World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:3, No:12, 2009

# Video-based Face Recognition: A Survey

Huafeng Wang, Yunhong Wang, And Yuan Cao

International Science Index, Computer and Information Engineering Vol:3, No:12, 2009 waset.org/Publication/15131

***Abstract*—**During the past several years, face recognition in video has received significant attention. Not only the wide range of commercial and law enforcement applications, but also the availability of feasible technologies after several decades of research contributes to the trend. Although current face recognition systems have reached a certain level of maturity, their development is still limited by the conditions brought about by many real applications. For example, recognition images of video sequence acquired in an open environment with changes in illumination and/or pose and/or facial occlusion and/or low resolution of acquired image remains a largely unsolved problem. In other words, current algorithms are yet to be developed. This paper provides an up-to-date survey of video-based face recognition research. To present a comprehensive survey, we categorize existing video based recognition approaches and present detailed descriptions of representative methods within each category. In addition, relevant topics such as real time detection, real time tracking for video, issues such as illumination, pose, 3D and low resolution are covered.

***Keywords*—**Face recognition, video-based, survey

## I. Introduction

SINCE proposed for the first time in 1880s, face recognition has received significant attention and became one of the most successful subareas of pattern recognition. Most of the algorithms demonstrate promising research while dealing with still facial images, which include Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Elastic Graph Matching (EGM) and so on. Compared with still images, video can provide more information, such as spatio-temporal information. Therefore, video-based face recognition gained more attention recently.

This paper surveys video-based face recognition which is on the scene for more than 20 years [1]. In the very beginning, most of the methods were based on the still-to-still techniques which aimed at selecting good frame and did some relative processing. Recently researchers began to truly solve such problems by spatio-temporal representations [2]. Most of the existing systems address video-based face recognition problems as follows: First, detect face and track it over time. Sometimes selecting good frames which contain frontal faces or valued cues is necessary. Next, when a frame satisfying certain criteria (size, pose, illumination and etc.) is acquired, recognition is performed, sometimes, by using still-to-still recognition technique. Figure 1 shows the whole process. In addition, some methods also utilize combination cues, such as audio, gait and so on, to make a comprehensive analysis and take decision.

In this paper, we review the current development of video-based face recognition. It is organized as follows: Section 2 introduces the typical techniques of face detection in video, real-time & multi-view methods. In section 3 the typical face tracking methods are discussed. Section 4 compares with still images, and lists the advantages and disadvantages of face recognition in video. The ways especially for solving the problems such as illumination, pose and low resolution and 3D are introduced. Section 5 presents several well-known video-based databases. A summary and discussion are presented in section 6.

Huafeng Wang is with the School of Computer Science And Engineering, Beihang University, China, and doing PhD research on pattern recognition (e-mail:wanghuafeng@)buaa.edu.cn).
Yunhong Wang is with the School of Computer Science And Engineering, Beihang University, China (e-mail:yhwang@buaa.edu.cn).
Yuan Cao is with the Civil Aviation Medicine Center (CAMC),China(e-mail: cawyly@gmail.com).

## II. Face Detection

Face detection is the first stage of a face recognition system. A lot of research has been done in this area, most of that is efficient and effective for still images only. So could not be applied to video sequences directly. In the video scenes, human faces can have unlimited orientations and positions, so its detection is of a variety of challenges to researchers.

Generally, there are three main processes for face detection based on video. At first, it begins with frame based detection. During this process, lots of traditional methods for still images can be introduced such as statistical modeling method[3], neural network-based method[4], SVM-based method[5], HMM method[6], BOOST[7] method and color-based face detection[8], etc. However, ignoring the temporal information provided by the video sequence is the main drawback of this approach. Secondly, integrating detection and tracking, this says that detecting face in the first frame and then tracking it through the whole sequence. Since detection and tracking are independent and information from one source is just in use at one time, loss of information is unavoidable. Finally, instead of detecting each frame, temporal approach exploits temporal relationships between the frames to detect multiple human faces in a video sequence. In general, such method consists of two phases, namely detection and prediction and then update-tracking. This helps to stabilize detection and to make it less sensitive to thresholds compared to the other two detection categories.

### A. Typical Approaches

In 2000, Zhu Liu and Yao Wang [9] presented a fast template matching procedure by iterative dynamic programming(DP) to detect frontal faces and track non-frontal faces with online adapted face models. Meanwhile, a fact was observed that a higher edge concentration appeared in the vicinity of facial features but less edge concentration appeared when slightly outside of facial features. Base on this fact, Li Silva et. al. [10] proposed a method, named edge pixel counting, to detect and track facial features in video sequences. In [11], Han et al. accomplished tasks of detecting and tracking multiple objects of unknown and varying number by using a graph structure that maintains multiple hypotheses. And in [12], automatic appearance models were built based on appropriate clustering over video segments. In addition, some approaches combined Edge Orientation Features to enhance the efficiency of detection [13] . In order to fully use the temporal information provided by video, [14] proposed a face detection method which made use of local histograms of wavelet coefficients represented with respect to a coordinate frame fixed to the object. What is more, Zhenqiu Zhang [15] et al. proposed Floatboost based face detection to make a local decision, and then utilized temporal information to confirm and validate the results.

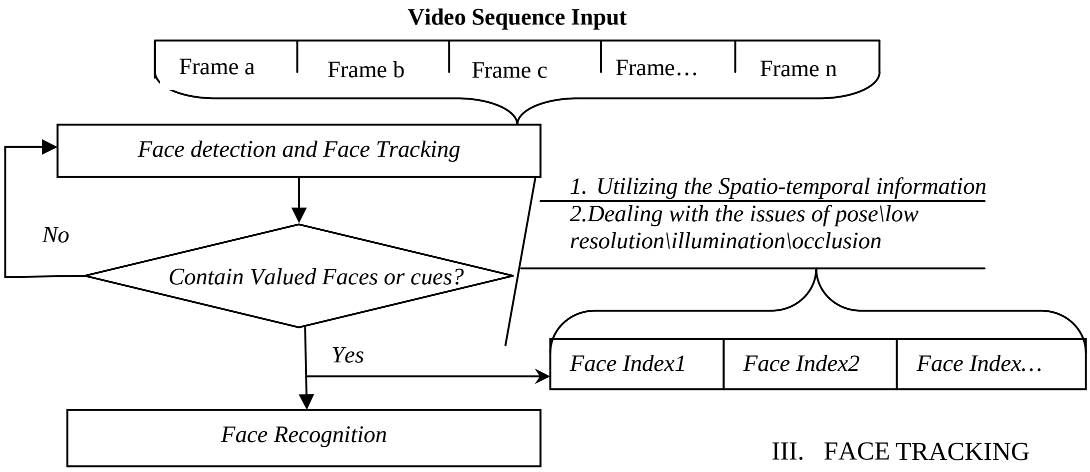


Fig. 1 Process of face recognition in video

### B. Real-time and multi-view approaches

Real-time and multi-view face detection is important in the video-based face detection. A real-time face detection methods could be classified into two groups: the first group consists of Cascade AdaBoost approach proposed by Viola and Jones[7], and the second group uses color information to detect and validate human face[16]. A non-parametric statistical technique is exploited by Bradski et al. to detect faces in 3D [17].

Schneiderman and Kanade[18] claimed their system to be the first one in the world for multi-view face detection. Over past years, much progress had been made in the area. There exist two main methods:(A) The method concerns building a single detector to deal with all views of faces ;(B)The method focuses on building several detectors response for different views.

Feraud et al.[19] used an array of 5 detectors with each detector for one view. The detectors rotates to perceive the pose changes, based on trajectories of faces in linear PCA feature spaces, [20] provided a useful mechanism for investigating these changes. In addition, detector-pyramid architecture was presented in [21], which adopted an integrated strategy of coarse-to-fine view decomposition[22], and simple-to-complex face or nonface classification. For achieving the minimum error rate, Li and Zhang[23] proposed an algorithm by integrating the principle of both Cascade AdaBoost and detector array.

However, as far as most of these approaches are concerned, a serious problem occurs because in-class variability of multi-view faces dataset is larger than that of front-view faces dataset. Though Detector-Pyramid Architecture AdaBoost (DPAA)[24] is able to handle this problem, the complexity increased leads to high computational load and over-fitting in training. Over-fitting has been addressed in [25], however, robust approaches are still required.

## III. FACE TRACKING

ınt procedure in face recognition system. Tracking exploits the temporal correspondence between frames. In general, face tracking can be divided into three categories, head tracking (color-based, model-based and shape-based), facial feature tracking and combination of head and facial feature tracking. For video processing, real-time is its foremost feature to be considered.

### A. Typical Approaches

First category for tracking is model-based, they usually include statistical models [26] and exemplar-based [27]. Hongli et al. applied effective boundary saliency map to the subsequent tracking process on the basis of previous segmented result [28]. Generally, the major steps in tracking stage are the boundary matching and connection. The former construct to determine the face boundary and the latter is employed to extract the region between two key points.

1) *Model-based Approaches.*
   Hager and Belhumeur[29] used a parametric model for tracking. Active Appearance Model (AAM) was

introduced by Cootes et al. [30], which contains a statistical model of the shape and grey-level appearance of the object of interest to encode shape and texture information. Based on the AAM, Jorgen Ahlberg [31] presented an Active Model for facial feature tracking. Adaptive template tracking [32] extended the idea of AAM and gained better results. Isard and Blake combined condensation algorithm with active contours which parameterized by low dimensional vectors [33]. The similar algorithm was generalized by Seo et al.[34], which took advantage of active contour with adaptive condensation as well and utilized color information.

2) *Color and shape based Approaches.*
Color and shape are important cues for tracking, based on which many methods are proposed. In [35], a robust face tracking method based on condensation algorithm which combined skin color with facial shape was presented. Moreover, skin color and facial shape trackers were constructed as the observation measure respectively. The result showed that the method was even robust in the situation of sudden change of lighting. Similarly, F.M Noguer and Alberto Sanfeliu [36] proposed the use of a new colorspace method based on Linear Discriminate Analysis method which integrated color and shape into condensation algorithm. Ying Ren and Chin Seng Chua integrated tracking procedure with the spatial domain and proposed a Bilateral Learning (BL) approach [37]. Their algorithm mainly selected reliable samples to update the color and spatial model within EM framework. Moreover, it didn't need accurate shape model. In [38], an enhanced mean-shift tracking approach using joint spatial-color feature and a novel similarity measure function was proposed.

Second and recently, facial feature tracking comes into focus. Though the approach is very sensitive to occlusion, scale or resolution changes, it is precise and reliable under planar movement. Yan Tong and YangWang[39] applied a model to simultaneously characterize the global shape constraints and the local structural details of human faces. Meantime, gabor wavelet jets and gray-level profiles are integrated for effective and efficient feature representation. In addition, a multi-model tracking approach is used to estimate the feature point positions dynamically. In [40], tracking was fitted into Kalman filtering framework in which the non-linear system was approximated by a local linear model. Christian Kublbeck and Andreas Ernst [41] presented tracking by means of continuous detection and introduced illumination invariant local structure features within Kalman filter.

The third category is tracking based on the combination of head and feature. Brais Martinez and Xavier Binefa [42] presented a method for tracking several features of a non-planar target undergoing an arbitrary 3D movement. Multiple kernel tracking methods can track objects undergoing parametric transformations. Moreover, the method extends the result to more general situations. Indra Sulistijono and Naoyuki Kubota [43] proposed an evolutionary method of SSGA and Particle Swarm Optimization (PSO) to perform the multiple human heads tracking. Besides, it could also reduce the computational cost. In [44], a data fusion particle filter for head tracking method was proposed based on the color and edge cues. A Boosted Adaptive Particle Filter (BAPF) was presented [45] to enable estimation of the proposal distribution and the posterior distribution with a much higher degree of accuracy. Besides, Birchfield[46], J. MacCormick [47] and Raja et al.[48] made their contributions separately.

### B. *Real-time tracking--methods*

Real-time tracking catches much attention recently. Existing real-time tracking techniques include: CAMSHIFT [49], condensation [50] and adaptive Kalman filtering. When the object is far away from the camera, the algorithm fails to track. Due to the efficiency in tracking performance and computational capacity, chroma distribution-based face tracking approaches have been presented. For example, Yao and Gao [51] proposed a face tracking algorithm based on the skin and lip chroma transform. Huang and Chen [52] built a statistical color model and deformable template for tracking multiple faces. In [53], it exploited a statistical approach based on the mean-shift algorithm, which is consisted of a gradient ascent search over the skin color distribution. In [54], the head is modeled as a texture mapped cylinder and tracking was formulated as an image registration problem in the cylinder's texture map image.

## IV. Face Recognition

Face recognition is the most significant stage in the whole system. Parts of the video based algorithms utilize approaches on the basis of still-to-still techs. However, videos are capable of providing more information than still image. There are four major advantages for using video: First is the possibility of employing redundancy contained in the video sequence to improve still images recognition performance. Second, recent psychophysical and neural studies [55] have shown that dynamic information is very crucial in the human face recognition process. Third, more effective representations, such as a 3D face model [56] or super-resolution images [57], can be acquired from the video sequence and be used to improve recognition effects. Fourth, besides those motivations mentioned above, video-based recognition allows learning or updating the subject model over time [58].Though the advantages are obvious, there also exits some disadvantages. For example, poor video quality, low image resolution, and other influence factors (such as illumination, pose change, motion, occlusion, decoration, expression, large distance from camera, etc). In spite of all those advantages and disadvantages, there are various aspects of approaches for video based face recognition.

### A. *Spatio-temporal information based approaches*

Most of the recent approaches utilize spatio-temporal information for face recognition in video. Typically, some [59] use temporal voting to improve identification rates. There are also several algorithms which extract 2D or 3D face structure from the video[60]. Other than simple voting approaches, Li et al.[61] proposed a method based on shape and texture models and kernel feature extraction as well. However, such method doesn't fully use the coherence information. Zhou and Chellappa[62] presented a method for incorporating temporal information in a video sequence for the task of human recognition. A state space model with tracking state vector and recognizing identity variable was used to characterize the identity. This probabilistic approach aimed to integrate motion and identity information over time though sequential importance sampling algorithm (SIS); it nevertheless considered only identity consistency in temporal domain and thus it may not work well when the target is partially occluded.[63] compared PCA, LDA and ICA in multiple images with those in video sequences, it is proved that weighed probabilistic approach can solve the problems, namely occlusion errors of localization, existed in the single still image. In [64], Krueger and Zhou selected representative face images as exemplars from training videos by on-line version of radial basis functions. This model is effective in capturing small 2D motion but it may not deal well with large 3D pose variation or occlusion. Li et al. [65] applied piecewise linear models to capture local motion. And a transition matrix among these models is taken to describe nonlinear global dynamics. Similar method was proposed by Kuang-Chih Lee [66], which took the way of propagating the probabilistic likelihood of the linear models through the transition matrix. The condensation algorithm could be used as an alternative to model the temporal structures [67].

The methods based on spatio-temporal representations for face recognition in video have some drawbacks: (i) though the local information is very important to facial image analysis, it is not well exploited; (ii) personal specific facial dynamics are useful for discriminating between different persons, however the intra-personal temporal information which is related to facial expression and emotions is also encoded and used; and (iii) equal weights are given to the spatiotemporal features despite the fact that some of the features contribute to recognition more than others;(iv) a lot of methods can only handle well aligned faces thus limiting their use in practical scene[68].

### *B. Statistic model based approaches*

Zhou et al. [69] obtained statistical models from video by using low level features (e.g., by PCA) contained in sample images, which was used to perform matching between a single frame and the video stream or between two video streams. Satoh [70] matched two video sequences by selecting the pair of frames those were closest across the two videos, which is still-to-still matching inherently. A few methods use video sequence to train a statistical model face for matching. The mutual subspace method in [71] took the video frames for each person separately to compute many individual eigenspaces, considering the angle between input and reference subspaces formed by the principal components of the image sequences as the measure of similarity. In [72], a method was proposed by using kernel principal angles on the original image space and using a feature space as the measure of similarity between two video sequences. For the sake of improvement, in [73],the author proposed simple algorithm based on facial features and positions to select the representative frames, then dimensional analyses were applied to transform them into new spaces. In [74], the proposed scheme achieved better performance to learn a sparse representation from video clips for online face recognition in an unconstrained environment. In [75], a new classification algorithm, namely principle component null space analysis (PCNSA),is designed that is suitable for the problem in which different classes have unequal and nonwhite noise covariance matrices.

Recently, the Auto-Regressive and Moving Average (ARMA) model [76] was used to model a moving face as a linear dynamical system and perform recognition. The widely used Hidden Markov models (HMM) have also been applied to face recognition in video. Liu et al. [77] used HMM and ARMA models for direct video level matching. In [78], it showed that the problem of visual constraints could be solved by HMM-based recognition framework.

### *C. Hybrid cues based approaches*

Video can provide more information than still image. Some methods utilize other cues obtained from video sequences, such as voice, gait, motion etc. For example, Shan et al. investigated the fusion of face and gait at feature level and gained performance increase by combining the two cues [79]. In [80], presented a new approach based on integrating information from side face and gait at the feature level by PCA and MDA. [81] adopted a face and speaker recognition techniques for audio-video biometric recognition. The paper combined histogram normalization, boosting technique and a linear discrimination analysis to solve the problem like illumination, pose and occlusion and proposes an optimization of a speech denoising algorithm on the basis of Extended Kalman Filter(EKF). In [82], an approach was presented by radial basis function neural networks, which is used to recognize a person in video sequences by using face and mouth modalities.

### *D. Advanced Topics*

For the past several years, more popular areas of video-based face recognition technology are as follows:

1) *Illumination*

   Many factors influence face recognition, among them the major two challenges are: illumination and pose. It is difficult for system to make recognition of individuals when change in light is larger. Adini, Moses, and Ullman[83] first observed it. But, Zhao and Chellappa[84] gave a theoretical proof of it on the basis of eigenface system projection. To handle such problems, the researchers propose various approaches.

   To handle such problems, the researchers have already proposed various approaches during these years. Belhumeur et al. [85] and Bartlett et al. [86] adopted the

PCA by discarding the first few principal components and achieved better performance for images under different lighting conditions. Their assumption is that first principal components capture only variations due to lighting. Consequently, some important discarded can influence the recognition under normal lighting conditions. In addition, some approaches are presented based on thoroughly mining the features of image. In [87], the Discrete Cosine Transform was employed by Chen et al. to compensate for illumination variations in the logarithm domain. Jacobs et al. [88] presented a method based on the fact that, for point light sources and objects with Lambertian reflectance, the ratio of two images from the same object is simpler than the ratio of images from different objects. Nanni et al. [89] proposed local based methods based on the Gabor filter. Liu et al. [90] used a ratio image to solve the illumination variation. Similar method had been proposed by Wang, et al. [91], which aimed to acquire an illumination-invariant face feature image for a group of images of the same subject. In [92], a hybrid approach based on the use of PCA and correlation filters was proposed. In Du et al. [93] a wavelet based normalization method was presented.

Local Binary Pattern (LBP) has attracted much attention since first proposed by Ojala et al. Some other researches also made respective contributions to this method, for instance, multi-resolution LBP[94] was presented where neighborhoods of different sizes are considered to deal with textures at different scales, and the uniform LBP, characterized by at most one 0–1 or 1–0 transition, to better represent primitive structural information such as edges and corners. Zhang et al. [95] proposed to couple the LBP representation with Gabor phases. Local ternary pattern (LTP)[96] was proposed by Tan and Triggs, which was also an extension of LBP.

Recently, in [97], an effective method of handling illumination variations was presented by using illumination cone. This method also dealt with shadowing and multiple lighting conditions which was on the basis of 3D linear subspace. The main side effect of this method is that the training set requires more than 3 aligned images per person.

2) *pose issues*

Pose is another most important factor for face recognition system. Current approaches can be divided into three groups: multiple images approaches, hybrid approaches and single image based approaches. In multiple images approaches, illumination cone and 3D surface based methods have been proposed to solve both illumination and pose problems. The hybrid approaches might be the most practical solution up to now, including linear class based method[98] which is on the basis of assumption of linear 3D object classes and extension of linearity to images, graph matching based method[99] with EBGM and view-based eigenface method[100] by constructing eigenfaces for each pose. The last group was proposed, but it is hard to apply currently due to high computational cost and complexity.

New AAM methods [101] have been proposed to handle both varying pose and expression. In [102], Eigen light-fields and Fisher light-fields method was proposed to do pose invariant face recognition. A method by 3D model of the entire head for exploiting features like hairline, which handled large pose variations in head tracking and video-based face recognition was presented [103]. Computing the Kullback-Leibler divergence between testing image sets and a learned manifold density was the other thought [104]. In [105] learns manifolds of face variations for face recognition in video. In [106], the research said they achieved pose robustness by decomposing each appearance manifold into semantic Gaussian pose clusters, comparing the corresponding clusters and fusing the results by RBF network.

*3) 3D researches*

Face recognition based on 3D is a hot research topic. Generally, reprehensive methods can be divided into three main categories, namely, 2D images based, 3D images based and multimodal systems.[107] The differences among these three categories are as follows: the first category includes approaches which use 2D images and 3D generic face model to improve the robustness and recognition rate. And for the second one, the methods work directly on 3D datasets. While the last group means those which utilize both 2D and 3D information.

An example is given by Blanz and Vetter in [108] that proposed a method to create 3D face models from a single image. Zhang and Cohen morphed 3D generic model from multi-view images by way of using a cubic polynomial [109]. But, it is still doubtable that whether 3D facial reconstruction from a single view image or multi-view images can be considered good enough.

Since 2000, more and more multimodal approaches have been proposed to improve face recognition performance. Dalong Jiang et al. [110] proposed an efficient and fully automatic 2D-to-3D integrated face reconstruction method in an analysis-by-synthesis manner. 3D face shape was reconstructed according to the feature points and a 3D face database. Then the face model was texture-mapped by projecting the input 2D image onto the 3D face shape. The author synthesizes virtual samples with variant PIE to represent the 2D face image space. In [111], the system was based on real-time quasi-synchronous color and 3D image acquisition was based on the color structured-light approach. 3D information made the segmentation and detection straightforward with mixture of Gaussians assumption. The parameters were estimated by Expectation Maximization algorithm. Besides, it also made the pose and illumination compensate for each other, which led to the improvement of face recognition.

*4) Low Resolution*

It is difficult to recognize human faces in video of low resolution. With the widely use of camera (surveillance etc.), solutions which solve such problems achieve more and more attention. The main two methods are Super Resolution (SR) and Multiple Resolution-faces (MRF) approach. The former can be applied to estimate high-resolution facial image from low-resolution ones. However, the disadvantage is that the multiple facial images that belong to the same subject captured from same scene are required. MRF[112] overcame such drawback, it increases the complexity and requires more memory storage in face recognition system. Recently, researchers improved existing methods of SR & MRF and proposed some new methods.

In [113], color invariance was applied to face recognition. Their results showed that color invariants do have substantial discriminative power and increase the robustness and accuracy for low resolution facials. In [114], the author proposed an approach to provide a feature subspace to be directly compatible with randomly changeable low-resolution of probe at application/testing stage and overcome dimension mismatch problem. In [115], the authors constructed high-resolution frames from a video sequence by using both spatial and temporal information present in a number of adjacent low-resolution ones. A new technique named face scoring was given by Tse-Wei Chen et al.[116]. The method included eight scoring functions based on feature extraction technique, integrated by a single layer neural network training system to obtain an optimal linear combination to select high-resolution faces.

## V. EVALUATION AND DATABASES

It is recommended to use a standard test dataset to benchmark an algorithm. Table 1 lists some famous video based face databases and some databases dealing with issues of illumination, pose, etc.

TABLE 1 FACE DATABASES

| Database Name | Purpose | No. of Subject | Image Resolution |
|---|---|---|---|
| Face In Action (FIA) | Illumination/Pose | 200 | 640*480 pixels at 30 frames per second |
| | http://amp.ece.cmu.edu/projects/FIADataCollection/ | | |
| The Extended M2VTS | Speech Sequence Head /Rotation Sequence | 295 | 720*576 |
| | http://www.ee.surrey.ac.uk/CVSSP/xm2vtsdb/ | | |
| Max Planck Institute (MPI) | Facial action/Pose | 246 | 786x576 at 25 frames per second |
| | http://vdb.kyb.tuebingen.mpg.de/ | | |
| VidTIMIT | head rotation sequence/ Speech Sequence | 43 | 512 x 384 pixels |
| | http://www.itee.uq.edu.au/~conrad/vidtimit/ | | |
| EQUINOX HID | Illumination/Facial Expression/ Speech Sequence | 91 | 240*320 pixels |
| | http://www.equinoxsensors.com/products/HID.html | | |
| Texas Video Database | Pose Facial expression | 284 | 720*480 pixels |
| | http://www.utdallas.edu/dept/bbs/FACULTY PAGES/otoole/database.htm | | |
| The Yale Face Database B | Illumination Pose | 10 | 640*480 pixels |
| | http://cvc.yale.edu/projects/yalefacesB/yalefacesB.html | | |
| PIE | Illumination Pose | 68 | 512 x 384 pixels |
| | http://www.ri.cmu.edu/research_project_detail.html?project_id=418&menu_id=261 | | |
| AR | facial expressions illumination, occlusions | 126 | 768 x 576 pixels |
| | http://rvl1.ecn.purdue.edu/~aleix/aleix_face_DB.html | | |
| CAS-PEAL | Illumination Pose Facial Expression | 1040 | 360 x 480 pixels |
| | http://www.jdl.ac.cn/peal/index.html | | |

## VI. DISCUSSIONS AND REMARKS

In this paper, we presented some major issues on video based face recognition. These fall into four groups:

Face detection: For the constrained conditions, many face detection methods for static image are not directly suitable to the task in video. We classified current approaches into three groups, and summarized their pros and cons.

Face tracking: it is a significant procedure in video-based face recognition. It usually exploits statistical model, examplar-based model, and skin color information to accomplish the tracking task. In addition, for these methods it also exploits CAMSHIFT, condensation, adaptive Kalman filter algorithms.

Face recognition: Since the spatio-temporal information plays a significant role in face recognition, how to fully exploit redundancy information in the video sequence is a key issue for video based recognition. In order to comprehensively understand the development on face recognition in video, in the first half of the paper, we classified the current approaches into two categories: methods without additional cues and methods with hybrid cues. In the later part of paper we thoroughly reviewed some of the developing topics, such as illumination and pose issues, 3D and low resolution.

Video database for face recognition: In this part, we list some famous video based face databases in tabular form for reference.

Out of several issues associated with the current systems, we discuss the following two issues:

1) Up to now, the databases used in many systems have been still small. Most of them contain about 200 subjects. This is partially due to the tremendous amount of storage space needed for video sequences. Before someone claims that the facial image processing & analysis system is reliable and robust, rigorous testing and verification on real-world datasets must be performed, including databases for face analysis and tracking in digital video. Fortunately, relatively large video databases exist, for example, the XM2VTSDB database, and the addition of video into the FERET databases. However, large-scale systematic evaluations are still lacking.
2) One of the chief advantages of video over still frames is that fact accumulation over multiple frames can provide better face recognition performance. However, surveillance videos are generally of low resolution containing faces mostly in non-frontal poses. Consequently, face recognition in video possesses more challenges to the current face recognition systems. Use of 3D face models has been suggested as a way to compensate for low resolution, poor contrast and non-frontal pose. By the way of constructing a 3D face model from multiple non-frontal frames in a video, and then generating a frontal view from the derived 3D model, and finally using a 2D face recognition algorithm to recognize the synthesized frontal view, the spatio-temporal information can be fully employed. Meantime, it will help solve the problem of occlusion, pose variance and illumination issues caused by video frame's poor quality.

In summary, we present a comprehensive survey on video-based face recognition. We have tried our best to provide researchers in the field with the up-to-date information of research on video-based face recognition.

## Acknowledgment

This work was supported by Program of New Century Excellent Talents in University, National Natural Science Foundation of China (No. 60575003, 60332010,60873158). Joint Project supported by National Science Foundation of China and Royal Society of UK (60710059), Hi-Tech Research and Development Program of China (2006AA01Z133), and the opening funding of the State Key Laboratory of Virtual Reality Technology and Systems (Beihang University).

## References

[1] W.Y. Zhao, R. Chellappa, A. Rosenfeld, P.J. Phillips, Face Recognition: A Literature Survey, ACM Computing Surveys,Vol:35, 2003.
[2] G. Aggarwal, A. R. Chowdhury, R. Chellappa, A system identification approachfor video-based face recognition, in: 17th International Conference on PatternRecognition, Vol. 4, 2004, pp. 175- 178.
[3] B. Moghaddam and A. Pentland, "Probabilistic Visual Learning for Object Representation," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 696-710, July 1997.
[4] H.A. Rowley, S. Baluja, and T. Kanade, "Neural Network-Based Face Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 1, pp. 23-38, Jan. 1998.
[5] B. Heisele, T. Poggio, and M. Pontil, "Face Detection in Still Gray Images," A.I. memo AIM-1687, Artificial Intelligence Laboratory, MIT, 2000.
[6] Ti-QiongXu; Bi-Cheng Li; Bo Wang, Face detection and recognition using neural network and hidden Markov models, Proceedings of the 2003 International Conference on Neural Networks and Signal Processing, pp. 228 - 231,2003.
[7] P. Viola and M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," Proc. Conf. Computer Vision and Pattern Recognition, pp. 511-518, 2001.
[8] R.L. Hsu, M. Abdel-Mottaleb, and A.K. Jain, "Face Detection in Color Images," Proc. Int'l Conf. Image Processing, pp. 1046-1049, 2001.
[9] Z. Liu and Y. Wang, "Face Detection and Tracking in Video Using Dynamic Programming" Proc. Int'l Conf. Image Processing, 2000.
[10] L. Silva, K. Aizawa and M. Hatori. "Detection and Tracking of Facial Features". Proc. Of SPIE Visual Communications and Image Processing, Taiwan. May, 1995.
[11] M. Han, A. Sethi, and Y. Gong. A detection-based multiple object tracking method. In Proc. Int. Conf. Image Process. (ICIP), pages 3065-3068, Singapore, Oct. 2004.
[12] D. Ramanan and D. Forsyth. Using temporal coherence to build models of animals. In Proc. Int. Conf. Comp. Vision(ICCV), pages 338-346, Nice, France, Oct. 2003.
[13] D. Maio and D. Maltoni. Real-time face location on grayscale static images. Pattern Recognition, 33:1525-1539, September 2000.
[14] Mikolajczyk, K.; Choudhury, R.; Schmid, C. ,Face detection in a video sequence-a temporal approach, Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol 2, pp: 96-101, 2001.
[15] Zhenqiu Zhang; Potamianos, G.; Ming Liu; Huang, T., Robust Multi-View Multi-Camera Face Detection inside Smart Rooms Using Spatio-Temporal Dynamic Programming, 7th International Conference on Automatic Face and Gesture Recognition, 2-6 April 2006 Page(s):407 - 412.
[16] J. Yang and A.Waibel. A real-time face tracker. In Proceedings of the Third IEEE Workshop on Applications of Computer Vision, pages 142-147, Sarasota, FL, 1996.
[17] G. Bradski. Computer vision face tracking for use in a perceptual user interface. Technical Report Q2, Intel Corporation, Microcomputer Research Lab, Santa Clara, CA, 1998.
[18] H. Schneiderman and T. Kanade. "A statistical method for 3d object detection applied to faces and cars". In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2000.
[19] J. Feraud, O. Bernier, and M. collobert. "A fast and accurate face detector for indexation of face images". In Proc. Fourth IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pages 52-59, 1998.
[20] S. Gong, S.McKenna, and J.Collins. "An investigation into face pose distribution". In Proc. IEEE International Conference on Face and Gesture Recognition, Vermont,1996.
[21] ZhenQiu Zhang; Long Zhu; Li, S.Z.; HongJiang Zhang,Real-time multi-view face detection, Fifth IEEE International Conference on Automatic Face and Gesture Recognition, pp:142-147,2002.
[22] P. Viola and M.J. Jones, "Robust real-time object detection", IEEE ICCV Workshop on Statistical and Computational Theories of Vision. Vancouver, Canada. July 13, 2001.
[23] S. Z. Li, Z. Q. Zhang, "FloatBoost Learning and Statistical Face Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. 26, NO. 9, September, 2004.
[24] Yan Wang; Yanghua Liu; Linmi Tao; Guangyou Xu ,Real-time multi-view face detection and pose estimation in video stream, 18th International Conference on Pattern Recognition ,pp: 354 - 357,2006.
[25] M. Nakamura, H. Nomiya and K. Uehara, "Improvement of boosting algorithm by modifying the weighting rule", Annals of Mathematics and Artificial Intelligence, 41:95-109,2004.
[26] C.J. Edward, C.J. Taylor, and T.F. Cootes, "Learning to Identify and Track Faces in an Image Sequence," Proc. Int'l Conf. Automatic Face and Gesture Recognition, pp. 260-265, 1998.

[27] K. Toyama and A. Blake, "Probabilistic Exemplar Based Tracking in a Metric Space," Proc. Int'l Conf. Computer Vision, vol. 2, pp. 50-57, 2001.
[28] Hongliang Li, King N. Ngan, Saliency model-based face segmentation and tracking in head-and-shoulder video sequences, Visual Communication & Image Representation, 2008,Vol.19(No.5).
[29] G.D. Hager and P.N. Belhumeur, "Efficient Region Tracking with Parametric Models of Geometry and Illumination", IEEE Transactions on Pattern Analyvsis and Machine Intelligence, vol. 20, 1998, pp 1025-1039.
[30] T. Cootes, G. Edwards and C.Taylor, "Active Appearance Models",IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, 2001, pp 681-685.
[31] Ahlberg J ,An Active Model for Facial Feature Tracking EURASIP Journal on Applied Signal Processing 2002,pp: 566-571.
[32] J Saragih, R Goecke, Monocular and Stereo Methods for AAM Learning from Video CVPR '07. IEEE Conference on Computer Vision and Pattern Recognition, pp:1-8,2007.
[33] M. Isard and A. Blake, "Condensation - Condition Density Propagation for Visual Tracking", Int. J. Coo puter Vision, vol. 29, 1998, pp. 5-28.
[34] K.-H. Seo, J-Y. Lee and J.J. Lee, "Adaptive Color Snake Tracker using Condensation Algorithm", IEEE 5th Asian Control Conference, 2004.
[35] Hyung-Soo Lee, Daijin Kim, Robust face tracking by integration of two separate trackers: Skin Color and facial shape, Pattern Recognition 40 (2007) ,pp:3225 - 3235
[36] Francesc Moreno-Noguer, Alberto Sanfeliu, Dimitris Samaras, Integration of deformable contours and a multiple hypotheses Fisher color model for robust racking in varying illuminate environments, Image and Vision Computing 25 (2007),pp: 285-296.
[37] Ying Ren, Chin Seng Chua,Bilateral learning for color-based tracking, Image and Vision Computing 26 (2008),pp:1530-1539.
[38] Jwu-Sheng Hu, Chung-Wei Juan, Jyun-Ji Wang,A Spatial-color mean-shift object tracking algorithm with scale and orientation estimation, Pattern Recognition Letters 29 ,2008 pp:2165-2173.
[39] Yan Tong, YangWang, Zhiwei Zhu, Qiang Ji ,Robust facial feature tracking under varying face pose and facial expression, Pattern Recognition 40 (2007), pp: 3195 - 3208.
[40] Tu JL, Tao H, Huang T;Face as mouse through visual face tracking; Computer Vision And Image Understanding; Vol:108(1-2),pp: 35-40, 2007.
[41] Christian Kublbeck, Andreas Ernst, Face detection and tracking in video sequences using the modified census transformation, Image and Vision Computing 24 (2006), pp: 564-572.
[42] Brais Martinez, Xavier Binefa, Piecewise affine kernel tracking for non-planar targets, Pattern Recognition 41 (2008), pp: 3682 - 3691.
[43] Indra Adji Sulistijono, Naoyuki Kubota, Evolutionary Robot Vision and Particle Swarm Optimization for Multiple Human Heads Tracking of A Partner Robot; 2007 IEEE Congress On Evolutionary Computation, Vol:1-10, pp: 1535-1541,2007.
[44] Yonggang Jin, Farzin Mokhtarian Data Fusion for Robust Head Tracking by Particles, Proceedings 2nd Joint IEEE International Workshop on VS-PETS, 2005.
[45] Wenlong Zheng, Suchendra M. Bhandarkar, Face detection and tracking using a Boosted Adaptive Particle Filter, Journal Of Visual Communication And Image Representation; Vol:20(1), pp:9-27,2009.
[46] S. Birchfield, "Elliptical Head Tracking Using Intensity Gradients and Color Histograms," Proc. Conf. Computer Vision and Pattern Recognition, pp. 232-237, 1998.
[47] J. MacCormick and A. Blake, "A Probabilistic Exclusion Principle for Tracking Multiple Objects," Proc. Int'l Conf. Computer Vision, 1995.
[48] Y. Raja, S.J. McKenna, and S. Gong, "Tracking and Segmenting People in Varying Lighting Conditions Using Color," Proc. Int'l Conf. Automatic Face and Gesture Recognition, pp. 228-233, 1998.
[49] D. Comaniciu, V. Ramesh, and P. Meer. Real-time tracking of non-rigid objects using mean shift. In Proc. of the CVPR, Hilton Head Island, S.C., U.S.A., 2000. Vol. 2, pp. 142-149.
[50] M. Isard and A. Blake, "Condensation-conditional density propagation for visual tracking", Int. Journal, Computer Vision, vol. 29, no. 1, pp. 5-28, 1998.
[51] H. Yao and W. Gao, "Face locating and tracking method based on chroma transform in color images", Signal Processing Proc.2000 , vol 2. pp. 1367-1371, 2000.
[52] F. Huang and T. Chen, "Tracking of multiple faces for human-computer interfaces and virtual environments," Int. Conf. Multimedia, Expo, pp. 1563-1566, vol. 3, 2000.
[53] Terrillon, J.-C.Pilpre, A.Niwa, Y.Yamamoto, K. ; DRUIDE : A Real-Time System for Robust Multiple Face Detection, Tracking and Hand Posture Recognition in Color Video sequences, Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004.
[54] M. L. Cascia and S. Sclaroff. Fast, reliable head tracking under varying illumination. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, Fort Collins, CO,June 1999.
[55] A.J. OToole, D.A. Roark, H. Abdi, Recognizing moving faces: a psychological and neural synthesis, Trends in Cognitive Science 6 (2002) 261-266.
[56] A. Roy Chowdhury, R. Chellappa, R. Krishnamurthy and T.Vo, "3D Face Recostruction from Video Using A Generic Model", In Proc. of Int. Conf. on Multimedia and Expo, Lausanne, Switzerland, August 26-29, 2002.
[57] S. Baker and T. Kanade, "Limits on Super-Resolution and How to Break Them", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 9, September 2002, pp.1167-1183.
[58] X. Liu, T. Chen and S. M. Thornton, "Eigenspace Updating for Non-Stationary Process and Its Application to Face Recognition", To appear in Pattern Recognition, Special issue on Kernel and Subspace Methods for Computer Vision, September 2002.
[59] G. Shakhnarovich, J. W. Fisher, and T. Darrell. Face recognition from long-term observations. In Proc. European Conf. on Computer Vision, volume 3, pp: 851-865, 2002.
[60] W. Y. Zhao and R. Chellappa. Symmetric shape-fromshading using self-ratio image. Int'l. J. Computer Vision, 45(1) pp:55-75, 2001.
[61] Y. Li, S. Gong, and H. Liddell. Constructing facial identity surface in a nonlinear discriminating space. In Proc. IEEE Conf. on Computer Vision and Pattern Recognition, volume 2, pp:258-263, 2001.
[62] S. Zhou and R. Chellappa. Probabilistic human recognition from video. In Proc. European Conf. on Computer Vision, volume 3, pp: 681-697, 2002.
[63] Zhang YB, Martinez AM, A weighted probabilistic approach to face recognition from multiple images and video sequence, Image And Vision Computing, Vol 24(6), pp: 626-638,2006.
[64] V. Kr¨ueger and S. Zhou. Exemplar-based face recognition from video. In Proc. European Conf. on Computer Vision, volume 4, pp: 732-746.
[65] Y. Li, T. Wang, and H.-Y. Shum. Motion textures: A two level statistical model for character motion synthesis. In Proc. SIGGRAPH, pp: 465-472, 2002.
[66] Kuang-Chih Lee; Ho, J.; Ming-Hsuan Yang; Kriegman, D., video-based face recognition using probabilistic appearance manifolds, Proceedings. 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 313-320, 2003.
[67] S. Zhou, V. Krueger, R. Chellappa, Face recognition from video: A condensation approach, in: IEEE Int. Conf. on Automatic Face and Gesture Recognition, 2002, pp. 221-228.
[68] Abdenour Hadid; Matti Pietik?inen; Combining appearance and motion for face and gender recognition from videos, Pattern Recognition, Vol:42(11),2009,pp: 2818-2827.
[69] Chellappa, R.; Kruger, V.; Shaohua Zhou, Probabilistic recognition of human faces from video, 2002 International Conference on Image Processing, Vol 1, 2002, pp. 41-45.
[70] S. Satoh. Comparative evaluation of face sequence matching for content-based video access. In IEEE International Conference on Automatic Face & Gesture Recognition, pp:163-168, 2000.
[71] Gregory Shakhnarovich, Baback Moghaddam Face Recognition in Subspaces, Handbook of Face Recognition, 2004.
[72] L. Wolf and A. Shashua. Kernel principal angles for classification machines with applications to image sequence interpretation. In Proc. of Intl. Conf. on Computer Vision and Pattern Recognition, 2003.
[73] Topkaya, I.S.; Bayazit, N.G.; Improving Face Recognition from Video with Preprocessed Representative Faces, 23rd International Symposium on Computer and Information Sciences, 2008. ISCIS '08. Page(s):1 - 4.
[74] Tangelder JWH, Schouten BAM ;Learning a Sparse Representation from Multiple Still Images for online face recognition in an Unconstrained Environment; 18th International Conference on Pattern Recognition, Vol 3, pp: 1087-1090,2006.

[75] Vaswani N, Chellappa R; Principal Components Null Space Analysis for Image and Video Classification; IEEE Transactions on Image Processing, 2006; pp: 1816-1830.
[76] S. Soatto, G. Doretto, Y. Wu, Dynamic textures, in: International Conference on Computer Vision, Vol. 2, Vancouver, BC, Canada, 2001, pp. 439-446.
[77] X. Liu and T. Chen, "Video-based face recognition using adaptive hidden markov models", Proc. IEEE Conference on Computer Vision and Pattern Recognition, Vol. 1: 340-345, 2003.
[78] Minyoung Kim; Kumar, S.; Pavlovic, V.; Rowley, H.; Face Tracking and Recognition with Visual Constraints in Real-World Videos, IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008, 23-28 June 2008 Page(s):1 - 8.
[79] C. Shan, S. Gong, P. McOwan, Learning gender from human gaits and faces,IEEE International Conference on Advanced Video and Signal based Surveillance, 2007, pp:505-510.
[80] Zhou X, Bhanu B ;Feature fusion of side face and gait for video-based human identification; Pattern Recognition,Vol:41(3); pp: 778-795,2008.
[81] Christian Micheloni , Sergio Canazza , Gian Luca Foresti; Audio-video biometric recognition for non-collaborative access granting; Visual Languages and Computing,2009.
[82] M. Balasubramanian , S. Palanivela, and V. Ramalingama; Real time face and mouth recognition using radial basis function neural networks; Expert Systems with Applications, Vol:36(3), pp: 6879-6888.
[83] Y Adini, Y Moses, S Ullman," Face recognition: the problem of compensating for changes in illumination direction" ,IEEE Transactions on Pattern Analysis and Machine, Vol.19: 721-732,1997.
[84] WY Zhao, R Chellappa "Illumination-insensitive face recognition using symmetricshape-from-shading" IEEE Conference on Computer Vision and Pattern Recognition, 2000.
[85] PN Belhumeur, JP Hespanha, DJ Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection" IEEE Transactions on pattern analysis and machine, Vol 19:711-720,1997.
[86] Bartlett, M. Stewart, & Sejnowski, T., 1997. Viewpoint invariant face recognition using independent Component analysis and attractor networks. In M. Mozer, M. Jordan, & T. Petsche, Eds.,Advances in Neural Information Processing Systems 9. Cambridge, MA: MIT Press: 817-823.
[87] Chen, W., Er, M. J., & Wu, S.. Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain. IEEE Transactions on Systems, Man and Cybernetics - Part B: Cybernetics, 36(2), 458-466,2006.
[88] Jacobs, D. W., Belhumeur, P. N., & Barsi, R.. Comparing images under variable illumination. In Proceedings, IEEE conference on computer vision and pattern recognition pp. 610-617. 1998.
[89] Nanni, L., & Maio, D. Weighted sub-Gabor for face recognition. Pattern Recognition Letters, 28(4), 487-492 2007.
[90] Liu, D. H., Shen, L. S., Lam, K. M., & Kong, X.. Illumination invariant face recognition. Pattern Recognition, 38, 1705-1716.2005.
[91] Wang, H., Li, S. Z., & Wang, Y. Face recognition under varying lighting conditions using self quotient image. In Proceedings of the IEEE international conference on automatic face and gesture recognition pp. 819-824,2004.
[92] Savvides, M., Kumar, B. V., & Khosla, P. K. Corefaces - robust shift invariant PCA based correlation filter for illumination tolerant face recognition. In Proceedings of the international conference on computer vision and pattern recognition Vol. 2, pp. 834-841,2004.
[93] Du, S., & Ward, R. Wavelet based illumination normalization for face recognition. In Proceedings of international conference on image processing Vol. 2, pp. 954-957,2005.
[94] Ojala, T., Pietikainen, M., & Maenpaa, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(7), 971-987. 2002.
[95] Zhang, W., Shan, S., Chen, X., & Gao, W. Are Gabor phases really useless for face recognition? In Proceedings of international conference on pattern recognition Vol. 4, pp. 606-609,2006.
[96] Tan, X., & Triggs, B. Enhanced local texture feature sets for face recognition under difficult lighting conditions. In Proceedings of the IEEE international workshop on analysis and modeling of faces and gestures pp. 168-182.2007.
[97] Georghiades, A., Kriegman, D., & Belhumeur, P. From few to many: Generative models for recognition under variable pose and illumination. IEEE Transactions Pattern Analysis and Machine Intelligence, 40, 643-660. 2001.
[98] T Vetter, T Poggio, Linear object classes and image synthesis from a single example image, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 19: 733-742, 1997.
[99] L. Wiskott, J. Fellous, N. Kruger, and C. V. malsburg. "Face recognition by elastic bunch graph matching". IEEE Transactions on Pattern Analysis and Machine Intelligence,19(7):775-779, 1997.
[100] Freeman, W.T.; Tenenbaum, J.B., Learning bilinear models for two-factor problems in vision, 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1997, pp. 554 - 560.
[101] T. Cootes, K. Walker, and C. Taylor. View-based active appearance models. In Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, pages 227-232, 2000.
[102] I. Matthews, R. Gross, and S. Baker, Appearance-based face recognition and light-fields, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 4, pp. 449-465,2004.
[103] M. Everingham and A. Zisserman, Identifying individuals in video by combinig 'generative' and discriminative head models, in Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV '05), vol. 2, pp. 11031110, Beijing,China, October 2005.
[104] O. Arandjelovic, G. Shakhnarovich, J. Fisher, R. Cipolla, and T. Darrell, Face recognition with image sets using manifold density divergence, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05), vol. 1, pp. 581-588, San Diego, Calif, USA.
[105] O. Arandjelovic and R. Cipolla, An illumination invariant face recognition system for access control using video, in Proceedings of the British Machine Vision Conference (BMVC '04), pp.537-546, Kingston, Canada, September 2004.
[106] Arandjelovic O, Cipolla R A pose-wise linear illumination manifold model for face recognition using video; Computer Vision And Image Understanding,Vol:113(1), pp: 113-125,2009.
[107] Abate AF, Nappi M, Riccio D, et al.; 2D and 3D face recognition: A survey; Pattern Recognition Letters, Vol:28(15),pp: 1885-1906,2007.
[108] V. Blanz, T. Vetter, A morphable model for the syntheses of 3D faces, in: Proceedings of International Conference on Computer Graphics, 1999, pp. 187-194.
[109] C. Zhang, F. Cohen, 3-D face structure extraction and recognition from images using 3-D morphing and distance mapping, IEEE Trans. Image Process. 11 (11) (2002).
[110] Dalong Jiang,Yuxiao Hu, Shuicheng Yan, Lei Zhang, Hongjiang Zhang, Wen Gao, Efficient 3D reconstruction for face recognition, Pattern Recognition 38 (2005), pp:787 - 798
[111] Sotiris Malassiotis, Michael G. Strintzis, Robust face recognition using 2D and 3D data: Pose and illumination compensation, Pattern Recognition 38 (2005) ,pp:2537 - 2548.
[112] Y.Cheng, K.Liu, J.Yang, Y.Zhuang and N.Gu, "Human Face Recogntion Method Based on the Statistical Model of Small Sample Size," in Intelligent Robots and Computer Vision X: Algorithms and Techniques, pp.85-95, 1991.
[113] Arandjelovic O., Cipolla R., Colour invariants for machine face recognition, 8th IEEE International Conference on Automatic Face & Gesture Recognition, 2008. FG '08. pp: 1-8.
[114] Jae Young Choi, Yong Man Ro, Konstantinos N. Plataniotis, Feature Subspace Determination in Video-based Mismatched Face Recognition, 8th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2008).
[115] Xiaoli Zhou; Bhanu B ,Human Recognition Based on Face Profiles in Video, IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, 2005. CVPR Workshops. Pp:15.
[116] Tse-Wei Chen, Shou-Chieh Hsu, Shao-Yi Chien, Feature-based Face Scoring in Surveillance Systems, Ninth IEEE International Symposium on Automatic Multimedia, 2007. pp: 139-146.

**Huafeng Wang** received an M.S. degree in Computer Science from Beihang University in 2004. He was born on May 17,1974, in ShanXi Province, China. Now he is working towards the Ph.D. degree at School of Computer Science and Engineering in Beihang University. His major fields are image processing and pattern recognition.

**Yunhong Wang** received B.S. degree from Northwestern Polytechnical University, Xian, China and received M.S. degree and Ph.D. degree from Nanjing University of Science and Technology. She was Born December 4, 1968.Her research interest include image processing, pattern recognition and computer vision.

**Yuan Cao** is a staff of Civil Aviation Medicine Center (CAMC),China. Her research interests include knowledge discovery in large databases, data mining, machine learning theory and applications and pattern recognition. She received her Masters' degree in 2009 from Beihang University,China, majoring in Software Engineering and Management.