

EVALUATION OF STATE-OF-THE-ART ALGORITHMS FOR REMOTE FACE RECOGNITION

Jie Ni and Rama Chellappa

Department of Electrical and Computer Engineering and Center for Automation Research, University of Maryland, College Park, MD 20742, USA

ABSTRACT

In this paper, we describe a remote face database which has been acquired in an unconstrained outdoor environment. The face images in this database suffer from variations due to blur, poor illumination, pose, and occlusion. It is well known that many state-of-the-art still image-based face recognition algorithms work well, when constrained (frontal, well illuminated, high-resolution, sharp, and complete) face images are presented. In this paper, we evaluate the effectiveness of a subset of existing still image-based face recognition algorithms for the remote face data set. We demonstrate that in addition to applying a good classification algorithm, consistent detection of faces with fewer false alarms and finding features that are robust to variations mentioned above are very important for remote face recognition. Also setting up a comprehensive metric to evaluate the quality of face images is necessary in order to reject images that are of low quality.

Index Terms— Remote, Face Recognition.

1. INTRODUCTION

During the past two decades, face recognition (FR) has received great attention and tremendous progress has been made. Currently, most of the FR algorithms are applied to databases which are collected at close range (less than a few meters) and under different levels of controlled environments, such as in CMU PIE [1], FRGC/FRVT [2], FERET [3] data sets. Yet, in many scenarios in real life applications, we cannot control the acquisition of face images; the images we get can suffer from poor illumination, blur, occlusion etc. which are great challenges to current FR algorithms. In [4], Yao et al. describe a face video database, UTK-LRHM, acquired from long distances and with high magnifications. They address the magnification blur to be the major degradation. Huang et al. [5] presented a database named "Labeled Faces in the Wild" (LFW) which has been collected from the web. Although it has "natural" variations in pose, lighting,

expression, etc., there is no guarantee that such a set accurately captures the range of variation found in the real world [6]. Besides, most objects in LFW only have one or two images which may be not enough to evaluate different FR experiments.

In order to study and develop more robust algorithms for FR, we have put together a remote face database in which a significant number of images are taken from long distances and under unconstrained outdoor environments. The quality of the images differs in the following aspects: the illumination is not controlled and is often pretty bad in extreme conditions; there are pose variations and faces are also occluded as the subjects are not cooperative [7]; finally, the effects of scattering [7] and high magnification resulting from long distance contribute to the blurriness of face images. We manually cropped and labeled the face images according to different illumination conditions (good, bad and really bad), pose (frontal and non-frontal), blur or no-blur etc in a systematic way so that users can conveniently select the desired images for their experiments.

We evaluated two state-of-the-art FR algorithms on this remote face database including a baseline algorithm and the recently developed algorithm based on sparse representation [8]. Based on our limited experiments using the remote face data set, we make the following observations: detection of faces and subsequent extraction of robust features is as important as the recognition algorithms that are used. The performance of recognition algorithms improves gradually as the number of gallery images increases. The recognition accuracy varies from low thirties to mid nineties depending on the quality of images and the number of available gallery images. It is important to design a quality metric so that face images that have low quality can be rejected.

The organization of this paper is as follows; In Section 2, we describe the remote face database collected by the authors' group. Section 3 briefly describes the algorithms that are evaluated and corresponding recognition results. Finally conclusions are given in Section 4.

This work was partially supported by the ONR MURI Grant N00014-08-1-0638.

2. REMOTE FACE DATABASE DESCRIPTION

The distance from which the face images were taken varies from 5m to 250m under different scenarios. Since we could not reliably extract all the faces in the data set using existing state-of-the-art face detection algorithms and the faces only occupied small regions in large background scenes, we manually cropped the faces and rescaled to a fixed size. The resulting database for still color face images contains 17 different individuals and 2106 face images in total. The number of faces per subject varies from 48 to 307. All images are 120 by 120 pixel png images. Most faces are in frontal poses.

We manually labeled the faces according to different illumination conditions, occlusion, blur and so on. In total, the database contains 688 clear images, 85 partially occluded images, 37 severely occluded images, 540 images with medium blur, 245 with sever blur, and 244 in poor illumination condition. The remaining images have two or more conditions, such as poor lighting and blur, occlusion and blur etc. These face images are not used in the following experiments. Figure 1 shows some sample images from the database: These face images show large variations, some of which are not easily recognizable even for humans.

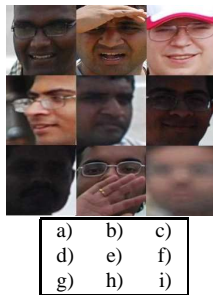


Fig. 1. Sample images from the remote face database: a) clear; b) and c) partially occluded; d) and e) have pose variations; f) and g) poorly illuminated; h) severely occluded; i) severely blurred.

3. ALGORITHMS AND EXPERIMENTS

In this section, we evaluate two state-of-the-art FR algorithms on the remote face database, and compare their performance.

3.1. Experiments with a Baseline Algorithm

This experiment involves using clear images from the database as gallery images. We gradually increase the number of gallery of faces from one to fifteen images per subject. Each time the gallery images are chosen randomly; and we repeat the experiments five times and take the average to arrive at the final recognition result.

3.1.1. Baseline Algorithm

A baseline recognition algorithm involving Kernel Principle Component Analysis (KPCA) [9], Linear Discriminate Analysis (LDA) [10] and a Support Vector Machine (SVM) [11] is used in this experiment.

The LDA is a well-known method for feature extraction and dimensionality reduction in pattern recognition and classification tasks. The basic idea is to maximize the between-class distance and minimize the within-class distance. In order to make the within-class scatter matrix nonsingular, we used the KPCA as a dimensionality reduction method to project the raw data onto a feature space with much lower dimension. Yet LDA can still fail when the number of samples is small. Especially, LDA does not work when there is only one image per subject. Hence we use the Regularized Discriminate Analysis (RDA) [12] to eliminate this effect. Also we added the mirror reflection images when there is only one image per subject in the gallery. The resulting low-dimensional discriminate features are fed into SVM for classification.

3.1.2. Handling illumination variation

Even for clear images, changes induced by illumination can make faces images of the same subject far apart than images of different subjects [13]. Hence we used estimates of albedo in the hope of mitigating the illumination effect. Albedo is the fraction of light that a surface point reflects when it is illuminated. It is an intrinsic property that depends on the material properties of the surface [7], and is invariant to changes in illumination conditions which makes it useful for illumination-insensitive matching of objects. The albedo is estimated using the method of minimum mean square error criterion [14]. The illumination-free albedo image is then used as input to the baseline algorithm. Figure 2 shows the results of albedo estimation for two face images acquired from 50 meters [7].



Fig. 2. Results of albedo estimation. Left: original images; Right: Estimated albedo images.

3.1.3. Experimental Results

In the first experiment, all the remaining clear images except the gallery images are selected for testing. To make a comparison, we used both albedo maps and intensity images as inputs for this experiment. The results are given in figure 3. All the parameters for KPCA, LDA and SVM are well tuned.

It is found that intensity images outperform albedo maps although the albedo map is intended to compensate for illumination variations. One reason may be that, the face images in the database are sometimes a bit away from frontal. As albedo estimation needs a good alignment between the observed images and the ensemble mean, the estimated albedo map is erroneous. Besides, extreme illumination conditions resulting in especially "dark" faces, also creates challenges as we cannot get a good initial estimate of the albedo. On the other hand, intensity images contain texture information which can partly counteract variations induced by pose.

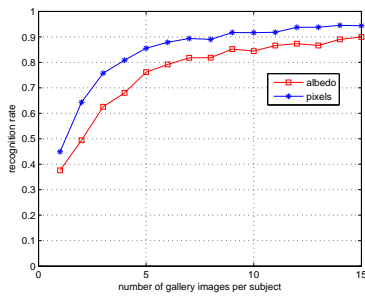


Fig. 3. Experiment 1: comparison between FR using albedo maps and intensities in the baseline algorithm.

Next, we changed the test images to be poorly illuminated, medium blurred, severely blurred, partially occluded and severely occluded respectively. The gallery still contains clear images as in experiment 1, the number varying from 1 to 15 images per subject. We used intensity images as input. The results are shown in figure 4, and the results from experiment 1 are also added for comparison.

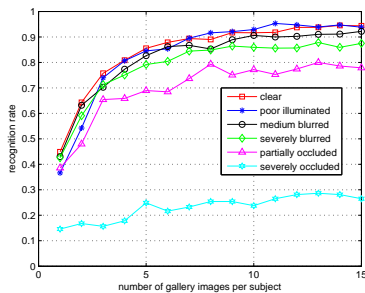


Fig. 4. Experiment 2: performance of baseline as the condition of test images varies.

From figure 4, it is clearly seen that the degradations in the test images decreases the performance of the system, especially when the faces are occluded and severely blurred.

3.2. Experiments using Sparse Representation

A sparse representation-based FR algorithm was proposed in [8] which is robust to occlusion. For evaluating this algorithm in this experiment, we used the implementation by Pillai et al. [15] which is a modification of [8]. It uses a modified BPDN (Basis Pursuit DeNoising) algorithm to get a sparser coefficient vector to represent the test image. For each test image, we compute its SCI (Sparsity Concentration Index) [8] value and reject the image if it is below certain threshold.

For this experiment, 14 subjects, 10 clear images per subject were selected to form the gallery set, and the test images were selected to be clear, blurred, poorly illuminated and occluded respectively. The experiment was repeated several times and the average was taken. We compare the results using sparse representation and the baseline algorithm in figure 5. To make a fair comparison, we use the same feature from KPCA and LDA in the baseline for sparse representation.

It turns out that when no rejection is allowed, the recognition accuracy of sparse representation-based method is low which may be due to the fact that the gallery does not have as much variation as the test set. As we increase the threshold of SCI, more test images with low quality are rejected and hence the recognition rate increases; the rejection rates in figure 5 are 6%, 25.11%, 38.46% and 17.33% when the test images are clear, poorly lighted, occluded and blurred respectively. Based on the results, the sparse representation-based FR algorithm has an obvious advantage than the baseline algorithm when there is occlusion in the test images.

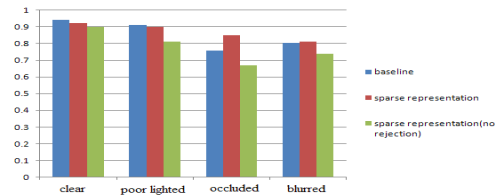


Fig. 5. Experiment 3: comparison between sparse representation and baseline algorithms: clear, poorly lighted, occluded and blurred stand for the conditions of test images.

3.3. Adding Degraded Images in the Gallery

In this experiment, we selected test images to be blurred, poorly illuminated and occluded, and added corresponding type of degraded images into the gallery set. To make a comparison with the result in experiment 3, we first kept the 140 clear images in the gallery, and moved one third of the test images into the gallery set for each case; also we divided the test images from experiment 3 into two for each case, using one half as gallery and the other half for testing. The result is

shown in figure 6. The baseline algorithm is used for recognition.

The result shows that for the recognition of degraded images, adding the corresponding type of variation into the gallery can improve the performance.

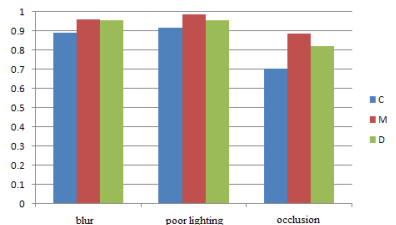


Fig. 6. Experiment 4: C, M, and D stand for using all clear, mixture of clear and degraded, all degraded images respectively as gallery images. Blur, poor lighting and occlusion represent the type of degradation that test images have in each case.

4. CONCLUSIONS AND FUTURE WORK

In this study, we described a remote face database we built and described the performance of state-of-the-art FR algorithms on it. The results demonstrate that recognition rate decreases as the face images acquired remotely are degraded. The evaluations reported here can provide guidance for further research in remote face recognition.

In our future work, we plan to address the following problems: 1) use image restoration/denoising algorithms to improve the quality of the image; 2) incorporate other robust texture features or obtain a better estimate of albedo for recognition; 3) develop a more comprehensive quality metric to reject low quality images in order to make the recognition system more effective in practical acquisition condition.

5. REFERENCES

- [1] T. Sim, S. Baker, and M. Bsat, "The cmu pose, illumination, and expression database," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 1615–1618, Dec. 2003.
- [2] P.J. Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, San Diego, CA, June 2005, pp. 947–9546.
- [3] P.J. Phillips, H. Wechsler, J. Huang, and P.J. Rauss, "The feret database and evaluation procedure for face-recognition algorithms," *Image and Vision Computing*, vol. 16, pp. 295–306, 1998.
- [4] Y. Yao, B. Abidi, N. Kalka, N. Schmid, and M. Abidi, "Improving long range and high magnification face recognition: database acquisition, evaluation, and enhancement," *Computer Vision and Image Understanding*, vol. 111, pp. 111–125, 2008.
- [5] G. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," *University of Massachusetts, Amherst, Technical Report 07-49*, 2007.
- [6] N. Pinto, J. DiCarlo, and D. Cox, "How far can you get with a modern face recognition test set using only simple features?," in *Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, Miami, FL, June 2009, pp. 2591–2568.
- [7] R. Chellappa, "Annual progress report: Muri on remote multi-modal biometrics for maritime domain," *University of Maryland, College Park, MD, Technical Report*, 2009.
- [8] J. Wright, A. Ganesh, A. Yang, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, pp. 210–227, Feb. 2009.
- [9] M.-H. Yang, "Kernel eigenfaces vs. kernel fisherfaces: face recognition using kernel methods," in *IEEE International Conference on Automatic Face and Gesture Recognition*, Washington, DC, October 2002, pp. 215–220.
- [10] K. Etemad and R. Chellappa, "Discriminant analysis for recognition of human face images," *Journal of the Optical Society of America*, vol. 14, pp. 1724–1733, August 1997.
- [11] G. Guo, S.Z. Li, and K. Chan, "Face recognition by support vector machines," in *IEEE International Conference on Automatic Face and Gesture Recognition*, Grenoble, France, October 2000, pp. 196–201.
- [12] J. Friedman, "Regularized discriminant analysis," *Journal of the American Statistical Association*, vol. 84, pp. 165–175, 1989.
- [13] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: the problem of compensating for changes in illumination direction," *IEEE Transactions on pattern Analysis and Machine Intelligence*, vol. 31, pp. 721–732, July 1997.
- [14] S. Biswas, G. Aggarwal, and R. Chellappa, "Robust estimation of albedo for illumination-invariant matching and shape recovery," in *Proc. Intl. Conf. Computer Vision*, Rio de Janeiro, Brazil, October 2007, pp. 1–8.
- [15] J. Pillai, V. Patel, and R. Chellappa, "Sparsity inspired selection and recognition of iris images," in *IEEE Third International Conference on Biometrics: Theory, Applications and Systems*, Crystal City, VA, Sept. 2009, pp. 1–6.