Memory-based Face Recognition for Visitor Identification

Terence Sim[1] Rahul Sukthankar[1,2] Matthew Mullin[2] Shumeet Baluja[1,2]

[1]The Robotics Institute
Carnegie Mellon Univ.
Pittsburgh, PA 15213

[2]Just Research
4616 Henry Street
Pittsburgh, PA 15213

{tsim,rahuls,mdm,baluja}@justresearch.com

Abstract

We show that a simple, memory-based technique for appearance-based face recognition, motivated by the real-world task of visitor identification, can outperform more sophisticated algorithms that use Principal Components Analysis (PCA) and neural networks. This technique is closely related to correlation templates; however, we show that the use of novel similarity measures greatly improves performance. We also show that augmenting the memory base with additional, synthetic face images results in further improvements in performance. Results of extensive empirical testing on two standard face recognition datasets are presented, and direct comparisons with published work show that our algorithm achieves comparable (or superior) results. Our system is incorporated into an automated visitor identification system that has been operating successfully in an outdoor environment since January 1999.

1. Introduction

The problem of visitor identification consists of the following: a security camera monitors the front door of a building, acquiring images of people as they enter; an automated system extracts faces from these images and quickly identifies them using a database of known individuals. The system must easily adapt as people are added or removed from its database, and the system must be able to recognize individuals in near-frontal photographs. This paper focuses on the face recognition technology that is required to address this real-world task.

Face recognition has been actively studied [7, 12], particularly over the last few years [9]. The research effort has focused on the subproblem of frontal face recognition, with limited variance in illumination and facial expression. In this domain, techniques based on Principal Components Analysis (PCA) [10] popularly termed *eigenfaces* [26, 16], have demonstrated excellent performance. This paper introduces a simple, memory-based algorithm for face recognition, termed ARENA, that satisfies the requirements outlined above and also significantly outperforms PCA-based methods on two standard face recognition datasets.

2. Image Datasets and Preprocessing

Our results use human face images from two standard datasets: Olivetti-Oracle Research Lab (ORL) [22] and FERET [17, 19]. ORL consists of 400 frontal faces: 10 tightly-cropped images of 40 individuals with only minor variations in pose ($\pm 20°$), illumination and facial expression. The faces are consistently positioned in the image frame, and very little background is visible. FERET contains over 1100 faces; however many of them are unsuitable for our experiments since they are partial or full profiles, or the individuals were only photographed twice. Therefore, from FERET, we selected the subset of images that satisfied the following two constraints: (1) near-frontal poses; (2) images of individuals with more than five such images (our tests require several images for each person). The resulting 275 images consist of 40 individuals, with greater variation in pose and lighting than in the ORL dataset. For instance, many of these images were taken over different days and display significant differences in hairstyles, eyewear, and illumination. Unlike the ORL images, the FERET faces are of non-uniform size and do not always appear in the same location of the image. We use the FERET images as provided to explore the potential limitations of our template-based face recognition technique. Figure 1 shows two images for each of two individuals from the two datasets.

Figure 1. Top row: Two sample images each, of two subjects from ORL (left), and FERET (right). Note the difference in facial orientation, expression and accessories between the two images of the same individual. FERET images tend to exhibit greater variation in appearance (including hairstyle and clothing). Bottom row: the corresponding ARENA reduced-resolution images (16×16 pixels).

3. The ARENA Face Recognition Algorithm

ARENA is a memory-based [1] algorithm that employs reduced-resolution images (16×16) and the L_0^* similarity measure (described below). The reduced-resolution images are created by simply averaging over non-overlapping rectangular regions in the image. The distance from the query image to each of the stored images in the database is computed, and the label of the best match is returned.

3.1. L_p^* similarity measures

Our results show that the obvious choice for ARENA's similarity measure, the Euclidian distance, performs poorly. In this section we present alternatives. The L_p norm is defined as: $L_p(\vec{a}) \equiv (\sum |a_i|^p)^{\frac{1}{p}}$. Thus, the Euclidian distance is simply: $L_2(\vec{x}-\vec{y})$. Note that since we are not interested in the actual distances, but only in the ordering, we can equivalently employ the similarity measure $L_p^*(\vec{a}) \equiv (\sum |a_i|^p)$.

Robust statistics literature shows that L_2^*, despite its convenient analytic properties, overly penalizes outliers [11]. For this reason, the L_1^* similarity measure is often used in noisy environments. For ARENA, we have explored several L_p^* similarity measures (see Figure 2). The L_0^* similarity measure is defined as $L_0^*(\vec{a}) \equiv \lim_{p \to 0^+} L_p^*(\vec{a})$. Intuitively, $L_0^*(\vec{x} - \vec{y})$ counts the number of components in $\vec{x}$ and $\vec{y}$ that differ in value. Our experiments indicate that the best performance on this task is achieved with $p \leq 1$.

In our application, each reduced-resolution image is converted into a vector, $\vec{x}$, where each pixel in the image is represented as a component of the vector. In practice, since individual pixel intensities are noisy, we relax the definition of L_0^* to be:

$$L_0^*(\vec{x} - \vec{y}) \equiv \sum_{|x_i - y_i| > \delta} 1$$

where δ is a threshold, such that pixels whose intensities differ by less than δ are considered equivalent.

4. Principal Components Analysis (PCA)

The most widely used baseline for face recognition, *eigenfaces* [26, 16] employs Principal Components Analysis (PCA), which is based on the discrete Karhunen-Loève (K-L), or Hotelling Transform [10], is the optimal linear method for reducing redundancy, in the least mean squared reconstruction error sense. Points in $\mathcal{R}^d$ are projected into $\mathcal{R}^m$, (where $m \leq d$, and typically $m \ll d$). PCA has become popular for face recognition with the success of *eigenfaces* [26]. For face recognition, given a dataset of N training images (full-resolution originals, each with d pixels), we create N d-dimensional vectors, $\vec{x}_1, \vec{x}_2, ..., \vec{x}_N$, where each pixel is a unique dimension. The principal components of this set of vectors is computed as described in [10, 26] to obtain a $d \times m$ projection matrix, W.

Now, the image $\vec{x}_i$ may be compactly represented as *weights*, $\vec{\theta}_i = (\theta_{i1}, \theta_{i2}, \ldots, \theta_{im})^T$, such that $\vec{z}_i = \vec{\mu} + W\vec{\theta}_i$ approximates the original image, where $\vec{\mu}$ is the mean of the $\vec{x}_i$ and this reconstruction is perfect when $m = d$. The columns of W form an orthonormal basis for the space spanned by the training images.

Two variants of PCA for face recognition are evaluated in this paper, termed PCA-1, and PCA-2. For both algorithms, each training image is first projected into the eigenspace, and represented as a weight vector $\vec{\theta}_i$:

$$\vec{\theta}_i \quad = \quad W^T(\vec{x}_i - \vec{\mu}) \qquad (1)$$

In PCA-1, the centroid of the weight vectors for each person's images in the training set is computed and stored [26] — PCA-1 assumes that each person's face images will be

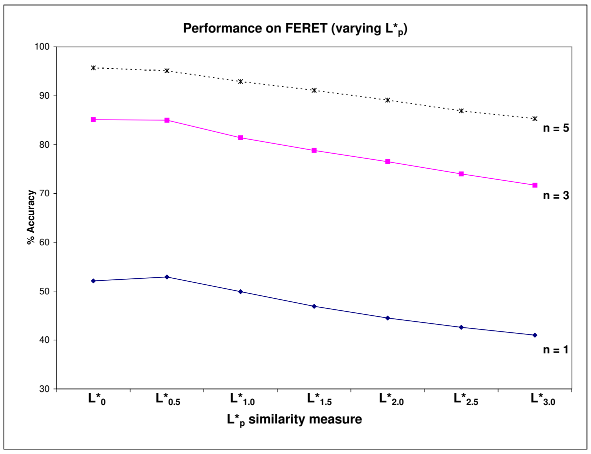

Figure 2. This figure shows how ARENA performs for different L_p^*. For L_0^*, δ was set to 10 (pixels range in value from 0 to 255). Note that the $p \leq 1$ norms perform significantly better regardless of the number of training images (n). Experiments were conducted on the FERET database with the original images subsampled to 16×16.

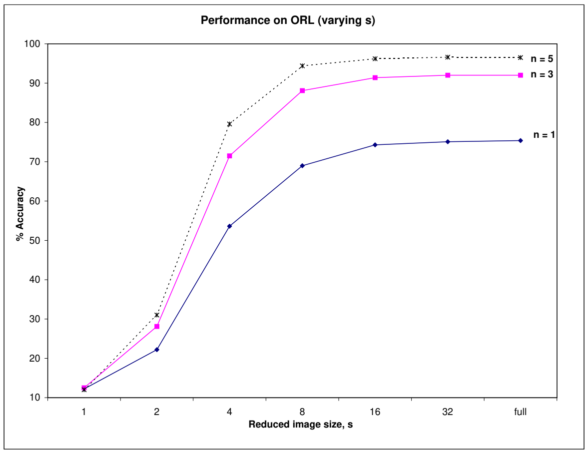

Figure 3. ARENA performance on the ORL dataset as s, the size of the reduced-resolution images is varied. "Full" indicates that the full-resolution image was used. ARENA's performance improves rapidly with s, and plateaus by $s = 16$.

clustered in the weight space, so the actual training data is not needed. In PCA-2, a memory-based variant of PCA, each of the weight vectors is individually stored [13] — requiring more storage space, but providing PCA-2 with a richer representation. When a test image is presented to the system, it is first projected into the eigenspace (by Equation 1), and its weight vector $\vec{\theta}_{\text{new}}$ is computed. $\vec{\theta}_{\text{new}}$ is then compared against the stored weight vectors, Θ, and the $\vec{\theta}_k$ that is closest $\vec{\theta}_{\text{new}}$ is located:

$$\vec{\theta}_{\text{best}} = \arg\min_{\vec{\theta}_k \in \Theta} L_2^*(\vec{\theta}_{\text{new}} - \vec{\theta}_k)$$

The label of $\vec{\theta}_{\text{best}}$ is returned as the identity of the face represented by $\vec{\theta}_{\text{new}}$.

5. Results

In the experiments described in this paper, $n \in \{1, 3, 5\}$ randomly-selected images for each individual in the dataset were placed in the training set, and the remaining images were used for testing. Multiple runs for each n were performed with different, random partitions between training and testing images, and the results were averaged.[1] The experiments were performed on both ORL and FERET images, and the results are reported separately so that they may be directly compared with other published results.[2]

5.1. Experiment: varying resolution

Here, we examine how ARENA's performance changes as the dimension of reduced-resolution images is varied. Each original face image is reduced to $s \times s$ using simple local averaging. Figure 3 shows experiments with $s \in \{2^k | k = 0, \ldots, 5\}$, and for 92×112 full resolution images. Performance improves rapidly as s increases (over all training set sizes) and shows no significant improvement beyond $s = 16$. In the remainder of this paper, we present results with 16×16 ARENA images.[3]

[1] In testing ARENA, we exploit the fact that the distances between any two images in the dataset are independent of the test/train split, and contain sufficient information to efficiently enumerate the number of test/train splits that result in a correct identification for each image in the dataset. This allows us to effectively compute the average performance of ARENA over *all* possible test/train splits, without suffering the combinatorial explosion (as detailed in [14]).

[2] FERET is available from `jonathon@nist.gov`. ORL is available at `www.cam-orl.co.uk/facedatabase.html`. The list of FERET images used in our experiments is available at `www.cs.cmu.edu/~rahuls/Research/ARENA`.

[3] Low-resolution images of similar dimensions are commonly used in the neural network literature [20].

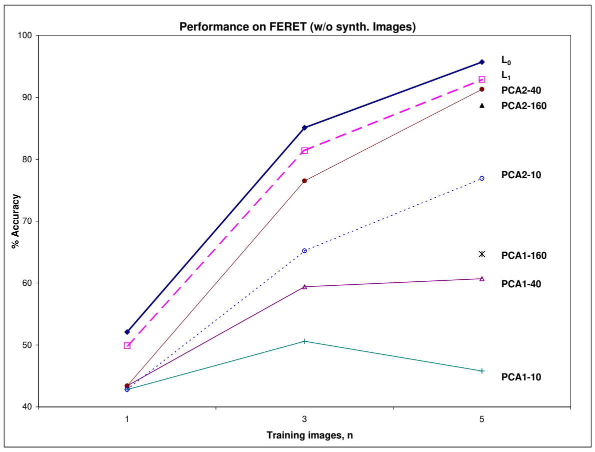

Figure 4. Performance of ARENA ($p \in \{0, 1\}$, 16×16 images), compared to PCA-nearest-centroid (PCA-1) and PCA-nearest-neighbor (PCA-2) on FERET. For PCA algorithms, number of eigenvectors $m \in \{10, 40, 160\}$. Since m is limited by the rank of the training set matrix, $m = 160$ can only be used when there are more than 160 training images (more than 4 training images for each of the 40 individuals, $n \geq 4$).

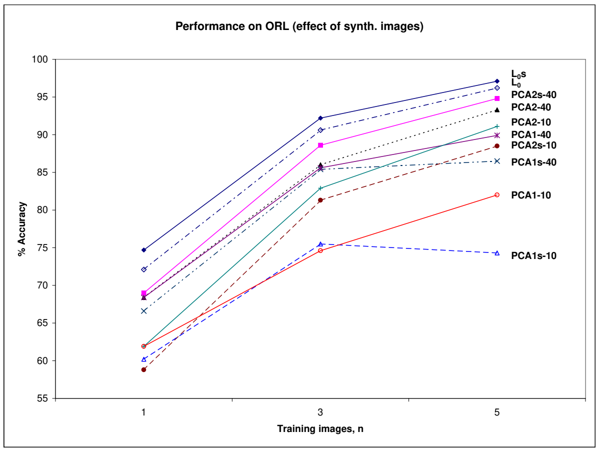

Figure 5. This figure shows the effect of adding 10 synthetic training images per original image to PCA-1 and PCA-2 ($m \in \{10, 40\}$). Results for ARENA ($p = 0$) are given for comparison. In all cases, PCA is outperformed by ARENA (even without synthetic images). Tests with synthetic images are marked with an 's' above.

5.2. Experiment: comparisons with PCA

PCA performance depends on the number of eigenvectors, m, that are stored. If m is too low, important information about the identity is likely to be lost. However, if m is too high, the weights corresponding to small eigenvalues will be noisy. This is analogous to selecting the appropriate subsampling ratio in ARENA.

Figure 4 shows that ARENA ($p \in \{0, 1\}$) outperforms both PCA-1 and PCA-2, as the number of dimensions, m is varied. It is interesting to observe that, on the FERET dataset, the accuracy for PCA-1 ($m = 10$) *drops* when the number of training images, n is increased from 3 to 5. This may be because the training faces for a given individual are not well-represented by the centroid of a single cluster. The memory-based techniques (PCA-2 and ARENA) are not adversely affected.

5.3. Experiment: adding synthetic images

Because we wish to perform recognition with the fewest number of training images per person, we augment the training set with additional, synthetically-generated face images. Since the task addressed in this paper is near-frontal face recognition, these images can be synthesized with simple geometric transformations (i.e., translation, rotation and scaling); more complex transformations to account for out-of-plane rotations have been explored in [4, 27]. Incorporating synthetic training images into a memory-based model generally improves performance because it can increase the likelihood that an unknown query image will be matched to a correct instance in the memory base. Note that methods such as normalized correlation [5, 18] automatically account for translation, but do not address either rotation or scaling. The process is similar in concept to the supplementary images used for neural network training for autonomous navigation [20] and automatic digit recognition [2, 23].

Figure 5 shows the effect of adding synthetic data to both ARENA and PCA algorithms. Note that ARENA's accuracy improves by 2-3%, while most of the PCA methods are affected detrimentally. This is because PCA is very sensitive to registration error, which is exacerbated by the synthetic images.

6. Additional Experiments

In addition to the comparisons described above, with standard PCA techniques, we have extensively compared ARENA with other state-of-the-art face recognition algorithms. Due to space limitations, only a brief summary of these experiments is presented.

6.1. Standard PCA variants

Recently, many modifications to the standard eigenface algorithm have been proposed and have been shown to work better in limited situations. We have duplicated two common such variants.

The first variant uses Mahalanobis distance [8] rather than standard Euclidian distance: PCA is initially used to reduce dimensionality by discarding eigenvectors corresponding to the lowest-magnitude eigenvalues (these are assumed to be noise). The remaining eigenvectors are then scaled such that their contributions to the distance are effectively equal. Unfortunately, in our experiments on this task, Mahalanobis-PCA does not consistently improve performance: Mahalanobis-PCA-1 is inferior to PCA-1 for low n or m, but slightly better in other cases; Mahalanobis-PCA-2 is uniformly inferior to PCA-2.

The second variant is motivated by the observation that the eigenvectors corresponding to the greatest eigenvalues often encode variations in illumination rather than the identity of the individual [3]. Consequently, if these eigenvectors are discarded (typically the top three [15]), then projecting the query image along the remaining eigenvectors should result in weights that do not encode these illumination effects. While there is some variation in lighting in both ORL and FERET datasets, we have found that PCA performance drops with this variant. It appears that the top three eigenvectors are (at least partially) encoding important information (supported by [3]).

In our experiments, even the best PCA algorithm,[4] which achieved an accuracy result of 94.8% in its best run, was significantly outperformed (in identical experiments) by ARENA (L_0^* without synthetic images: 96.2%, with synthetic images, 97.1%).

6.2. Comparisons with other algorithms

We have also duplicated some experiments reported in [13]. They examined the performance of four algorithms, “Eigenfaces - average per class” (identical to PCA-1), “Eigenfaces - one per image” (identical to PCA-2), “PCA+CN” (PCA combined with a convolutional network classifier), and “SOM+CN” (Self-Organizing Map combined with a CN), on the ORL dataset with n ranging from 1 to 5. Table 1 summarizes these results. The last two rows of the table present results obtained with ARENA, augmented with the synthetic images ($p \in \{0, 1\}$). Both variants of ARENA outperform all of the reported results.

Face recognition using Hidden Markov Models (HMM) on the ORL database is reported in [22]. Their best algorithm, with $n = 5$, obtained an accuracy of 88%, putting it between Lawrence’s implementations of PCA-1 and PCA-2. This is inferior to any ARENA variant.

[4]This was PCA-2 with Euclidian distance, $m = 40$, $n = 5$ (without discarding top eigenvectors), and 10 synthetic images when tested on the ORL dataset.

Images per person	1	3	5
Eigenface - avg per class	61.4%	71.1%	74.0%
Eigenface - one per img	61.4%	81.8%	89.5%
PCA+CN	65.8%	76.8%	92.5%
SOM+CN	70.0%	88.2%	96.5%
ARENA ($p = 0$,$s = 16$)	74.7%	**92.2%**	**97.1%**
ARENA ($p = 1$,$s = 16$)	**75.1%**	92.0%	96.8%

Table 1. Comparison of ARENA with results reported in [13].

7. Computational Complexity and Storage

In this section, we examine the computational complexity of PCA-1, PCA-2, and ARENA, and compare their storage requirements. We also discuss an important practical consideration: whether the algorithms can support incremental updates. We define the following terms:

- c The number of people in the training set.
- n The number of training images per person.
- N The total number of training images: $N = cn$.
- d Each image is represented as a point in $\mathcal{R}^d$, where d is the dimensionality of the image.
- m The dimension of the reduced representation: number of stored weights (PCA), or number of pixels (s^2) in reduced-resolution ARENA. Normally, $d \gg m$.

The asymptotic behavior of the various algorithms is summarized in Table 2. The following observations are noteworthy. First, the training time for ARENA scales linearly with N, while both PCA-1 and PCA-2 training times scale poorly (due to the eigenvector computations inherent in the PCA algorithm). Second, the classification times for PCA-2 and ARENA are asymptotically slower than the corresponding time for PCA-1; however, ARENA avoids the dm term which is required for both PCA algorithms. Third, the storage space for ARENA is typically smaller than that of either PCA algorithm: ARENA always requires substantially less storage than PCA-2, and unless N is very large, ARENA also requires less storage than PCA-1. This is because ARENA performs all computations in the reduced dimensional space, and does not need to store any vectors of size d whereas any variant of PCA must store the projection matrix (m vectors of dimension d). ARENA achieves further savings in storage space by quantizing the pixel values in the reduced-resolution image to a single byte (compared to the 8-byte double-precision values used for every element in the PCA methods). Our experiments show that this quantization does not significantly reduce ARENA's accuracy.

Finally, ARENA can learn incrementally in constant time: given a new training sample, learning simply involves subsampling and storing, while PCA requires recomputing the projection matrix, W.[5] Although techniques for incremental PCA have been developed [6], these are still computationally expensive since all of the old training images need to be stored at full resolution, so that they can be re-projected using the new W. By contrast, ARENA never requires the full-resolution images. Incremental training is crucial for applications that must adapt to changes. For instance, a visitor identification system must readily incorporate additional training data as new visitors are added, or when known individuals radically change their appearance (e.g., facial hair or headgear).

8. Discussion

Why does ARENA perform so well? Let us consider the behavior of the ARENA algorithm from two perspectives: (1) ARENA is performing a dimensionality reduction which, although non-domain-specific, is well-suited for face recognition since it reduces noise and compensates for small changes in the image; (2) it is performing a variant of template-matching (correlation) on reduced-resolution images.

The first viewpoint indicates that ARENA is transforming high-dimensional points into a space that is manageable for nearest-neighbor algorithms. ARENA uses local averaging, which unlike PCA, is more robust to small image registration errors. The synthetic images further help the nearest-neighbor algorithm by populating the space with positive instances. Using synthetic images for standard PCA-based methods is expensive because the order-of-magnitude expansion in the training set results in high memory usage during the training phase (not to mention training time, as shown in Table 2).

From the second viewpoint, ARENA uses a large number of static templates (from several training images, augmented by the synthetic images). Template-matching has been used in early face recognition research [12], for facial-feature-detection; many recent approaches to face recognition can also be considered to be sophisticated versions of template-matching [5]. There are several significant differences between these correlation algorithms and ARENA. First, the correlation algorithms use high-resolution images and are therefore sensitive to small details in the image. Second, the use of the L_2^* similarity measure further exacerbate this sensitivity to unimportant differences. Consequently, two images of the same person with slightly different orientation or facial expression may be difficult to match.

ARENA has been integrated into a visitor identification system. The system obtains images from a security camera that monitors the front door of a building. Faces are extracted from these images using a neural-network-based face tracker [21], histogram-equalized and sent to ARENA. ARENA attempts to recognize the visitor and the system notifies interested parties of the visitor's arrival. ARENA is particularly well-suited for this application because it supports incremental training: a human operator can label incorrect guesses and these are immediately incorporated into the training set. This system has been operational (24 hours a day) since January 1999. Note that the images gathered often display significant out-of-plane rotation, occlusion and extreme lighting conditions (half-faces); therefore, the images for a given individual can look very different. However, by acquiring many images for each frequent visitor, the system is able to robustly recognize these individuals in a variety of situations. Under these challenging conditions, we are pleased to report overall accuracies of 55% for an image set containing 50 individuals (more than 1000 training images, added incrementally over a period of several months). To test the system further, we added 1500 "distractor" images of faces collected from the web and tagged them with the single label "stranger". There has been no noticeable drop in classification performance of known visitors, but unknown visitors are often correctly classified as "stranger". For detailed performance statistics, see [24].

9. Conclusions and Future Work

This paper demonstrates that ARENA, a very simple algorithm, can significantly outperform established face recognition algorithms on standard datasets. Unlike the standard PCA-based algorithms, ARENA easily handles incremental updates to the face recognition database and has been shown to scale well. Given the algorithm's simplicity, ARENA's accuracy is somewhat surprising. We invite other researchers to independently confirm our findings.

We are extending the work described here in several directions. First, we are comparing ARENA against Fisher Discriminant Analysis (FDA) [3, 25] approaches to frontal face recognition. Most FDA methods require a dimensionality-reduction step, traditionally performed using PCA. We are exploring whether reduced-resolution images (as used in ARENA) can perform this role. We also plan to investigate the effects of using different L_p^* similarity measures in such algorithms.

Finally, a useful byproduct of the visitor identification system is the collection of a labelled face dataset, with unposed images captured in the natural lighting of an outdoor environment. This dataset will be made available on our website to researchers in the near future.

[5] One could use PCA without recomputing W; however, this improves efficiency at the expense of accuracy.

Method	Training time	Classification time	Storage space	Incremental update cost Recomputed W	Unchanged W
PCA-1	$O(N^3 + N^2 d)$	$O(cm + dm)$	$O(cm + dm)$	$O(N^3 + N^2 d)$	$O(md)$
PCA-2	$O(N^3 + N^2 d)$	$O(Nm + dm)$	$O(Nm + dm)$	$O(N^3 + N^2 d)$	$O(md)$
ARENA	$O(Nd)$	$O(Nm + d)$	$O(Nm)$	$O(d)$	

Table 2. Comparison of asymptotic behavior. ARENA displays clear advantages over both PCA-based techniques.

Acknowledgments

We would like to thank: Henry Rowley, for providing us with the eye positions needed to extract the faces from the FERET database; Erich Greene and Keiko Hasegawa, for ensuring that the statistical significance tests were rigorous; and Henry Rowley and Gita Sukthankar for valuable feedback.

References

[1] C. Atkeson, W. Moore, and S. Schaal. Locally weighted learning. *AI Review*, 11, 1997.

[2] S. Baluja. Making templates rotationally invariant: An application to rotated digit recognition. In *Advances in Neural Information Processing Systems*, 1998.

[3] P. Belhumeur, J. Hespanha, and D. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on PAMI*, 19(7), 1997.

[4] D. Beymer and T. Poggio. Face recognition from one example view. In *Proceedings of ICCV*, 1995.

[5] R. Brunelli and T. Poggio. Face recognition: Features versus templates. *IEEE Transactions on PAMI*, 15(10), 1993.

[6] S. Chandrasekaran, B. Manjunath, Y. Wang, J. Winkeler, and H. Zhang. An eigenspace update algorithm for image analysis. In *Proceedings of CVGIP*, 1997.

[7] R. Chellappa, C. Wilson, and S. Sirohey. Human and machine recognition of faces: A survey. *Proceedings of the IEEE*, 83(5), 1995.

[8] R. Duda and P. Hart. *Pattern Classification and Scene Analysis*. John Wiley and Sons, 1973.

[9] T. Fromherz. Face recognition: A summary of 1995–1997. Technical Report ICSI TR-98-027, International Computer Science Institute, Berkeley, 1998.

[10] R. Gonzales and R. Woods. *Digital Image Processing*. Addison-Wesley, 1992.

[11] P. Huber. *Robust Statistics*. Wiley, 1981.

[12] T. Kanade. *Picture Processing by Computer Complex and Recognition of Human Faces*. PhD thesis, Kyoto University, 1973.

[13] S. Lawrence, C. Giles, A. Tsoi, and A. Back. Face recognition: A hybrid neural network approach. Technical Report UMIACS-TR-96-16, University of Maryland, 1996.

[14] M. Mullin and R. Sukthankar. An efficient technique for calculating exact nearest-neighbor classification accuracy. Technical Report JPRC-TR-1999-002-1, Just Research, 1999.

[15] A. O'Toole, H. Abdi, K. Deffenbacher, and D. Valentin. Low dimensional representation of faces in high dimensions of the space. *Journal of the Optical Society of America A*, 10, 1993.

[16] A. Pentland, B. Moghaddam, and T. Starner. View-based and modular eigenspaces for face recognition. In *Proceedings of Computer Vision and Pattern Recognition*, 1994.

[17] P. Phillips, H. Moon, P. Rauss, and S. Rizvi. The FERET evaluation methodology for face-recognition algorithms. In *Proceedings Computer Vision and Pattern Recognition*, 1997.

[18] P. Phillips, H. Moon, P. Rauss, and S. Rizvi. The FERET september 1996 database and evaluation procedure. In *Proceedings of Audio and Video-based Biometric Person Authentication*, 1997.

[19] P. Phillips, H. Wechsler, J. Huang, and P. Rauss. The FERET database and evaluation procedure for face-recognition algorithms. *Image and Visual Computing*, 16(5), 1998.

[20] D. Pomerleau. *Neural Network Perception for Mobile Robot Guidance*. PhD thesis, Carnegie Mellon University, February 1992.

[21] H. Rowley, S. Baluja, and T. Kanade. Neural network-based face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1), 1998.

[22] F. Samaria and A. Harter. Parametrisation of a stochastic model for human face identification. In *Proceedings of IEEE Workshop on Applications on Computer Vision*, 1994. ORL database is available at: `<www.cam-orl.co.uk/facedatabase.html>`.

[23] B. Schölkopf, C. Burgess, and V. Vapnik. Incorporating invariances in support vector learning machines. In *Artificial Neural Networks — ICANN'96*, 1996.

[24] T. Sim, R. Sukthankar, M. Mullin, and S. Baluja. High-performance memory-based face recognition for visitor identification. Technical Report JPRC-TR-1999-001-1, Just Research, 1999.

[25] D. Swets and J. Weng. Using discriminant eigenfeatures for image retrieval. *Transactions on PAMI*, 18(8), 1996.

[26] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1), 1991.

[27] T. Vetter. Synthesis of novel views from a single face image. *International Journal of Computer Vision*, 28(2), 1998.