Improving In Video Face Recognition under Occlusion Using HaarCascading Classifier

Swati P. Kamble, Prof. R. K. Krishna

Abstract— **Identifying faces in images iseasier but face identification in videos is more difficult than that in images because of low resolution, occlusion, non-rigid deformations, large motion, complex background and other uncontrolled conditions make the results of face detection and recognition unreliable. It is a challenging problem due to the huge variation in the appearance of faces in video to achieve accuracy.The main objective of proposed system is to efficiently identify faces even in case of occlusion like glasses, etc. which results in accuracy of system.Facial occlusions, due for example to sunglasses, hats, scarf, beards etc., can significantly affect the performance of any face recognition system. Unfortunately, the presence of facial occlusions is quite common in real-world applications especially when the individuals are not cooperative with the system such as in video surveillance scenarios. While there has been an enormous amount of research on face recognition under pose/illumination changes and image degradations, problems caused by occlusions are mostly overlooked. The focus of this paper is thus on facial occlusions, and particularly on how to improve the recognition of faces occluded by sunglasses and scarf. We propose an efficient approach which demonstrates state-of-the-art performance on streaming video face recognizing in various genres of videos and label them with the corresponding relevant names.**

Index Terms— **Face Detection, Face Recognition, Facial Occlusion sunglass, HAAR Cascading Method, Streaming Video**

I. INTRODUCTION

Due to the enormous growth in movies, video application, a huge amount of data is being generated every day. Automatic face identification of faces in videos has drawn significant research interests and led to many securitiesbased applications. A key issue in face identification in videos is more difficult than that in images because of low resolution, occlusion, non-rigid deformations, large motion, complex background and other uncontrolled conditions make the results of face detection and tracking unreliable. In presence of occlusion on faces in videos, the situation is even worse. It is a challenging problem due to the huge variation in the appearance of each character in streaming videos. The proposed schemes demonstrate state-of-the-art performance on streaming video face recognizing in various genres of videos and label them with the corresponding relevant names.

Face is a complex multidimensional structure .The face is our primary and first focus of attention in social life playing an important role in identity of individual. Faces could be applied to a wide variety of practical applications including criminal identification, security systems, identity verification, video surveillance, etc.

Face recognition is an integral part of biometrics. In biometrics basic traits of human is matched to the existing data and depending on result of matching identification of a human being is traced. There are various Issues in face recognition, Pose variation, Occlusion, Facial expression, Imaging condition, Low resolution, Complex background, uncontrolled condition ,etc hence It will be unreliable to identify the n faces this degrades the system accuracy. So this study is motivated by these challenges.

II. RELATED WORK

A. Advances in Face Detection and Recognition Technologies.

This paper describes advances in the authors' face detection and recognition technologies. For face detection, a hierarchical scheme for combined face and eye detection has been developed based on the Generalized Learning Vector Quantization method to achieve precise face alignment. For face recognition, the perturbation space method has been improved to reduce the adverse effects of illumination changes as well as pose changes by using a standard face model.

B. Face Recognition under Occlusions and Variant Expressions with Partial Similarity.

The goal of this paper is to deal with one class of facerecognition problem where some of facial appearances in a given face image are badly deformed by such variations as large expression changes or partial occlusions (or disguise) due to sunglasses, scarves, mustaches and so on. Such variations in facial appearance are commonly encountered in uncontrolled situations and may cause big trouble to the face-recognitionbased security system but are less studied in literatures [5]. Notice that in this paper, we don't intend to deal with other commonly encountered variations in uncontrolled conditions like lighting changes and ageing effect, which are of interest but usually change people's facial appearance in a moreholistically way. By contrast, the facial appearance changes caused by variant expressions and partial occlusions are mostlylocal in nature, i.e., only parts of facial appearance change largely while others are less affected. The challenge lies in that such local deformations or occlusions in facial appearance can be anywhere and in any size or shape in a give face image and we don't have any prior knowledge about it.

Manuscript received July 20, 2015
Swati P. Kamble, Department of Computer Science and Engineering. Rajiv Gandhi College of Engg, Research & Technology, Chandrapur, India
Prof. R. K. Krishna, Department of Computer Science and Engineering. Rajiv Gandhi College of Engg, Research & Technology, Chandrapur, India

C. *Robust Face-Name Graph Matching for Movie Character Identification*

Automatic face identification of characters in movies has drawn significant research interests and led to many interesting applications. It is a challenging problem due to the huge variation in the appearance of each character. Although existing methods demonstrate promising results in clean environment, the performances are limited in complex movie scenes due to the noises generated during the face tracking and face clustering process. In this paper we present two schemes of global face name matching based framework for robust characteridentification.In this paper focus is on annotating charactersin the movie and TVs, which is called movie characteridentification [22]. The objective is to identify the faces of thecharacters in the video and label them with the correspondingnames in the cast. The textual cues, like cast lists, scripts,subtitles and closed captions are usually exploited. Fig.1 showsan example in our experiments. In a movie, characters are thefocus center of interests for the audience. Their occurrencesprovide lots of clues about the movie structure and content.Automatic character identification is essential for semantic movie index and retrieval [23], [24], scene segmentation [25], summarization [26] and other applications [27].

D. Glasses Removal from Facial Image Using Recursive Error Compensation

In this paper [1], we propose a new method of removing glasses from a human frontal facial image. We first detect the regions occluded by the glasses and generate a natural looking facial image without glasses by recursive error compensation using PCA reconstruction. The resulting image has no trace of the glasses frame or of the reflection and shade caused by the glasses. The experimental results show that the proposed method provides an effective solution to the problem of glasses occlusion and we believe that this method can also be used to enhance the performance of face recognition systems.

E. Face Recognition under Occlusions and Variant Expressions with Partial Similarity

Recognition in uncontrolled situations is one of the most important bottlenecks for practical face recognition systems. In particular, few researchers have addressed the challenge to recognize non-cooperative or even uncooperative subjects who try to cheat the recognition system by deliberately changing their facial appearance through such tricks as variant expressions or disguise (e.g. by partial occlusions). This paper [2] addresses these problems within the framework of similarity matching. A novel perception inspired non-metric partial similarity measure is introduced, which is potentially useful in deal with the concerned problems because it can help capturing the prominent partial similarities that are dominant in human perception. Two methods, based on the general *golden section* rule and the *maximum margin* criterion, respectively, are proposed to automatically set the similarity threshold. The effectiveness of the proposed method in handling large expressions, partial occlusions and other distortions is demonstrated on several well-known face databases.

III. 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, neural network-based method, SVM-based method, HMM method, BOOST method and color-based face detection, 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. Thishelps to stabilize detection and to make it less sensitive to thresholds compared to the other two detection categories.

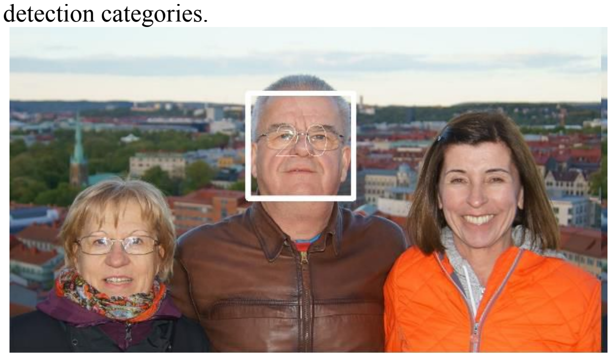

Fig no. 1 Face Detection shows in white rectangular box

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 have shown that dynamic information is verycrucial in the human face recognition process. Third, more effective representations, such as a 3D face model or super-resolution images, 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. Though the advantages are obvious, there also exits some disadvantages. For example, poor videoquality, 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 use temporal voting to improve identification rates. There are also several algorithms which extract 2D or 3D face structure from the video. Other than simple voting approaches, Li et al. 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 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. 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 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. 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, 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. 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 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 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 took the video frames for each person separately to compute many individual eigen spaces, considering the angle between input and reference subspaces formed by the principal components of the image sequences as the measure of similarity. In 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 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 the proposed scheme achieved better performance to learn a sparse representation from video clips for online face recognition in an unconstrained environment.In 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 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. used HMM and ARMA models for direct video level matching. In it showed that the problem of visual constraints could be solved by HMM-based recognition framework

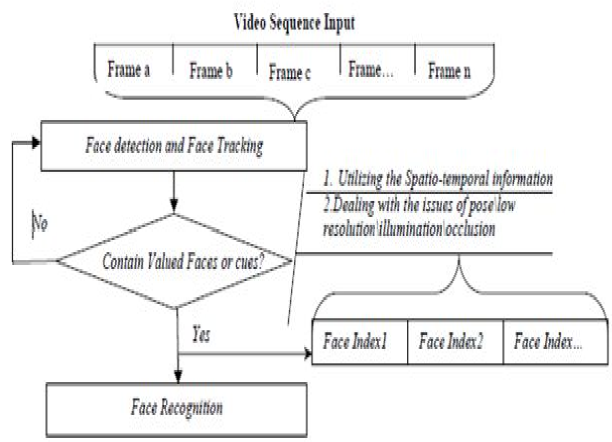

Fig no. 2 Detection and Recognition process

C. Advanced Topics

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

1) Illumination
2) Pose issues
3) Low Resolution

V. OCCLUSION

Facial occlusions may occur for several intentional or undeliberate reasons (See Fig. 2). For example, football hooligans and ATM criminals tend to wear scarf and/or sunglasses to prevent their faces from being recognized. Some other people do wear veils for religious convictions or cultural habits. Other sources of occlusions include medical masks, hats, beards, moustaches, hairs covering the face, make up, etc. Undoubtedly, occlusions can Significantly affect the performance of even most sophisticated face recognition systems, if occlusion analysis is not specifically taken into account. The focus of this paper is on how to improve face recognition performance under occlusions, particularly caused by sunglasses and scarf.

Fig no. 3 Examples of occluded face images from different sources

VI. PRAPOSED WORK

The proposed system efficiently identifies faces even in case of occlusion like glasses, etc.which results in accuracy of system. Flow of propose work will be as follows:

Phase 1: Streaming video is nothing but the array of interrelated frames It will be difficult or non-feasible to process the video in real time hence in this phase system will process the frames and send it for further processing.

Phase 2: Once the system get the individual processing frame system will find out the face area and store it in File.

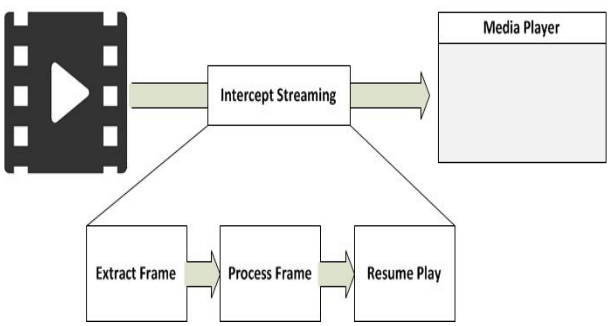

Fig no. 4 Streaming Process

Phase 3: Further each frame is analyzed from where the faces are extracted names are assigned to it. Meanwhile each processing of frame faces extraction and name association follows.

Phase 4: Streaming video is still running multiple faces occurs so the system will recognize the faces display its identification. if face are not identified then it is added to the dataset for future identification .The process carries until the media file stops playing meanwhile if occluded face occurs so the face still recognize by reconstructing the original face.

VII. IMPLEMENTED WORK

a) Face Detection Process

HAAR CASCADING METHOD

The proposed approach, at first, they detect the face using Viola and Jones' Boosting algorithm [10] and a set of Haarlikecascade features. Using Haar-like features method system is able to detect the faces from running video. Haar feature which are used to detect the presence of that feature in the present frame.Each feature result in a single value, which is calculates by subtracting the sum of pixels under white rectangle from sum of pixels under black rectangle. There are different typesof HAAR CASCADING METHOD, method for eyes detection, nose detection, and face detection. In our project we are using face method for detecting face from the video.Voila and Jones algorithm is used for face detection. Where it is uses in both creating database and face recognition process. Where in case of creating database it takes input image through a web camera continuously. Captured image undergoes face detection. Detected face will be cropped and stored in database. Where in case of face recognition if there is any movement in video surveillance will be uses to detect the moving object. The captured image undergoes face detection and further processed later by face recognition. We can see the cropped images in Fig No. 6

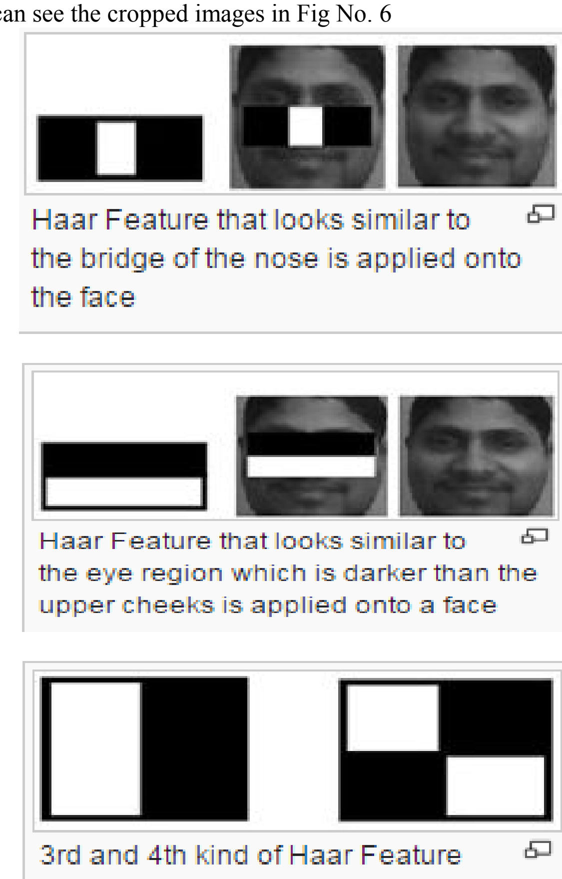

Fig no. 5Haar Feature Pixels

b) Face Recognition Process Under Occlusion

Through the same HAARCASCADDING CLASIFIER method we are able to assign a name to detected face. Haarcascading method trained our system for detecting face and also provides a method for recognising process it means; it displays a name of detected faces. When a video is running and detecting face it generate rectangular box around the face, then it capture the rectangular face area at the same time we are assigning a name to the detected face.

The captured face images are stored in file and assigned names are also stored in text file in the system or in database.The system will recognised or displaying name of the detected face if the person face is already stored and name is already assigned. In this way we trained our system for detection and for recognition which improves our system accuracy.

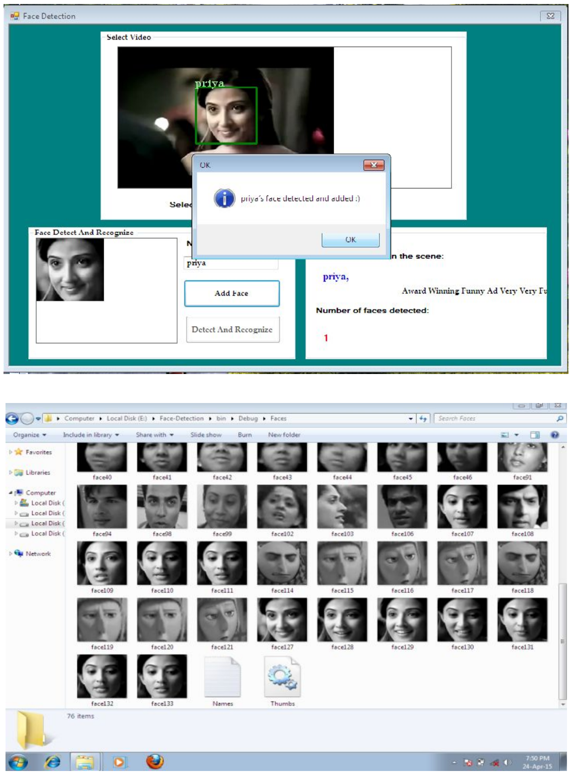

Fig no. 6Project Screen Short

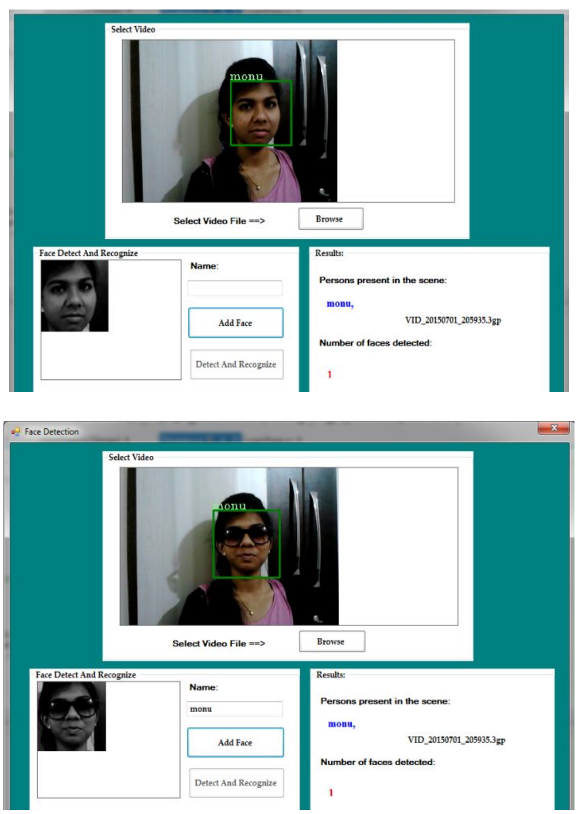

Fig no. 7 Face with occluded and Non-occluded

VIII. RESULT ANALYSIS AND DISCUSSION

We performed numerous experiments and comparedthe proposed algorithm with other methods in the context of video-based recognition. Since there is no standard database that contains cropped images from video frames for video-based face recognition, we collected a set of videos of different people for experiments. A variant of the eigen-subspace tracker [2] was used to locate the face, and the results were inspected by humans. Each image was then downsampled to 30 _ 30 pixels for computational efficiency.. Some cropped images from the videos are shown in Figure 6. Which makes system trained for face detection and recognition even faces are under occlusion.

The system is efficiently detect the faces from video and recognized if the faces are already trained. Voila and Jones algorithm is used for face detection. Where it is uses for bothcreating database and face recognition process. The proposed approach, at first, they detect the face using Viola and Jones' Boosting algorithm [10] and a set of Haarlike cascade features. Using Haar-like features method system is able to detect and recognized the faces from running video, which results in accuracy of system. Table 1 shows Theoretical comparison of several existing methods in terms of key parameters with Feature base face detection. Table 2 shows pros & cons of several existing methods.

Approach /Parameter	Haar like Feature base Face Detection	Geometric Base Face Detection
Precision	High	Low
Execution Time	Low	High
Learning Time	High	High
Ratio between detection rate & false alarm	High	Low

Table 1
Theoretical Comparison of Several Existing Methods In Terms of Key Parameters with Feature Base Face Detection

Technique	Merits	Demerits
Feature base Face Detection	• More accurate • Low execution time	• High learning time
Geometric Base Face Detection	• Effective approach • Easy to implementation	• Low accuracy • More false alarm
Haar Like Feature Base Face Detection	• Improved feature extraction part • Less false alarm	• High execution time • Complex to implement

Table 2
Pros & Cons of Several Existing Methods

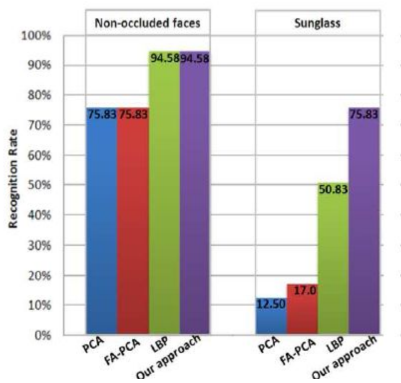

Fig no. 7Recognition Performance Different Methods On Two Test Sets: Non-occluded Faces, Face Occluded with Sunglasses

The results in Fig. 7 clearly show that our proposed Approach significantly outperform all other methods.On the non-occluded faces, our approach and LBP yielded equal performance (94.83%) while the Eigenface method (with and without occlusion detection) yielded much lower performance (75.83 %). On the test set of faces with scarves, our proposed approach gave best results (92.08%), followed by LBP (60.83%), and then by PCA based methods (34.17% and 5.42%). Note that LBP performed quiteeven under occlusion, thus confirming the earlier findings stating that local feature-based methods are more robust against occlusions than holistic methods. Comparing the results on the test sets of faces with sunglasses and scarves, we notice that most methods are more sensitive to sunglasses than to scarf. This is an interesting conclusion which is in agreement with the psychophysical findings indicating that the eye regions play the most important role in face recognition.

CONCLUSION

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 trackingtask. 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.

REFRENCES

[1] Jeong-Seon Park, You Hwa Oh, Sang ChulAhn , and Seong-Whan Lee, Sr. Memb “Glasses Removal from Facial Image Using Recursive Error Compensation” .IEEE Transaction on Pattern Analysis And Machine Intelligence.Vol. No.27 2012

[2] Maria De Marsico, Member, IEEE, Michele Nappi, and Daniel Riccio “Face Recognition Against Occlusions and Expression Variations . IEEE Transaction on System , Man, And Cybernetics And Humans, VOL. 40, NO. 1, January 2010

[3] Zafar G. Sheikh, V. M. Thakare, S. S. Sherekar “Advances in Face Detection Techniques in Video”2012 2nd IEEE International Conference on Parallel, Distributed and Grid Computing.

[4] Minglan Sheng, ZhangliLan “Face Detection in Video Sequence with Complex Background” Proceedings of 2008 IEEE International Conference on electronics and Automation.

[5] R. Gross, J. Cohn, and J. Shi, “Quo vadis face recognition,” in Working Notes of the 3rd Workshop on Empirical Evaluation Methods in Computer Vision, Kauai, HI, 2001.

[6] Face Detection System on Ada boost Algorithm Using Haar Classifiers M. GopiKrishna,A. Srinivasulu, Prof (Dr.) T.K.Basak International Journal of Modern Engineering Research (IJMER) Nov-Dec. 2012

[7] “Improving the Recognition of Faces Occluded by Facial Accessories “ Rui Min Multimedia Communications Dept. AbdenourHadid Machine Vision Group from University of Oulu, Finland.

[8] B.G. Park, K.M. Lee, and S.U. Lee, “Face recognition using face-ARG matching,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 12, pp. 2005**.**

[9] Ali Javed ,Face Recognition Based on Principal Component Analysis J. Image, Graphics and Signal Processing, Feb 2013 .

[10] Ole Helvig Jensen KongensLyngby “Implementing the Viola-Jones Face Detection Algorithm”. KongensLyngby 2008 in Proc. ICME, 2011, pp. 1–6.

[11] Zafar G. Sheikh, V. M. Thakare, S. S. Sherekar 2nd IEEE International Conference on Parallel, Distributed and Grid Computing “Advances in Face Detection Techniques inVideo”.

[12] W. Zhang, S. Shan, X. Chen, and W Gao, "Local Gabor Binary Patterns Based on Kullback–Leibler Divergence for Partially Occluded Face Recognition," Signal Processing Letters, IEEE , vol.14, no.11, pp.875- 878, Nov. 2007.

[13] 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.

[14] J. Yang and A.Waibel. A real-time face tracker. In Proceedings of theThird IEEE Workshop on Applications of Computer Vision, pages 142-147, Sarasota, FL, 1996.

[15] 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.

[16] 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.

[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 objectdetection 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 posedistribution". In Proc. IEEE International Conference on Face and Gesture Recognition, Vermont,1996.
[21] ZhenQiu Zhang; Long Zhu; Li, S.Z.; HongJiangZhang,Real-timemulti-view face detection, Fifth IEEE International Conference onAutomatic Face and Gesture Recognition, pp:142-147,2002.
[22] J. Sang, C. Liang, C. Xu, and J. Cheng, "Robust movie characteridentification and the sensitivity analysis," in ICME, 2011, pp. 1–6.
[23] Y. Zhang, C. Xu, H. Lu, and Y. Huang, "Character identification in feature-length films using global face-name matching," IEEE Trans.Multimedia, vol. 11, no. 7, pp. 1276–1288, November 2009.
[24] M. Everingham, J. Sivic, and A. Zissserman, "Taking the bite out of automated naming of characters in tvvideo," in Jounal of Image andVision Computing, 2009, pp. 545–559.
[25] J. Sang and C. Xu, "Character-based movie summarization," in ACM MM, 2010.
[26] R. Hong, M. Wang, M. Xu, S. Yan, and T.-S. Chua, "Dynamiccaptioning: video accessibility enhancement for hearing impairment,"in ACM Multimedia, 2010, pp. 421–430.

AUTHORS PROFILE

Swati Kambleis a student of Rajiv Gandhi College of Engineering, Research and Technology, Chandrapur. She is currently working on her thesis as a student of MTech (CS) at RCERT, Chandrapur. Her research interests include Digital Image Processing and Software Engineering. She completed her BE (IT) degree from GHRCE, Nagpur in 20013.

Prof. R. K. Krishna has been theAssistant Professor at Department of Electronics Rajiv Gandhi College of Engineering, Research and Technology, Chandrapur. Prof. R. K. Krishna has more than 7 years of experience including teaching graduate and undergraduate classes.