

## **A Comparative Evaluation of Average Face on Holistic and Local Face Recognition Approaches**

### **Author**

Zhao, S, Zhang, X, Gao, Y

### **Published**

2008

### **Conference Title**

Proceedings - International Conference on Pattern Recognition

### **DOI**

<https://doi.org/10.1109/ICPR.2008.4761430>

### **Copyright Statement**

© 2008 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

### **Downloaded from**

<http://hdl.handle.net/10072/22905>

### **Link to published version**

<http://ieeexplore.ieee.org/servlet/opac?punumber=4740202>

### **Griffith Research Online**

<https://research-repository.griffith.edu.au>

# A Comparative Evaluation of Average Face on Holistic and Local Face Recognition Approaches

Sanqiang Zhao, Xiaozheng Zhang, and Yongsheng Gao  
School of Engineering, Griffith University, Nathan Campus, QLD 4111, Australia  
{s.zhao, x.zhang, yongsheng.gao}@griffith.edu.au

## Abstract

*This study focuses on a recent paper "100% Accuracy in Automatic Face Recognition" published on Science, in which an "Average Face" is proposed and claimed to be capable of dramatically improving performance of a face recognition system. To reveal its working mechanism, we perform the averaging process using pose-varied synthetic images generated from 3D face database and conduct a comparative study to observe its effectiveness on holistic and local face recognition approaches. Two representative methods, i.e. Eigenface and Local Binary Pattern (LBP) are employed to perform the experiments. It is interesting to find from our experiments that the performance of the "Average Face" is not independent of the face recognition approaches. Although face averaging increases the recognition accuracy of Eigenface method, it impairs the performance of LBP method.*

## 1. Introduction

As one of the most active applications in image analysis and computer vision, automatic face recognition has received significant attention from both academic and industrial communities. Despite the large numbers of commercial systems and forensic applications, face recognition is still far from perfect, especially under unconstrained conditions where practical interferences from facial expression, illumination and pose variations have to be counteracted [1].

Existing face recognition techniques are roughly classified into holistic approaches and local approaches. The representative holistic approaches include Eigenface [2] and Fisherface [3]. Deriving feature information from the whole face image for classification, holistic approaches usually suffer from environmental variations in practice. The local approaches, such as Elastic Bunch Graph Matching (EBGM) [4] and Local Binary Pattern (LBP) [5], extract information from

local facial features to distinguish faces, and have the advantage of robustness to environmental changes. Evaluations of face recognition techniques have been reported in the literature, both from academic [1] and commercial societies [6]. Recently, psychological researchers from University of Glasgow proposed an "Average Face" [7] to improve face recognition approaches. The "Average Face" is constructed from real-world images collected from diverse sources, covering a natural range of variations in pose, expression, age and hairstyles. Although the conditions of these images vary, pose variation is confined within approximately  $\pm 30^\circ$  to make all feature points visible in the image [7s]. The researchers experimented with a gallery set comprising an average of 9 different photos for each of 3628 individuals, and a probe set composed of 20 photos for each of 25 persons. They reported that averaging the 20 images of each person can increase the recognition accuracy from 54% to 100% [7]. However, the authors did not provide the details of the algorithm employed in the experiment. The theory behind the face averaging process is also unclear.

As an attempt to reveal the working mechanism of the average face, this paper is to find out what types of face recognition approaches are benefited most from face averaging. We only consider the process of averaging face images with different poses because pose variation is one of the most prominent issues in face recognition [8]. Other variations such as illumination and expression changes can be tested separately. Synthetic images generated from the USF Human ID 3D database [9] are utilized to conduct the averaging process. Two representative face recognition methods, Eigenface from holistic approaches and LBP from local approaches, are selected in our experiments to evaluate the effectiveness of averaging face images with pose variations. Interestingly, experimental results demonstrate that face averaging increases the recognition accuracy of Eigenface method, but decreases that of LBP method. It is possible that the process of face

averaging may not necessarily improve all the face recognition systems, but is to a certain extent dependent on the face recognition approaches being employed in practice.

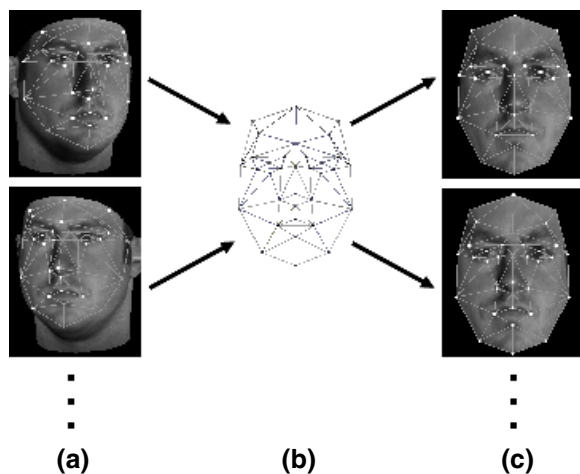
The remainder of this paper is organized as follows. Section 2 introduces the process of constructing an average face. Section 3 presents the experimental settings to reveal the effectiveness of face averaging. Section 4 describes our comparative experiments using Eigenface and LBP methods. The last section concludes the paper.

## 2. Average Face

This section provides a brief description of constructing an average face from multiple different face images of a same person. The details of the image averaging procedure are reported in [7s].

### 2.1. Morphing individual images

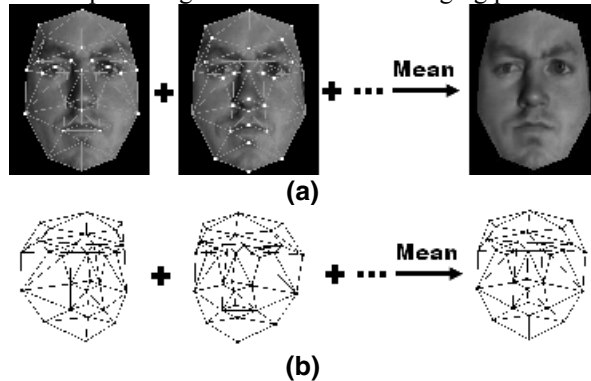
All the face images from the same person are first labeled with 34 feature points (xy-coordinates) for shape capturing. These feature points have exact physical meanings on a face image (see Fig. 1(a)). Some feature points are located at corners of the eyes, tip of the nose, corners of the mouth, etc. A fixed 34-point template, illustrated in Fig. 1(b), is pre-defined for individual image morphing. The bi-cubic interpolation is used to morph all the face images to the template, with the aid of labeled feature points. Fig. 1 displays this morphing process.



**Fig. 1. Face image morphing. (a) Original face images labeled with 34 feature points. (b) Fixed 34-point template. (c) Face images after morphing.**

### 2.2. Averaging textures and shapes

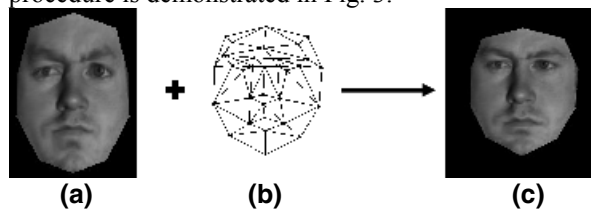
After all the face images are registered through morphing themselves to the template, we get a group of morphed textures and their specific shapes. All the face textures and face shapes are then averaged into a mean texture and a mean shape respectively. The average texture is obtained by calculating the mean gray-level value at each pixel, and the average shape is obtained by calculating the mean xy-coordinates of each feature point. Fig. 2 illustrates this averaging process.



**Fig. 2. Face texture and shape averaging. (a) Texture averaging. (b) Shape averaging.**

### 2.3. Constructing average face

Finally, the average face of a person is produced through morphing the person's average texture to his/her average shape. Image adjustment methods can be employed to enhance the final photograph. This procedure is demonstrated in Fig. 3.



**Fig. 3. Average face morphing. (a) Average texture. (b) Average shape. (c) Average face from morphing.**

## 3. Experimental settings

### 3.1. Face database

To reveal the working mechanism of the average face against pose variations, we use the USF Human ID 3D database [9] in our experiments. This database contains the 3D coordinates of each face, as well as its associated texture map. The reason of using a 3D database is two folded. First, by performing a one-off labe-

ling of 3D face scans with 34 feature points directly (see the middle image in Fig. 4), we can get both accurate 2D feature points and pose-varied 2D images through 3D-to-2D projection. This not only avoids the efforts of labeling lots of 2D images, but also minimizes the labeling errors that could be involved during labeling 2D images otherwise. Second, the selection of a 3D database can also exclude possible illumination, expression and aging variations, and thus ensure our experiments exclusively dealing with pose variations only.

Synthetic 2D images of 50 people in the database with pose variations but subject to the same uniform illumination are generated using OpenGL rendering. Fig. 4 presents three samples of individuals from the database. For each person, the facial pose is varied by incrementing the yaw angle (left-right) within  $\pm 30^\circ$  range in steps of  $1^\circ$ . This yields a total of  $61 \times 50 = 3050$  images from 50 individuals. All the images are normalized (in scale and rotation) and cropped to  $100 \times 100$  pixels based on the positions of two eyes. Fig. 5 provides examples of normalized images.



**Fig. 4. Three sample individuals from the USF Human ID 3D database [9]. The individual in the middle is marked with 34 feature points.**



**Fig. 5. Examples of normalized images.**

### 3.2. Evaluation protocol

In our experiments, the face images of frontal view (i.e. yaw angle is  $0^\circ$ ) are used as the gallery set. For the probe set, we randomly select 20 angles that are evenly distributed within the pose range  $[-30^\circ, +30^\circ]$ . The face images of the selected angles are collected to form a probe set. This random selection is repeated for ten times. Eventually we get ten groups of probe sets of face images. In each group, two recognition rates are recorded. The first recognition rate  $R_{all}$  is obtained from matching all the probe images in the group against the gallery set. The second recognition rate  $R_{avg}$  is obtained using each person’s average face as the probe image, which is synthesized via the method in

Section 2 using 20 images of that person. The comparison of  $R_{all}$  and  $R_{avg}$  can reflect the effectiveness of the average face for face recognition.

## 4. Experimental results

### 4.1. Eigenface results

As a representative holistic approach, Eigenface [2] is commonly considered as the baseline algorithm for face recognition. The comparative recognition rates of Eigenface method are listed in Table 1. The number of eigenvectors in this experiment is 40. It can be clearly observed from the table that the average faces obtain a consistently higher recognition rate than using all the faces with pose variations. The average improvement of close to 10% recognition accuracy is in compliance with the results from [7].

**Table 1. Recognition rates of Eigenface**

Probe sets	$R_{all}$ (%)	$R_{avg}$ (%)
Group 1	69.22	78.00
Group 2	75.53	78.00
Group 3	64.78	80.00
Group 4	68.78	72.00
Group 5	79.79	80.00
Group 6	65.33	86.00
Group 7	75.58	78.00
Group 8	70.11	82.00
Group 9	71.44	80.00
Group 10	65.88	80.00
<b>Average</b>	<b>70.64</b>	<b>79.40</b>

### 4.2. LBP results

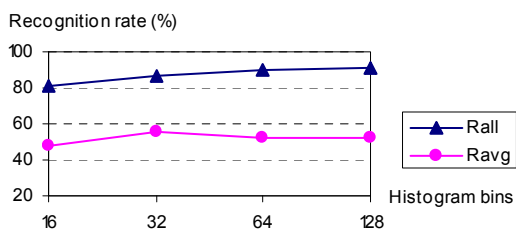
The LBP operator [5] thresholds the pixels of a local neighborhood around each position with the value of the center pixel, and binomially encodes the results as a local image descriptor. In order to preserve spatial information, a face image is divided into several non-overlapping rectangular sub-regions. A spatial histogram, which concatenates the histograms of all the sub-regions, is employed to represent the face. In this experiment, we use the radius ( $R=1$ ) and the sampling points ( $P=8$ ), with  $10 \times 10$ -sized sub-regions and 32 histogram bins. The comparative recognition rates of LBP method are listed in Table 2. It is interesting to find from the table that the performance of average faces is much worse than that of original faces. The reduction of recognition rates in some probe sets are even over 30% (see Group 1, 2, 4, 5 and 7).

It is possible that the parameters of LBP are not finely tuned in favor of average faces. To rule out this possibility, we vary the size of partitioned sub-regions

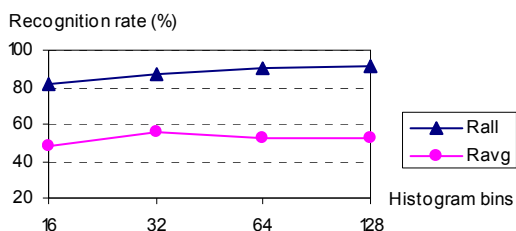
and the number of histogram bins and perform the experiments again on the probe set of Group 1. Fig. 6 and 7 illustrate the recognition rates of LBP with different sub-region sizes and different histogram bins respectively. In all these settings, the results of average faces are still inferior to those of original faces.

**Table 2. Recognition rates of LBP**

Probe sets	$R_{all}$ (%)	$R_{avg}$ (%)
Group 1	86.67	56.00
Group 2	87.41	56.00
Group 3	80.89	58.00
Group 4	83.89	48.00
Group 5	90.00	56.00
Group 6	83.22	72.00
Group 7	88.84	52.00
Group 8	82.33	64.00
Group 9	88.89	62.00
Group 10	85.18	62.00
<b>Average</b>	<b>85.73</b>	<b>58.60</b>



**Fig. 6. Comparative LBP results with 32 histogram bins on the set of Group 1.**



**Fig. 7. Comparative LBP results with 10x10-sized sub-regions on the set of Group 1.**



**Fig. 8. Averaging makes subtle features diluted or vanished. The two moles marked by the white boxes are indiscernible in the average face (far right).**

Our experiments demonstrated that the performance of the average face is dependent on the face recognition approaches being applied. The low recognition accuracy of LBP on average faces may stem from the face averaging process, in which salient local features

(such as eyes, nose and mouth) are distorted from their original appearances (as this is a pure 2D averaging). Especially, some subtle features (e.g., moles) are either diluted or vanished because of averaging (see Fig. 8). These local features are more important in local face recognition approaches than in holistic approaches. Any modifications on local features will significantly affect the performance of local approaches.

## 5. Conclusions

This paper conducted a preliminary investigation on the working mechanism of the "Average Face" recently published on Science. Experimental results show that the average face does not necessarily improve the performance of all face recognition approaches. It is dependent on the face recognition algorithms being employed in practice. At this stage, the pose variation in our experiments is only confined in yaw rotations. The performance of the average face with other in-depth rotations and different illuminations will be our investigations in the future.

## Acknowledgement

This research was supported by the Australian Research Council (ARC) Discovery Grants DP0451091 and DP0877929.

## References

- [1] W. Zhao, R. Chellappa, P.J. Phillips, and A. Rosenfeld, "Face Recognition: A Literature Survey," *ACM Computing Surveys*, 35:399-459, 2003.
- [2] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Journal of Cognitive Neuroscience*, 3:71-86, 1991.
- [3] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *TPAMI*, 19:711-720, 1997.
- [4] L. Wiskott, J.M. Fellous, N. Krüger, and C. von der Malsburg, "Face Recognition by Elastic Bunch Graph Matching," *TPAMI*, 19:775-779, 1997.
- [5] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face Description with Local Binary Patterns: Application to Face Recognition," *TPAMI*, 28:2037-2041, 2006.
- [6] P.J. Phillips, W.T. Scruggs, A.J. O'Toole, P.J. Flynn, K.W. Bowyer, C.L. Schott, and M. Sharpe, "FRVT 2006 and ICE 2006 Large-Scale Results," *NISTIR 7408*, 2007.
- [7] R. Jenkins and A.M. Burton, "100% Accuracy in Automatic Face Recognition," *Science*, 319:435-435, 2008.
- [7s] R. Jenkins and A.M. Burton, "Supporting Online Material for '100% Accuracy in Automatic Face Recognition'," *Science*, 319, 2008.
- [8] S. Zhao and Y. Gao, "Automated Face Pose Estimation Using Elastic Energy Models," *Proceedings of ICPR*, 4:618-621, 2006.
- [9] V. Blanz and T. Vetter, "A Morphable Model for the Synthesis of 3D Faces," *SIGGRAPH'99*, 187-194, 1999.