の顔の評価値分布は，顔の大きさを基準として変化する．そのため，分布を評価する際にも検出する際の大きさを考慮する必要がある．式 (2) における分布の評価もすべて顔の大きさを基準とした座標において分布を評価している．本研究では， $w(\xi - u, \eta - v, \theta/s)$ の分布を統計的に処理し，設計した．まず，学習に用いた画像とは別に 50 個の顔画像を収集し，それぞれの評価値分布を求める．これらの画像は評価値分布を得るために少し大きめのものを用いている．そして，分布の平均を一次関数で近似したものを真の顔の分布とした．近似に利用したデータは経験的に評価値が 0 以上

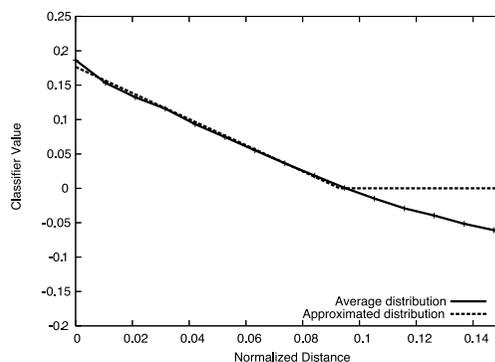


図 3 近似した評価値分布 (位置): 実線は顔の評価値分布の平均．破線は近似した分布

Fig. 3 Approximated distribution over location: Solid line is the average distribution of typical face images. Dashed lines is the approximated distribution.

のものに限定した．図 3 に分布の平均と一次関数で近似結果を示す．

4. 実験結果

本章では，入力シーンからの顔検出結果により，提案の分布ベースの顔検出と，従来のサブウィンドウベースの顔検出の比較を行う．入力シーンに対して顔検出を行い，顔検出結果とあらかじめ手動により設定した顔の位置と大きさが一定の範囲内の場合に正しく検出されたとし，検出率を算出した．

実験では，511 個の顔を含んだ 170 枚のシーン画像を用いた．このシーン画像は MIT+CMU テストセット³⁾と WWW から集めたシーン画像から構成される．比較対象であるサブウィンドウベースのグルーピング処理には，Rowley らの手法³⁾を適用した．

図 4 にこの実験における分布ベースおよびサブウィンドウベースの ROC カーブを示す．図 4 の ROC カーブを確認すると，提案手法である分布ベースの顔検出が従来のサブウィンドウベースの手法を上回っており，同一の誤検出数のときに最大で約 5% 検出率が改善されている．

図 5 は MIT+CMU テストセットにおける顔の検出結果である．正しい顔を検出した結果は白い枠で，非顔画像を誤検出した結果は黒い枠でそれぞれ示してある．閾値は使用したテスト画像全体の検出率が 80% になるよう図 4 から算出した．サブウィンドウベースでの検出結果である図 5 の (a) を見ると，4 個の誤検出が確認できる．一方で，分布ベース (b) では誤検出数が 0 個になっていることが分かる．

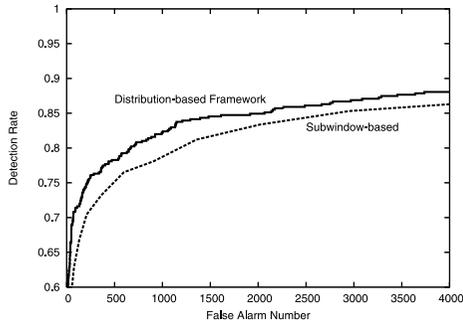


図4 実シーンを使った際の ROC カーブ：横軸は誤検出数・縦軸は検出率

Fig.4 ROC curves using real-world scenes: the horizontal axis shows false alarm numbers and the vertical axis shows the detection rate.

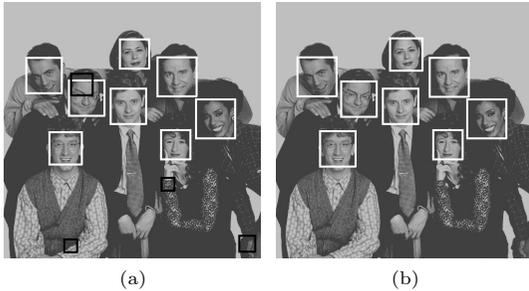


図5 実画像における顔検出結果：(a) サブウィンドウベース，(b) 分布ベース

Fig.5 Face detection results from the MIT+CMU dataset: (a) subwindow-based (b) distribution-based.

5. おわりに

本研究では、識別器による評価値の分布を陽に利用した顔検出の枠組みを提案した。提案手法は位置と大きさにおける分布に着目し、近傍の分布を比較することで、より精度の良い顔検出を行う。分布を利用することで、顔画像と非顔画像の差が強調されるため、検出率が改善されると考えられる。

今後の課題としては、提案した分布ベースの枠組みを他の識別器に適用した際の比較・検討があげられる。特に、AdaBoost アルゴリズムを高速化したカスケード型 AdaBoost 識別器⁴⁾ に対する適用は実用面で重要であると考えている。また、分布のより適切な評価方法も検討していきたい。

参考文献

1) Friedman, J., Hastie, T. and Tibshirani, R.: Additive logistic regression: A statistical view of boosting, *The Annals of Statistics*, Vol.38, No.2, pp.337-407 (2000).

- 2) Okuma, K., Taleghani, A., de Freitas, N., Little, J. and Lowe, D.: A boosted particle filter: Multitarget detection and tracking, *European Conference on Computer Vision*, pp.28-39 (2004).
- 3) Rowley, H.A., Baluja, S. and Kanade, T.: Neural network-based face detection, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.20, No.1, pp.23-38 (1998).
- 4) Viola, P. and Jones, M.: Robust real-time face detection, *International Journal of Computer Vision*, Vol.57, No.2, pp.137-154 (2004).
- 5) Yang, M.H., Kriegman, D.J. and Ahuja, N.: Detecting Faces in Images: A Survey, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.24, No.1, pp.34-58 (2002).

(平成 19 年 4 月 14 日受付)

(平成 19 年 7 月 31 日採録)

(担当編集委員 中井 宏章)



高塚 皓正

2005 年東京工業大学制御システム工学科卒業。2007 年同大学大学院理工学研究科機械制御システム専攻修士課程修了。



田中 正行

1998 年東京工業大学工学部制御システム工学科卒業。2000 年同大学大学院理工学研究科制御工学専攻修士課程修了。2003 年同大学院理工学研究科機械制御システム専攻博士課程修了。同年アジレント・テクノロジー(株)入社。2004 年東京工業大学大学院理工学研究科機械制御システム研究員。工学博士。



奥富 正敏(正会員)

1981 年東京大学工学部計数工学科卒業。1983 年東京工業大学大学院理工学研究科制御工学専攻修士課程修了。同年キヤノン(株)入社。1987~1990 年カーネギーメロン大学コンピュータサイエンス学科客員研究員。1994 年東京工業大学大学院情報理工学研究科情報環境学専攻助教授。2002 年同大学院理工学研究科機械制御システム専攻教授。工学博士。