A METHODOLOGY FOR EVALUATING ROBUSTNESS OF FACE RECOGNITION ALGORITHMS WITH RESPECT TO VARIATIONS IN POSE ANGLE AND ILLUMINATION ANGLE

Greg Little, Sreekar Krishna, John Black and Sethuraman Panchanathan

Center for Cognitive Ubiquitous Computing
Arizona Sate University
ASU Main, Tempe, AZ - 85287

ABSTRACT

In this paper, we present a methodology for precisely comparing the robustness of face recognition algorithms with respect to changes in pose angle and illumination angle. For this study, we have chosen four widely-used algorithms: two subspace analysis methods (Principle Component Analysis (PCA) and Linear Discriminant Analysis (LDA)) and two probabilistic learning methods (Hidden Markov Models (HMM) and Bayesian Intra-personal Classifier (BIC)). We compare the recognition robustness of these algorithms using a novel database (FacePix) that captures face images with a wide range of pose angles and illumination angles. We propose a method for deriving a robustness measure for each of these algorithms, with respect to pose and illumination angle changes. The results of this comparison indicate that the subspace methods perform more robustly than the probabilistic learning methods in the presence of pose and illumination angle changes.

1. INTRODUCTION

Face Recognition has been an active area of research for the last decade, due to the availability of fast computing systems and an increased level of security requirements in public places. The research has not only lead to the development of improved algorithms, but also the deployment of access control and identity verification systems, based on face recognition. Although there are numerous algorithms today that can achieve acceptable recognition rates on idealized image sets, there exists no algorithm capable of adequately recognizing people in real-world situations.

This paper proposes a methodology for evaluating the robustness of face recognition algorithms with respect to variation in pose and illumination angles. We measure the robustness of recognition algorithms using a face database called FacePix [1], which includes images of faces recorded at precisely measured pose and illumination angles. This comprehensive database, along with the proposed methodology, provides a useful evaluation tool for Face Recognition Systems (FRS).

As discussed in the Section 2, though many attempts have been made to measure robustness of face recognition algorithms, there has been no method that has proven adequate. The research described here demonstrates a methodology for evaluating the robustness of face recognition algorithms. An analysis is carried out on four face recognition algorithms that were identified by [2] as being the most basic and widely used.

The rest of this paper is organized as follows. Section 2 reviews the research that has gone into comparing the robustness of various face recognition algorithms. Section 3 explains in detail the FacePix database that we have used, along with the details of the methodology. The results of the experiments are reported in Section 4 and a discussion of these results is provided in Section 5. Section 6 concludes the paper with a discussion of extensions to our work.

2. RELATED WORK

Computer vision researchers [3] [4] have emphasized the need to find methodologies to characterize the performance of face recognition algorithms, which deal with a huge amount of high-dimensional data.While Micheals et al. [5] and Beveridge et al. [6] employed elaborate statistical techniques in an attempt to obtain evaluation estimates, the FERET face database that they used did not include a wide range of pose and illumination angles, thus limiting the application of their results. In fact, virtually all attempts to measure recognition rates, including [7] [8] and [9] make no attempt to study the robustness of face recognition algorithms with respect to changes in pose angle or illumination angle. In view of the importance of this unanswered question, we evaluate the robustness of face recognition algorithms under varying pose and illumination conditions by using a comprehensive database of face images with precisely measured pose and illumination angles.

greg.little@asu.edu; Ph: 480 854 1681

3. METHODOLOGY

3.1. The FacePix Database

Earlier works on Face Recognition algorithms have used databases such as FERET, XM2VTS, the CMU PIE Database, AT&T, Oulu Physics Database, Yale Face Database, Yale B Database and MIT Database for evaluating the recognition performance of algorithms. Some of these databases provide face images with a wide variety of pose angles and illumination angles. However, none of them use a precisely calibrated mechanism for acquiring pose and illumination angles. To achieve a precise measure of recognition robustness this study uses a database called FacePix, which contains face images with pose and illumination angles annotated in 1 degree increments. Figure 1 shows the apparatus that is used for capturing the face images. A video camera and a spot light are mounted on separate annular rings which rotate independently around a subject seated in the center. Angle markings on the rings are captured simultaneously with the face image in a video sequence, from which the required frames are extracted.

Fig. 1. FacePix capture apparatus

The FacePix database consists of two sets of face images: a set with pose angle variations, and a set with illumination angle variations. Each set may be conceptualized as a 2D matrix of face images, where each matrix has 30 rows (representing 30 different subjects), and 181 columns (representing angles from $-90°$ to $+90°$ at 1 degree increments) [1]. All the face images (elements) in each matrix are 128 pixels wide and 128 pixels high. These images are normalized, such that the eyes are centered on the 57^{th} row of pixels from the top, and the mouth is centered on the 87^{th} row of pixels. The pose angle images appear to rotate such that the eyes, nose, and mouth features remain centered in each image. Also, although the images are down sampled, they are scaled as much horizontally as vertically, thus maintaining their original aspect ratios. Figure 2 provides two examples extracted from the database, showing pose angles and illumination angles ranging from $-90°$ to $+90°$ in steps of $10°$.

Fig. 2. FacePix Database: Pose and Illumination sets

3.2. The Experimental Procedure

We run several experiments on the FacePix database, and combine the results of these experiments to gauge the robustness of the various face recognition algorithms. Each experiment measures the degradation in recognition rate as an algorithm attempts to recognize probe (test) images farther and farther (in terms of pose or illumination angle) from the gallery (training) set. Each such experiment may be conceptualized as a function, with the following inputs:

- Algorithm to test: PCA [10], LDA [11], BIC [12], or HMM [13]
- Database set: Pose angle, or Illumination angle
- Gallery set list: One or more columns from the given database set, e.g., all the images at pose angles $-90°$, $0°$, and $+90°$ (NOTE: In this scheme, each gallery set contains only one image of each subject. However, some of the algorithms require multiple versions of each image in each gallery set. In such cases, we artificially manufacture 3 additional versions of each gallery image. One of these images is a low-pass filtered version of the original image, while two of these images are noisy versions of the original image.)
- Probe set: The entire 2D matrix of the database set

The output of this function is the "distance" of each probe image to the "nearest" gallery image of the gallery set(s).

Using these distances, we produce a rank ordering of the 30 subjects for each probe image (A rank of 0 indicates correct recognition of the probe image). These ranks then provide a basis for computing a measure of robustness (R) for the algorithm trained with the chosen gallery sets. The robustness at a particular angle θ is given by

$$R(\theta) = 1 - \frac{2}{N-1}\left(\frac{1}{N}\sum_{i=1}^{N} r_{\theta i}\right) \quad (1)$$

where,
N is the number of subjects in the database.

[1] The illumination set is captured with the subject looking directly into the camera while the light source is moved around the subject

$r_{\theta i}$ is the rank that is assigned for the i^{th} subject at the pose or illumination angle θ (this value ranges from 0 to $N-1$).

A Robustness value of 1 means that the recognition was accurate, while a value of 0 means that the recognition was no better than guessing randomly.

4. RESULTS

Figure 3 shows the robustness curves for all four face recognition algorithms, as a function of pose and illumination angles. The solid line illustrates the pose angle robustness, while the dotted line illustrates the illumination angle robustness. Each row in Figure 3 corresponds to one algorithm, and each column corresponds to a different training set. The first column shows the results when the algorithms were trained with just the $0°$ images, while the second column shows the results when trained with $-90°$, $0°$, and $90°$ images. The third column shows the results when trained with $-90°$, $-45°$, $0°$, $+45°$, and $+90°$. Table 1 and Table 2 show the average robustness across pose and illumination variations respectively, while Table 3 and Table 4 show the average recognition rate across pose and illumination variations.

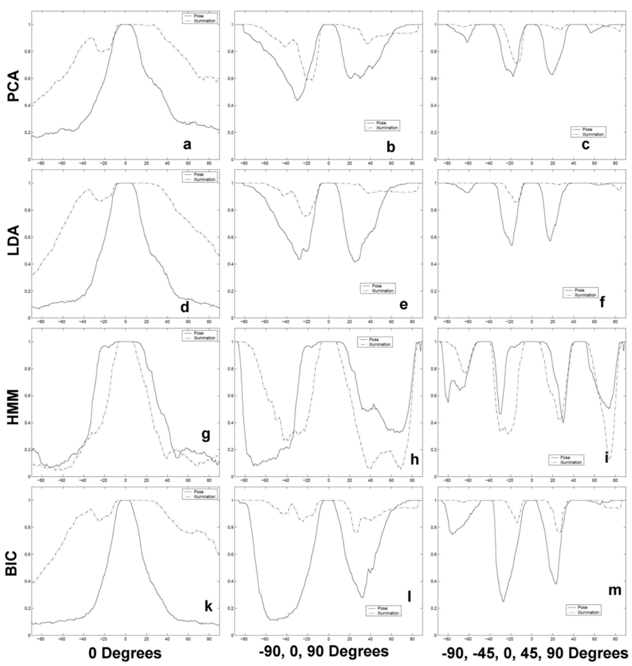

Fig. 3. Robustness curves

5. DISCUSSION OF THE RESULTS

The robustness graphs are a direct measure of the recognition reliability of the algorithms. This is evident from the

	$0°$	$-90°, 0°, 90°$	$-90°, 0°, 90°$ $45°$ $-45°$
PCA	0.4306	0.8039	0.9178
LDA	0.3485	0.7985	0.9245
HMM	0.4391	0.5764	0.8317
BIC	0.3003	0.6250	0.8443

Table 1. Average Robustness for Pose

	$0°$	$-90°, 0°, 90°$	$-90°, 0°, 90°$ $45°$ $-45°$
PCA	0.7715	0.9136	0.9735
LDA	0.7820	0.949	0.9888
HMM	0.3278	0.5483	0.7735
BIC	0.8108	0.9388	0.9740

Table 2. Average Robustness for Illumination

	$0°$	$-90°, 0°, 90°$	$-90°, 0°, 90°$ $45°$ $-45°$
PCA	20.74%	50.53%	71.66%
LDA	20.70%	56.92%	78.67%
HMM	31.68%	41.27%	63.50%
BIC	18.42%	45.19%	69.47%

Table 3. Recognition rate for Pose

	$0°$	$-90°, 0°, 90°$	$-90°, 0°, 90°$ $45°$ $-45°$
PCA	48.84%	71.71%	90.33%
LDA	53.04%	79.52%	94.92%
HMM	19.26%	37.38%	59.37%
BIC	49.80%	79.10%	93.54%

Table 4. Recognition rate for Illumination

direct relationship between the average robustness and the average recognition rate for both pose and illumination variations. Comparing the robustness of HMM with PCA, LDA and BIC, it is clear that HMM is the poorest performing algorithm.

From the roll off regions of the robustness curves, it is clear that the two subspace methods (PCA and LDA) have a more gradual roll off than the probabalistic methods (HMM and BIC); accordingly they have a better recognition rate across changes in both pose and illumination angles.

The roll off rate is higher near 0° (i.e. frontal views) than at the edges (i.e. profile views). This suggests that better overall robustness might be achieved by using a more densely spaced gallery set around the frontal region than towards the profile regions.

Comparing the results from each of the algorithms with respect to pose angle variance, LDA ranks first, followed by PCA, close behind is BIC with HMM being the last. For illumination angle variance, LDA performs the best, followed by BIC, PCA and HMM. Thus, LDA performs best with respect to changes in both pose angle and illumination angle.

The robustness curves have a symmetrical structure, as expected, except near the -30° to -10° region of the illumination curves (For example, see Figure 3(a)-PCA 0°). There seems to be a anomalous drop in the robustness curve. Since this is consistent across 3 of the 4 algorithms, we suspect an anomaly in our FacePix database. This is currently under investigation.

6. CONCLUSION AND FUTURE WORK

In this paper, we have presented a novel methodology for evaluating the recognition robustness of face recognition algorithms with respect to changes in pose angle and illumination angle. The FacePix database has been used to measure the robustness of four popular face recognition algorithms. Work is currently in progress to increase the size of the FacePix database. As a part of future research in this direction, we intend to apply statistical analysis techniques to the robustness data to further explore the recognition robustness of face recognition algorithms.

7. REFERENCES

[1] Black J; Gargesha M; Kahol K; Kuchi P; Panchanathan S;, “A framework for performance evaluation of face recognition algorithms,” *ITCOM, Internet Multimedia Systems II, Boston*, July 2002.

[2] Zhao W; Chellappa R; Rosenfeld A; Phillips P;, “Face recognition: A literature survey,” *UMD CfAR Technical Report CAR-TR-948*, 2000.

[3] Haralick R;, “Performance characterization in computer vision,” *CVGIP*, vol. 60.

[4] Bowyer K W; Phillips J; (editors), “Emperical evaluation techniques in computer vision,” *IEEE Computer Society*, 1998.

[5] Micheals R J; Boult T E;, “Efficient evaluation of classification and recognition systems,” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001. CVPR 2001.*, vol. 1, pp. I50–I57, Dec. 2001.

[6] Beveridge J R; She K; Draper B A; Givens G H;, “A nonparametric statistical comparison of principal component and linear discriminant subspaces for face recognition,” *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001. CVPR 2001.*, vol. 1, pp. I–535 – I–542, Dec. 2001.

[7] Jian Li; Shaohua Zhou; Shekhar C;, “A comparison of subspace analysis for face recognition,” *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003.*, vol. 3, pp. III 121–4, April 2003.

[8] Matas J; Hamouz M; Jonsson K; Kittler J; Li Y; Kotropoulos C; Tefas A; Pitas I; Teewoon Tan; Hong Yan; Smeraldi F; Bigun J; Capdevielle N; Gerstner W; Ben-Yacoub S; Abeljaoued Y; Mayoraz E;, “Comparison of face verification results on the xm2vtfs database,” *Proceedings. 15th International Conference on Pattern Recognition, 2000.*, vol. 4, pp. 858–863, Sep 2000.

[9] Phillips P J; Hyeonjoon Moon;, “Comparison of projection-based face recognition algorithms,” *1997 IEEE International Conference on Systems, Man, and Cybernetics, 1997. 'Computational Cybernetics and Simulation'.*, vol. 5, pp. 4057–4062, Oct. 1997.

[10] Turk M; Pentland A;, “Face recognition using eigenfaces,” *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 586–591, 1991.

[11] Etemad K; Chellappa R;, “Discriminant analysis for recognition of human face images,” *Journal of Optical Society of America*, pp. 1724–1733, 1997.

[12] Moghaddam B; Pentland A;, “Probabilistic visual learning for object representation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19(7), pp. 696–710, July 1997.

[13] Nefian A V; Hayes III M H;, “Hidden markov models for face detection and recognition,” *IEEE International Conference on Image Processing*, vol. 1, pp. 141–145, October 1998.