On Multi-scale differential features for face recognition

S. Ravela
Center for Intelligent Information Retrieval

Allen R. Hanson
Vision Laboratory

Dept. of Computer Science, University of Massachusetts at Amherst, MA, 01002
{ravela,hanson}@cs.umass.edu*

Abstract

This paper describes an algorithm that uses multi-scale Gaussian differential features (MGDFs) for face recognition. Results on standard sets indicate at least 96% recognition accuracy, and a comparable or better performance with other well known techniques. The MGDF based technique is very general; its original application included similarity retrieval in textures, trademarks, binary shapes and heterogeneous gray-level collections.

1 Introduction

Face recognition technologies can significantly impact authentication, monitoring and image indexing applications. This paper presents an algorithm to compute similarity of faces as a whole. The task is to query a database using the image of a face and then have the system either ascertain its identity, or retrieve the top N similar matches. As such, the technique is general and has hitherto been used successfully in image retrieval tasks such as finding similar scenes, trademarks, binary shapes and textures [23, 24, 25]. The approach is based on the two hypotheses; first that visual appearance of a face plays an important role in judging similarity and second, multi-scale differential features of the image brightness surface form effective appearance features.

The first hypothesis is based on the observation that visual appearance is an important cue with which we judge similarity. We readily recognize objects that share a visual appearance as similar, and in the absence of other evidence, are likely to reject those that do not. A precise definition of visual appearance is difficult. The physical and perceptual phenomena that define appearance are not well known, and even when there is agreement, such as the effect of object (3D)shape, surface texture, illumination, albedo and viewpoint, it is non-trivial to decompose an image along these components. However, early recognition algorithms [9, 16, 32] brought forward the notion that the similarity between computational representations of imaged brightness surfaces, in many cases, correlates with similarities in visual appearance of objects. Therefore, it is not unreasonable to develop appearance representations and similarity measures to suit the semantics of the retrieval or recognition task.

In this paper, an appearance representation for face recognition using distributions of local features of the image brightness surface is constructed. Local features are obtained by applying operators to the image that, equivalently, can be thought of as tunable spatial-frequency filters, statistical descriptors of the brightness surface, or approximations of the local shape of the image brightness surface. Specifically, multi-scale differential features are used [3, 5, 7, 11, 15, 23, 24, 25, 26, 21, 28, 29] and this choice is motivated by arguments [3, 7] that the local structure of an image can be represented in a stable and robust manner by the outputs of a set of multi-scale Gaussian derivative filters (MGDFs) applied to an image. In order to deduce global similarity between two face images, multi-scale differential features are composed into histograms and correlated.

The first part of this paper begins with a brief review of scale-space theory underlying MGDFs and ends with an algorithm to deduce global similarity. In the second part, this algorithm is applied to face recognition. Using the databases and protocol for evaluation described by Sim et. al. [30], this paper demonstrates that the algorithm presented here is at least as effective when compared to several other methods.

*This material is based on work supported in part by the following sponsors: 1. Department of Commerce under cooperative agreement number EEC-9209623, 2. Defense Advanced Research Projects Agency/ITO under ARPA order number D468, issued by ESC/AXA contract number F19628-95-C-0235, 3. Air Force Office of Scientific Research under grant number F49620-99-I=0138, 4. National Science Foundation IRI-9619117 and Multimedia CDA-9502639. Additionally this work has been made possible due to the support of the Library of Congress. Any opinions, findings and conclusions or recommendations expressed in this material are the author's and do not necessarily reflect those of the sponsors.

1.1 Related Work

Face recognition has received significant attention and it is beyond the scope of this paper to fully investigate the available techniques. Instead we describe techniques that are most relevant to our approach. Sim et al [30] use a relatively simple technique of matching decimated images with extremely good results. Although our approach is completely different we use their evaluation methodology. Other techniques for face recognition have also been developed using projection profiles [33], deformable surfaces [14], hidden Markov models (HMM) [27], and self-organizing maps [10]. None of these techniques are related to the ones presented here, but comparisons can be made by reading the results presented here and those results presented by Lawrence and Sim [10, 30]. Results on the FERET collection with other techniques may also be found in Phillips [19].

From an appearance representation standpoint, principal component analysis (PCA) based techniques are more relevant. PCA was pioneered by Kirby and Sirovich [9] as a representation for faces which was also developed into an effective face recognition system by Turk and Pentland [32], with generalizations to multiple views [16, 18], illumination changes [16], and replicated on other objects [16]. Since the success of eigen decomposition depends on the objects being correlated an attempt was made to overcome this restriction by Swets et. al [31, 36]. They extend the traditional PCA method to multiple classes of objects using Fischer's discriminant analysis [1]. The approach presented in this paper is different because Eigen decompositions are not used to characterize appearance. Further, the method presented here uses no learning and does not require constant sized images. In fact, one of the conclusions drawn from this paper is that a scale-space decomposition (rather than an eigen one) performs equivalently well. That is, an unbiased representation performs as well (if not better) than the learned representation.

Appearance features can also be extracted in the frequency domain and in this sense are commonly related to texture features. In the context of image retrieval Ma et. al. [12] use Gabor filters to retrieve images with similar texture. Gabor jets [34] have also been used for face recognition. We find that a comparison between Gaussian and Gabor filters is instructive. Gabor filters are sine modulated Gaussian functions, which can be tuned to respond to a bandwidth around a certain center frequency. They exhibit compactness in space and frequency, are optimal in the sense of the uncertainty principle (time-bandwidth product) and are complete. Gabor filters are not equivariant with rotations, and separable implementations are expensive. In contrast, Gaussian derivatives exhibit the same time-bandwidth property and although they have infinite support, they can be safely truncated at around four standard deviations. While Gaussian derivatives have coupled bandwidth and center frequency, in practice separate tuning is not necessary. Rather, the derivatives provide a "natural" sampling of the frequency space, because they represent the orders of approximation in a Taylor series sense. The significant advantage of using the Gaussian derivatives is that, they are equivariant with rotations [4] eliminating the need for explicitly oriented filters and also support the formulation of rotational invariants. Gaussian derivatives are separable and efficient implementations are possible. There are several other interesting properties and the reader is referred to [6, 23] for a more basic review.

2 Computing Global Similarity

The steps involved in deducing similarity between a query face image and a database image are as follows: Database images are filtered *a priori* with Gaussian derivatives, and then, at each pixel, the gradient orientation and surface curvature is computed. A query image is filtered the same way and multi-scale histograms of curvature and orientation are correlated to measure similarity. In the authentication task the identity of the best matching image in the database is ascribed to the query and in the monitoring task, the top N are presented to the user. Below, the use of differential features and the steps in the algorithm are discussed.

2.1 Differential features:

The simplest differential feature is a vector of spatial derivatives. For example, given an image I, and some point, p, the first two orders of spatial derivatives can be used as a feature (vector). This vector approximates the shape of the local intensity surface in the sense of a second order Taylor approximation. Including higher orders produces a more precise approximation. Derivatives capture useful statistical information about the image. The first derivatives represent the gradient or "edgeness" of the intensity and the second derivatives can be used to represent curvatures (bars, blobs and so on).

However it is important that derivatives be computed in a stable manner. Derivatives will be stable if, instead of using just finite differences, they are computed by filtering an image with normalized Gaussian derivative filters (actually any C^∞ function will do [3]). In two dimensions, a Gaussian derivative is the derivative of the function

$$G(.,\sigma) = \frac{1}{2\pi\sigma^2} exp - \frac{x^2+y^2}{2\sigma^2}$$

In the frequency domain, a Gaussian derivative filter is a band-pass filter, as shown in Figure 1 (one-dimensional case). Computing derivatives by filtering with a Gaussian derivative at a certain scale, therefore, implies that only

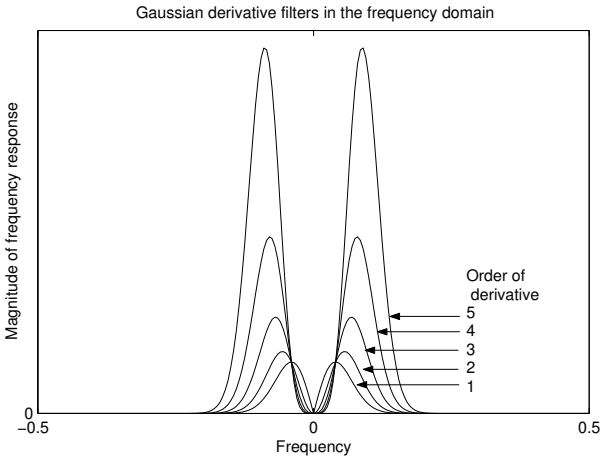

Figure 1: Gaussian derivative filters in the frequency domain.

a limited band of frequencies are being observed. Thus, in order to describe the original image more completely, a multi-scale representation is necessary. Sampling the scale-space of the image becomes essential.

2.2 Gaussian scale-space:

The necessity of a multi-scale representation described above can be concluded for *any* smooth band-limiting filter by using the commutativity of differentiation and convolution. The Gaussian happened to be a convenient function; it has natural scale parameterization, smoothness and self-similarity across scales. However, the Gaussian is more than just convenient. There are compelling theory and implementation related arguments for using multi-scale Gaussian derivatives to form appearance features. In particular, it has been shown by several authors [3, 5, 11, 26, 35], that under certain general constraints, the (isotropic) Gaussian filter forms a unique operator for representing an image across the space of scales. Structures (such as edges) observed at a coarser scale can be related to structures already present at a finer scale and not as an artifact of the filter. In general, the Gaussian (linear) scale-space serves as an unbiased (without using any other information) front end (pre-processor) for representing the image from which differential features may be computed. It is beyond the scope of this document to engage in a full discussion about the scale-space image representation and, instead, the reader is referred to the following papers [3, 11, 26, 35, 23]. Other reasons for choosing the Gaussian are presented in Section 1.1.

2.3 Curvature and Orientation:

Several differential features can be constructed from derivatives and several representations and methods have been developed [21, 28, 25, 29, 24, 23, 22] for recognition and retrieval. The choice of these features depends on several factors, primary (among these) is tolerance to rotation, illumination, scale since variations in these affects appearance. Here we argue for two particular features.

Since the task is to robustly characterize the 3-dimensional intensity surface (X, Y, Intensity), local curvatures are appropriate because the surface is uniquely determined from them. In particular, two principal curvatures, namely the isophote and flowline curves can be computed at a point, and represent the curvatures of the iso-intensity contours and the gradient integral curves. In fact, principal curvatures are nothing more than the second order spatial derivatives expressed in a coordinate frame (gauge [3]) determined by the orientation of the local intensity gradient. The principal curvatures of the intensity surface are invariant to image plane rotations, monotonic intensity variations and further, their ratios are, in practice, quite tolerant to scale variations of the entire image. The isophote (N) and flowline (T) curvatures are defined as [8, 3]:

$$N = A \star \left[2 I_x I_y I_{xy} - I_x^2 I_{yy} - I_y^2 I_{xx}\right] \quad (1)$$

$$T = A * \left[I_{xy}(I_x^2 - I_y^2) + I_x I_y (I_{yy} - I_{xx})\right] \quad (2)$$

$$A = (I_x^2 + I_y^2)^{-\frac{3}{2}} \quad (3)$$

$I_x = I_x(p, \sigma)$ and $I_y = I_y(p, \sigma)$ are the first order partial spatial derivatives of image I around point p, computed using Gaussian derivative at scale σ. Similarly, I_{xx}, I_{xy}, and I_{yy} are the corresponding second derivatives. The isophote curvature N and flowline curvature T are then combined into a ratio called the shape index, expressed as follows [8, 2, 15]: $C = [0.5 - \frac{1}{\pi} * atan\frac{N+T}{N-T}]$. The index value C is undefined when either N and T are both zero, and is, therefore, not computed. This is interesting because very flat portions of an image (constant or constant slope in intensity) are eliminated. The shape index is in the range [0,1]. Nastar [15] also uses the shape index for recognition and retrieval. However, his approach uses curvatures computed at a single scale. Clearly, as the experiments suggest (see Section 3), this is not enough.

The second feature used is local orientation. Local orientation is the direction of the local gradient. Orientation is independent of curvature, is stable with respect to scale and illumination changes. The orientation is simply defined as $P = atan2(I_y, I_x)$ Note that P is defined only at those locations where C is and ignored elsewhere. As with the shape index P is rescaled and shifted to lie between the interval [0,1].

Feature Histograms: Histograms of the shape index and orientation are used to represent the distributions of features over an image. Histograms form a global representation because they capture the distribution of local features and they are the simplest ways of estimating a non parametric distribution. In this implementa-

tion, curvature and orientation are generated at several scales and represented as a one dimensional record or vector. The representation of the image I is the vector $V_i = < H_c(\sigma_1) \ldots H_c(\sigma_n), H_p(\sigma_1) \ldots H_p(\sigma_n) >$. H_c and H_p are the curvature and orientation histograms respectively. It should be noted that [28] use histograms of various differential features. However, the difference between the two approaches is that their method uses multi-dimensional histograms of features that does not include curvature. Further, their representations are computed at a single scale. Multi-dimensional histograms tend to be very sparse, and further, are computationally more expensive to match. We believe that using one dimensional histograms at several scales (and stringing them together) provides a sufficiently rich representation of the image.

Figure 2: Examples of the FERET(first pair) and ORL(next four pairs) sets.

Matching feature histograms: Two representations are compared using normalized cross-covariance defined as $d_{ij} = \frac{V_i^{(m)} \cdot V_j^{(m)}}{||V_i^{(m)}|| ||V_j^{(m)}||}$ Where $V_i^{(m)} = V_i - mean(V_i)$. There are other possible measures, such as the Kulback-Leibler [1] and Mahalanobis [13] distances which could be used. The query histogram vector V_q is compared with each database histogram vector V_i. The corresponding images are ranked by their score. We call this algorithm the 1D curvature/orientation or CO-1 algorithm.

3 Face Recognition

Two variations of the algorithm are compared for face recognition. The first is CO-1, where histograms are built over the entire image (CO-1). The second is PCO-1, where the image is partitioned into three tiles roughly covering a third of the image and histograms for each tile are generated separately and concatenated (PCO-1). Assuming the images are roughly face segmented to begin with, the top tile corresponds to the forehead region, the middle tile to the mid-face and the bottom tile corresponding to the chin region.

Datasets: The following three datasets are used for evaluations. 1. ORL Set [17]: the ORL (Olivetti Research Lab) collection is a publicly available collection of 400 faces. This collection contains 40 individuals. The database contains small view, gesture, and intensity variation. See the second through fourth face pair of Figure 2. 2. FERET Set [19]: The FERET dataset is maintained by NIST and the CDROM contains 3737 images. However, our tests were repeated in exactly the same configuration as Sim [30] and therefore we only used 275 images of 40 individuals. These images contain bust photographs with varying bust coverage, and small facial gesture and image illumination changes. See first face pair in Figure 2. 3. UMASS TeaCrowd Set [20]: The UMass Tea Crowd set consists of 119 images of faces extracted from a live video feed of cameras monitoring a Tea Party. There are total of 15 people in this collection. These faces contain gesture, illumination, and view variations, in addition to motion blur and occlusion. See Figure 3.

Evaluation: The evaluation methodology follows the one described by Sim et. al. [30]. During each trial a database is randomly split into a training set and a test set. The configurations of training set per trial uses either 5 exemplars per person or the greatest number less than half the number of faces available for that person, whichever is smaller . The remaining faces for the person become the test set. Each of these test set images becomes a query. A query is matched with all of the training set and the identity of the best matching training set image is ascribed to the query. Over a large (100) number of trials the proportion of correctly identified people is reported as the recognition rate. For example, in the ORL set a trial will consist of 200 training and test images each. Thus, over 100 trials 20,000 queries (test set) are matched with a random training/test pick at every trial.

Examples: In Figure 2, queries and corresponding exemplar images (selected during some trial) they match to are shown. The first face pair is drawn from the FERET set. Note that these images were not processed to localize the face portion alone. The remaining four pairs in Figure 2 show results from the ORL set. Note that the second pair in the second row in Figure 2 is a mismatch. The correct identity is not recovered, but qualitatively both these faces share a significant similarity in appearance.

In Figure 3, several examples from the TeaCrowd set are shown from a retrieval perspective. Each "row" of this Figure contains six images, the first being the query and the

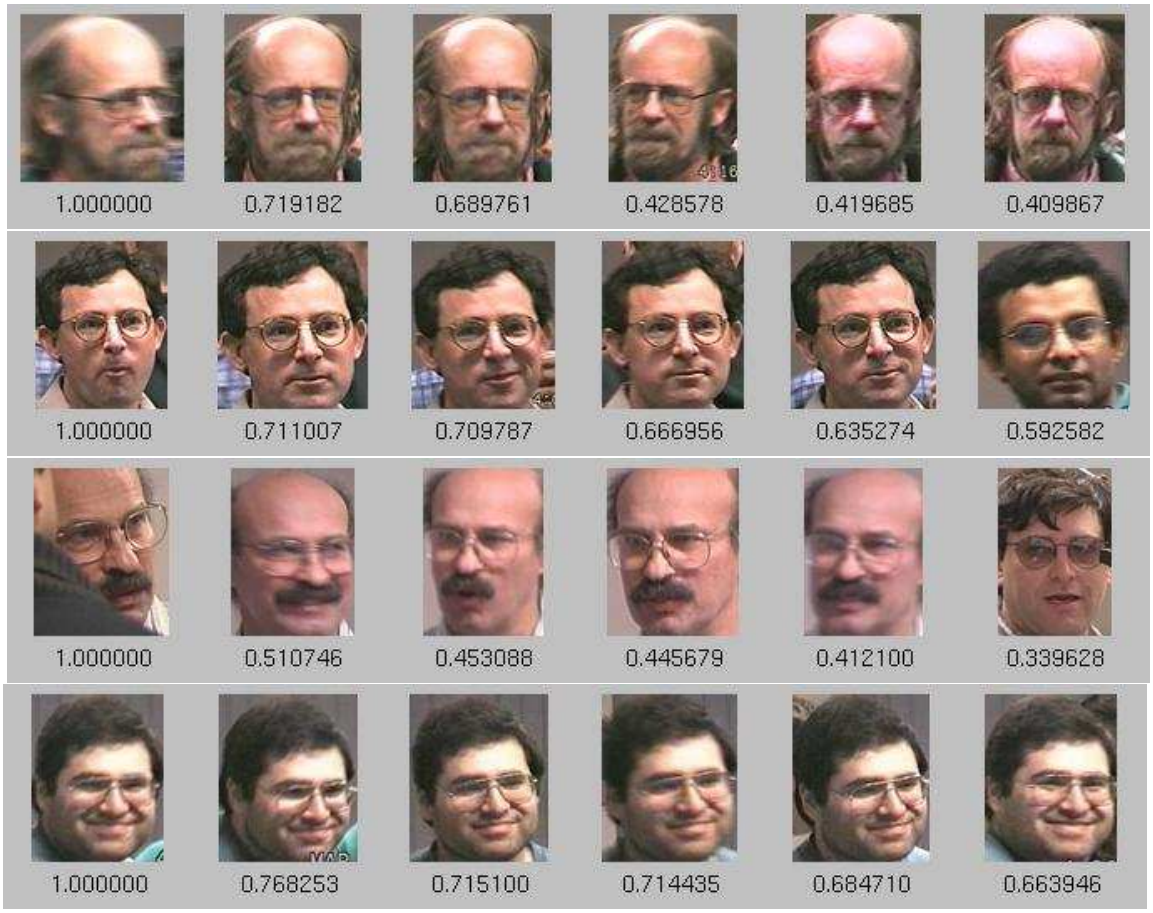

Figure 3: Examples of the Tea Crowd set from a retrieval point of view.

remainder being the images matched in rank order. Each image is labeled by its match score to the query (1.0 is maximum). These examples show recognition from a retrieval point of view. The queries include gesture variations, scale variations, occlusions, motion blur and view variations.

Analysis: The performance of the algorithm is depicted in Table 1. On all three sets the performance is very good and comparable to other algorithms, specifically, those based on Principal component analysis [32] and CMUs [30] technique. The reader is referred to Sim's paper [30] for additional comparisons with other techniques (they perform worse than CMUs technique). In Table 1, column 2 indicates the evaluation parameters used. In all methods 5 exemplars are used and when it is not possible to do so, only half the available are used. In our technique nothing is done to the images in terms of intensity stretching, warpings, face extraction or generating synthetic images. In contrast in Sim's technique based on matching thumbnails, synthetic images are generated from exemplars (rotated and slightly scaled versions) and these become part of the training set. A query's score against a database individual is the mean over the scores that it gets for all training samples of the individual. We pick the maximum. The implementation of Eigenfaces reported in the same paper also uses synthetic images from the exemplars, 40 Eigen values and the L2 norm to compare the query vector. In this case, like our method, the identity of the best matching image is ascribed to the query. Note that the results reported here for Eigenfaces are the best of the results reported by Sim et. al. [30](also see Lawrence's comparisons [10]).

The algorithm presented here has two principal parameters; scales and the bin sizes of the histograms. The graph in Figure 4, depicts the performance of the system with variation in scale for the ORL set using the CO-1 algorithm (other sets have similar results). For this graph the number of curvature and orientation bins were each fixed at 40. The X-axis of this graph is a byte-encoded number that indicates the scales used. The LSB means a scale value of 1, the next least significant bit corresponds to a scale value of $\sqrt{2}$ and so on through steps of $\sqrt{2}$, to an MSB value representing $8\sqrt{2}$. The valid numbers for this byte are 1-255,

Technique	Evaluation Parameters	ORL	FERET	TeaCrowd
UMASS PCO1	5 samples, 0 synthetic	98%	96%	96%
CMU	L0, 5 samples,10 synth	97%	96%	.IP.
UMASS CO1	5 samples, 0 synth	95%	90%	90%
Eigen-face	40 vector, L2, 5 samples,10 synth	95%	90%	.IP.

Table 1: The performance of MGDF methods with PCA and CMUs techniques

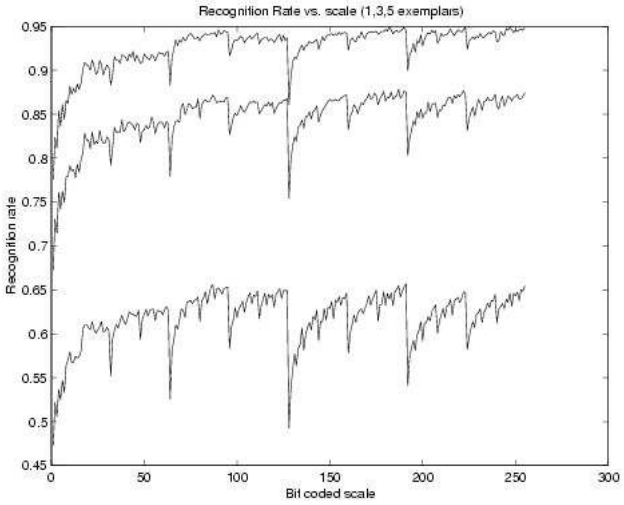

Figure 4: The performance on the ORL set. For this graph 40 bins were used in the histogram.

1 implying the use of only scale 1, 255 implying the use of all 8 scales. The Y-axis of this graph depicts recognition rate over 100 trials. Thus the recognition performance with respect to scales is exhaustively plotted. There are three plots in Figure 4. The lower one corresponding to the use of 1 exemplar, the middle one corresponding to 3 exemplars, and the top one corresponding to 5 exemplars.

Several conclusions can be drawn from this figure. First, the performance improves categorically with increase in exemplars, and this is true for all variations of the algorithms presented here. Second, a single scale, which is characterized by large dips in the plot is indicative of poor performance, and shows the necessity for multiple scales. Third, all eight scales are not necessary. It can be observed for example that a packed set of scales of smaller extent (such as bit code 96) give approximately the same performance as using all scales (such as bit code 255). Finally, a dense packing of scales is not essential either. A sequence of scales that is densely packed, such as $\ldots 1111 \ldots$, causes only marginal changes in accuracy in relation to one that is coarser, such as $\ldots 1010 \ldots$. In most cases we find that an octave spacing is sufficient, and two octave separation results in less than 1% drop in recognition accuracy. This suggests that the multi-scale representation can have a somewhat large sample width across scales. This is good news because it implies that significant "compression" in the representation is possible. The shape of this graph repeats itself for various bin combinations.

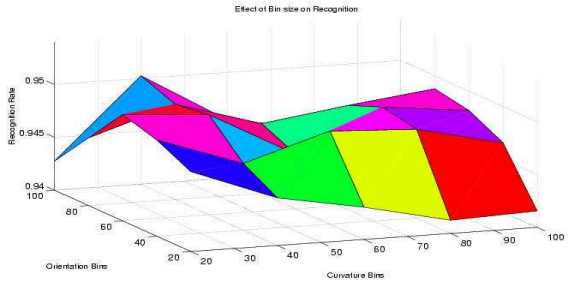

Figure 5: Recognition Performance with variation in Bin sizes for CO-1 on ORL set. All scales were used.

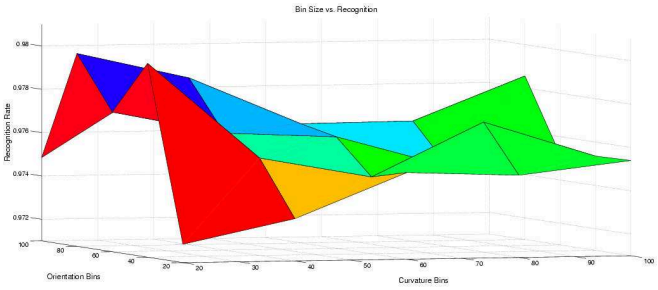

Figure 6: Recognition Performance with variation in Bin sizes for PCO-1 on ORL set. All scales are used.

The second factor that was varied is the bin size. For the experiments conducted all the scales were used with 5 exemplars and the bin sizes were systematically varied from 10 to 100 for curvature and orientation independently, therefore giving a matrix of 100 combinations. Surprisingly, the recognition rates held very stable: PCO-1 varied between 97.2% and 98.2% (see Figure 6; and CO-1 (see Figure 5) between 94.1% and 95.2%. The variance for any given observation over the trials was less than 1%. Finally, in terms of computation, it takes a few milli-seconds to recognize approximately 200 images from the database, and in contrast it takes about 0.4 seconds on a $400MHz$

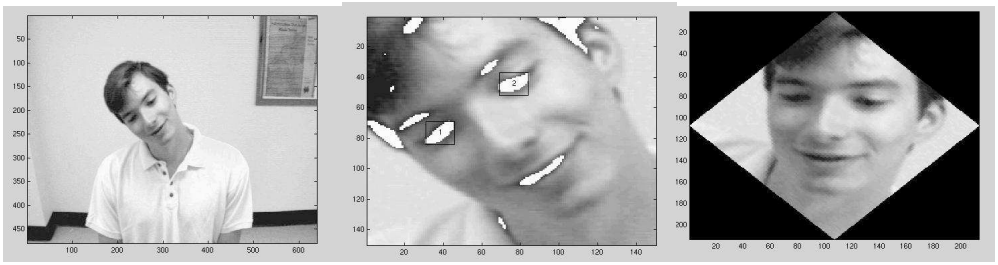

Figure 7: Face localization and rectification for recognition in a kiosk.

Pentium II processor with sufficient memory.

4 Summary and Conclusions

The results presented in this paper are very exciting for the following reasons. First, the curvature and orientation based method performs well; especially because there is no learning involved with respect to any of the parameters. Arguably, a representation based on the differential decomposition of the image at multiple scales is giving comparable performance to one based on learning a compact representation from the data, namely PCA. Thus, we find these features to be good from an appearance similarity point of view. Second, while scale is important, it seems in faces, the change of the feature (blur) with scale is rather slow. This is why dense sampling of scales is not necessary. This is good for a multi-scale representation. Third, the application of a "spatial" partition dramatically improves the results, suggesting that explicit representation of space may be necessary and might be the principal reason why the recognition rates improve. In conclusion, we believe that the representation presented here is turning out to be quite versatile.

We are extending this work towards constructing a *kiosk* that can be used for authentication using inexpensive cameras (QuickCams). Our present approach is to pre-process acquired images by localizing faces and detecting facial features. Once detected facial features can be used to establish a coordinate basis from which partitions can be computed for PCO-1. One way to do this is to simply rectify the face for orientation and scale. Further, facial feature detection provides coarse inference of facial view and thus, matching can be speeded up to nearby views in the database. For example, in Figure 7 three images taken at the kiosk are shown. The first is the full image taken by the camera, the second the detected face with an overlay of facial features, and boxes around the (final) localization of eyes. The third is the orientation rectified view of the face that simultaneously uses the orientation histogram and the inter-eye angle to rectify the face. While complete experimentation is forthcoming, in the context of this paper, it may be noted that facial features are localized using multi-scale differential features with natural scale selection.

Acknowledgement

The authors thank James Allan, W. Bruce Croft, R. Manmatha, and Edward Riseman for support and feedback through various stages of this work. The authors also thank Michael Remillard and Jerod Weinman for implementing portions of the kiosk.

References

[1] C. M. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1994

[2] Chitra Dorai and Anil Jain, "COSMOS - A representation scheme for free form surfaces", ICCV 95, pp. 1024-1029, 1995.

[3] L M J Florack, The Syntactic Structure of Scalar Images, PhD Dissertation, University of Utrecht, 1993

[4] W. T. Freeman and E. H. Adelson, The design and use of steerable filters, IEEE Trans. Patt. Anal. and Mach. Intel., 13(9):891-906, 1991

[5] J. J. Koenderink, The Structure of Images, Biological Cybernetics, 50:363-396, 1984.

[6] Gosta Granlund and Hans Knutsson, Signal Processing for Computer Vision, Kluwer Academic Publishers, 1995

[7] J. J. Koenderink and A. J. van Doorn, Representation of Local Geometry in the Visual System, Biological Cybernetics, 55:367-375, 1987

[8] J. J. Koenderink and A. J. Van Doorn, Surface Shape and Curvature Scales, Image and Vision Computing, 10(8), 1992

[9] M. Kirby and L. Sirovich, Application of the Kruhnen-Loeve Procedure for the Characterization of Human Faces, IEEE Trans. Patt. Anal. and Mach. Intel., 12(1): 103-108, Jan. 1990.

[10] S. Lawrence, C. Giles, A. Tsoi and A. Back, Face Recognition: A Hybrid Neural network approach. Tech. Report UMIACS-TR-96-16, University of Maryland, 1996.

[11] T. Lindeberg, Scale-Space Theory in Computer Vision, Kluwer Academic Publishers, 1994

[12] W. Y. Ma and B. S. Manjunath, Texture-Based Pattern Retrieval from Image Databases, Multimedia Tools and Applications, 2(1):35-51, Jan. 1996

[13] P. C. Mahalanobis, On the Generalized Distance in Statistics, Proceedings of the National Institute of Science, India, 12:49-55, 1936.

[14] B. Moghaddam, C. Nastar, and A. Pentland, Bayesian face recognition using deformable intensity surfaces, In Proc. Comp. Vision and Patt. Recognition 96, pp. 638-645, 1996.

[15] C. Nastar, The image shape spectrum for image retrieval, Technical Report 3206, INRIA, June 1997.

[16] S. K. Nayar and H. Murase and S. A. Nene, Parametric Appearance Representation, Early Visual Learning, Oxford University Press, Feb. 1996.

[17] Olivetti Research Labs, Face Dataset, $http : //www.cam - orl.co.uk/facedatabase.html$

[18] A . Pentland, B. Moghaddam, and T. Starner, View-based Modular Eigenspaces for face recognition, Proc. Comp. Vision and Patt. Recognition, 1994.

[19] P. Jonathon Phillips, H. Moon, S. A. Rizvi and P. J. Rauss, The FERET Evaluation Methodology for Face-Recognition Algorithms, Proc. Computer Vision and Patt. Recognition, 1997.

[20] S. Ravela, The Tea Crowd Dataset", 2000, Contact: ravela@cs.umass.edu

[21] Rajesh Rao and Dana Ballard, Object Indexing Using an Iconic Sparse Distributed Memory, Proc. International Conference on Computer Vision, pp. 24-31, 1995.

[22] S. Ravela, S. and C. Luo, Appearance-based Global Similarity Retrieval of Images. In Advances in Information Retrieval, W. Bruce Croft (Ed), Kluwer Academic Publishers 2000.

[23] S. Ravela, R. Manmatha, Gaussian Filtered Representations of Images, Encyclopedia of Electrical and Electronic Engineering, John Webster (Editor), John Wiley, 1999

[24] S. Ravela, S. and R. Manmatha, Retrieving Images by Appearance, Proc. of the International Conf. on Computer Vision, (ICCV), Bombay, India, Jan 1998.

[25] S. Ravela, R. Manmatha, and E. M. Riseman, Retrieval from Image Databases using Scale Space Matching. Proc. of the European Conf. on Computer Vision ECCV '96, Cambridge, U.K., pages 273–282,Springer, April 1996

[26] B. M. ter Har Romeny, Geometry Driven Diffusion in Computer Vision, Kluwer Academic Publishers, 1994

[27] F. Samaria and A. Harter, Parameterization of a stochastic model for human face identification. In Proc. of Workshop on Applications of Computer Vision, 1994.

[28] Bernt Schiele and James L. Crowley, Object Recognition Using Multidimensional Receptive Field Histograms, Proc. 4th European Conf. Computer Vision, Cambridge, U.K., April 96.

[29] Schmid, R. Mohr, Local Grayvalue Invariants for Image Retrieval, PAMI (19), No. 5, pp. 530-535, May 1997.

[30] T. Sim, R. Sukthankar, M. Mullin, and S. Baluja, High-Performance Memory-based Face Recognition for Visitor Identification, 1999 (see $http : //www.ri.cmu.edu/pubs/pubs_2772.html$)

[31] D. L. Swets and J. Weng, Using Discriminant Eigen Features for Retrieval, IEEE Trans. Patt. Anal. And Mach. Intel. 18(8): 831-836, 1996. 30.

[32] M. Turk and A. Pentland, Eigen Faces for Recognition, Jrnl. Cognitive Neuroscience, 3:71-86, 1991.

[33] J. Wilder, Face Recognition using transform coding of grayscale projection projections and the neural tree network, In Artificial Neural Networks with Applications in Speech and Vision, R. J. Mammone(ed), pp.520-536, Chapman Hall, 1994

[34] L. Wiskott, J.-M. Fellous, N. Kruger and C. von der Malsburg, Face recognition by elastic bunch graph matching, IEEE Trans Patt. Anal. and Mach. Intell. 17(7):775-779, 1997.

[35] A. P. Witkin, Scale-Space Filtering, Proc. Intl. Joint Conf. Art. Intell., pp. 1019-1023, 1983

[36] W. Zhao, A. Krishnaswamy, R. Chellapa, D. Swets, and J. Weng, Discriminant analysis of principal components for face recognition, In Face Recognition: From Theory to Applications, H. Weschler, P. J. Phillips et. al. (ed), pp 73-85, Springer-Verlag, Berlin, 1998.