NAÏVE BAYES FACE/NONFACE CLASSIFIER: A STUDY OF PREPROCESSING AND FEATURE EXTRACTION TECHNIQUES

Son Lam Phung, Abdesselam Bouzerdoum, Douglas Chai, and Anthony Watson
Emails: {s.phung, a.bouzerdoum, d.chai, a.watson}@ecu.edu.au
Edith Cowan University, Perth, AUSTRALIA

ABSTRACT

This paper presents a classifier of face and nonface patterns that is based on the naïve Bayes model. Using this classifier as a tool, we analyze the effects on classification performance of preprocessing, feature extraction and classifier combination techniques. Our analysis shows that image normalization techniques that reduce the effects of different lighting conditions improve face/nonface classification significantly. In addition, techniques such as background masking and combining classifiers that use different feature vectors are shown to enhance classification performance. Over a test set of 12,000 patterns, the combined classifier using four feature vectors has correct detection rates (CDRs) of 96.2% and 99.2% at false detection rates (FDRs) of 1% and 5%, respectively.

1. INTRODUCTION

Automatic face detection in still images and video has attracted strong interests in recent years [2, 3, 5, 7-9]. It has applications in face recognition, intelligent human computer interface, video retrieval, and surveillance. Recently, we proposed a new color-based face detection algorithm [4]. This algorithm combines both analytic and holistic approaches to face detection. The algorithm consists of two important components: (i) face candidate selection using analytic facial features, and (ii) face candidate verification using holistic face/nonface pattern classifiers.

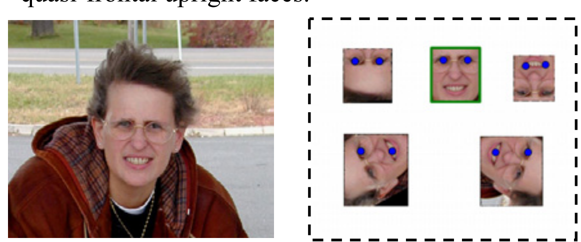
This paper addresses the naïve Bayes face/nonface classifier that we use in the face detection algorithm. Using this classifier architecture as a tool, we also investigate important issues in face/nonface classification, namely the effects on classification performance of image normalization, background masking, feature extraction schemes, and classifier combination. The paper is organized as follows. A brief overview of the face detection algorithm is given in Section 2. The naïve Bayes face/nonface classifier is addressed in Section 3. In that section, preprocessing, feature extraction, and classifier combination techniques are described. Experimental results and analysis are presented in Section 4. Conclusions are given in Section 5.

2. OVERVIEW OF COLOR-BASED FACE DETECTION ALGORITHM

Skin regions in the color input image are first located using a Bayesian classifier. Within each skin region, eye regions are identified using their color characteristic. For each detected eye pair, two face candidates are formed using a geometric face model, which specifies the face region boundary in terms of the two eye locations. All face candidates are preliminarily checked against a number of heuristics about the face. The remaining candidates are rotated to the upright position where the eye-line is horizontal. Examples of *normalized* candidates generated by this face candidate selection scheme are shown in Fig. 1. These normalized face candidates are then processed by a face/nonface classifier to ensure that they are quasi-frontal upright faces.

The face candidate selection scheme has the following advantages:

- It eliminates the need for scanning the input image exhaustively window-by-window; face candidates are generated quickly.
- Quasi-frontal faces with arbitrary in-plane rotation can be detected. However, due to normalization of the eye line, the face/nonface classifier needs to focus only on quasi-frontal upright faces.



(a) input image (b) normalized face candidates

Figure 1. Face candidate selection scheme in [4]. The eye points for each face candidate are also marked. These candidates will be verified by the face/nonface classifier.

3. NAÏVE BAYES CLASSIFIER

For a given input window, a feature vector **x** is extracted. This feature vector is classified into face or nonface class according to the Bayesian decision rule [1]. Let $p(\mathbf{x}|\omega_{\text{face}})$ and $p(\mathbf{x}|\omega_{\text{nonface}})$ be the class-conditional pdfs of face and nonface classes, respectively. The feature vector **x** is classified as a face pattern if the following log-likelihood score exceeds a threshold:

$$\mathcal{L}(\mathbf{x}) = \log\left(p(\mathbf{x} \mid \omega_{face})\right) - \log\left(p(\mathbf{x} \mid \omega_{nonface})\right). \quad (1)$$

In our approach, the class-conditional pdfs are estimated using the naïve Bayes model, which has been observed to work well in practice [1]. The naïve Bayes model assumes statistical independence among elements of the feature vector **x**. Under this assumption, we obtain:

$$p\left(\mathbf{x} \mid \omega_{face}\right) = \prod_{i=1}^{N} p\left(x_i \mid \omega_{face}\right), \quad (2)$$

$$p\left(\mathbf{x} \mid \omega_{nonface}\right) = \prod_{i=1}^{N} p\left(x_i \mid \omega_{nonface}\right). \quad (3)$$

The marginal pdfs $p(x_i|\omega_{\text{face}})$ and $p(x_i|\omega_{\text{nonface}})$ are calculated using the histogram technique. The naïve Bayes classifier has been used by Schneiderman and Kanade [6]. In their approach, an input window of size 64×64 is divided into overlapping subregions, from which a feature vector is extracted; subregion appearance and position are taken into consideration. In our work, we experiment with several feature extraction techniques.

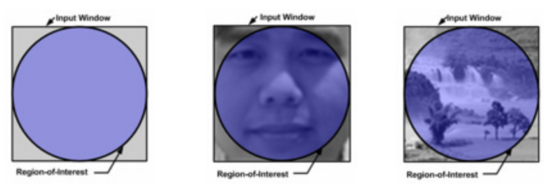
3.1. Preprocessing Techniques

A challenge in face/nonface classification is coping with variations, both intrinsic and extrinsic, in the face pattern. Intrinsic variations are due to different people, facial expressions, and natural facial features such as moustaches and beards. Extrinsic variations are caused by different lighting conditions, viewing angles, scales, and artificial accessories such as eye glasses. In practice, large training sets are used to cater for intrinsic variations, whereas preprocessing techniques are used to reduce extrinsic variations. Preprocessing techniques includes image normalization for reducing the influence of the lighting conditions, and background masking for excluding irrelevant image portion.

We study the effects of common image normalization techniques on face/nonface classification performance. These techniques include:

- *Mean normalization*: removing the mean intensity from input window;
- *Range normalization*: stretching pixel intensities linearly to occupy the full range [0, 255];
- *Standard deviation normalization*: scaling linearly pixel intensities so that the window has a zero mean, and a standard deviation of 1;
- *Histogram equalization*: applying an intensity mapping on the window so that its histogram becomes approximately uniform.
- *Illumination gradient correction* [7]: subtracting from the window a linear function of pixel coordinates that best fits the window intensities; the new window is then histogram-equalized. This normalization reduces the effect of directional lighting.
- *Left-right histogram equalization* [6]*:* the left half and the right half of the input window are separately histogram-equalized. This normalization addresses the situation when, due to directional lighting, the face's left half appears dimmer or brighter than the right half.

With background masking, the pixels near the four corners of the rectangular window are excluded from classification. These pixels either belong to the image background or lie near the face contour, which is highly variable from one person to another. We shall investigate also the effects of background masking on face/nonface classification. The circular mask as shown in Fig. 2 is used as an example.



(a) circular mask (b) face (c) nonface

Figure 2. Background masking.

3.2. Feature Extraction Techniques

The size of base input window is 64×64. However, this input window is resized to smaller windows as needed by a particular feature extraction technique. In our work, four feature extraction techniques are studied:

- *Intensity*: **x** consists of pixel intensities taken directly from a 16×16 input window through lexicographic ordering;
- *Projection on face subspace* (PFS): **x** is the projection of the 16×16 input window onto a face subspace spanned by 100 eigenvectors. These eigenvectors are found through principal component analysis of 6,000 training face patterns.
- *Edge*: **x** is formed through lexicographic ordering a binary edge map of the 64×64 input window. The edge map is found through the Canny edge detector.
- *Profile*: **x** is formed by concatenating the projections of an 32×32 input window on horizontal and vertical image axes.

3.3. Classifier Combination

It is possible to use only one naïve Bayes classifier for face/nonface classification. However, we suspect that classification will improve if an ensemble of many classifiers, which are based on diverse feature extraction techniques, is used. Hence, we investigate this hypothesis by examining the performance of an ensemble of four individual classifiers. These classifiers use the feature extraction techniques described in the previous section. Let $\mathcal{L}_{\text{int}}(\mathbf{x})$, $\mathcal{L}_{\text{pfs}}(\mathbf{x})$, $\mathcal{L}_{\text{edge}}(\mathbf{x})$, and $\mathcal{L}_{\text{prof}}(\mathbf{x})$ be the log-likelihood scores of the four individual classifiers. These scores are clipped and mapped linearly to the range of [0, 1], and considered as pseudo scores. The four pseudo scores are averaged to give the face score of the classifier ensemble. Face/nonface classification is done by thresholding this final face score.

4. EXPERIMENTAL RESULTS AND ANALYSIS

4.1 Data Preparation

We collected about 12,000 face windows and 1,800 landscape photos. The face windows were manually cropped from Web images and images in existing face databases (e.g. BioID, UMIST, AT&T and AR). The faces are of different people and skin types (see Fig. 3). The nonface windows were randomly extracted from the landscape photos. Each window has a size of 64×64 pixels. The set of 12,000 face windows was partitioned into two parts: 6,000 face windows for training, and 6,000 face windows for testing. For training purpose, each training face window was rotated, scaled and shifted by small random amount to generate 10 extra face windows. Although we focus on quasi-frontal upright faces, these artificial variations would make the classifier more robust to small variations in orientation, scale and position. The final data sets were:

- Training: 60,000 face and 120,000 nonface patterns;
- Testing: 6,000 face and 6,000 nonface patterns.

4.2 Analysis

Analysis of preprocessing techniques

The intensity feature vector extracted from window of size 16×16 was used in this experiment. The Receiver Operating Characteristic (ROC) curves of different preprocessing techniques are plotted in Fig. 4. The bottom two curves in Fig. 4 were obtained when no image normalization was applied: one curve with background masking, the other without. These two curves show that background masking did improve classification slightly. For the other ROC curves in Fig. 4, background masking was applied in addition to different image normalization techniques. The ROC curves show that classification is greatly improved if image normalization is applied. For example, for FDR between 5-20%, there was a difference of more than 10% in CDR when image normalization techniques were applied, compared to no image normalization. Among the tested normalization schemes, the two histogram equalization techniques performed significantly better than the rest. Lastly, the left-right histogram equalization performs slightly better than the standard histogram equalization.

Figure 3: Examples of face patterns used in our work.

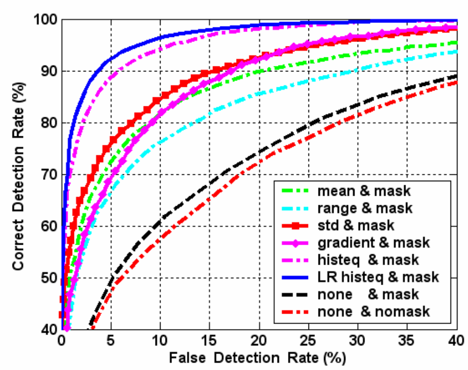


Figure 4. Comparison of preprocessing techniques using the intensity feature vector.

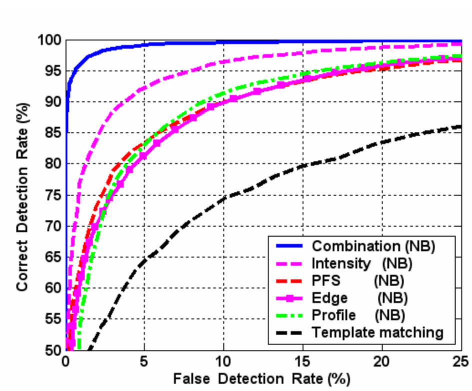


Figure 5. Comparison of feature extraction techniques and classifier combination.

Analysis of feature extraction techniques and classifier combination

We compared the four feature extraction techniques and the classifier ensemble. The face/nonface classifier based on template matching and cross-correlation was included as a baseline for comparison. Results in Fig. 5 show that the intensity feature vector performed better than the other three feature vectors.

There was a significant improvement in classification rates when the four classifiers were combined. The combined classifier using four feature vectors had CDRs of 96.2% and 99.2% at FDRs of 1% and 5%, respectively. In comparison, the classifier using intensity feature vector had CDRs of 77.8% and 92.2% at FDRs of 1% and 5%, respectively. The improvement can be attributed to the diversity in the feature vector pool. For example, we notice that while the intensity feature vector had the best overall performance, it sometimes failed to reject nonface patterns that have uniform or smooth textures. In comparison, the edge-based feature vector was very successful in rejecting such nonface patterns. Fig. 5 also shows that the naïve Bayes classifiers all performed better than the correlation-based template-matching classifier.

The naïve Bayes classifier performs well when the feature vector $\mathbf{x}$ satisfies the assumption of statistical independence among its elements. At presence, we have examined only four feature vectors. We plan to use the naïve Bayes classifier with more diverse feature vectors (e.g. independent component analysis and wavelets).

In our face detection algorithm in [4], the detection is further improved using contextual information. A face candidate is kept only if its face score exceeds the face scores of all face candidates that overlap it. This technique helps removing false alarms because in the proposed face candidate selection, many face candidates tend to overlap each other.

5. CONCLUSIONS

A face/nonface classifier based on the naïve Bayes model is presented. Using this classifier, we show that image normalization techniques for lighting correction can improve classification performance significantly. The results demonstrate that the left-right histogram equalization technique performs better compared to the other tested image normalization techniques. We also confirm that classification is enhanced if background portion of the input window is excluded from classification. In this paper, four different feature extraction techniques are proposed for face/nonface classification. It is shown that the classification is also improved if the four feature vectors are combined.

6. ACKNOWLEDGMENT

This work is supported in part by a grant from the Australian Research Council.

7. REFERENCES

[1] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York: John Wiley & Sons, Inc., 2001.

[2] R.-L. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face detection in color images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 696-707, 2002.

[3] C. Liu, "A Bayesian discriminating features method for face detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 725 - 740, June 2003.

[4] S. L. Phung, A. Bouzerdoum, D. Chai, and W. Kuczborski, "A color-based approach to automatic face detection," *Proc. IEEE International Symposium on Signal Processing and Information Technology*, Dec. 2003.

[5] H. A. Rowley, S. Baluja, and T. Kanade, "Neural network-based face detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, pp. 23-38, 1998.

[6] H. Schneiderman and T. Kanade, "Probabilistic modeling of local appearance and spatial relationships for object recognition," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp. 45-51, 1998.

[7] K. K. Sung and T. Poggio, "Example-based learning for view-based human face detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, pp. 39-51, 1998.

[8] H. Wu, Q. Chen, and M. Yachida, "Face detection from color images using a fuzzy pattern matching method," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, pp. 557-563, 1999.

[9] M.-H. Yang, D. Kriegman, and N. Ahuja, "Detecting faces in images: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 34-58, Jan. 2002.