



UNIVERSITY
OF WOLLONGONG
AUSTRALIA

University of Wollongong
Research Online

Faculty of Informatics - Papers (Archive)

Faculty of Engineering and Information Sciences

2009

Face recognition from single sample based on human face perception

Ce Zhan

University of Wollongong, czhan@uow.edu.au

Wanqing Li

University of Wollongong, wanqing@uow.edu.au

Philip Ogunbona

University of Wollongong, philipo@uow.edu.au

Publication Details

Zhan, C., Li, W. & Ogunbona, P. (2009). Face recognition from single sample based on human face perception. *International Conference Image and Vision Computing New Zealand* (pp. 56-61). Wellington, New Zealand: IEEE.

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library:
research-pubs@uow.edu.au

Face recognition from single sample based on human face perception

Abstract

Although research show that human recognition performance for unfamiliar faces is relatively poor, when the sample is always available for analysis and becomes "familiar", people are able to recognize a previous unknown face from single sample. In this paper, a method is proposed to deal with the one sample per person face recognition problem based on the process how unfamiliar faces become familiar to people. Particularly, quantized local features which learnt from generic face dataset are used in the proposed method to mimic the prototype effect of human face recognition. Furthermore, a landmark-based scheme is introduced to quantify the distinctiveness of each facial component for the sample face, then the difference between the sample and the average face is emphasized by weighting face regions according to the gained distinctiveness. The experiments on ORL and FERET face databases demonstrate the efficiency of the proposed method.

Keywords

single, sample, human, recognition, face, perception

Disciplines

Physical Sciences and Mathematics

Publication Details

Zhan, C., Li, W. & Ogunbona, P. (2009). Face recognition from single sample based on human face perception. International Conference Image and Vision Computing New Zealand (pp. 56-61). Wellington, New Zealand: IEEE.

Face Recognition from Single Sample Based on Human Face Perception

Ce Zhan, Wanqing Li, and Philip Ogunbona
 School of Computer Science and Software Engineering
 University of Wollongong, Australia
 Email: {cz847, wanqing, philipo}@uow.edu.au

Abstract—Although research show that human recognition performance for unfamiliar faces is relatively poor, when the sample is always available for analysis and becomes "familiar", people are able to recognize a previous unknown face from single sample. In this paper, a method is proposed to deal with the one sample per person face recognition problem based on the process how unfamiliar faces become familiar to people. Particularly, quantized local features which learnt from generic face dataset are used in the proposed method to mimic the prototype effect of human face recognition. Furthermore, a landmark-based scheme is introduced to quantify the distinctiveness of each facial component for the sample face, then the difference between the sample and the average face is emphasized by weighting face regions according to the gained distinctiveness. The experiments on ORL and FERET face databases demonstrate the efficiency of the proposed method.

I. INTRODUCTION

Significant progress have been achieved in automatic face recognition during the last 20 years. Numerous commercial face recognition systems have been released and a great number of papers have been published in journals and conferences dedicated to the related area of face recognition. However, reliable face recognition under unconstrained conditions still offers a great challenge to computer vision and pattern recognition researchers. Recently, most of the efforts are made to allow face recognition systems more robust against different variations such as pose, illumination, expression and occlusion.

One common basis of existing face recognition systems dealing with above mentioned variations is a large, representative training dataset. Unfortunately, in many real-world applications, sample images can be used for training are very limited. More specifically, in many application scenarios, especially in large-scale identification applications, such as law enforcement, driver license or passport identification, there is usually only one training sample per person in the database.

This realistic one sample per person problem severely challenges existing face recognition algorithms, especially their robustness performances under possible variations. Take the most famous face recognition technique, eigenface (PCA), for example. When tested by ORL [1] database which mainly addresses the pose variations (reported by Tan et al. in [2]), if only one training sample per person is used, the average recognition rate of eigenface falls to below 65%, with a 30% drop from 95% when 9 training samples per person are given.

This serious performance drop mainly due to intra-class variation cannot be obtained from one training sample. For the same reason, most of state of the art face recognition methods suffer from the one sample problem. Subspace methods such as Linear Discriminant Analysis (LDA), and Bayesian matching methods may even fail to work when there is only one training sample per class.

Intuitively, it is felt that human beings do not need many views (images) of a person in order to develop a proper model of his appearance. A good illustration of this common impression is the photo ID verification process such as in border control, where travelers are identified by custom officers based on single passport photo. However, psychological studies [3] have found that although people are excellent at recognizing faces familiar to us, even from very low quality images across large variations, we are not so good at recognizing unfamiliar faces. In [4] [5] subjects learned previously unknown faces from a single image and are asked to recognize the face from a second novel image of the person after a delay. Low level of recognition accuracy was recorded even for frontal face images in the experiments, and when there are changes in the viewpoint between the sample and test images, the accuracy further declines.

The above mentioned experimental scenario is not quite similar as the one sample problem, since the memory component is involved. In the case of one sample machine learning, the sample image is always available, comparison and analysis are allowed all the time between the sample image and test image, thus the sample face could becomes somehow "familiar". This case is similar to perceptual matching experiments. In the matching experiments [6] [7], subjects are given enough time and asked to determine if the persons pictured in two simultaneously presented images are the same or different. The matching accuracy is much higher than the recognition accuracy reported in [4] [5], however variations in viewpoint and illumination between the sample image and test image still affect the accuracy.

From the two observations, we can suggest that although human recognition performance for unfamiliar faces is poor, when the sample is always available for analysis and becomes "familiar", people are able to handle the one sample problem. This finding gives us a good hint for solving the one sample problem based on the process how unfamiliar faces become familiar to people. On the other hand, in both of the experiments,

the differences between learning and test sample affect the recognition performance, which reminds us that the essence of one sample problem is not a problem concerning how many training samples each person has, but that concerning how to improve robustness performance against different variations under this extremely small sample size condition. However, although several methods have been proposed in the literature dealing with the one sample problem (that will be briefly reviewed in Section II), the variation issue is far from solved. In some of the methods, none of the variation factors are explicitly addressed, and the performance of these methods were evaluated using mainly frontal face dataset with limited variations. For the methods that addressed the variation issue, most of them are only robust against some variations while not against others.

In this paper, a method is proposed to deal with the one sample per person face recognition problem based on the process how unfamiliar faces become familiar to people. In addition, factors of human familiar face perception are incorporated into the proposed model so as to enhance the robustness performance against different variations. Particularly, quantized local features which learnt from generic face dataset are used in the proposed method to mimic the prototype effect of human face recognition. Furthermore, a landmark-based scheme is introduced to quantify the distinctiveness of each facial component for the sample face, then the difference between the sample and the average face is emphasized by weighting face regions according to the gained distinctiveness.

The rest of the paper is organized as follows: In Section II we briefly review the existing works dealing with one sample problem. Detail of the proposed method is described in Section III. Section IV presents one implementation of the proposed method and the experimental results. Conclusions are drawn in Section V.

II. RELATED WORKS

The challenges and significance for practical applications of the one sample per person problem have created increasing interests in the computer vision community, several methods have been proposed to solve this problem. One natural solution to the problem is to artificially generate extra samples for each person, so that the traditional methods can be used based on the new samples. Yang et al [8] described an approach known as symmetrical PCA, which employs a standard PCA algorithm for feature extraction based on two virtual image sets, i.e., the even and odd symmetrical image sets. Martinez [9] introduced a perturbation-based approach to generate new samples for PCA. While in [10], the training set is enlarged by constructing new presentations for the single training image. The problem of this kind of methods is that generated facial images and representations are highly correlated and should not be considered as truly independent training samples, i.e., the created variations are usually not large enough to cover those observed in reality.

Some methods solve the one sample problem by extending the PCA method. Wu et al. [11] introduced a method called

$(PC)^2A$ to enrich the information of face space at pre-processing stage before the standard PCA. They combine the original image with its first-order projection map, thus salient facial features are emphasized and unimportant features are faded out after the pre-processing. Following the $(PC)^2A$ framework, Chen [12] proposed an enhanced $(PC)^2A$ solution by including a second-order projection map while Zhang et al. [13] introduced a SVD perturbation at the pre-processing stage. Rather than extending PCA in pre-processing, Yang et al. [14] focused on the covariance matrix estimation under one sample training. They proposed a 2DPCA method which uses straightforward 2D image matrices to estimate the covariance matrix instead of 1D vectors. All of the above mentioned methods reported better performance compared to that of the standard PCA. However, they actually handle the one sample problem in an indirect way, that is, the variations of expression, illumination or pose are not explicitly addressed. Therefore, their robustness performance is somehow predictable.

One possible way to handle the variation issue in one sample scenario is to use local facial representations, due to the observation that local features are generally not as sensitive as global features to appearance changes. Martinez [9] proposed a local probabilistic approach, where the subspace of each individual is learned and represented by a separate Gaussian distribution. The method is then extended by Tan et al. [15], who proposed an alternative way of representing the face subspace with self-organizing maps (SOM). In [16], face image is divided into rectangular blocks and each block is represented by histogram of the local binary patterns (LBP). Comparing with other existing method, the local based approach provides additional flexibility to recognize a face based on its parts, thus it seems more suitable for handling the one sample problem. However, some common problems are still unsolved in these local based methods. One of the problems is how to weight the local regions. Most of the methods use stable predefined weight for combining local classifiers, such as in [16], where eyes area are given the highest weight since they seem to be the most important cue in human face recognition. Although experiments showed that this kind of weighting mechanisms do improve the recognition rates, bias could be involved in the predefined weights and more sophisticated and adaptive weighting scheme is expected for the local approach. In addition, as methods based on local features, the spatial (configural) information of the face are sometimes ignored or not properly considered. However, the configural features also carry important information for identifying faces. Thus, how to incorporate the spatial information is also a crucial problem to be solved for the local based approach. Finally, what kind of local representation should be used for the face is always the key question to answer. In the following section, we introduce a local based method that address the above problems based on observations of human face perception.

III. THE PROPOSED METHOD

A. Overview

Previous psychological studies claimed that human perception of faces is a nondecomposable holistic process, which is used as a fundamental assumption of a great number of existing face recognition methods. However, recent findings challenged the assumption, and suggested that nameable facial parts (facial components such as mouth, eyes and nose) play an important role in face perception especially when variations occur [17], [3]. The proposed method follows the parts-based theory and represents the face hierarchically. After face normalization, first the facial image is divided into different regions, each region R_i contains one key facial component. Then the region R_i is further divided into sub-blocks, and one feature F_k is extracted from each sub-block. Thus during the recognition, test images are compared with the sample image, the best match is output as recognition result by searching the minimum integrated distance:

$$D = \sum_i w_i d_i$$

where i is the index for different facial region; w_i is the weight for each region; and d_i is the similarity measure between the corresponding regions. This hierarchical representation shares the common advantages of local based method, to further solve the one sample problem, as mentioned in Section II, the local representation used to calculate d_i and the weight w_i need to be properly defined.

B. Regional Similarity Measure Based on Quantized Local Features

In human face recognition, the prototype effect refers to the discovery that prototype faces which are composed of typical facial features (components) extracted from previously viewed faces can be confused with truly familiar ones [18]. Researchers explain the prototype effect as that human tend to respond to the central values of exemplars, and suggested that our internal representation for faces may be based on commonly experienced sensory phenomena (features). Following this finding, in the proposed method, each local feature F_k is quantized according to a codebook to mimic the prototype effect. The codebook is constructed using clustering approach based on generic face dataset which consists of images from subjects other than those under consideration. Thus, each code vector in the codebook represents a kind of local pattern that frequently appears in human face images. Ideally, if the code vectors only encode inter-person variations, when faces are projected into the quantized feature space, different faces from the same person will converge at the same location in the space.

Rather than directly using the quantized local features, each facial region R_i is represented by a regional histogram H_i , which is obtained by counting the occurrence of the quantized features within R_i . Formally:

$$H_{j,i} = \sum_k I\{Q(F_k) = j\} I\{k \in R_i\}, j = 0, \dots, n - 1$$

Where j is the index of code vector, $Q()$ is the quantization process mapping F_k into the learnt quantized feature space, k is the index of sub-block, n is the size of codebook, and $I\{A\} = 1$ if A is true, $I\{A\} = 0$ if A is false. This regional histogram representation focuses on the the distribution of the local features and ignores the spatial information within the region, thus it is relatively robust against viewpoint variations. On the other hand, by partition the face region meaningfully based on facial components, the configural information for features from different facial components is remained, which makes the regional histogram representation more efficient.

Then Chi square statistic can be used to measure the distance d_i ,

$$d_i(S_i, M_i) = \sum_j \frac{(s_{i,j} - m_{i,j})^2}{s_{i,j} + m_{i,j}}$$

where S_i and M_i are two corresponding regions in the sample and the test image; $s_{i,j}$ and $m_{i,j}$ is the frequency of a code vector j that belongs S_i and that belongs to M_i , respectively.

C. Weighting by Distinctiveness

Numerous studies have shown that unusual faces are better remembered and recognized than prototypical faces. Where a prototypical face refers to a blend of typical or homogeneous features that are “normal” or “average”. In [19] O’Toole et al. suggest that faces are retained in a manner that enhances the prototypical features of the face. Thus an exemplar is encoded for recognition. Accuracy in recognition occurs where the facial detail differ from the exemplar. When unusual features such as a big nose, small eyes or thick eyebrows found on a face, a specific memory that could be easily recalled would be created. Thus the face becomes more familiar than usual faces. This finding also suggests a reason for the superior recognition of caricatures over veridical faces. Studies [20] have shown that caricatures which emphasize the differences between a face and the average are recognized more accurately and more quickly than veridical versions of the faces, even when the faces are unfamiliar.

Inspired by the above observations, in the proposed method, the facial region is weighted according to the distinctiveness of corresponding facial component for the sample face. During the distinctiveness measuring, facial parts are compared with the average face as suggested by human face perception, and the coordinates of key facial points are used for quantitative analysis. Each facial component is represented by a coordinate vector, which is formed by orderly listing the x and y coordinates of each key facial point. Since faces are normalized, the coordinate vectors contain not only the shape information but also the size information of facial components. Thus, the distinctiveness of each facial part can be measured by the Mahalanobis distance between its coordinate vector and the corresponding mean coordinate vector, which is calculated based on a generic face database. Note that the distinctiveness of the mouth is not measured in the proposed method. Many studies [3] have shown that due to shape variations caused by expressions, the mouth is not recognized as good as the



Fig. 1. The normalization process

upper-face features. Thus, when weighting the facial regions, the lowest weight is always given to the mouth part and other face regions are weighted according to the distinctiveness:

$$w_i = \frac{c_i}{\sum_i c_i + \min(c_i)}$$

where i is the index for different facial regions except mouth part and c_i is the distinctiveness for different facial region. For the mouth region:

$$w_m = \min(w_i)$$

IV. EXPERIMENTAL RESULTS

A. The Implementation

The face region of the image is first detected by the Viola-Jones face detection method [21]. Then a modified version of the Viola-Jones face detection method is employed to find the areas of mouth and eyes within the detected face. For details on the method, readers are referred to our previous work [22]. Once the eyes and mouth have been localized, using the differences between the x and y coordinates, the original image is rotated so that the centers of eyes and mouth are at the same pixel coordinates in all images. Then the face area is cropped and resized to a 64×64 “standard” face image. Figure 1 shows the whole normalization process.

The normalized face is divided into 6 regions based on facial components as showed in Figure 4 a. Note that the eyebrows are divided separately, due to the study from O’Toole et al. [19], who claim that of the different facial components, eyebrows are among the most important for recognition. Then the facial region is further divided into 8×8 pixel sub-blocks with overlaps. The degree of overlap affects the recognition performance, we will evaluate the effects of varying sub-block overlap in Section IV-B.

In the implementation, we calculate the 2D-DCT to extract the feature F_k in each sub-block, this results in a 8×8 DCT coefficient matrix with 64 coefficients. Only a subset of the DCT coefficients are used to represent each sub-block. Ideally, the coefficient subset should be insensitive to illumination variations and contains as few elements as possible while represent all the necessary information (the contributions of different subsets to the recognition performance is investigated later).

The codebook for all the sub-blocks is generated using “k-means clustering” [23] from BioID face database [24]. To capture all possible inter-person variations, the training datasets should be large enough with all kinds of diversities, and the size of codebook (that is the k in k-means clustering) cannot



Fig. 2. Examples of original face image and the corresponding reconstructed images

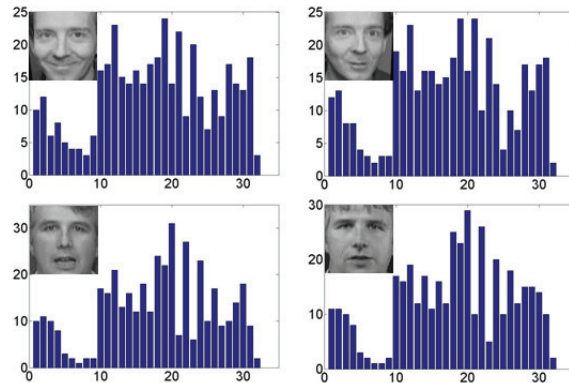


Fig. 3. Typical examples for regional histograms of left eye region

be too small. On the other hand, to avoid code vectors capture intra-person variations, the codebook size cannot be too large. The effects of different codebook sizes are investigated in Section IV-B. After the codebook has been trained, all the sub-blocks from the face are mapped to the code vectors in quantized feature space by a nearest neighbor strategy, and regional histograms are calculated. Figure 2 shows examples of original face images and their corresponding reconstructed images. Figure 3 shows typical examples for regional histograms of left eye, where codebook of the size 32 is used. As can be seen, histograms of different persons are clearly different while histograms of the same person are resembled, though there is a small difference in detail.

41 key facial points are located for distinctiveness measuring in the implementation, as showed in Figure 4 b. Quite a few reliable methods have been proposed to detect the key facial points, such as [25] and [22]. For this implementation, these facial points are labeled manually to avoid the measuring error. The mean coordinate vector of each facial component is calculated based on frontal view images from BioID face database. Some typical examples of facial region weights gained based on distinctiveness measuring can be seen from Table I.

B. Parameter Selection Based on ORL Database

Experiments are conducted based on ORL face database [1] to investigate the effect of different parameters. The ORL face database is composed of 400 grayscale images with 10 images for each of 40 individuals. The variations of the images are across pose, time, illumination and facial expression. During




Sample image	Weight of each facial region gained based on distinctiveness measuring					
	Left eye-brow	Right eye-brow	Left eye	Right eye	Nose	Mouth
	0.2264	0.1987	0.0989	0.0647	0.3465	0.0647
	0.3149	0.2027	0.1154	0.0770	0.2130	0.0770
	0.2035	0.1135	0.2445	0.1298	0.1953	0.1135

TABLE I
EXAMPLES OF FACIAL REGION WEIGHTS GAINED BASED ON
DISTINCTIVENESS MEASURING

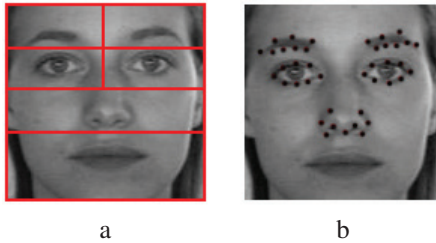


Fig. 4. Facial region partition and key facial points

the test, each time, one image is selected as the training sample, all the other images in the database are used as testing samples. Considering the accuracy of distinctiveness measurement, only the selected images with frontal view are used as training samples.

1) *DCT Subsets*: When the DCT coefficients are ordered according to a zig-zag pattern [26], we know that the first few coefficients which correspond to low-frequency components always contain the most important information. In particular, the first DCT coefficient (DC component) reflects the average pixel value inside each sub-block and hence would be affected the most by any illumination change. The second and third coefficients ((0, 1), (1, 0) component) represent the average horizontal and vertical pixel intensity change, respectively. As such, they would also be significantly affected by illumination changes. Thus, to find the optimal subset mentioned above, following DCT subsets are examined in the experiment:

- S1, the first 16 coefficients.
- S2, removing the DC component, and selecting the first 16 coefficients from the remaining ones.
- S3, removing the DC, (0, 1) and (1, 0) components, and selecting the first 16 coefficients from the remaining ones.
- S4, the first 8 coefficients.
- S5, removing the DC component, and selecting the first 8 coefficients from the remaining ones.
- S6, removing the DC, (0, 1) and (1, 0) components, and selecting the first 8 coefficients from the remaining ones.

In the experiment, for different DCT features, corresponding codebook with fixed size of 64 are generated and used. To

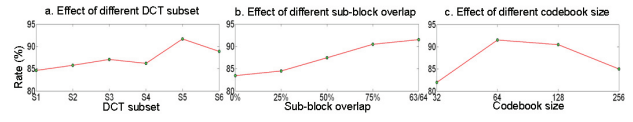


Fig. 5. Effect of different parameters

ensure adequate representation of the face, face images are divided by sliding the 8×8 dividing-partition one pixel by one pixel, thus, sub-blocks with 63/64 overlap are used. The recognition performance for different DCT subsets is showed in Figure 5 a. As can be observed, DCT subsets with less elements achieved better results. By discarding the elements affected by illumination variations, the recognition rate is improved. The best results are obtained by using subset S5, which only removes the DC component.

2) *Sub-block Overlap*: One direct effect of varying the overlap of sub-block is that the number of feature vectors extracted from an face image grows in an exponential manner as the overlap is increased. Experiment is conducted here to investigate the indirect effect of overlaps on the final recognition performance. In the experiment, DCT subset S5 is adopted for feature extraction and the corresponding codebook with 64 code vectors is used for quantization. The experiment results are showed in Figure 5 b. As can be seen from the curve, by increasing the percentage of overlap, the recognition performance is improved. This may be explained as that the code vector distribution for one image can be better estimated based on more extracted features. Another observation from the figure is that increasing the overlap from 75% to 63/64 had little effect on the recognition performance at the expense of extracting significantly more feature vectors.

3) *Codebook Size*: To find out the proper codebook size, in the experiment, 4 different codebooks are generated based on the same training data but only varying the codebook size. These codebooks are tested while keeping the other two parameter unchanged (DCT subset S5 and 63/64 sub-block overlap is used), the result is showed in Figure 5 c. As we analyzed before, codebooks with middle range of the size achieved better recognition rates.

C. Experiments on FERET Database

The FERET face database is used to evaluate the proposed method according to the standard FERET evaluation protocol [27] with the gallery set including 1196 frontal images of 1196 persons and four probe sets: **fafb** (1195 images with expression variations); **fafc** (194 images with illumination variations); **dup.I** (722 images taken in less than 18 months); **dup.II** (234 images taken about 18 months later). In the evaluation, a subset of 500 images from the gallery set are randomly selected as training samples, and all the four probe sets are tested. The rest experimental setup is the same as introduced in Section IV-B, with selected parameters: codebook size of 64; sub-block overlap of 63/64; DCT subset of S5. The performance of the proposed method is showed in Table II, including some reported results for the one sample

Methods	fafb	fafc	dup.I	dup.II
Fisherface [28]	0.94	0.73	0.55	0.31
EBGM Optimal [29]	0.90	0.42	0.46	0.24
LBP [16]	0.97	0.79	0.66	0.64
SIFT_GRID [30]	0.94	0.35	0.53	0.36
Proposed with predefined weight	0.92	0.75	0.68	0.64
Proposed	0.92	0.77	0.75	0.71

TABLE II
THE RECOGNITION RATES OF DIFFERENT METHODS ON FERET PROBE
DATA SETS

problem based on the same dataset. “Proposed with predefined weight” denotes predefined weighting in [16] is used for the proposed method instead of the weighting scheme based on distinctiveness. It can be seen that the proposed method achieved comparable recognition rate on **fafb** and **fafc** sets while outperformed all the other methods on **dup.I** and **dup.II** sets, in which images are taken in different conditions and different time. The results also showed the efficiency of the proposed weighting scheme.

V. CONCLUSION

Among a number of methods proposed in the literature to deal with one sample face recognition problem, the local based methods seems more suitable than others, since they provide additional flexibility to recognize a face based on its parts. However, some common problems of local methods are still unsolved, such as how to weight the local regions and what kind of local representation should be used for the face. In this paper, we proposed a method to solve the one sample problem within the local based framework and address the issues of local method from human perception point of view. Particularly, quantized local features which learnt from generic face dataset are used in the proposed method to mimic the prototype effect of human face recognition. Furthermore, a landmark-based scheme is introduced to quantify the distinctiveness of each facial component for the sample face, then the difference between the sample and the average face is emphasized by weighting face regions according to the gained distinctiveness. Experiments conducted on benchmark face databases with variations of expressions, illuminations and viewpoints demonstrated the efficiency of the proposed method. In particular, the proposed method performed well on recognizing face images that are taken in different conditions and different time.

REFERENCES

- [1] F. Samaria and A. Harter, “Parameterisation of a stochastic model for human face identification,” *Proceedings of 2nd IEEE Workshop on Applications of Computer Vision*, pp. 138–142, 1994.
- [2] X. Tan, S. Chen, Z.-H. Zhou, and F. Zhang, “Face recognition from a single image per person: A survey,” *Pattern Recognition*, vol. 39, pp. 1725–1745, September 2006.
- [3] S. S. Rakover and B. Cahlon, *Face recognition: cognitive and computational processes*, illustrated ed. John Benjamins Publishing Company, 2001.
- [4] N. Troje and H. Bulthoff, “Face recognition under varying pose: The role of texture and shape,” *Vision Research*, vol. 36, pp. 1761–1771, 1996.
- [5] A. O’Toole, S. Edelman, and H. Bulthoff, “Stimulus-specific effects in face recognition over changes in viewpoint,” *Vision Research*, vol. 38, pp. 2351–2363, 1998.
- [6] S. Edelman and H. Bulthoff, “Orientation dependence in the recognition of familiar and novel views of three-dimensional objects,” *Vision Research*, vol. 32, no. 12, pp. 2385–2400, 1992.
- [7] W. L. Braje, D. Kersten, M. Tarr, and N. Troje, “Illumination effects in face recognition,” *Psychobiology*, vol. 26, pp. 371–380, 1999.
- [8] Q. Yang and X. Ding, “Symmetrical PAC in face recognition,” in *Proceedings of ICIP’02*, vol. 2, 2002, pp. 92–100.
- [9] A. Martinez, “Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 6, pp. 748–763, 2002.
- [10] F. D. la Torre, R. Gross, S. Baker, and B. V. Kumar, “Representational oriented component analysis (roca) for face recognition with one sample image per training class,” in *Proceedings of CVPR’05*, vol. 2, 2005, pp. 266–273.
- [11] J. Wu and Z.-H. Zhou, “Face recognition with one training image per person,” *Pattern Recognition Letters*, vol. 23, no. 14, pp. 1711–1719, 2002.
- [12] S. C. Chen, D. Zhang, and Z.-H. Zhou, “Enhanced $(PC)^2A$ for face recognition with one training image per person,” *Pattern Recognition Letters*, vol. 25, no. 10, pp. 1173–1181, 2004.
- [13] D. Zhang, S. Chen, and Z.-H. Zhou, “A new face recognition method based on svd perturbation for single example image per person,” *Applied mathematics and computation*, vol. 163, pp. 895–907, 2005.
- [14] J. Yang, D. Zhang, A. F. Frangi, and J. Yu Yang, “Two-dimensional pca: A new approach to appearance-based face representation and recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 131–137, 2004.
- [15] X. Tan, S. Chen, Z.-H. Zhou, and F. Zhang, “Recognizing partially occluded, expression variant faces from single training image per person with som and soft knn ensemble,” *IEEE Transactions on Neural Networks*, vol. 16, no. 4, pp. 875–886, 2005.
- [16] T. Ahonen, A. Hadid, and M. Pietikinen, “Face recognition with local binary patterns,” in *Proceedings of ECCV’04*, 2004, pp. 469–481.
- [17] A. Schwaninger, J. S. Lobmaier, and S. M. Collishaw, “Role of featural and configurational information in familiar and unfamiliar face recognition,” in *Lecture notes in computer science 2525*, 2002, pp. 643–650.
- [18] R. Solso and J. McCarthy, “Prototype formation of faces: A case of pseudo-memory,” *British Journal of Psychology*, vol. 72, pp. 499–503, 1981.
- [19] A. O’Toole, K. Deffenbacher, D. Valentin, and H. Abdi, “Structural aspects of face recognition and the other-race effect,” *Memory & Cognition*, vol. 22, pp. 208–224, 1994.
- [20] P. J. Hancock and V. Bruce, “Recognition of unfamiliar faces,” *Trends in Cognitive Sciences*, vol. 4, no. 9, pp. 330–337, 2000.
- [21] P. Viola and M. Jones, “Robust real-time object detection,” *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004.
- [22] C. Zhan, W. Li, P. Ogunbona, and F. Safaei, “Real-time facial feature point extraction,” *Advances in Multimedia Information Processing - PCM2007, LNCS 4810*, pp. 88–97, 2007.
- [23] J. A. Hartigan and M. A. Wong, “A k-means clustering algorithm,” *Applied Statistics*, vol. 28, no. 1, pp. 100–108, 1979.
- [24] R. Frischholz and U. Dieckmann, “Bioid: A multimodal biometric identification system,” *Computer*, vol. 33, no. 2, pp. 64–68, 2000.
- [25] A. R. Mirhosseini, H. Yan, and K.-M. Lam, “Adaptive deformable model for mouth boundary detection,” *Optical Engineering*, vol. 37, pp. 869–875, 1998.
- [26] E. Ifeachor and B. Jervis, *Digital Signal Processing: A Practical Approach*. Prentice Hall, 2001.
- [27] P. Phillips, H. Wechsler, J. Huang, and P. Rauss, “The feret database and evaluation procedure for face recognition algorithms,” *Image and Vision Computing*, vol. 16, pp. 295–306, 1998.
- [28] P. N. Belhumeur and J. Hespanha, “Eigenfaces vs fisherfaces: recognition using class specific linear projection,” *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 20, no. 7, pp. 711–720, 1997.
- [29] L. Wiskott, J. Fellous, N. Kuiger, and C. von der Malsburg, “Face recognition by elastic bunch graph matching,” *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 775–779, 1997.
- [30] M. Bicego, A. Lagorio, E. Grosso, and M. Tistarelli, “On the use of sift features for face authentication,” *Proceedings of CVPR Workshop 2006*, vol. 17, no. 22, pp. 35–35, 2006.