Pose insensitive Face Recognition Using Feature Transformation

Hongzhou Zhang, Yongping Li, Lin Wang, Chengbo Wang

Shanghai Institute of Applied Physics, Chinese Academy of Science

Summary

Face recognition has diverse applications especially as an identification solution which can meet the crying needs in security areas. Pose problem is a big challenge applying this technology under real world conditions. Appearance based approach was proposed. Face recognition was implemented by reconstructing frontal view features using linear transformation. Experiments on popular FERET database proved that the proposed method can cope with the head rotation roughly within half profile view. Compared with algorithms model based approaches, feature transformation method is not dependent on heavy computation and has merit of easy implementing in live conditions. Popular feature extractions, least square (LS) and total least square (TLS) solution in calculating were compared as well as..

Key words:
face recognition, pose, and feature transformation.

1. Introduction

Face recognition, an effective biometric method, has diverse applications especially as an identification solution which can meet the crying needs in security areas. It involves image processing, pattern recognition, intelligent learning and so on. Considerable achievements of face recognition have been attained in recent years [15], but big challenges still remain, such as pose variation [15, 18].

Many algorithms are developed to overcome pose effect. A mainstream is to generate frontal image from side view image inputted utilizing techniques in computer vision [6, 7, 8, 9, 10, 11]. Among these works, one way is model-based, such as 3D Morphable Model [6] or Active Appearance Model (AAM) [10] which tries to directly reflect geometric structure of subject's head. Another way is to use affine transformation or wrap technology [8, 9] in virtual frontal image synthesis. Model based methods are powerful in posed face recognition, but fitting a face model for an input image is time-consuming. Some affine transformation technologies reduced computation in a great extent, but geometrical aligning input image to standard view model still needs sets of feature points which mainly marked by hand. Recently, Gross proposed Eigen Light-fields algorithm [11] by estimating light-field of the subject's head from face images; but precise computation of plenoptic function is difficult. In one word, algorithms based computer vision are effective, but their disadvantages are also standout, especially many of them depend on heavy computation. It limits their applications under real world condition.

Techniques based on face subspace analysis [16, 17] or statistical properties of face images are successful in frontal view recognition. With no time-consuming model fitting and too many fiducial feature points, these algorithms are more suitable for applications under live conditions. An early work of extending frontal view subspace recognition approaches to non-frontal view is Pentland's multi-view work [17]. In multi-view subspace, all images are represented by view-dependent subspace corresponding to their pose and face recognition is performed within same view subspace. Apparently such multi-view recognition is duplicating of conventional frontal view works, associations between different poses are discarded. Murase and Nayar also proposed a generic object visual learning method which is called as parametric subspace work [16]. In parametric subspace, a unique subspace is used as feature extractor, objects are represented by their feature manifolds varying with pose, and face recognition is converted into manifolds identification. A recent face recognition work suing parametric subspace was reported in [19] of which identification is to compute the shortest Euclidean distance from a given feature point to the point on the manifolds corresponding with testing view. To form subject's manifold utilizing B-spline interpolation technique, images at many different views are required [16, 19]. Association between views is actually considered into these manifolds, but this consideration is quite redundant.

A limitation of above two subspace algorithms is that client's frontal and non-frontal face images are both required to enroll or form manifolds. It is too rigid for some applications where only frontal view images are enrolled conventionally, but the pose of probing image is usually uncontrollable. Recently some scientists have turned their attentions to extending statistical work to face recognition under similar scenarios, such as [11, 22].

In this paper we simulate the scenario of applications that frontal view images are used for gallery and only non-frontal view images are available for probing. We develop a subspace analysis method of face recognizing across poses based on the idea of transformation matrix. To focus our attention on feature transformation, we also suppose that each image's pose has been given out (by a pose estimator) at present. Different pose estimators could be found in [10, 19, 25]. Experiments show that transformed

Manuscript received February 5, 2007
Manuscript revised February 25, 2007

feature is equal to face recognition across poses with a large rotation.

The rest of paper is organized as follows. In Section 2, we explain our feature transformation method. In Section 3, transformation matrix is generated by solving equations composed of feature vectors. Different feature extractions are then introduced in Section 4. The experimental results are reported in the followed section. Discussions are given in Section 6 finally.

2. Linear Transformation matrix

Human has capability to associate stranger's photos under different poses together. The fact shows that correlation between views is helpful to improving posed face image recognition. Recent work [11, 21] has proved this from the point view of statistics. Traditional subspace recognition, such as Eigenface [20] or Fisherface [5], is actually view-dependent due to images at the single (frontal) view are used to train the representations. When a posed image is represented by the frontal view subspace, wrongly image representing will lead to system's failure. By introducing additional view subspaces, Pentlad et al. [17] avoided wrong image representation. To accomplish face recognition across poses successfully, we train the similar view-dependent subspaces as [17].

Thus, the key issue is to learn correlations between views which were discarded in the early work [17]. Because only frontal view images are available in enrollment procedure, knowledge on the correlations must be learned from separate training data, which can be collected quite freely by imaging a group of people under simulated applying conditions before set up a recognition system. We call this partition of images as generic training set. The scheme of our transformation work is illustrated Fig. 1.

In generic training set, subjects' one frontal view and one side view images comprise image pairs; they are used to train two view-dependent subspaces respectively. Images in the same pair are taken as simultaneously as possible. $V_F = \{v^i{}_F \mid i = 1, 2, ..., N\}$ and $V_P = \{v^i{}_P \mid i = 1,2, ..., N\}$ are projections of frontal and profile view images on their corresponding subspaces. Superscript "i" denotes the projection of images in the i-th pair, N is total image pair number. We suppose that frontal view vectors V_F could be recovered from side view vectors V_P by a transformation T:

$$V_F = T(V_P) \tag{1}$$

Generally speaking, transformation denoted by Eq. (1) is non-linear. But a linear transformation is more preferred because linear transformation is convenient for problem description and well researched by scientists in different areas. Fortunately, different works have proved that linear transformation could give out satisfying results. Lanitis et al. [12] have showed that linear model is sufficient to simulate considerable pose variation as long as overlap does not seriously take place. Based on the theory of linear class [23], Beymer et al. [8] have proved that frontal view information can be recovered from side view images utilizing the prior information learned from images at several example views.

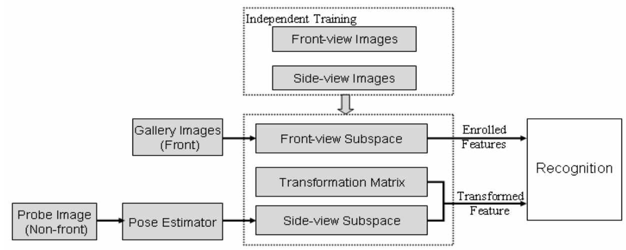

Fig. 1. A flow chart of the proposed approach.

For linear transformation, matrix W connects frontal view and side view feature vectors together:

$$(v_F^1, v_F^2, ..., v_F^N) = W \cdot (v_P^1, v_P^2, ..., v_P^N) \tag{2}$$

$$V_P{}^T w_k = v_k{}^F \tag{3}$$

where w_K and $v^F{}_K$ are the k-th column vector of W^T and $V_F{}^T$ respectively. $(\cdot)^T$ denotes matrix transpose. We call W as *transformation matrix*. Eq. (2) or (3) could be treated as matrix equations which can be resolved by means of linear algebra [24]. Once Eq.(2) is resolved, a probe feature vector in side view subspace, v_{test}, can be converted into frontal view subspace by the following Eq. (4). Then recognition will be performed in frontal view subspace using transformed feature v'_{test}.

$$v'_{test} = W \cdot v_{test} \tag{4}$$

3. Equation solving and similarity metrics

3.1 Least Square solution

The solution of Eq.(3) depends on property of the coefficients matrix, a detail of derivation could be found in [24]. Because we trained two independent subspaces, generally $rank(V_P{}^T \mid V^F{}_k) \neq rank(V_P{}^T)$. For face recognition, $V_P{}^T$ is a N-by-m matrix, N is the total image pairs in generic training set and m is dimensionality selected in each recognition experiment, generally $N > m$. So, Eq. (2) is a inconsistent and over-determined system of linear equations which is exactly unsolvable. According to theory of matrix, inconsistent system has approximation solutions under 2-norm constraint, which is called least square (LS) solution. LS solution minimizes square error.

$$w_k = \min_{w_k \in R^n} \| V_P^T w - V_F^T \| \quad (5)$$

where $\|\cdot\|$ denotes Euclidean norm. Among least square solutions, an optimum approximation solution is minimal linear least square solution. A particular optimum solution of Eq. (2) is expressed as:

$$W^T = (V_P^T)^+ \cdot (V_F^T) \quad (6)$$

$(V_P{}^T)^+$ is the Moore-Penrose pseudoinverse of $V_P{}^T$ which could be calculated by means of singular value decomposition (SVD).

3.2 Total least square solution

In above LS solution, one question we have actually not answered. Whether there are some noises in the feature vectors in Eq. (3)? In another words, we treat V_F^T and $v^F{}_K$ as precise measurements in above section. A more general and reasonable situation is there are some noises in sample images. Algebra mathematics tells that LS solution of Eq. (3) is unbiased only if there is no noise in Eq. (3) or noise is only in the right hand of Eq. (3). When both hands of equations are contaminated, LS solution will no longer be the optimal solution from a statistical point of view and LS approach should be replaced by the total least square (TLS) which is a generalized least square technique [1, 2, 24]. TLS technique for an over-determined system tries to compensate for arbitrary noise in both sides of equations using perturbations. Such solutions can be derived by Lagrange multipliers directly form the problem definition or by use of SVD. For equation

$$(A + E)x = b + e \quad (7)$$

where E and e are noise evolved into both sides of equation $Ax = b$, an equivalent form of above equation is:

$$(B + D)z = 0 \quad (8)$$

B=[-b, A] and D=[-e, E] are augmented matrix and disturbing matrix respectively. Equation of which form is same with Eq. (8) has a TLS solution expressed as [24]:

$$x_{TLS} = \frac{1}{v(1, n+1)} \begin{bmatrix} v(1, n+1) \\ \vdots \\ v(1, n+1) \end{bmatrix} \quad (9)$$

where v is the right singular vectors of augmented matrix B, and $v(i, n+1)$ is the entry of v located in i-th row and (n+1)-th column.

3.2 Comparisons on LS and TLS

TLS solution of Eq.(7) can also be expressed algebraically as the following [1, 2, 24]:

$$x_{TLS} = (A^T A - \sigma_{n+1} I)^+ A^T b \quad (10)$$

In above expression, σ_{n+1} is the smallest singular value of matrix B. The above Eq. (10) shows that TLS is a kind of LS methods which eliminated noise factor, $\sigma_{n+1}I$, introduced by matrix A.

Theorem: *if B is m × (n-1) matrix by deleting one column from m × n matrix A.* σ_A and σ_B *are the smallest singular value of two matrix respectively, then* $\sigma_A < \sigma_B$ [24].

For face recognition, feature matrix V_F or V_P with different dimensionalities in Eq. (3) is equivalent to deleting some rows from feature matrixes in full subspace. When dimensionality increases, the portion subtracted from Eq. (10) becomes smaller. So it was expected that TLS will trend to LS solution when selected subspace augment to full subspace:

$$x_{TLS} \longrightarrow x_{LS}\big|_{m \to n} \quad (11)$$

4. Feature Extractions

4.1 Feature Extractions Used in Experiment

Popular feature extractions, i.e. principal components analysis (PCA) [20], linear discriminant analysis (LDA) [5] and independent components analysis (ICA) [3] are used in our experiments.

PCA, which is closely related to the Karhunen-Loeve Transform (KLT), is a Maximum Expression Feature (MEF) extraction and widely used in data reduction and image reconstruction. ICA, a generalization of PCA [3], is derived form blind sources separation and its components are designed to be non-Gaussian. We executed ICA using fix-point fast ICA calculation algorithm [4] on Bartlett's [3] "Architecture One", where images, not pixels, are treated as independent random variables.

Linear discriminant analysis (LDA), also known as Fisher linear discriminant (FLD), is a Maximum Discriminating Feature (MDF) extracting method. It can distinguish within-class and between-class scatters. However, a drawback of LDA is larger number images per subject are required for training to ensure a good generalization, otherwise, "small sample size (SSS)" problem is encountered and recognition score will be deteriorated. Various revised LDA algorithms, i.e. Regularized-LDA [14] we used besides traditional LDA in present work, are developed to overcome the SSS problem.

4.2 Subspace Selection

Carefully selected subspace could enhance Signal Noise Ratio of the representation. Before calculating the transformation matrix, it needs further select a subset of dimensions from full subspace learned by different methods. Subspace selecting criteria we taken in present work is base on discriminability of each dimension in itself subspace. PCA and LDA-based subspaces could be selected conventionally according eigenvalues of their basis vectors in solving the eigen problems. ICA basis vectors have no such special order. We calculate discriminability of each IC on the generic training data, subspace spanned by top discriminibale ICs is selected what was suggested in [3].

5. Experimental results

5.1 Database and Experiment Setting

FERET database [22], which is a standard database in this area, was used to evaluate our algorithm. The merits of the large scale and plenty of side views in b-subset of FERET make it possible to explore more characters of the proposed approach.

For LDA based subspace requires two simples each class to train the representation, images in right-side views were mirrored to their corresponding left-side views, frontal view images were same mirrored. Thus there are 4 side-views remained (two full profile views are discarded for serious overlap), but two images for each subject are available under each of poses. The inter-oculars distance was set identical before all images were aligned according to their eyes coordinates. Then they were cropped into a size of 131-by-181 (row-column) pixels. Histogram equalization was applied to reduce illumination effect (Fig. (2)).

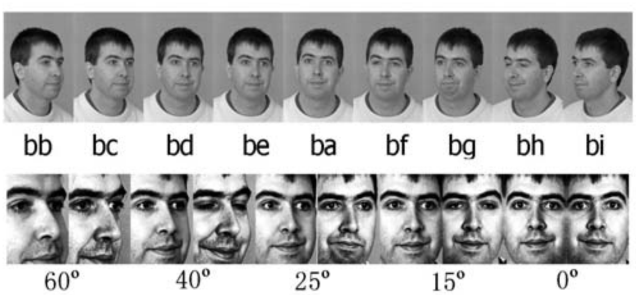

Figure 2. Samples of FERET database. From left to right roughly are: -60(bb), -40(bc), -25(bd), -15(be), 0(ba, front), 15(bf), 25 (bg), 40(bh), 60(bi) degrees. The first row is snapshots of original images in database; the second row is normalized images in size of 131-by-181.

Training set and test set are separate. The database was firstly divided into two portions according to subjects' identities. The first portion was selected as generic training set, bases of feature subspaces and transformation matrixes were learned from this part of data. The remainder, test set, was further divided: all frontal view images were used as gallery set and non-frontal images were used as probe set. To minimize errors attributed to the database itself, all experiments were "leave-one-out" test by dividing database into 40 smaller parts. The simplest Nearest Neighbor (NN) classifier was used in our experiments.

5.2 Experimental Result

To explore transformation recognition rate, both training and test set had a size of 100 persons.

In the first experiment, two different solution, LS and TLS solution were compared. It was found that LS solution is more stable than TLS. One of such results on 15 degree ("be" subset) view was illustrated in Figure 3. As expected in Section 3, when selected subspace was nearing full, two solutions gave out same recognition score. But TLS was not superior to LS solution in the most selected dimensionalities except some low dimensions: in the case of dimensionality was less than 25, TLS method gave out higher scores about three to five percent than LS; from 25 to roughly 65, approximation error of TLS solution and corresponding recognition rate had an oscillation. This phenomenon may due to that TLS is based on assumptions of noise distribution but such assumptions are too rigid for present application. Before full information on noise was known and LS solution is an ideal choice. Only LS solution was used in following experiments for this reason.

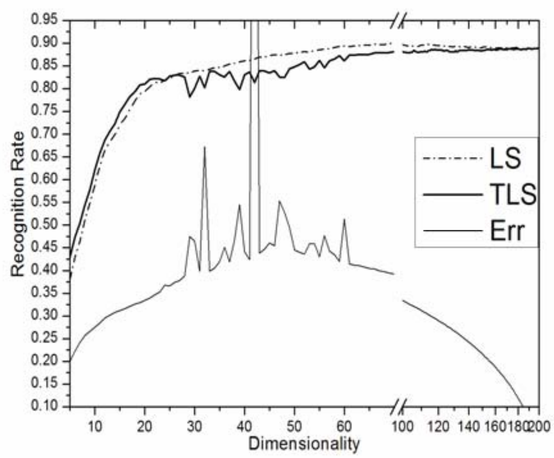

Fig. 3. A comparison on TLS and LS solution in transformation matrix calculation using PCA representation. Approximation error has divided by a constant for convenience.

Face recognition rate across different poses was tested in the second experiment, four feature extractions and different metrics were compared as well (see Table 1). From the experiment, it concluded that the proposed transformation matrix improved recognition rate in all test side views. Within 45 degrees, the scores were satisfying, but the converting ability of Matrix W decreased as face turned away from frontal view. Euclidean distance (ED) had average performances across different feature extractions and was the best metric for PCA. Angle measurement (noted as AD, also known as normalized correlation or cosine) had similar performances and they were more suitable for LDA-based feature extractions.

To improve performance of ICA, PCA was used as data reduction prior to it. In such experiment on "be" subset of FERET database, the top score of 90.2% was attained at the PCA dimensionality is 75 or 80. Less than half ICs were calculated, and the scores were higher than the one obtained without PCA data reduction (refer to Table 1), for some noise was eliminated in data reduction.

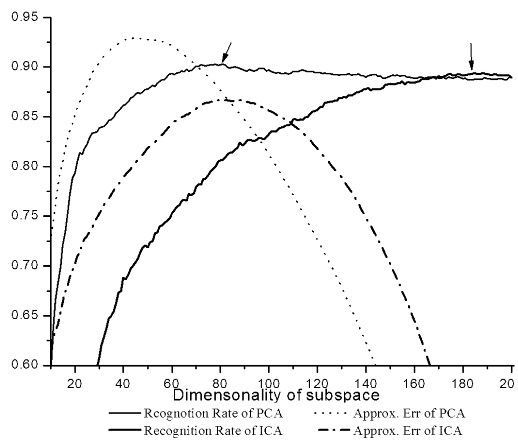

Fig. 4. Dependence of Recognition rate (solid line) and approximation error (dot-line). Arrows points out the maximum recognition rate. For clarity, only ICA (thick lines) and PCA (thin lines) feature are illustrated, the other features have same dependence. Two approximation errors have divided by different constants respectively for convenience in drawing.

Different dimensions contribute differently to recognition. To test this property, different dimensions form several to full subspace were selected, transformed recognition score and approximation error (see the right-hand of Eq. (5)) were same time calculated, see Figure 4. It showed that the some lower order dimensions contribute more for transformation recognition than higher order dimensions. The case was similar to conventional frontal view recognition. There was certain relation lying between transformation recognition and approximation error when different dimensions were selected. Though more dimensions gave out more accurate image representing, larger selected subspace not produced less approximation error before a maximum approximation error occured. After the maximum error arrived, the highest recognition score followed. We argue that such a phenomenon could be looked as a clue to subspace selection in transformation recognition. The reason may be derived from approximation in equation resolving and different weigh of each dimension in transformed feature recognition.

In the last experiment, test set kept a size of 50 persons while the size of generic training set varied. The respective best similarity measurements for different features are used (see Table 2, only result on 'be' subset is given out). Recognition rate depending on scale of training data was learned. Credible transforming performance was attained when size of generic training set and test set were equal (this is the case of the former experiment). However, training data with larger scale brought more satisfactory scores due to the statistical inherence of the method. As the size of training set reaches triple, the R-LDA feature gave out a top score similar to what ICA or PCA feature gave. This indicates the fact that training set with larger size should be employed to train LDA-based representation and transformation matrix.

6. Conclusion and Discussion

Based on above presentation, It concluded that proposed feature transformation work was an effective solution for recognition across poses. Comparing our work with others, two typical computer vision based works [6, 9] were listed in last two rows of Table 1. They all used 200 subjects' image to test their algorithm. In our experiment, 100 subjects' images were used to evaluate the proposed feature transformation recognition. Because different works differed in experiment setting and image preprocessing, it is hard to do comparison directly. Nevertheless, such comparison gave some reference. It showed that 3D model [6] was powerful in all tested views and there was only five percent score decreased when probe view turned from near frontal to near full profile. The reported affine transformation work [9] was also satisfying in 15 degree view, but it deteriorated with pose quickly. Within 45 degree pose rotation, the proposed approach still was comparable with others. It must be pointed out that the proposed work is basically real-time and also do not depend on too many fiducial points.

So far, we took the linear transformation assumption. The recognition rates were not high enough when pose is out of 45 degree. The reason is that overlap could not be ignored and the linear assumption was broken when face turned far away from the front. In such case, 3D-model based approaches are more suitable. We have observed that different dimensions contributed differently to feature transformation; a more elaborate subspace selecting strategy shall improve performance of transformed recognition. As for TLS solution, its property should be explored fully on the other database. In addition, if nonlinear part and more sophisticated classifier were concerned, near frontal view face recognition would be further improved. All above will be our future work.

We noticed that similar linear transformation matrix was used to generate virtual frontal image using PCA-type features in [13]. In present work, face recognition was directly performed on transformed features without image synthesis; also our experiments showed Maximum

Discriminant Feature (MDF), such as R-LDA feature, has the same performance with Maximum Expression Feature (MEF), such as PCA features. It also proved that image synthesis was not necessary. Furthermore, different solutions as well as some properties of transformation matrix were also explored.

Acknowledgement

We acknowledge authors of the database providing their database for our experiments. This work is supported by Hundreds Talent Program of CAS and Pujiang Program under the contract No.05PJ14111.

Table 1. Recognition rates on the FERET database. Numbers in brackets denote the dimension of subspace where corresponding score achieved. “Dir.” means directly recognition; “trans.” denotes the transformed recognition.

			be(15º)	bd(25º)	bc(45º)	bb(60º)
LDA	Dir.	ED	1.0	1.2	1.3	1.3
(99)	Trans.	ED	19.8 (98)	10.8 (96)	4.5 (99)	3.7 (78)
		AD	38.1 (99)	23.2 (99)	9.6 (99)	5.3 (78)
RLDA	Dir.	ED	1.5	1.2	1.3	1.3
(99)	Trans.	ED	86.4 (99)	74.9 (99)	47.7 (99)	28.1 (99)
		AD	87.0 (99)	78.1 (99)	52.0 (99)	29.3 (99)
PCA	Dir.	ED	7.6	2.18	1.30	1.25
(199)	Trans.	ED	90.3 (78)	83.6 (106)	63.2 (104)	35.4 (134)
		AD	88.9 (76)	82.0 (128)	61.4 (103)	39.0 (94)
ICA	Dir.	ED	1.13	1.13	1,15	1.15
(200)	Trans.	ED	89.1 (197)	82.2 (190)	60.1 (196)	34.1 (191)
		AD	89.1 (198)	82.1 (191)	59.8 (196)	34.2 (192)
Ref. 4			99.5	96.9	95.4	94.8
Ref. 7			77.5	55.5	N/A	N/A

Table 2. Recognition rate (RR) varying with model set size (M-Size) on “be” part of FERET. Dimension of subspaces where best score occurred is in row of “dim”. AD-2 as similarity measurement for RLDA and ED for the others.

M-Size		15	20	25	50	75	100	125	150
PCA	RR	69.1	75.3	78.4	87.2	90.7	91.9	92.7	92.7
	dim	22	34	38	60	76	79	90	71
ICA	RR	70.15	76.2	80.23	86.6	89.7	91.0	91.7	92.2
	dim	30	40	49	99	134	170	212	230
RLDA	RR	55.7	61.6	66.6	82.6	88.9	90.6	91.8	92.3
	dim	14	19	24	49	73	99	124	149

Reference

[1] Theagenis J. Abatzoglou, Jerry M. Mendel, “Constrained Total Least Squares”, *in Proc. 1987 IEEE ICASSP,* (Dallas, TX), pp. 1485-1488, 1987.

[2] Theagenis J. Abatzoglou, Jerry M. Mendel, Gail A. harada, “the Constrained Total Least Squares Technique and Its Appliacation to Harmonic Superreolution”, IEEE trans. Signal Processing, Vol. 39, no. 5, pp. 1070-1087, 1991.

[3] M. S. Bartlett，J. R. Moellan, T. J. Sejnnowski. “Face Recognition by Independent Component Analysis.” *IEEE trans. Neural Networks*, Vol. 13, no. 6, pp. 1450-1464, 2002.

[4] E. Bingham, A. Hyvarinen. “A Fast Fixed-Point Algorithm for Independent Component Analysis, ” *Int. J. of Neural Systems*, Vol. 10, no. 1, pp.1–8, 2000.

[5] P. N. Belhumeur, J. P.Hespanha, D. J. Kriegman, “Eigenface vs. Fisherface: Recognition Using Class Specific Linear Projection,” *IEEE trans. Pattern Analysis and Machine Intelligence*, Vol. 19, no. 7, pp. 711-720, 1997.

[6] Blanz, T. Vetter. “Face Recognitin Based on Fitting a 3D Morphable Model,” *IEEE trans. Pattern Analysis and Machine Intelligence,* Vol. 25, no. 9, pp. 1-12, 2003.

[7] D. Beymer. “Face recognition under varying pose,” *Technical Report 1461, MIT AI Laboratory*, 1993

[8] D. Beymer and T. Poggio. “Face Recognition from One Example View.” *A.I. Memo No. 1536, MIT AI Lab*, 1995

[9] X. Chai, S. Shan, W. Gao, X. Liu. “Pose Normalization Using Generic 3D Face model as a Priori for Pose-Insensitive Face Recognition”, pp. 144-152, *Sino-Biometrics*, 2004

[10] T. Cootes, G. Wheeler, K. Walker, and C. Taylor. “View-based Active AppearanceModels,” *Image and Vision Computing*, Vol. 20, pp. 657-664, 2002

[11] R. Gross, S. Baker, I. Matthews, Takeo Kanade. “Face Recognition Across Pose and Illumination,” *Handbook of Face Recognition*, page 193-216, ISBN 0-387-40590-x, Springer, 2004

[12] Lanitis, C. J. Taylor and T. F. Coots, Automatic Interpretation and Coding of Face Images Using Flexible Models, *IEEE trans. Pattern Analysis and Machine Intelligence,* Vol. 19, no. 7, pp.743-976, 1997.

[13] Hyung-Soo Lee, Daijin Kim, “Generating frontal view face image for pose invariant face recognition,” *Pattern Recognition Letters*, Vol. 27 pp. 747–754, 2006.

[14] J. Lu, K.N. Plataniotis, A.N. Venetsanopoulos. “Regularization Studies of Linear Discriminate Analysis in Small Sample Size Scenarios with Application to Face Recognition,” *Pattern Recognition Letter*, Vol. 26, no. 2, pp. 181-191, 2005

[15] X. Lu. “Image Analysis for Face recognition,” *http:/www.cse.msu.edu/*

[16] H. Murase and S. Nayar. “Visual Learning and Recognition of 3-D Objects from Appearance.” *International Journal of Computer Vision*, Vol. 14, pp. 5-24, 1995.

[17] Pentland, B. Moghaddam, and T. Starner. “View-based and Modular Eigenspaces for Face Recognition.” *In Proc. IEEE 1994 CVPR*, pp. 84-91, 1994

[18] P.Jonathon Phillips, “Face Recognition Vendor test 2002: Overview and summary”, *http://www.frvt.org*.

[19] Frank Y. Shih, Camel Y. Fu, Kai Zhang, “Multi-view face identification and pose estimation using B-spline interpolation”, *Information Sciences*, Vol.169, pp. 189–204, 2005.
[20] M. Turk and A. Pentland. “Face Recognition Using Eigenfaces.” *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* 1991
[21] C. Sanderson, S. Bengio, “Statistical Transformation techniques for Face Verification Using Faces Rotation in Depth”, *IDAP-RR 04-04*, 2004
[22] http://www.itl.nist.gov/iad/humanid/feret/
[23] Shimon Ullman and Ronen Barsi, Recognition by Liinear combinations of models, *IEEE trans. Pattern Analysis and Machine Intelligence*, Vol. 13, no. 10, pp. 992-1006, 1991.
[24] X. D. Zhang, *Matrix Analysis and Applications*, Tsinghua and Springer Publishing house, Beijing, 2004
[25] Z. Zhong, J. Li, D. Zhao, and Y. Hong, Effective Pose Estimation from Point Pairs, *Image and Vision Computing*, Vol. 23, pp. 651-660, 2005.

Hongzhou Zhang received MS degree in Physics in 2004 from Northwest University. He is a PhD candidate of Graduate school of Chinese Academy of Sciences now. His research is focused on digital image processing and pattern recognition.

Yongping Li received MS degree in the Electrical Engineering in 1989 from the Graduate School of Chinese Academy of Sciences. He worked as research fellow at Shanghai Institute of Applied Physics for various research projects on computerized instrumentations from 1989-1995. He joined the Center for Vision Speech and Signal Processing, Univ. of Surrey, UK in 1996 with a Royal Society fellowship, and received PhD in Pattern Recognition, with emphasis on Biometrics, in 2000 under the supervision of Professor J.Kittler. He worked on speech signal processing at a startup company in Silicon Valley, California from 2001 to 2002. From 2003 to present, he is a research professor and PhD student supervisor at Shanghai Institute of Applied Physics, and a part-time PhD student supervisor at the Center for Biometrics and Security Research (CBSR), Institute of Automation, Chinese Academy of Sciences. His current research interests include biometric algorithms for face and fingerprint recognition, biometric solution integration, biometrics system performance evaluation, testing, and standardization. Dr. Li is a member of the IEEE, IEEE computer society and IEEE instrumentation and measurement