Face Verification Using Gabor Wavelets and AdaBoost

Mian Zhou and Hong Wei
School of Systems Engineering, University of Reading
Whiteknight, Reading, RG6 6AY, United Kingdom
{m.zhou,h.wei}@reading.ac.uk

Abstract

This paper presents a new face verification algorithm based on Gabor wavelets and AdaBoost. In the algorithm, faces are represented by Gabor wavelet features generated by Gabor wavelet transform. Gabor wavelets with 5 scales and 8 orientations are chosen to form a family of Gabor wavelets. By convolving face images with these 40 Gabor wavelets, the original images are transformed into magnitude response images of Gabor wavelet features. The AdaBoost algorithm selects a small set of significant features from the pool of the Gabor wavelet features. Each feature is the basis for a weak classifier which is trained with face images taken from the XM2VTS database. The feature with the lowest classification error is selected in each iteration of the AdaBoost operation. We also address issues regarding computational costs in feature selection with AdaBoost. A support vector machine (SVM) is trained with examples of 20 features, and the results have shown a low false positive rate and a low classification error rate in face verification.

1. Introduction

The task of face verification is to verify a claimed identity by comparing a claimed image of the individual with other images belonging to the individual in a database. A set of images is divided into classes which are either clients or impostors. A client is a registered person with claimed identity. Impostors are all other persons except of the client. The face verification process consists of two phases: feature selection and classification. Feature selection not only reduces the dimension of the data, but also makes verification more accurate. Classification verifies a new face image as a client or an impostor. In this paper, we present a novel face verification algorithm based on Gabor wavelets and AdaBoost (Adaptive Boosting).

The Gabor wavelets perform desirable characteristics of spatial locality and orientation selectivity. In [8, 10, 11], it has been declared that the Gabor wavelet representation of face images is robust against variations due to illumination and facial expression changes. Two-dimensional Gabor wavelets were introduced by Daugman [1] for human iris recognition. Lades *et al.*[4] employed Gabor wavelets for face recognition using the Dynamic Link Architecture (DLA) framework. Wiskott *et al.*[14] expanded on DLA by developing a Gabor wavelet-based Elastic Bunch Graph Matching (EBGM) algorithm to label and recognise human faces. Liu and Wechsler [6] applied the Enhanced Fisher linear discriminant Model (EFM) to an augmented Gabor feature vector derived from the Gabor wavelet representation of face images. Wu *et al.*[15] used a boosting algorithm for glass detection by applying two types of wavelet features, Haar and Gabor, and the results have shown that the Gabor performed better than the Haar. AdaBoost was formulated by Freund and Schapire [2]. It is a relatively efficient, simple, and easy learning strategy for improving the performance of classification algorithms. It was first applied to face detection by Viola and Jones [13] to select Haar wavelet features and train a cascade of classifiers.

The rest of the paper is organised as follows. In section 2, we describe the Gabor wavelet features. In section 3, the AdaBoost algorithm for feature selection is given. The experimental method of classification and results are shown in section 4. Section 5 gives the conclusions.

2. Gabor Wavelet

A Gabor wavelet $\psi_{\mu,\nu}(z)$ is defined as [5]

$$\psi_{\mu,\nu}(z) = \frac{\|k_{\mu,\nu}\|^2}{\sigma^2} e^{-\frac{\|k_{\mu,\nu}\|^2\|z\|^2}{2\sigma^2}} [e^{ik_{\mu,\nu}z} - e^{-\frac{\sigma^2}{2}}] \quad (1)$$

where $z = (x, y)$ is the point with the horizontal coordinate x and the vertical coordinate y. The parameters μ and ν define the orientation and scale of the Gabor kernel, $\|\cdot\|$ denotes the norm operator, and σ is related to the standard derivation of the Gaussian window in the kernel and determines the ratio of the Gaussian window width to the wavelength. The wave vector $k_{\mu,\nu}$ is defined as follows

$$k_{\mu,\nu} = k_\nu e^{i\phi_\mu} \tag{2}$$

where $k_\nu = \frac{k_{max}}{f^\nu}$ and $\phi_\mu = \frac{\pi\mu}{8}$ if 8 different orientations have been chosen. k_{max} is the maximum frequency, and f^ν is the spatial frequency between kernels in the frequency domain.

In our approach, five different scales and eight orientations of Gabor wavelets were used, *i.e.*, $\nu \in \{0,\ldots,4\}$, and $\mu \in \{0,\ldots,7\}$. Gabor wavelets were chosen with related to $\sigma = 2\pi$, $k_{max} = \frac{\pi}{2}$, and $f = \sqrt{2}$ [14, 6, 5, 15].

The Gabor wavelet representation $O_{\mu,\nu}(z)$ is the convolution of the image $I(z)$ with a family of Gabor kernels $\psi_{\mu,\nu}(z)$. The response $O_{\mu,\nu}(z)$ to each Gabor kernel is a complex function, so that the magnitude response $\|O_{\mu,\nu}(z)\|$ is used to represent the features. Therefore, a Gabor wavelet feature j is configured by the three key parameters: position z, orientation μ, and scale ν, defined as

$$j(z,\mu,\nu) = \|O_{\mu,\nu}(z)\|. \tag{3}$$

3. AdaBoost

For a given image $I(z)$ with $M \times N$ pixels, the number of Gabor wavelet features is of the order of $M \times N \times 40$. They reside in a very high dimensional space which is 40 times larger than the original image space. We use AdaBoost to select significant features from the pool of Gabor wavelet features, hence to reduce the dimension.

The algorithm of AdaBoost is presented in Table 1. It maintains a probability distribution of weights ω_t over the training set. The initial weight $\omega_{1,i}$ for each example is given according to the proportion of client's examples to impostors' examples. All weights within the same class (either client or impostor) are set equally. In the course of iteration, the value of error ε_j varies in each iteration with updated weights $\omega_{t,i}$. In the step 4 of Table 1, the updating of the weights is controlled by two parameters: β_t and e_i. The parameter β_t is determined by the lowest error ε_t in iteration t. For preventing the loss of any generality, the weak classifier h_t has better classification performance than random guessing [2]. This requires $\varepsilon_t < 1/2$. It is because if a classifier performs in random guessing, statistically the classification error ε is equal to or greater than $1/2$. By employing this concept, the computational cost can be reduced (see Section 4). According to Equation (4), when the lowest error ε_t in each iteration is less than 1/2, obviously β_t will be less than 1. Therefore, in step 4, if example x_i is classified correctly, the weight $\omega_{t+1,i}$ for the next iteration is decreased. If example x_i is misclassified, $\omega_{t+1,i}$ remains constant. At the beginning of iteration $t+1$, the weights are normalised and kept as a distribution *i.e.*, $\sum_{i=1}^{n} \omega_{t,i} = 1$. Therefore, the ratio of allocated weights corresponding to each example remain unchanged. In general, the weights are decreased or increased over time relatively, not absolutely. Each Gabor wavelet feature j corresponds to a weak classifier h_j. A Linear Fisher Discriminant (LFD) classifier is adopted as the weak classifier in feature selection. It determines the optimal threshold for the classification function, such that the minimum number of examples is misclassified. The final strong classifier H takes the form of a combination of weighted weak classifiers h_t followed by a threshold. In each iteration of training, one feature is selected by choosing the corresponding weak classifier which has the lowest error in each iteration.

Table 1. The AdaBoost algorithm for feature selection.

* Given the training set $(x_1, y_1), \ldots, (x_n, y_n)$, where x_i is the data of the ith example, and $y_i = 0, 1$ for impostors and clients respectively.
* Initialize weights $\omega_{1,i} = 1/2m, 1/2l$ for $y_i = 0, 1$ respectively, where m and l are the number of impostors and clients respectively.
* For $t = 1, \ldots, T$:
  1. Normalize the weights, $\omega_{t,i} \leftarrow \frac{\omega_{t,i}}{\sum_{i=1}^{n} \omega_{t,i}}$ so that ω_t is a probability distribution.
  2. For each feature j, train a classifier h_j which uses a single feature. The error is evaluated with respect to ω_t, $\varepsilon_j = \sum_i \omega_{t,i} |h_j(x_i) - y_i|^2$.
  3. Choose the classifier h_t, with the lowest error ε_t.
  4. Update the weights:

     $$\omega_{t+1,i} = \omega_{t,i} \beta_t^{1-e_i} \tag{4}$$

     where $e_i = 0$ if example x_i is classified correctly, $e_i = 1$ otherwise, and $\beta_t = \frac{\varepsilon_t}{1-\varepsilon_t}$.
* The final strong classifier is the combination of classifiers with the lowest error found in each iteration.

  $$H(x) = \text{sign}(\sum_{t=1}^{T} \alpha_t h_t(x))$$

  where $\alpha_t = \log \frac{1}{\beta_t}$

4. Experiments and Results

The XM2VTS database [7] is used in the experiments for feature selection, classifier training, and testing. XM2VTS contains 295 subjects consisting of 200 clients and 95 impostors with 8 up-right and frontal view face images per subject. The database was divided into 2 sets: training set and testing set. The training set is used to select features and train a classifier, which includes the first 4 images of each client. The testing set is used to evaluate the algorithm performance, which includes the other 4 images of each client, and 8 images of each impostor. There are 2360 images in the database: 800 for training, and 1560 for testing. All images were captured with moderate difference in illumination, expressions and facial details. Using the manually detected centres of the two eyes on each face image, all images are properly rotated, translated, segmented and scaled to fit a grid size of 25×24. A similar approach has been used in [3, 12], where positive examples are intra-personal differences, and negative examples are extra-personal differences. Both approaches require a large number of training data which are demanded by the nature of the AdaBoosting algorithm. However it leaves the ratio between positive and negative examples unbalanced, *e.g.* few positive examples and many negative examples.

4.1. Feature Selection

Each image is convolved with 40 Gabor wavelets according to Equation (1) with $\nu \in \{0, \ldots, 4\}$, and $\mu \in \{0, \ldots, 7\}$. Consequently, the number of features for each image is $25 \times 24 \times 40 = 24,000$. In the training set, we have all 200 clients in the XM2VTS with 4 images as positive examples and other 796 images as negative examples. AdaBoost (in Table 1) trains weak classifiers corresponding to each feature from all 24,000 features. One hypothesis [2] for AdaBoost is that if each weak classifier is slightly better than random guessing, the error of the final classifier drops down exponentially. The hypothesis requires the error ε_t (in Table1) from each weak classifier be less than 1/2, if the weak classifier can contribute to the performance of the final classifier. In our experiment, AdaBoost training excludes those classifiers whose errors are equal or greater than 1/2. By applying this method, the computational costs are reduced greatly. For client 2, without adopting this technique, 455,810 classifiers in total are learned in 20 iterations, while with the technique, only 247,178 classifiers are learned in 20 iterations. The computational time is reduced to 54.23% of the original time. In the 20th iteration, only 6,195 features from 24,000 are remained in AdaBoost, while 17,785 features are rejected by the algorithm. The selected Gabor wavelet features after AdaBoost training are shown in Figure 1. The positions of the top 20 Gabor wavelet features selected from client 1 to client 8 are shown in the top row of Figure 1. Some features are overlapped because they are sharing the same position, but different orientation or scale. The first Gabor wavelet features selected from client 1 to client 8 are shown in the bottom row of Figure 1. The selected features are randomly distributed in the face area rather than concentrated on some regions of the faces. The first Gabor wavelet feature for each client is varied with different position z, orientation μ, and scale ν.

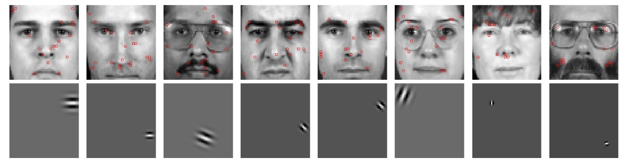

Figure 1. The AdaBoost algorithm selects the top 20 features and the first Gabor feature

4.2. Classification

With the 20 Gabor wavelet features selected for each client, a face image is represented by a vector with 20 components. Classification is performed on the training set and the testing set by using a support vector machine (SVM) [9] for each client. The idea of SVM is to maximise the margin between classes and minimise a quantity proportional to the number of misclassification errors. Experiments were carried out on eight clients from the XM2VTS. A non-linear SVM classifier with a polynomial kernel of degree 3 is constructed from the 800 training examples, which are the first four images across all clients. The testing set is from the rest four images across all clients and the eight images across all impostors in XM2VTS. Table 2 shows the classification results from client 1 to client 8 (C1 to C8) with 1,560 testing examples (4 positive examples and 1,556 negative examples from 199 clients with 4 images and 95 impostors with 8 images per subject) in XM2VTS. By adjusting the bias of each SVM classifier for each client, the false positive rate is set to 0%, 25%, 50%, 75% or 100%, and the corresponding classification error (Error rate) and false negative rates are presented. The optimal boundary for a SVM classifier gives a low error rate but a high false positive rate because the ratio between positive examples and negative examples remains unbalanced. Therefore, the decision surface gets close towards the actual boundary of the negative class, but far away from the place where the positive class resides in the feature space. Moreover, SVMs are only concerned with the trade off between margin and misclassification error, but not with the false positive rate. This leads the trained SVM classifiers having strong ability to recognise the negative examples, but relatively weak ability to recognise the positive examples. It also makes the false negative rate much closer to the classification error rate. By adjusting the bias of the SVM classifier, the false positive rate can be reduced, while the classification error rate will be increased, or *vice versa*, *e.g.* for client 4, the optimal boundary of the SVM classifier makes the false positive rate equal to 25%, and the error rate is 0.26%. However, when the bias of the classifier is decreased, no false positive is detected, and the error rate is increased to 2.95%. Table 2 shows that the classification to clients 1, 4 and 5 has better performance than to other clients. It indicates that the 20 selected features contribute well enough for some clients to verification, but they are not sufficient to verify some other clients *e.g.* clients 6 and 7. In this case, it may need more features selected by AdaBoost for client representation.

Table 2. The classification results from client 1 to client 8 (C1 to C8) in XM2VTS

False(%)	Error rate(%) / False Negative rate(%)			
Positive	C1	C2	C3	C4
0	4.94/4.95	13.72/13.75	8.53/8.55	2.95/2.96
25	1.73/1.67	12.12/12.08	11.10/8.29	0.26/0.20
50	0.83/0.71	8.72/8.61	9.58/7.13	0.32/0.19
75	0.38/0.19	1.41/1.22	8.30/6.10	0.32/0.13
100	0.26/0.0	0.58/0.32	2.12/1.35	0.26/0.0
	C5	C6	C7	C8
0	2.31/2.31	81.92/82.13	53.08/53.21	15.64/15.68
25	0.45/0.39	14.81/14.78	46.47/46.53	8.59/8.55
50	0.38/0.26	1.86/1.74	44.49/44.47	0.51/0.39
75	0.38/0.19	0.51/0.32	36.47/36.38	0.38/0.19
100	0.32/0.06	0.26/0.0	0.64/0.39	0.32/0.06

5. Conclusions

In this paper, a new face verification algorithm is presented. Gabor wavelets extract features from original face images. AdaBoost selects the top 20 significant features which distinguish a specified client from other subjects in the face database. The experiment has been carried out on the XM2VTS face database. Based on 20 selected Gabor wavelet features, a SVM classifier is built up for each client for verification. By adjusting bias in SVMs, we achieved face verification in a low false positive rate and a low false negative rate empirically. Gabor wavelet transform reflects salient changes between pixels. This makes it robust against to illuminance changes between images. AdaBoost is a well known time consuming online learning algorithm. We reduced the computational costs by removing weak classifiers whose errors are larger than that of random guessing. In 20 iterations of AdaBoost learning, the computational time is reduced to 54.23% of the original time. The algorithm is tested on 8 clients from XM2VTS. Experimental results have proved that the developed algorithm presented in this paper has high accuracy for face verification.

References

[1] J. Daugman. High confidence visual recognition of persons by a test of statistical independence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15:1148–1161, 1993.

[2] Y. Freund and R. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Computational Learning Theory: Eurocolt'95*, 1:23–37, 1995.

[3] M. Jones and P. Viola. Face recognition using boosted local features. Technical report, MERL, 2003.

[4] M. Lades, J. Vorbruggen, J. Budmann, J. Lange, C. Malsburg, and R. Wurtz. Distortion invariant object recognition on the dynamic link architecture. *IEEE Transactions on Computers*, 42:300–311, 1993.

[5] C. Liu. Gabor-based kernel PCA with fractional power polynomial models for face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26:572–581, 2004.

[6] C. Liu and H. Wechsler. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Transactions on Image Processing*, 11:467–476, 2002.

[7] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre. XM2VTSDB: The extended M2VTS databas. In *Proceedings of Second International Conference on Audio and Video-based Biometric Person Authentication*, volume 1, 1999.

[8] B. Olshasen and D. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381:607–609, 1996.

[9] E. Osuna, R. Freund, and F. Girosi. Training support vector machine: an application to face detection. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1997.

[10] R. Rao and D. Ballard. An active vision architecture based on iconic representations. *Artificial Intelligence*, 78:461–505, 1995.

[11] B. Schiele and J. Crowley. Recognition without correspondence using multidimensional receptive field histograms. *International Journal on Computer Vision*, 36:31–52, 2000.

[12] L. Shen, L. Bai, D. Bardsley, and Y. Wang. Gabor feature selection for face recognition using improved adaboost learning. In *Proceedings of International Workshop on Biometric Recognition System*, 2005. in conjuction with ICCV'05.

[13] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition*, pages 511–518, 2001.

[14] L. Wiskott, J. Fellous, N. Krüger, and C. Malsburg. Face recognition by elastic bunch graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (19):775–779, 1997.

[15] B. Wu, H. AI, and R. Liu. Glasses detection by boosting simple wavelet features. In *Proceedings of International Conference on Pattern Recognition*, Cambridge, 2004.