A Gabor Feature Classifier for Face Recognition

Chengjun Liu
Dept. of Math and Computer Science
University of Missouri
St. Louis, MO 63121
cliu@cs.umsl.edu

Harry Wechsler
Dept. of Computer Science
George Mason University
Fairfax, VA 22030
wechsler@cs.gmu.edu

Abstract

This paper describes a novel Gabor Feature Classifier (GFC) method for face recognition. The GFC method employs an enhanced Fisher discrimination model on an augmented Gabor feature vector, which is derived from the Gabor wavelet transformation of face images. The Gabor wavelets, whose kernels are similar to the 2D receptive field profiles of the mammalian cortical simple cells, exhibit desirable characteristics of spatial locality and orientation selectivity. As a result, the Gabor transformed face images produce salient local and discriminating features that are suitable for face recognition. The feasibility of the new GFC method has been successfully tested on face recognition using 600 FERET frontal face images, which involve different illumination and varied facial expressions of 200 subjects. The effectiveness of the novel GFC method is shown in terms of both absolute performance indices and comparative performance against some popular face recognition schemes such as the Eigenfaces method and some other Gabor wavelet based classification methods. In particular, the novel GFC method achieves 100% recognition accuracy using only 62 features.

1. Introduction

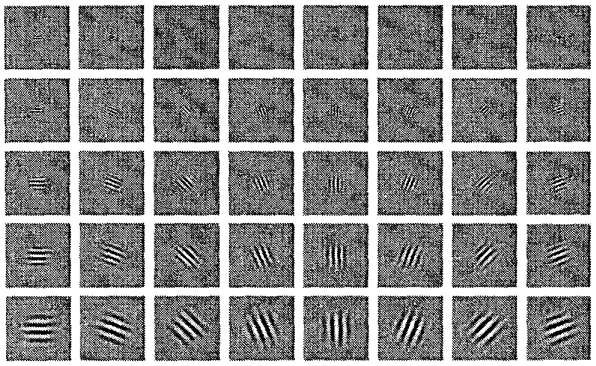
Face recognition has been largely motivated by the need for surveillance and security, telecommunication and digital libraries, human-computer intelligent interaction, and smart environments [17], [3], [6], [16], [12]. It usually employs various statistical techniques, such as PCA (principal component analysis) [19], [15], FLD (Fisher linear discriminant, a.k.a. linear discriminant analysis, or LDA) [18], [1], [8], [13], ICA (independent component analysis) [7], and Gabor based bunch graphs [20], to derive appearance-based models for classification.

This paper introduces a novel Gabor Feature Classifier (GFC) method for face recognition. The GFC method applies an enhanced Fisher model (EFM) [13] to an augmented Gabor feature vector, which is derived from the Gabor wavelet transformation of face images. The Gabor wavelets, whose kernels are similar to the 2D receptive field profiles of the mammalian cortical simple cells, exhibit desirable characteristics of spatial locality and orientation selectivity. The biological relevance and computational properties of Gabor wavelets for image analysis have been described in [4], [14], [5], [10]. As a result, the Gabor transformed face images yield features that display scale, locality, and differentiation properties. These properties are quite robust to variability of face image formation, such as the variations of illumination and facial expressions. To encompass all the features produced by the different Gabor kernels we concatenate the resulting wavelet features to derive an augmented Gabor feature vector. The dimensionality of the vector space is then reduced under the eigenvalue selectivity constraint of the EFM method to derive a low-dimensional feature representation with enhanced discrimination power. The feasibility of the new GFC method has been successfully tested on face recognition using 600 FERET frontal face images, which involve different illumination and varied facial expressions of 200 subjects. The effectiveness of the novel GFC method is shown in terms of both absolute performance indices and comparative performance against some popular face recognition schemes such as the Eigenfaces method and some other Gabor wavelet based classification methods. In particular, the novel GFC method achieves 100% recognition accuracy using only 62 features.

2. Gabor Feature Analysis

The Gabor wavelets, which capture the properties of spatial localization, orientation selectivity, spatial frequency selectivity, and quadrature phase relationship, seem to be a good approximation to the filter response profiles encountered experimentally in cortical neurons [4], [14], [9], [10], [2]. The Gabor wavelets have been found to be particu-

0-7695-1143-0/01 $10.00 © 2001 IEEE



(a)

(b)

Figure 1. Gabor Wavelets. (a) The real part of the Gabor kernels at five scales and eight orientations under the following parameters: $\sigma = 2\pi$, $k_{max} = \pi/2$, and $f = \sqrt{2}$. (b) The magnitude of the Gabor kernels at five different scales.

larly suitable for image decomposition and representation when the goal is the derivation of local and discriminating features. Most recently, Donato et al. [7] have experimentally shown that the Gabor filter representation is optimal for classifying facial actions. In this section, we review the basics on Gabor wavelets, describe the Gabor feature representation of images, and derive an augmented Gabor feature vector for face recognition.

2.1. Gabor Wavelets

Gabor wavelets are used for image analysis because of their biological relevance and computational properties [4], [14], [5], [10]. The Gabor wavelets (kernels, filters) can be defined as follows [4], [14], [11]:

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

where μ and ν define the orientation and scale of the Gabor kernels, $z = (x, y)$, $\|\cdot\|$ denotes the norm operator, and the wave vector $k_{\mu,\nu}$ is defined as follows:

$$k_{\mu,\nu} = k_\nu e^{i\phi_\mu} \quad (2)$$

where $k_\nu = k_{max}/f^\nu$ and $\phi_\mu = \pi\mu/8$. k_{max} is the maximum frequency, and f is the spacing factor between kernels in the frequency domain [11].

The set of Gabor kernels in Eq. 1 are all self-similar since they can be generated from one filter, the mother wavelet, by scaling and rotation via the wave vector $k_{\mu,\nu}$. Each kernel is a product of a Gaussian envelope and a complex plane wave, while the first term in the square brackets in Eq. 1 determines the oscillatory part of the kernel and the second term compensates for the DC value. The effect of the DC term becomes negligible when the parameter σ, which determines the ratio of the Gaussian window width to wavelength, has sufficiently high values.

In most cases one would use Gabor wavelets at five different scales, $\nu \in \{0, ..., 4\}$, and eight orientations, $\mu \in \{0, ..., 7\}$ [9], [10], [2]. Fig. 1 shows the real part of the Gabor kernels at five scales and eight orientations and their magnitude, with the following parameters: $\sigma = 2\pi$, $k_{max} = \pi/2$, and $f = \sqrt{2}$. The kernels exhibit desirable characteristics of spatial locality and orientation selectivity, making them a suitable choice for image feature extraction when one's goal is to derive local and discriminating features for (face) classification.

2.2. Gabor Feature Representation

The Gabor wavelet transformation of an image is the convolution of the image with a family of Gabor kernels as defined by Eq. 1. Let $I(x, y)$ be the gray level distribution of an image, and define the convolution output of image I and a Gabor kernel $\psi_{\mu,\nu}$ as follows:

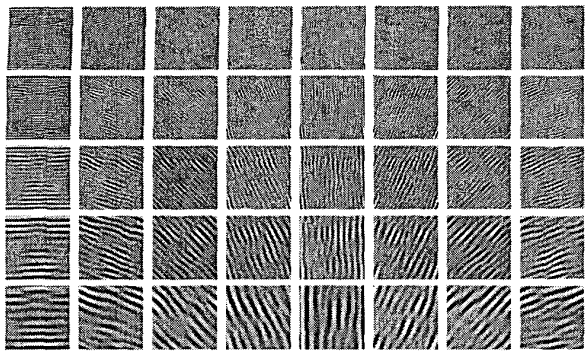
$$O_{\mu,\nu}(z) = I(z) * \psi_{\mu,\nu}(z) \quad (3)$$

where $z = (x, y)$, and $*$ denotes the convolution operator.

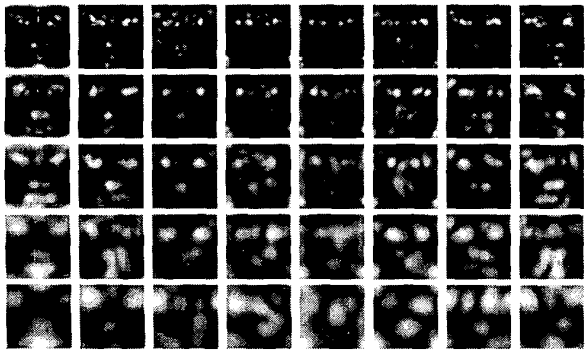
The convolution outputs (the real part and the magnitude) of a sample image and Gabor kernels (see Fig. 1) are shown in Fig. 2. The outputs exhibit desirable characteristics of spatial locality, scale and orientation selectivity similar to those displayed by the Gabor wavelets in Fig. 1. Since the outputs $O_{\mu,\nu}(z)$ consist of different local, scale and orientation features, we concatenate all those features together in order to derive an augmented feature vector $\mathcal{X}$. Before the concatenation, we first downsample each $O_{\mu,\nu}(z)$ by a factor ρ to reduce the space dimension, and normalize it to zero mean and unit variance. We then construct a vector out of the $O_{\mu,\nu}(z)$ by concatenating its rows (or columns). Now, let $O^{(\rho)}_{\mu,\nu}$ denote the normalized vector constructed from $O_{\mu,\nu}(z)$ (downsampled by ρ and normalized to zero mean and unit variance), the augmented Gabor feature vector $\mathcal{X}^{(\rho)}$ is then defined as follows:

$$\mathcal{X}^{(\rho)} = \left(O^{(\rho)t}_{0,0} \quad O^{(\rho)t}_{0,1} \quad \cdots \quad O^{(\rho)t}_{4,7} \right)^t \quad (4)$$

where t is the transpose operator. The augmented feature vector thus encompasses all the outputs, $O_{\mu,\nu}(z), \mu \in \{0, ..., 7\}, \nu \in \{0, ..., 4\}$, as important and discriminating information.



(a)



(b)

Figure 2. Gabor wavelet transformation of a sample image. (a) The real part of the transformation. (b) The magnitude of the transformation.

3. Gabor Feature Classifier

As our novel Gabor Feature Classifier (GFC) applies an enhanced Fisher discrimination model (EFM) [13] to the augmented Gabor feature vector $\mathcal{X}^{(\rho)}$ derived in Sect. 2.2, we briefly review the EFM method and discuss the similarity measures and the classification rule for GFC. The dimensionality of the resulting vector space is reduced, using the eigenvalue selectivity analysis of the EFM method, in order to derive a low-dimensional feature representation with enhanced discrimination power.

3.1. The Enhanced Fisher Discrimination Model (EFM)

The EFM method improves the generalization capability of the Fisher Linear Discriminant (FLD) criterion by decomposing it into a simultaneous diagonalization of the two within- and between-class scatter matrices [13]. Let Σ_w and Σ_b be the within- and between-class scater matrices, respectively. The EFM method first whitens the within-class covariance matrix:

$$\Sigma_w \Xi = \Xi \Gamma \quad and \quad \Xi^t \Xi = I \tag{5}$$

$$\Gamma^{-1/2} \Xi^t \Sigma_w \Xi \Gamma^{-1/2} = I \tag{6}$$

where $\Xi, \Gamma \in \mathbb{R}^{m \times m}$ are the eigenvector and the diagonal eigenvalue matrices of Σ_w. Note that m is the dimensionality of the reduced PCA space, and the choice of the range of principal components (m) for dimensionality reduction takes into account both the spectral energy and the magnitude requirements for adequate representation and good generalization, respectively [13].

The EFM method proceeds then to compute the between-class scatter matrix as follows:

$$\Gamma^{-1/2} \Xi^t \Sigma_b \Xi \Gamma^{-1/2} = K_b \tag{7}$$

Diagonalize now the new between-class scatter matrix K_b:

$$K_b \Theta = \Theta \Delta \quad and \quad \Theta^t \Theta = I \tag{8}$$

where $\Theta, \Delta \in \mathbb{R}^{m \times m}$ are the eigenvector and the diagonal eigenvalue matrices of K_b.

The overall transformation matrix of the EFM method is now defined as follows:

$$T = \Xi \Gamma^{-1/2} \Theta \tag{9}$$

3.2. Similarity Measures and Classification Rule for the Gabor Feature Classifier (GFC)

Let the overall transformation matrix be T, as defined be Eq. 9. The new feature vector, $\mathcal{U}^{(\rho)}$, of an image is defined as follows:

$$\mathcal{U}^{(\rho)} = T^t \mathcal{Y}^{(\rho)} \tag{10}$$

Let $\mathcal{M}_k^0, k = 1, 2, \ldots, L$, be the mean of the training samples for class ω_k after the overall transformation (Eq. 9). The GFC method employs a nearest neighbour (to the mean) classification rule using some similarity (distance) measure δ:

$$\delta(\mathcal{U}^{(\rho)}, \mathcal{M}_k^0) = \min_j \delta(\mathcal{U}^{(\rho)}, \mathcal{M}_j^0) \quad \longrightarrow \quad \mathcal{U}^{(\rho)} \in \omega_k \tag{11}$$

The image feature vector, $\mathcal{U}^{(\rho)}$, is classified as belonging to the class of the closest mean, $\mathcal{M}_k^0$, using the similarity measure δ.

The similarity measures used in our experiments to evaluate the efficiency of different representation and recognition methods include L_1 distance measure, δ_{L_1}, L_2 distance measure, δ_{L_2}, Mahalanobis distance measure, δ_{Md}, and cosine similarity measure, δ_{cos}, which are defined as follows:

$$\delta_{L_1}(\mathcal{X}, \mathcal{Y}) = \sum_i |\mathcal{X}_i - \mathcal{Y}_i| \tag{12}$$

Figure 3. Example FERET images used in our experiments (cropped to the size of 128 × 128 to extract the facial region). The figure shows in the top two rows the examples of training images used in our experiments, and in the bottom row the examples of test images.

$$\delta_{L_2}(\mathcal{X},\mathcal{Y}) = (\mathcal{X}-\mathcal{Y})^t(\mathcal{X}-\mathcal{Y}) \tag{13}$$

$$\delta_{Md}(\mathcal{X},\mathcal{Y}) = (\mathcal{X}-\mathcal{Y})^t\Sigma^{-1}(\mathcal{X}-\mathcal{Y}) \tag{14}$$

$$\delta_{cos}(\mathcal{X},\mathcal{Y}) = \frac{-\mathcal{X}^t\mathcal{Y}}{\|\mathcal{X}\|\,\|\mathcal{Y}\|} \tag{15}$$

where Σ is the covariance matrix, and $\|\cdot\|$ denotes the norm operator. Note that the cosine similarity measure includes a minus sign in Eq. 15, because the nearest neighbour (to the mean) rule of Eq. 11 applies minimum (distance) measure rather than maximum similarity measure.

4. Experiments

We assessed the feasibility and performance of our novel Gabor Feature Classifier (GFC) method on the face recognition task, using 600 face images corresponding to 200 subjects from the FERET database, which has become a standard testbed for face recognition technologies. Each subject has three images of size 256 × 384 with 256 gray scale levels. Fig. 3 shows some example images used in our experiments that are already cropped to the size of 128 × 128 in order to extract the facial region. Note that the images are acquired during different photo sessions, they capture both different lighting conditions and facial expressions. Two images are randomly chosen from the three images available for each subject for training, while the remaining image (unseen during training) is used for testing.

For comparison purpose, we first implemented the Eigenfaces method [19] on the original images as shown in Fig. 3, using the four different similarity measures introduced in Sect. 3.2. Fig. 4 shows that the Mahalanobis

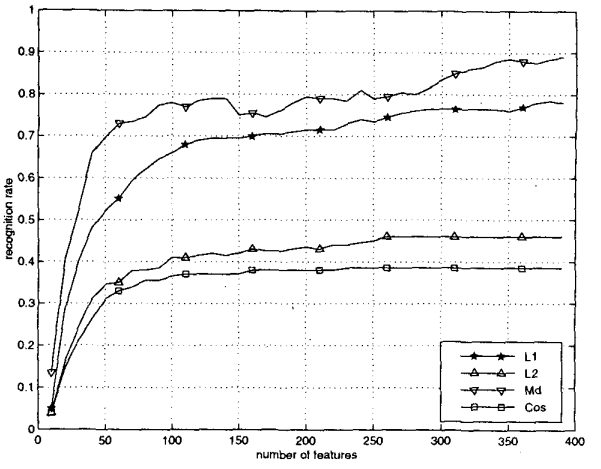


Figure 4. Face recognition performance of the Eigenfaces method on the original images as shown in Fig. 3, using four different similarity measures: L1 (L_1 distance measure), L2 (L_2 distance measure), Md (Mahalanobis distance measure), and Cos (cosine similarity measure).

distance measure performs better than the L_1 distance measure, which is followed in order by the L_2 distance measure and the cosine similarity measure. The reason that the Mahalanobis distance measure performs the best is that it counteracts the fact that L_1 and L_2 distance measures in the PCA space weight preferentially for low frequencies. Such behavior should be expected to be even more pronounced when faces are aligned and cropped, as the first several leading eigenvalues encode mostly for prototypical (norm) representational aspects rather than discrimination information. As the L_2 measure weights more the low frequencies than L_1 does, the L_1 distance measure should perform better than the L_2 distance measure, a conjecture validated by our experiments. The cosine similarity measure does not compensate the low frequency preference, and it performs the worst among all the measures. Actually, the superiority of the cosine similarity measure to the others can be revealed only when the discriminating features (derived by GFC) rather than the most expressive features (derived by PCA) are used for classification (see Fig. 6).

The next series of experiments used the Gabor convolution outputs, $O_{\mu,\nu}(z)$, derived in Sect. 2.2, and the L_1, L_2 and cosine similarity measures, respectively. (The Mahalanobis metric is not used here because it employs transformed data suitable for PCA-like schemes. The L_1, L_2 and cosine metrics are compared at different downsampling rates without further data transformations.) For the first

Table 1. Face recognition performance using the Gabor wavelet transformation with the three different similarity measures, respectively: L1 (L_1 distance measure), L2 (L_2 distance measure), and Cos (cosine similarity measure). $G_{\mu,\nu}^{(16)}$ represents the Gabor transformation downsampled by a factor of 16 and normalized to unit length, as suggested by [7]. $\mathcal{X}^{(4)}$, $\mathcal{X}^{(16)}$, and $\mathcal{X}^{(64)}$ represent the augmented Gabor feature vector $\mathcal{X}^{(\rho)}$ as defined by Eq. 4 using three different downsampling factors, respectively: ρ = 4, 16, or 64.

measure \representation	$G_{\mu,\nu}^{(16)}$	$\mathcal{X}^{(4)}$	$\mathcal{X}^{(16)}$	$\mathcal{X}^{(64)}$
L_1 distance measure	76%	76.5%	76.5%	76.5%
L_2 distance measure	73.5%	72%	72%	72%
cosine similarity measure	72%	70.5%	70.5%	70%

set of experiments, we downsampled the Gabor convolution outputs by a factor 16 to reduce the dimensionality and normalized them to unit length, $G_{\mu,\nu}^{(16)}$, [7]. The classification performance using such Gabor outputs is shown in Table 1. The best performance is achieved using the L_1 similarity measure. Comparing Table 1 with Fig. 4, we found that (i) under the L_2 and cosine similarity measures, the Gabor features carry more discriminating information than the PCA features do, a finding consistent with that reported by Donato et al. [7] on facial action classification; and (ii) the performance with the three similarity measures, L_1, L_2 and cosine, varies less drastically than that shown in Fig. 4. The second finding indicates that the Gabor feature representation is less likely affected by preferential low frequency weighting, which qualifies the Gabor feature representation as a discriminant representation method rather than an expressive representation method. We have also experimented on the augmented Gabor feature vector $\mathcal{X}^{(\rho)}$ as defined by Eq. 4 with three different downsampling factors: ρ = 4, 16, and 64, respectively. From the classification performance as shown in Table 1, we found that (i) the augmented Gabor feature vector $\mathcal{X}^{(\rho)}$ carries quite similar discriminating information to the one used by Donato et al. [7]; and (ii) the performance differences using the three different downsampling factors are not significant. As a result, we choose the downsampling factor 64 for the next series of experiments, since it reduces to a larger extent the dimensionality of the vector space than the other two factors do. (We experimented with other downsampling factors as well. When the downsampling factors are 256 and 1024, the performance is marginally less effective; when the downsampling factor is 4096, the recognition rate drops drastically.)

Even though the performance shown in Table 1 indicates that the Gabor based feature representation carries discriminating rather than expressive information, it is still not convenient to use directly such a representation for classification, since the dimensionality of the augmented Gabor feature vector space is very high. To reduce the dimensionality of the vector space, we applied PCA on the augmented Gabor feature vector $\mathcal{X}^{(\rho)}$, where the downsampling factor ρ is set to be 64. Fig. 5 shows the face recognition performance of PCA using the augmented Gabor feature vector $\mathcal{X}^{(\rho)}$. Our results indicate that (i) compared to Fig. 4, the recognition performance improves by a large margin for all the similarity measures, and this shows that the augmented Gabor feature vector carries more discriminating information than the original images do; and (ii) Mahalanobis and L_1 distance measures perform better than the other two similarity measures — an inherent disadvantage of the PCA method.

Our last series of experiments, performed using the novel Gabor Feature Classifier (GFC) method described in this paper, show that the GFC derives discriminating Gabor features with low dimensionality and enhanced discrimination power. Fig. 6 shows face recognition performance of GFC on the augmented Gabor feature vector $\mathcal{X}^{(\rho)}$ using L_1, L_2 and cosine similarity measures, respectively. Again, the downsampling factor ρ is set to be 64. The features derived by GFC do not preferentially weight low frequencies, since the GFC involves a whitening operation during the simultaneous diagonalization of the two within- and between-class scatter matrices. As a result, the superiority of the L_1 distance measure to the L_2 distance measure displayed by the PCA method as shown in Fig. 3 and Fig. 5 should not exist for the GFC method, a conjecture validated by our experiments. Fig. 6 shows that both the L_1 and the L_2 distance measures display similar recognition performance. The cosine similarity measure, however, performs the best, with 100% correct recognition rate using only 62 features (note that the curves in Fig. 6 were drawn with an interval resolution of 5 features, and it shows that 100% correct recognition rate happens when 65 features are used).

References

[1] P. Belhumeur, J. Hespanha, and D. Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear pro-

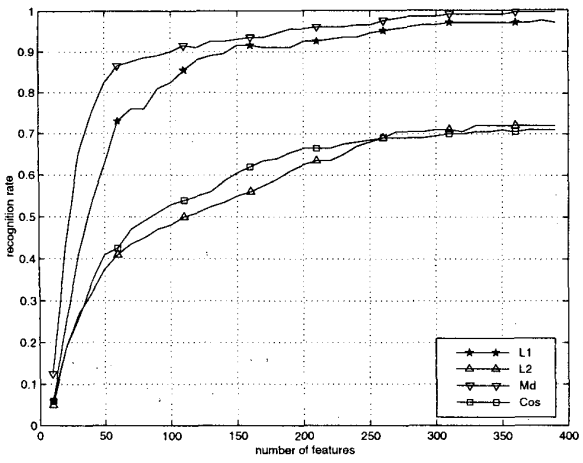


Figure 5. Face recognition performance of PCA on the augmented Gabor feature vector $\mathcal{X}^{(\rho)}$, ρ=64, using four different similarity measures: L1 (L_1 distance measure), L2 (L_2 distance measure), Md (Mahalanobis distance measure), and Cos (cosine similarity measure).

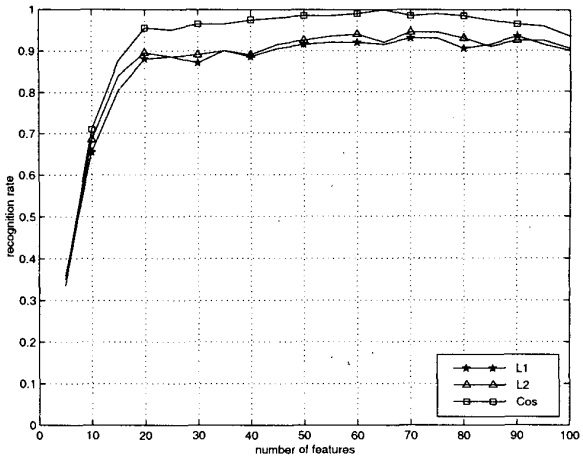


Figure 6. Face recognition performance of Gabor Feature Classifier (GFC) on the augmented Gabor feature vector $\mathcal{X}^{(\rho)}$, ρ=64, using three different similarity measures: L1 (L_1 distance measure), L2 (L_2 distance measure), and Cos (cosine similarity measure).

jection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(7):711–720, 1997.

[2] D. Burr, M. Morrone, and D. Spinelli. Evidence for edge and bar detectors in human vision. *Vision Research*, 29(4):419–431, 1989.

[3] R. Chellappa, C. Wilson, and S. Sirohey. Human and machine recognition of faces: A survey. *Proc. IEEE*, 83(5):705–740, 1995.

[4] J. Daugman. Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research*, 20:847–856, 1980.

[5] J. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional cortical filters. *Journal Opt. Soc. Amer.*, 2(7):1160–1169, 1985.

[6] J. Daugman. Face and gesture recognition: Overview. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(7):675–676, 1997.

[7] G. Donato, M. Bartlett, J. Hager, P. Ekman, and T. Sejnowski. Classifying facial actions. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21(10):974–989, 1999.

[8] K. Etemad and R. Chellappa. Discriminant analysis for recognition of human face images. *J. Opt. Soc. Am. A*, 14:1724–1733, 1997.

[9] D. Field. Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Amer. A*, 4(12):2379–2394, 1987.

[10] J. Jones and L. Palmer. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *J. Neurophysiology*, pages 1233–1258, 1987.

[11] M. Lades, J. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, W. R.P., and W. Konen. Distortion invariant object recognition in the dynamic link architecture. *IEEE Trans. Computers*, 42:300–311, 1993.

[12] C. Liu and H. Wechsler. Evolutionary pursuit and its application to face recognition. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(6):570–582, 2000.

[13] C. Liu and H. Wechsler. Robust coding schemes for indexing and retrieval from large face databases. *IEEE Trans. on Image Processing*, 9(1):132–137, 2000.

[14] S. Marcelja. Mathematical description of the responses of simple cortical cells. *Journal Opt. Soc. Amer.*, 70:1297–1300, 1980.

[15] B. Moghaddam and A. Pentland. Probabilistic visual learning for object representation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(7):696–710, 1997.

[16] A. Pentland. Looking at people: Sensing for ubiquitous and wearable computing. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(1):107–119, 2000.

[17] A. Samal and P. Iyengar. Automatic recognition and analysis of human faces and facial expression: A survey. *Pattern Recognition*, 25(1):65–77, 1992.

[18] D. Swets and J. Weng. Using discriminant eigenfeatures for image retrieval. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 18(8):831–836, 1996.

[19] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 13(1):71–86, 1991.

[20] L. Wiskott, J. Fellous, N. Kruger, and C. von der Malsburg. Face recognition by elastic bunch graph matching. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(7):775–779, 1997.