Towards Robust Face Recognition from Video

Jeffery R. Price *Timothy F. Gee*
Image Science and Machine Vision Group
Oak Ridge National Laboratory
Oak Ridge, TN 37831-6010
{pricejr, geetf}@ornl.gov

Abstract

A novel, template-based method for face recognition is presented. The goals of the proposed method are to integrate multiple observations for improved robustness and to provide auxiliary confidence data for subsequent use in an automated video surveillance system. The proposed framework consists of a parallel system of classifiers, referred to as observers, where each observer is trained on one face region. The observer outputs are combined to yield the final recognition result. Three of the four confounding factors – expression, illumination, and decoration – are specifically addressed in this paper. The extension of the proposed approach to address the fourth confounding factor – pose – is straightforward and well supported in previous work. A further contribution of the proposed approach is the computation of a revealing confidence measure. This confidence measure will aid the subsequent application of the proposed method to video surveillance scenarios. Results are reported for a database comprising 676 images of 160 subjects under a variety of challenging circumstances. These results indicate significant performance improvements over previous methods and demonstrate the usefulness of the confidence data.

1. Introduction

Despite the availability of commercial systems, face recognition continues to be an active topic in ongoing computer vision research, indicating that there exists significant room for improvement. Current face recognition systems perform quite well under good circumstances, but tend to suffer when variations in expression, illumination, decoration (i.e., glasses, facial hair), or pose are present. Most current face recognition research – including the work presented here – aims to improve recognition performance in the presence of such confounding factors. An additional goal of the work presented herein is to develop a recognition engine to serve as the foundation for a broader video surveillance system, primarily targeted at the secure facility monitoring scenario for personnel tracking and monitoring. Used in the facility monitoring scenario, however, such a system could also easily provide further capabilities such as an additional safeguard for access control and/or threat detection. We note that this paper does not address face detection [17], face landmark (eyes, nose, etc.) location, or face tracking [5, 13].

Although much of the current research is focused upon improving performance in the presence of confounding factors, current algorithms address no more than two of the four of the confounding factors – illumination, expression, decoration, and pose – simultaneously. Below we briefly survey some recent work aimed at addressing one or more of these four factors.

Illumination. Approaches for dealing with varying illumination are primarily based upon linear discriminant analysis (LDA), sometimes referred to as "Fisherfaces" [3, 29, 30]. A motivating principle behind these techniques is the approximation of a face as a Lambertian surface. As noted in [3], the images of a Lambertian surface under varying illumination lie in a three-dimensional (3D), linear subspace of the entire image space. Hence, the different face surfaces should be linearly separable using LDA.

Expression. Varying facial expression can be modeled to some degree by the active appearance models (AAMs) presented in [9]. AAMs characterize shape and texture information using a statistical point distribution approach.

Illumination and expression. Bayesian face recognition [21, 22, 23] has been proposed to improve robustness in the presence of varying illumination and expression. These approaches employ probabilistic models to characterize *intrapersonal* and *interpersonal* differences with a principal component analysis (PCA) or "eigenface" representation. In [21], it is noted that the Bayesian approach can be thought of as a general, nonlinear extension of LDA. With this in mind, it seems that LDA should also handle both illumination and expression to some degree.

U.S. Government Work Not Protected by U.S. Copyright

Facial decoration. There has been very little work towards explicitly handling facial decoration. In [23] and [24], it was shown that two "eigenfeature" images – the eyes and the nose – could be used for accurate recognition after a change in facial hair. However, no method for selecting the appropriate eigenfeatures was suggested. In the LDA approach described in [29], some promising results were obtained after artificially degrading face images, indicating that LDA might also provide a reasonable solution to handling some degree of decoration, assuming the registration landmarks can still be located.

Pose. One method to handle varying pose is the view-based eigenspace approach [24], which was recently shown to perform quite well [16]. Each pose is represented by its own subspace and the multiple subspaces act as independently trained "experts" or observers trying to explain the data. Similarly motivated techniques include characteristic eigenspace curves [15] and view-based active appearance models [6].

Pose and expression. Recently, new AAM methods [6, 7] have been proposed to handle both varying pose and expression. No recognition results, however, are presented to demonstrate performance.

Pose and illumination. Methods were presented in [14] and [31] to deal with varying pose and illumination. These methods rely upon generative models that can synthesize a given face under varying illumination from different viewpoints. Although the performance in [14] is quite remarkable, the proposed method employs seven training images for each subject under strictly controlled lighting and does not address expression or decoration.

Video. Although there has been some previous work in video-based face recognition [10, 11, 16, 20], we are unaware of any proposed methods that address all four of the confounding factors. In fact, [20] addresses none, [16] tackles only pose, and [10] and [11] address only pose and expression.

When we intuitively consider how humans recognize faces visually, we note that humans integrate information, often obtained at different times, by analyzing multiple observations of different face regions and discounting confounding factors when present. For example, a new beard might be confusing, but identity can still be revealed by the eyes and nose region. Or perhaps strange lighting can initially cause confusion, but accurate recognition is achieved as the subject moves through different lighting conditions. This naturally leads us to consider using multiple observations, such as those available from video sequences and/or multiple cameras, and integrating the results. It is also evident that no single algorithm in the above survey addresses more than two confounding factors at once. With these thoughts in mind, we now note the specific contributions of this paper. A face recognition algorithm is proposed that specifically addresses three of the four confounding factors – illumination, expression, and decoration; extending the approach to address pose is discussed in Section 3. Furthermore, the proposed approach provides a revealing confidence measure that will aid in the application of the algorithm to video surveillance scenarios.

The remainder of this paper is organized as follows. In Section 2, we review PCA and LDA methods for dimensionality reduction. Both PCA and LDA are used in our approach and are quite common in template-based face recognition in general. In Section 3, we describe our recognition framework. We discuss the results of some experiments in Section 4 and make some concluding remarks in Section 5.

2. Dimensionality Reduction

Template-based face recognition algorithms essentially use the actual image pixels (after geometry and grayscale normalization) as features. Without dimensionality reduction, however, the size of such a feature vector would be quite impractical. For example, the full frontal face images we employ are 92×115 pixels, which would imply a feature vector of dimension 10,580. Dimensionality reduction is employed to reduce such feature vectors to a more tractable size. There are two types of optimal dimensionality reduction that are widely used in face recognition: PCA [21, 26, 28] and LDA [3, 27, 29]. The methods that use LDA – including ours – generally employ PCA as an initial step. In this section, we briefly review the fundamentals of PCA and LDA. We begin with a set of N images, represented as vectors $\{x_1, x_2, \ldots, x_N\}$ with n pixels each, i.e., $x_k \in \mathbb{R}^n$. Each image belongs to one of C classes denoted by $\{\omega_1, \omega_2, \ldots, \omega_C\}$.

2.1. Principal Component Analysis (PCA)

PCA, which is the basis of the well known eigenface techniques, seeks a linear projection to maximize the scatter of all the sample vectors in the reduced dimensionality space. Each of the sample vectors is to be projected onto a p-dimensional subspace, where $p < n$, by

$$\mathbf{y}_k = \mathrm{P}^T \mathbf{x}_k, \tag{1}$$

where $\mathrm{P} \in \mathbb{R}^{n \times p}$. The total scatter (i.e., covariance) matrix in the original space is given by

$$\mathrm{S}_t = \sum_{k=1}^{N} (\mathbf{x}_k - \boldsymbol{\mu})(\mathbf{x}_k - \boldsymbol{\mu})^T. \tag{2}$$

In the reduced space $\mathbb{R}^p$, the total scatter matrix is simply $\mathrm{P}^T \mathrm{S}_t \mathrm{P}$. PCA seeks the linear transformation P_o to max-

imize the determinant of the total scatter matrix in the reduced space:

$$\mathrm{P}_o = \arg\max_{\mathrm{P}} |\mathrm{P}^T \mathrm{S}_t \mathrm{P}|. \tag{3}$$

The well known solution to this problem is the matrix $\mathrm{P}_o \in \mathbb{R}^{n \times p}$ whose columns are the p eigenvectors of S_t with the p largest eigenvalues. Note that S_t is an $n \times n$ matrix, where n is the number of image pixels which can be quite large (10,580 for our full frontal face images, for example). The direct computation of the eigenvectors and eigenvalues for such a large S_t is nontrivial. In the face recognition problem, however, the number of images N is generally much less than the dimensionality of original space n. In this small sample size scenario, S_t is not full rank and its eigenvectors and eigenvalues can actually be found from a much more tractable $N \times N$ matrix [12].

The drawback to using PCA for dimensionality reduction is that class membership is not considered. The projection is chosen to capture the most variation in the samples, even if that variation does not discriminate between different classes. This is illustrated for a two-dimensional, two class problem in Fig. 1. PCA selects the line along which the scatter of all the projected samples is maximum, even though the two classes will overlap along that line. As described in [27], PCA captures the *most expressive features.*

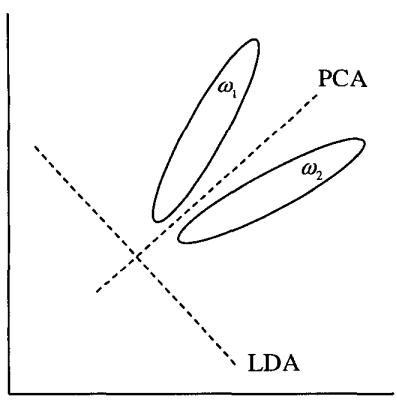


Figure 1. Reduction to one dimension for a two-dimensional, two class problem. Projection onto the line selected by PCA results in completely overlapping classes, even though most of the overall sample variation is retained. Projection onto the line found by LDA, however, preserves the class separation.

2.2. Linear Discriminant Analysis (LDA)

Assume that each sample $\mathbf{y}_k$ is now a p-vector and to be projected onto a d-dimensional subspace, where $d < p$ by

$$\mathbf{z}_k = \mathrm{D}^T \mathbf{y}_k, \tag{4}$$

where $\mathrm{D} \in \mathbb{R}^{p \times d}$ is the projection matrix being sought. In LDA, an effort is made to retain the *discriminative* features, rather than the *expressive* features, by defining within-class and between-class scatter matrices. The within-class scatter matrix, S_w, is given by

$$\mathrm{S}_w = \sum_{i=1}^{C} \sum_{\mathbf{x}_k \in \omega_i} (\mathbf{x}_k - \boldsymbol{\mu}_i)(\mathbf{x}_k - \boldsymbol{\mu}_i)^T, \tag{5}$$

where $\boldsymbol{\mu}_i$ is the sample mean for class ω_i. The between-class scatter matrix, S_b, is defined by

$$\mathrm{S}_b = \sum_{i=1}^{C} (\boldsymbol{\mu}_i - \boldsymbol{\mu})(\boldsymbol{\mu}_i - \boldsymbol{\mu})^T, \tag{6}$$

where $\boldsymbol{\mu}$ represents overall sample mean. The within-class scatter matrix in the reduced space is given by $\mathrm{D}^T \mathrm{S}_w \mathrm{D}$ while the between-class scatter matrix in the reduced space is $\mathrm{D}^T \mathrm{S}_b \mathrm{D}$. LDA seeks the projection matrix D_o to maximize the ratio of between-class scatter to within-class scatter in the reduced space:

$$\mathrm{D}_o = \arg\max_{\mathrm{D}} \frac{|\mathrm{D}^T \mathrm{S}_b \mathrm{D}|}{|\mathrm{D}^T \mathrm{S}_w \mathrm{D}|}. \tag{7}$$

The solution to this problem [8, 12] is the matrix $\mathrm{D}_o \in \mathbb{R}^{p \times d}$ whose columns are the d eigenvectors of $\mathrm{S}_w^{-1} \mathrm{S}_b$ with the d largest eigenvalues. Note that the matrix S_b defined in (6) is of rank $C - 1$ or less. This is evident by noting that S_b is the sum of C rank one vector outer products. Only $C - 1$ of these vectors are linearly independent, however, since $\boldsymbol{\mu}$ is a linear combination of the $\boldsymbol{\mu}_i$'s.

Methods that employ LDA generally employ PCA as an initial step to reduce the dimensionality of the LDA problem to a more tractable size. First, the optimal PCA matrix $\mathrm{P}_o \in \mathbb{R}^{n \times p}$ is computed from the original sample vectors $\{x_1, x_2, \ldots, x_N\}$. These sample vectors are then projected onto $\mathbb{R}^p$, as in (1), using P_o. LDA is then used to find $\mathrm{D}_o \in \mathbb{R}^{p \times d}$ from the reduced dimension vectors $\{y_1, y_2, \ldots, y_N\}$. The PCA reduced p-vectors are then projected onto $\mathbb{R}^d$, as in (4), using D_o. Recognition is performed in $\mathbb{R}^d$. When a new image is to be classified, it can, of course, be projected onto $\mathbb{R}^d$ directly using the single $n \times d$ projection matrix W defined by

$$\mathrm{W} = \mathrm{P}_o \mathrm{D}_o. \tag{8}$$

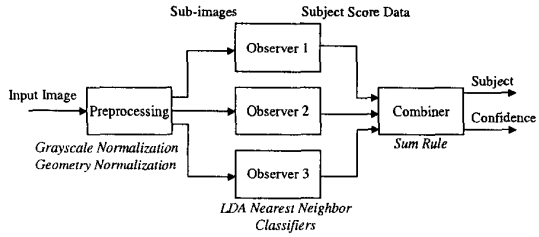


Figure 2. Example of the proposed framework with three observers.

3. Multiple Observers

Perhaps the earliest suggestions for the use of multiple observers can be found in [4] and also in the view-based and modular eigenspaces of [24]. It is noted in [24] that recognition from a full frontal face image is sensitive to changes in expression and decoration. By breaking the full face image into modular subregions, it is shown that improved accuracy can be obtained in the presence of expression and decoration variation. For example, a change in facial hair does not degrade the performance of the eye region classifier or perhaps the mouth region can be discarded in the presence of confounding facial expression. No method is provided, however, for fusing the information from the different experts nor is illumination addressed. An extension to the view-based eigenspaces of [24], along with more comprehensive experimental data showing much promise, is found in [16]. We note that the modular eigenspaces suggested in [24] are the primary motivation for our different LDA observers. Furthermore, we note that the view-based classifiers of [16] are combined – using a neural network – in a manner similar to how our LDA observers are combined.

Through the survey of previous work in Section 1 and the previous multiple observer efforts discussed above, we propose a face recognition framework to provide improved robustness to three of the four confounding factors – illumination, expression, and decoration – simultaneously. This framework will employ a parallel system of observers, each of which is trained on a specific region of the face from a specific viewpoint. Each such observer will be an LDA-based classifier and the outputs of all the observers will be combined using a simple classifier combination method [18]. An example of such a system with three observers is shown in Fig. 2. The multiple observers will provide some robustness to decoration, as will the use of LDA. Illumination will be handled by LDA. Expression will be addressed through the combination of multiple observers and LDA – some observers will be approximately invariant to expression, while LDA will discount expression variability, as it is not discriminatory, when constructing the projection matrices. Finally, additional robustness will be provided by integrating the responses of the observers to obtain a final classification. Based upon the success of previous work in [24] and particularly [16], we hypothesize that the extension of our approach to handle varying pose is quite straightforward, requiring only additional observers for additional viewpoints. In the following subsections, we describe the components of our current system in more detail.

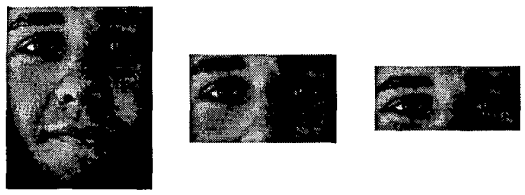
3.1. Preprocessing

The preprocessing stage indicated in Fig. 2 assumes a frontal face image as input with previously labeled landmarks. In our implementation, the centers of the eyes, the tip of the nose, and the center of the mouth are used. Landmarks can be located automatically by various methods, but have been found manually for our experiments. In the preprocessing stage, the full frontal image is warped to align landmarks according to each observer. Observer 1 employs the full frontal image and affine warping is used to align the center of each eye and the center of the mouth to predefined model. Observer 2 uses the eyes and nose region of the face; affine warping is applied to align the center of each eye and the tip of the nose to a fixed model. Observer 3 uses only the eyes region and linear conformal warping is used to align the center of each eye to a model. After this geometry normalization, a mask is applied to remove background information and then each image is normalized to have zero-mean and unit-variance to account for gross grayscale variation. The resulting images are raster scanned to create the feature vectors for input into the observers.

3.2. Observers

Observer 1 works with 92×115 pixel full frontal face images implying an original feature vector length of $n_1 = 10,580$. Observer 2 uses 92×56 pixel images of the eyes and nose face region, implying an original feature vector length of $n_2 = 5,152$. Observer 3 works with 92×40 pixel images of the eyes region, implying an original feature vector length of $n_3 = 3,680$. Examples of images corresponding to each observer are shown in Fig. 3. Each observer uses PCA to reduce the original dimensionality of n_o for $o = \{1,2,3\}$ to $p = 120$. LDA is subsequently applied to reduce dimensionality from $p = 120$ to $d = 50$. Note that during recognition, the reduction to dimension $d = 50$ is carried out in a single step by the projection matrix W given by (8).

For classification purposes, each observer compares each query image to every subject in its own database. For query image q and each observer o, a score α^o_{qs} is computed for each subject s and passed on to the combiner. The score for



Observer 1 Observer 2 Observer 3

Figure 3. Examples of the face image regions used by the three observers.

observer o is given by

$$\alpha_{qs}^{o} = \left(F_q^o d_{qs}^o\right)^{-1} \tag{9}$$

where d_{qs}^o is the Euclidean distance between the query image q and its nearest neighbor in subject class s and the normalization factor F_q^o is given by

$$F_q^o = \sum_{s=1}^{S} (d_{qs}^o)^{-1}, \tag{10}$$

where S is the number of subjects, so that

$$\sum_{s=1}^{S} \alpha_{qs}^{o} = 1 \tag{11}$$

With these definitions, the score α_{qs}^o can be thought of as an approximation to the probability that the query image q belongs to subject class s in the space seen by observer o. Without combining the observers, nearest neighbor classification is implemented for each observer individually by simply selecting the subject class with the highest score.

3.3. Observer Combination

For each query image q, the score data for each subject s from each observer o is passed to the combiner. Since the score data has been normalized so that it approximates a probability, a number of simple classifier combination strategies [18] can be employed. In our implementation, we employ the sum rule. As noted in [18], the assumptions required to make the sum rule optimal are quite restrictive. Despite this fact, the sum rule was reported to be the best performing in [18] and also found to be so in our own experiments. This is attributed in [18] to the sum rule being more resilient to estimation errors. According to the sum rule, a combined score c_{qs} for each subject s is computed by simply adding the scores for subject s reported by each observer:

$$c_{qs} = \sum_{o=1}^{3} \alpha_{qs}^{o}. \tag{12}$$

The subject with the highest combined score is then selected as the classification result for the query q.

In our implementation, we compute a confidence measure m_q as an additional combiner output. For a given query image q, let the best combined score be denoted c_{qs_1} and the next best combined score be denoted c_{qs_2}. The confidence measure m_q is computed as the logarithm of the ratio of the best score to the next best score:

$$m_q = \log\left(\frac{c_{qs_1}}{c_{qs_2}}\right). \tag{13}$$

4. Experimental Results

We report here on the performance of the proposed algorithm on a challenging face image database. The database we constructed for our tests contained 676 images of 160 subjects from four different databases [1, 2, 19, 25] with significant confounding factors including varying illumination, expression, and decoration (primarily eyeglass changes).

A series of 100 training and classification runs were performed. In each such run, 400 images were randomly selected for training and the remaining 276 were used for testing. The ratio of training images to testing images was actually set to 0.50 for each subject, but we had an odd number of images for the majority of our database subjects and rounding up to the nearest integer gave the actual ratio of about 0.59 (400/676) for training. The observers were tested both individually and combined using PCA and LDA. As the LDA spaces were of dimension 50, only 50 PCA components, excluding the first three, were retained when PCA-based classification was tested. The first three PCA components were excluded because it has been shown previously [3] that these often capture much of the variation due to illumination and/or other non-discriminative information. As in [3], we found that excluding the first three PCA components provided modest performance improvements. Recognition performance is reported in Table 1 for both LDA and PCA for each observer individually as well as for each of the four possible combination scenarios. Examining the results 1, we note that the {1,2,3} and {1,3} LDA combinations achieve the best overall rate of 94.2%, which is a 3.3% improvement over the best individual observer rate of 90.9%. This 3.3% improvement equates to 910 additional correct classifications of the 27,600 queries in the 100 separate runs. Also note that Observer 3 – the eyes region – provides the best individual performance for both LDA and PCA. Finally, we note the significant performance improvements of LDA over PCA in each case; the low PCA performances indicate the challenging nature of our database.

We now turn our attention briefly to the confidence measure discussed in Section 3.3 using our combined LDA system. Referring to Fig. 4, we note the striking difference in

Observers	PCA	LDA
1	68.5	90.5
2	72.0	85.7
3	80.9	90.9
{1,2,3}	83.1	**94.2**
{1,2}	78.1	92.5
{1,3}	83.8	94.2
{2,3}	82.1	92.2

Table 1. Recognition performance reported as correct classification percentage. The bottom four entries employ classifier combination.

the confidence measure between the successes and failures. Using this difference, we can actually reject decisions with low confidence. The result of using the confidence ratio to reject decisions is illustrated in Fig. 5 where the correct classification rate is plotted against the ratio of rejected images for both PCA and LDA (all three observers were combined). Note that with our parallel LDA observers, we can achieve 100% correct identification of the accepted images by rejecting only about 15% of the images. Using PCA, about 45% of the images would be rejected to achieve 100% correct classification on the remaining images.

5. Conclusions

In this paper, we have proposed a face recognition algorithm that addresses three of the four confounding factors – illumination, expression, and decoration – simultaneously. We additionally note how our approach can easily be extended to work in the presence of varying pose. The method proposed employs a parallel system of LDA-based classifiers, referred to as observers, followed by a simple classifier combination scheme. Results indicate significantly improved accuracy over other methods. A confidence measure is also reported and shown to be a strong indicator of correct classification versus incorrect classification.

Acknowledgements

Research sponsored by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory, managed by UT-Battelle, LLC, for the U.S. Department of Energy under Contract No. DE-AC05-00OR22725. Portions of the research in this paper use the FERET database of facial images collected under the FERET program.

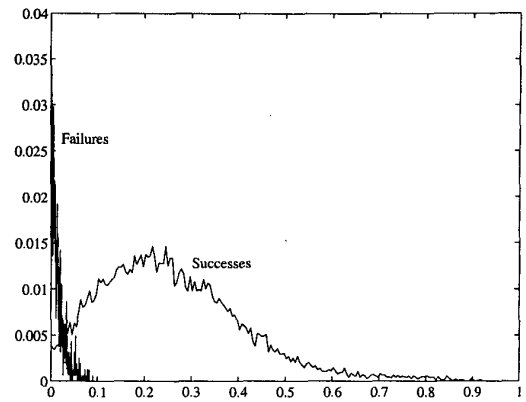


Figure 4. Normalized sample histograms of the confidence measure defined in (13) for successes and failures.

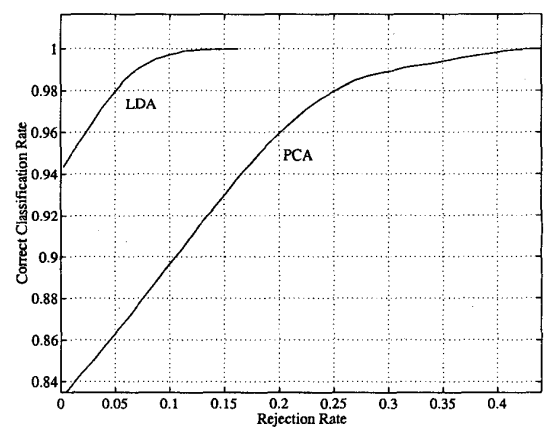


Figure 5. Plot of correct classification rate versus rate of images rejected due to low confidence. These curves were computed using the combination of all three observers.

References

[1] The CVL Face Database. www.lrv.fri.uni-lj.si.

[2] The Yale Face Database. cvc.yale.edu.

[3] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):711–720, July 1997.

[4] R. Brunelli and T. Poggio. Face recognition: Features versus templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(10):1042–1052, October 1993.

[5] M. L. Cascia, S. Sclaroff, and V. Athitsos. Fast, reliable head tracking under varying illumination: An approach based on registration of texture-mapped 3d models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(4):322–336, April 2000.

[6] T. Cootes, K. Walker, and C. Taylor. View-based active appearance models. In *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, pages 227–232, 2000.

[7] T. Cootes, G. Wheeler, K. Walker, and C. Taylor. Coupled-view active appearance models. In *Proceedings of the British Machine Vision Conference*, volume 1, pages 52–61, 2000.

[8] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley-Interscience, second edition, 2001.

[9] G. Edwards, T. Cootes, and C. Taylor. Face recognition using active appearance models. In *Proceedings of the European Conference on Computer Vision*, volume 2, pages 581–595, 1998.

[10] G. Edwards, C. Taylor, and T. Cootes. Learning to identify and track faces in image sequences. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 317–322, 1998.

[11] G. Edwards, C. Taylor, and T. Cootes. Improving identification performance by integrating evidence from sequences. In *Proceedings of the IEEE Computer Society Conference on Pattern Recognition*, pages 486–491, 1999.

[12] K. Fukunaga. *Statistical Pattern Recognition*. Morgan Kaufmann, 1990.

[13] T. F. Gee and R. Mersereau. Model-based face tracking for dense motion field estimation. In *Proceedings of the Applied Imagery and Pattern Recognition Workshop*, 2001.

[14] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):643–660, June 2001.

[15] D. B. Graham and N. M. Allinson. Face recognition from unfamiliar views: Subspace methods and pose dependency. In *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, pages 348–353, 1998.

[16] F. J. Huang, Z. Zhou, H.-J. Zhang, and T. Chen. Pose invariant face recognition. In *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, pages 245–250, 2000.

[17] D. Keren, M. Osadchy, and C. Gotsman. Antifaces: A novel, fast method for image detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(7):747–761, July 2001.

[18] J. Kittler, M. Hatef, R. P. Duin, and J. Matas. On combining classifiers. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3):226–239, March 1998.

[19] A. Martinez and R. Benavente. The AR face database. CVC Technical Report 24, Purdue University, June 1998.

[20] S. McKenna, S. Gong, and Y. Raja. Face recognition in dynamic scenes. In *Proceedings of the British Machine Vision Conference*, 1997.

[21] B. Moghaddam, T. Jebara, and A. Pentland. Bayesian face recognition. *Pattern Recognition*, 33(11):1771–1782, November 2000.

[22] B. Moghaddam, C. Nastar, and A. Pentland. Bayesian face recognition using deformable intensity surfaces. In *Proceedings of the IEEE Computer Society Conference on Pattern Recognition*, pages 638–645, 1996.

[23] B. Moghaddam and A. Pentland. Probabilistic visual learning for object representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):696–710, July 1997. face recognition.

[24] A. Pentland, B. Moghaddam, and T. Starner. View-based and modular eigenspaces for face recognition. In *Proceedings of the IEEE Computer Society Conference on Pattern Recognition*, pages 84–91, 1994.

[25] P. Phillips, H. Wechsler, J. Huang, and P. Rauss. The FERET database and evaluation procedure for face recognition algorithms. *Image and Vision Computing*, 16(5):295–306, 1998.

[26] L. Sirovich and M. Kirby. Low-dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America A (Optics and Image Science)*, 4(3):519–524, March 1987.

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

[28] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, January 1991.

[29] W. Zhao, R. Chellappa, and A. Krishnaswamy. Discriminant analysis of principal components for face recognition. In *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, pages 336–341, 1998.

[30] W. Zhao, R. Chellappa, and P. Phillips. Subspace linear discriminant analysis for face recognition. Technical Report CS-TR-4009, Center for Automation Research, University of Maryland, April 1999.

[31] W. Y. Zhao and R. Chellappa. SFS based view synthesis for robust face recognition. In *Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition*, pages 285–292, 2000.