ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

Volume 3, Special Issue 3, March 2014

2014 International Conference on Innovations in Engineering and Technology (ICIET'14)

On 21st & 22nd March Organized by

K.L.N. College of Engineering, Madurai, Tamil Nadu, India

High Performance View Based System for Face Recognition in Real Time

B.Vijayalakshmi [#1], V.Balambigai[#2],G.Gayathri[#3],M.Kokiladevi[#4]

Associate professor , M.E(Ph.D), Department Of ECE, K.L.N. College of Information Technology, India

Department Of ECE, K.L.N. College of Information Technology, India,

Department Of ECE, K.L.N. College of Information Technology, India,

Department Of ECE , K.L.N. College of Information Technology, India

ABSTRACT— Most of the existing face recognition systems are concerned with the espial and cognizance of single face with multiple views and it is protracted. This paper focuses on the detection and recognition of multiple faces with discretionary views. Real time samples of varying illumination, poses and complex background are taken as input and tested with multi-view face detection algorithm. Training data set is optimized by the introduction of face foreordination. Probabilistic classifier that evaluates whether a local feature cluster corresponds to a face is used to espy and to pinpoint multiple faces with discretionary views. The proposed multi-view face detection algorithm is based on Mamdani Fuzzy model to obtain a better performance compared to the avant-garde methodologies in terms of Mean Square Error, Face Registration Percentage and Area overhead.

INDEX TERMS: Image Registration, Face Foreordination, Espy, Mamdani Fuzzy model.

I.INTRODUCTION

Face detection is the most important step in many application that involves faces, including faces recognition. Face recognition is not concerned more with face recognition, but with face detection. The meticulous detection of human faces in discretionary scenes along with varying poses, illumination and complex background, are the most important process involved. The major breakthrough in the field of detection was Viola and Jones method [1] which used cascaded weak classifiers that works simple easy to calculate blubbery like features. Despite its very high speed , it was not capable of detecting multi-view faces, because running a separate detector for each class would take a lot of time.

An approach to simultaneously detect and localize multiple faces having arbitrary views and different scales was proposed by S.M.H. Anvar, W.Y. Yau and E.K. Teoh [2]. The main contribution of this paper is the introduction of a face foreordination, which enables multiview face detection and localization. To rectify the issues oriented with face alignment, patch-based algorithms have been proposed [3], [4]. However, a priori espial of faces are requisite in these methodologies. Furthermore, the contemporaneous algorithms do not percolate well with nonfrontal faces or in the presence of barricade. Furthermore, window- based face detection techniques, [5], [6] and probabilistic frameworks [7], [8] require a large amount of manual mediation to concoct a training set. Usually, these consist of thousands of manually mutated images that must be truncated and labeled. Viola and Jones [9] introduce AdaBoost with a cascade scheme and apply an integral image concept for face detection. They propose a two-class AdaBoost learning algorithm for training efficient classifiers and a cascade structure for rejecting none face images. At a later date, fine-tuning algorithms are required to crop and allineate the detected faces to convene the requirements of the face recognition module [10]. Most fine-tuning algorithms require locating the position of the eyes [11] or commensuration with template matching [12], [13], [14]. Thusly, the preparation training images is a very long-drawn-out and extortionate process.

Many algorithms have been proposed to deal with image-based recognition where both the training and test set consist of still face images. Recently, face recognition based on video has gained wide interest especially due to its role in surveillance systems. Video-based recognition has superior advantages over image- based recognition because a video contains image sequences as well as temporal information. However, surveillance videos are generally of low-resolution and contain faces mostly in non-frontal poses. In order to address these problems, we propose a multi-view, video-based face recognition algorithm using the fuzzy logic. The proposed system combines face detection, localization and feature selection tasks into a single framework The proposed method implements face recognition algorithm to handle multiple faces of arbitrary views to detect other object classes using fuzzy logic. A solution is provided to solve the problem of establishing correspondences when the intraclass variations are large.

The organization of this paper is as follows: Section 2 presents a literature survey of face detection techniques. Section 3 deals with different types of datasets used. Section 4 introduces the framework for constructing a multiview face constellation by finding distinctive correspondence points among the images and describes a method to register arbitrary views of faces from different persons and at different scales. Then a description of probabilistic classification algorithm that is used to locate the registered faces having multiple views and scales in highly cluttered background scenes. Section 5 provides the empirical study for each stage of the proposed method, and finally, Section 6 concludes the paper.

II.THE PROPOSED METHOD

We propose a system that combines face detection, localization and feature selection tasks into a single framework.The proposed method implements face recognition algorithm to handle multiple faces of arbitrary views to detect other object classes using fuzzy logic. We provide a solution to solve the problem of establishing correspondences when the intraclass variations are large.

Linguistic variable definition for face recognition. To find out the membership function for identify the error and changes in error(correspondence angle between the face). Rule development process(setting rules for selection). Process for face detection(FUZZY). Process for Controlling(DEFUZZY). Most commonly seen fuzzy methodology due to its simple structure of ‘mim-max’ operation. It is used for the image selection process. The selection of images is based on the triangle bell shape. The process starts with the evaluation of antecendents. It is followed by obtaining each rule’s conclusion. Then aggregate conclusions are obtained. After the aggregation process, there is fuzzy set of each o/p variable that needs defuzzification. It is intuitive, has wide spread acceptance and is well suited to human input.

III.DATASETS

For training purpose, generally used datasets are CMU's PIE dataset and the algorithm was tested on PIE dataset and , FDDB benchmark dataset and CMU profile face database. Both data set had rich collection pf faces in different poses.

3.1CMU's PIE dataset

PIE dataset is a huge face dataset maintained by CarnegieMellon University. This dataset has a rich collection of around 40,000 faces, categorized under different poses, expressions, illumination conditions etc. It has images of 68 different subjects, taken at 13 different poses,14 different illumination conditions, and 4 different expressions. They also provided the background images of the photos with faces, which is useful to segment faces from the images. Different poses of the faces are given below.

Fig 3.1 Different poses of faces

3.2 FDDB dataset

This is a more generic dataset which contains unconstrained images of faces taken from the “Faces in the Wild Dataset”. It contains about 5171 faces in a set of 2845 images. Some examples from the dataset is given below.

Fig 3.2 FDDB Datasets

3.3 CMU profile face dataset

This is an unconstrained face dataset, mainly featuring profile faces. It contains 209 images with more than 500 images.

Fig 3.3 CMU Datasets

3.4 Real time dataset

In real time dataset, there is no restriction for the number of training images, poses, expressions, illumination conditions etc. It constitutes of the dataset that has been taken in the real time. This suits better for all face recognition systems, as most systems are in public places such as banks, organisations etc. We used our real time dataset, recorded by a cctv camera in our institution.

Fig 3.4 Realtime Datasets

IV.DETECTING MULTIPLE FACES OF ARBITRARY VIEWS

In this section, we generalize the proposed face detection algorithm to handle multiple faces of arbitrary views. In our proposed approach, it is possible to approximately locate the features belonging to specific facial traits once a given image is registered. Therefore, to detect multiple faces, we need to cluster all the features that approximately form a face region. Probabilistic classifier is used for registering and to detect multiple faces. Probability based approaches are generally divided into two groups: Maximum Likelihood and Maximum a-Posterior. These approaches have the power of complete consideration of the a-prior knowledge and the details of the imaging model.

Real time input is taken in the video format. The video is first converted into frames and then further processed. Each cropped facial image was downsampled to 20×20 pixels and the pixels in each image were normalized to have zero mean and unit variance. The confidence threshold $\tau = 0.999$ was used. The recognition rate was about 92%. We also applied a frame-by-frame strategy using the same appearance model and likelihood measurement, and performed recognition in a video by temporal majority voting. The recognition rate was 83% in this case. When 40 dimensional features were used, our proposed algorithm achieved the 97.3% recognition rate by processing 37.95 frames on average. The following is the example for converting video into frames.

The proposed method implements face recognition algorithm to handle multiple faces of arbitrary views to detect other object classes using fuzzy logic. We provide a solution to solve the problem of establishing correspondences when the intraclass variations are large. The transformation of system

Fig 4.1 Frame conversion

inputs which are crisp numbers into fuzzy sets is done by fuzzification function. The transformation the fuzzy set obtained by the interference engine into a crisp value. Mamdani fuzzy logic is preferred for this purpose. It is most commonly seen fuzzy methodology due to its simple structure of ‘min-max’ operation. It is used for the image selection process. The selection of images is based on the triangle bell shape. The process starts with the evaluation of antecendents. It is followed by obtaining each rule’s conclusion. Then aggregate conclusions are obtained. After the aggregation process, there is fuzzy set of each output variable that needs defuzzification. It is intuitive, has wide spread acceptance and is well suited to human input.

V.EMPIRICAL STUDY

We evaluated our proposed algorithm using the real time color dataset. The dataset contained images of 6 people of various age, ethnicity and complexion. Each person had an average of 10 images, with different illumination conditions and pose variations. The number of images in each viewpoint was not uniform. For instance, a person in the dataset may have more than three frontal faces but only one full profile face. We resized all the images to 218 X 218 pixels. The image sizes and resolutions were not uniform. Some face images were very small. Faces in this dataset had high background clutter, and each image contained more than one face from different viewpoints. Faces in this dataset were very challenging, with cases of occlusion, varying poses, low resolution, and out of focus. For all the above datasets, we converted the color images into grayscale.

5.1 Experimental setup

We used SIFT features and applied the proposed method outlined in Section 4 on all the image pairs. The SIFT parameters such as the number of octaves of the Gaussian scale space and the number of scale levels within each octave were determined such that at least 1,000 features were extracted from each image. Potential correspondence points between the image pairs were found using Lowe's method [23]. Based on Matlab implementation and using a 2.66 GHz C2Q PC with 3 GB RAM running. Windows Vista 32-bit OS, the entire process to locate the correspondence points took less than a second for an image pair of size 512 X 512 pixels.

5.2 Finding correspondence points

Conversion of video into frames. Extraction of SIFT features from two images. Removal of common features from them using appearance descriptor of SIFT. Consideration of the potential correspondence point between two images obtained from the Lowe's matching algorithm. In this algorithm SIFT keypoint of objects are first extracted from a set of reference images and stored in a database. An object is recognized in a new image by individually comparing each feature from the new image by individually comparing each feature from the new image to this database and finding candidate matching feature. Normalization of all the geometric descriptors of potential points in the first image to the geometric descriptors of the reference point. Obtaining correspondence poits is shown below.

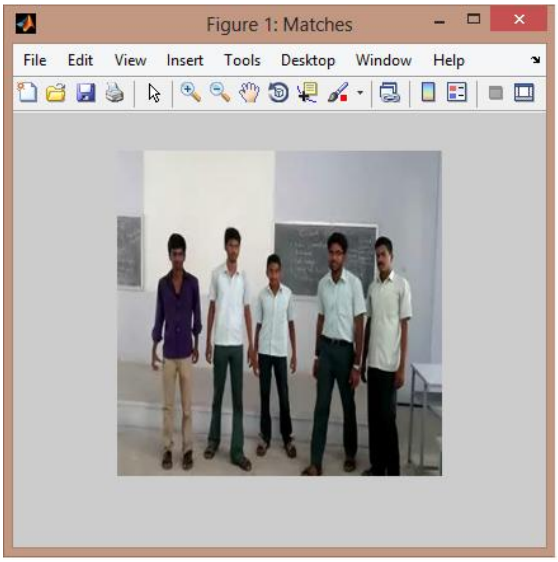

Fig 5.2.1 Input Frame

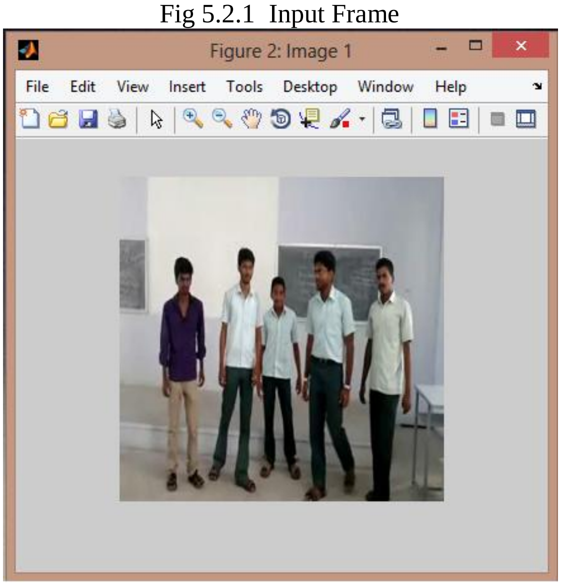

Fig 5.2.2 Reference Frame

Finding the cluster of consensus point with respect to the reference point, which are the point that their normalized geometric descriptors . Establishing the face constellation based on reference map indicating registration of images in training set with respect to reference image. Registering the faces based on the reference points for multiple face detection of arbitrary views using adaboost(existing)/fuzzy(proposed) logics. Applying this system for the real time video input. Comparision of the results of the proposed and the state of art systems. Face with cluster points are shown in the following figure.

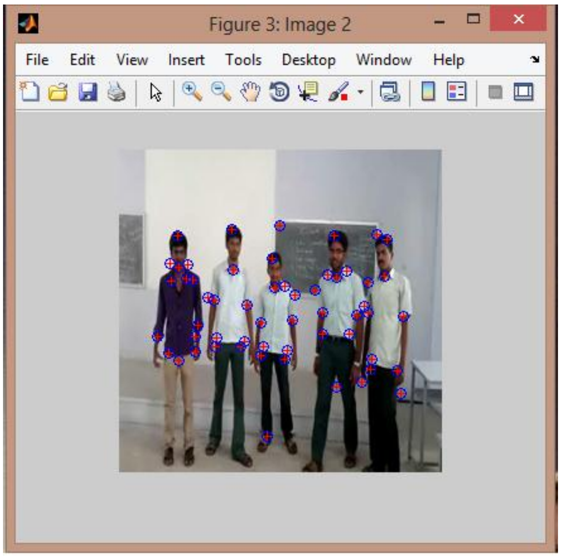

Fig 5.2.3 Corresspondance point for input frame

5.3 Face registration process

5.4

This experiment was carried out on the real time dataset. We randomly chose images of different people. We then applied our proposed method to register these images. We took a frontal view image as the reference image and manually selected the nose base and upper part of nose in the middle of two eyes as the reference points for registration. Marking these two points is the only manual intervention needed. Starting with 50 training images to form the initial constellation of links, we evaluated the percentage of images that were successfully registered, wrongly registered, and could not be registered over the total number of images available in the training dataset. An image was considered as correctly registered when the estimated reference vector falls within a circle centered at the base of the nose, with a radius of half of the nose length (distance between the nose base and upper part of the nose in the middle of the two eyes). We then repeated the procedure by increasing the number of training images until all the available images were used. To crop facial image patches in a image sequence manually fast, we have created a software named imageclipper and the software is available online. This software is useful not only for facial images but also for any kinds of images and works under multi-platform (Windows and Linux). Using this software, we can

(1) open images in a directory sequentially, (2) open a video, frame by frame, (3) crop (save) an image patch and go to the next image by pressing one button, (4) move and resize the rectangle region to crop by hotkeys or right mouse button, and (5) rotate and shear deform, i.e., affine transform, the rectangle region. A snapshot of the software is shown below.

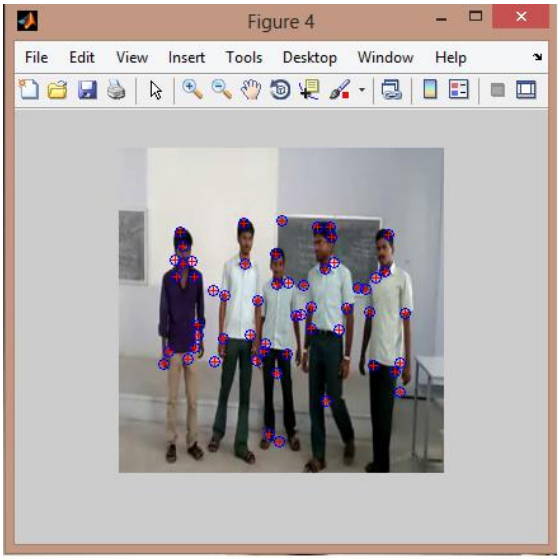

Fig 5.3 Correspondence points for reference frame

5.5 Face constellation:

Facial features are grouped into face-like constellations using more robust face models based on statistical analyses. Burl et al.[15] make use of statistical shape theory on the features detected from a multi-scale Gaussian derivative filter. A probabilistic model for the spatial arrangement of facial features enables higher detection flexibility. The algorithm is able to handle missing features and problems due to translation, rotation, and scale to a certain extent. Most detection failures are caused by significant rotation of the subject's head. Huang et al.[16] also apply a Gaussian filter for pre-processing in a framework based on image feature analysis. Pre-processed images are searched with a structure model, a texture model, and a feature model for face-like patterns, thus face constellations are formed.

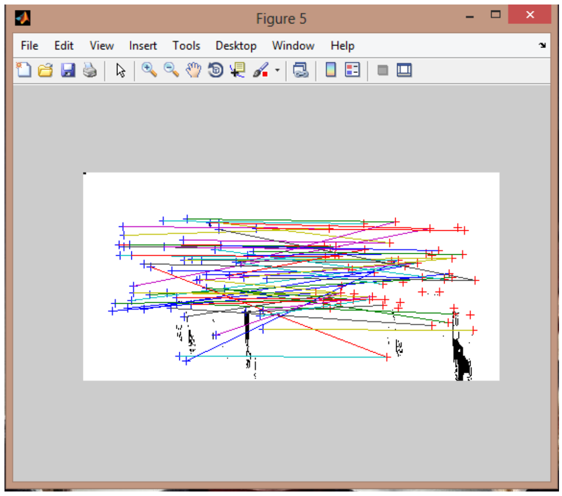

Fig 5.4 Obtaining face constellation

5.6 Detecting multiple faces of arbitrary views.

The main objective of the simultaneous framework is face recognition. The simultaneous framework does not aim to estimate a particular tracking state in a video sequence because face recognition, which is the main objective, is performed simultaneously unlike the tracking-then-recognition framework. Yet, a particular tracking state can also be estimated. Having multiple low-resolution input images of the same face, we have used the proposed registration algorithm, based on the probabilistic classifier and mamdani fuzzy algorithm. After randomly selected images of left profile and left oblique faces from the real time dataset to the training stage, such images can be successfully detected. Adding this amount of training images also increased the number of correctly registered images, while decreasing the number of nonregistered faces. We compared the faces detected with the ground-truth data, and the degree of match is obtained by calculating the ratio of an intersecting area using

$$S(D,G) = \frac{Area(D) \cap Area(G)}{Area(D) \cup Area(G)}$$

where *D* is the detected face and *G* is the ground-truth face. *Area()* indicates the area of face that typically is defined by an ellipse or a square window. Those detected faces with match values, *S(D,G)* greater than 0.5 are considered as a correctly detected face in the discrete score. In the continuous score, the direct value of *S(D,G)* is used.

This is complete as we perform post-processing to validate the face region, and drew an ellipse with parameters similar to the reference image and centered it at the midpoint of the two reference points to meet the requirement of the evaluation. Thus, for nonfrontal faces, the estimated ellipse will cover some background regions. In the extreme case of a profile face, the ellipse drawn
will only cover about half the face.

Fig 5.5.1 Ada Boost result

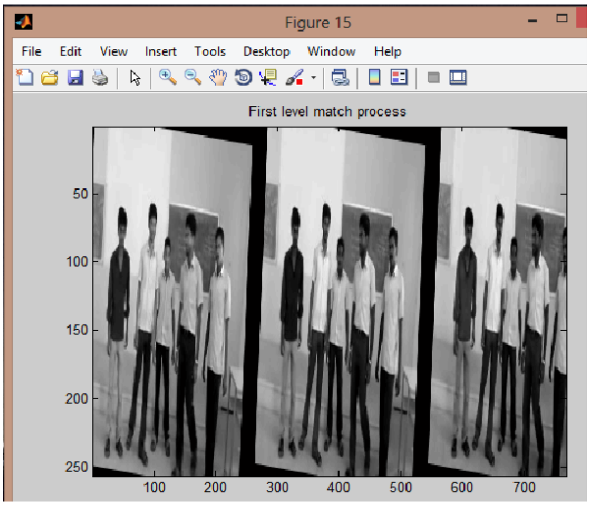

Fig 5.5.2 Fuzzy logic result

5.6 Parameters Estimated

We implemented the face detection and localization portion using MATLAB. The average time to detect and localize a face in an image of 512 X 512 pixels was about 540 ms using an ordinary PC (2.8 GHz C2D PC with 3 GB RAM and Linux OS using a single core). As a comparison, the elapse time for the window-based method of is 10.2743 seconds (using a Pentium 4 3.0 GHz) and 10.1326 seconds (using a Pentium 3 700 MHz), respectively. However, these methods yield only coarse detection requiring face alignment to localize the face. On the contrary, our approach performs both detection and localization, providing the position, scale, and amount of in-plane rotation of a detected face simultaneously. Thus, the elapse time of our method is not significantly slower. By contrast, our method is up to seven times faster than that of existing methodologies on the same machine. This is because our method is able to select more informative and distinctive features, resulting in a more compact and efficient model construction. The majority of the required time is spent on SIFT feature extraction (4.80 seconds). The time needed for comparing the model and the extracted features is not significant. Our proposed method yields a faster elapse of 7.19 seconds with a face registration percentage of 97%. Thus our method provides better face recognition.

Parameters	**Face registration %**	**Error**	**Elapse Time**
Ada Boost method	97.13	2.8681	9.275471

Fuzzy logic (Mamdani)	97.37	2.6261	7.192117

Table 5.6

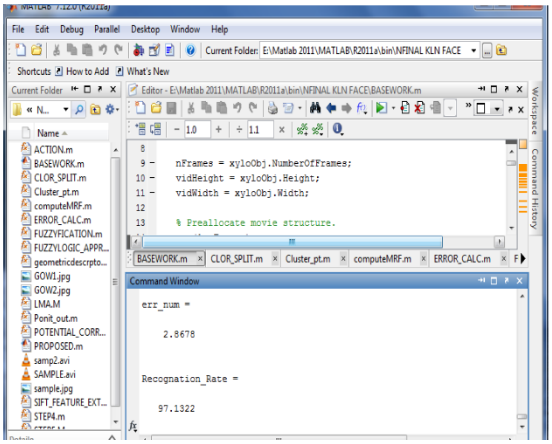

Fig 5.6.1

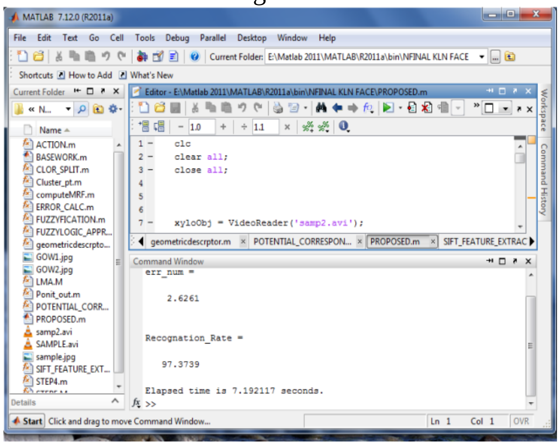

Fig 5.6.2

VI.CONCLUSION

In this paper, we have presented an approach to simultaneously detect and localize multiple faces from arbitrary views and different scale images. We have shown that the proposed method is able to handle arbitrary views, varying illumination conditions, and complex backgrounds. Advantageously, the probabilistic framework and fuzzy algorithm are able to detect multiple and occluded faces. To prevent any error propagation due to registering wrongly matched features, weak links are detected and removed. Unlike other approaches, the proposed approach does not require manually labeled faces apart from two points in the reference face image for training. Despite its simplicity, experimental results show that the performance is better than the other state-of-the-art approaches for multi view face detection. Note that the proposed method does not have any prior assumptions about the training images used. More diverse training images enrich the face constellation, allowing it to detect more complex variations of the face. Thus, the capability of the proposed approach will only grow with more diverse images so long as the intraclass variation is not so large that sufficient correspondence points cannot be found.

Our future work is to extend the proposed approach to reduce the memory allotted for training data set and to detect other object classes and to solve the problem of establishing correspondences when the intraclass variations are large.

REFERENCES

[1] P. Viola and M.J. Jones, "Robust Real-Time Face Detection," Int'l J. Computer Vision, vol. 57, no. 2, pp. 137-154, 2004, doi: 10.1023/B:VISI.0000013087.49260.fb.

[2] .M.H. Anvar, W.Y. Yau and E.K. Teoh , "Multiview Face Detection and Registration Requiring Minimal Manual Intervention" IEEE transactions on pattern analysis and machine intelligence, vol. 35, no. 10, october 2013

[3] A.B. Ashraf, S. Lucey, and C. Tsuhan, "Learning Patch Correspondences for Improved Viewpoint Invariant Face Recognition,Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 1-8,2008.

[4] L. Liang, R. Xiao, F. Wen, and E. Sun, "Face Alignment via Component-Based Discriminative Search," Proc. European Conf.Computer Vision, pp. 72-85, 2008

[5] H. Schneiderman and T. Kanade, "Object Detection Using the Statistics of Parts," Int'l J. Computer Vision, vol. 56, no. 3, pp. 151-177, 2004,doi:10.1023/B:VISI.0000011202.85607.00.

[6] C. Huang, H.Z. Ai, Y. Li, and S.H. Lao, "High-PerformanceRotation Invariant Multiview Face Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 29, no. 4, pp. 671-686, Apr.2007, doi: 10.1109/TPAMI.2007.1011.

[7] S.Z. Li and Z. Zhenqiu, "FloatBoost Learning and Statistical Face Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 26, no. 9, pp. 1112-1123, Sept. 2004, doi: 10.1109/TPAMI.2004.68.

[8] M. Toews and T. Arbel, "Detection, Localization, and Sex Classification of Faces from Arbitrary Viewpoints and under Occlusion," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 31, no. 9, pp. 1567-1581, Sept. 2009, doi: 10.1109/TPAMI.2008.233.

[9] P.Viola, M. Jones, "*Rapid Object Detection Using a Boosted Cascade of SimpleFeatures*," IEEE Computer Vision and Pattern Recognition, pp. 511-518, 2001

[10] G. Dedeoglu, T. Kanade, and S. Baker, "The Asymmetry of Image Registration and Its Application to Face Tracking," IEEE Trans.Pattern Analysis and Machine Intelligence, vol. 29, no. 5, pp.807823,May2007,doi:10.1109/TPAMI.2007.1054.

[11] A. Gee and R. Cipolla, “Determining the Gaze of Faces in Images,”Image Vision Computing, vol. 12, no. 10, pp. 639-647, 1994, doi:10.1016/0262-8856(94)90039-6.

[12] L. Xiaoming, “Generic Face Alignment Using Boosted Appearance Model,” Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 1-8, 2007.

[13] L. Gu and T. Kanade, “A Generative Shape Regularization Model for Robust Face Alignment,” Proc. European Conf. Computer Vision, pp. 413-426, 2008.

[14] Z. Jianke, G. Luc Van, and S.C.H. Hoi, “Unsupervised Face Alignment by Robust Nonrigid Mapping,” Proc. IEEE Int’l Conf.Computer Vision, pp. 1265-1272, 2009.

[15] M. C. Burl, T. K. Leung, and P. Perona”, Face localization via shape statistics”, in Int. Workshop on Automatic Face and Gesture Recognition, Zurich, Switzerland, June 1995.

[16] W. Huang, Q. Sun, C. P. Lam, and J. K. Wu,” A robust approach to face and eyes detection from images with cluttered background”, in Proc. of International Conference on Pattern Recognition, 1998