A Survey On Automated Systems for Face Recognition in Videos

Apoorva Biswal[1],Poonam Sonawale[2], Meghali Veer[3], Yash Sahay[4]

[1234]BE IT, Department of Information Technology, RMD Sinhgad School of Engineering, Pune, Maharashtra, India.

Mr. SanjeevkumarAngadi[5]

[5]Assistant Professor, Department of Information Technology, RMD Sinhgad School of Engineering, Pune, Maharashtra, India.

Abstract –*Videos have enabled to take face recognition to the next step, providing various audio, visual cues and 3D models. Images captured in each frame provide plentiful data consisting of numerous features, using which variations in pose, illumination, and occlusion can be addressed, thus increasing the accuracy. Being a dynamic object, the human face has a high variation in its appearance making face recognition a complex problem. With numerous surveillance and commercial applications, face recognition in videos is a challenging problem for which wide variety of methods have been proposed.In order to surpass these hurdles, an algorithm should process the face images efficiently, thus reducing the complexity. The proposed system aims at building a robust system that can handle the abundant data in an efficient manner. Various pre-processing methods are applied to make each face image suitable for further processing. After which, several algorithms are applied to categorize faces on the basis of gender, complexion and eye glasses.This paper provides a detailed survey of significant techniques used to overcome the various challenges faced while recognizing faces.*

***Key Words*:Face recognition, videos,pose, illumination, gender, complexion and eye glasses.**

1. INTRODUCTION

Human face is an important cue that is used by all people for recognition. This makes automatic face recognition an important aspect in law enforcement and commercial applications as well.Face recognition has been a widely explored area and extensive research has been done since decades, using which numerous methodshave been developed. Although significant work has been done, the available systems are still not close to efficiently and accurately identifying faces. Also, most face recognition research has been focused on recognizing faces from still images. In a video, the human face is susceptible several variations in pose and illumination. Face recognition from videos however is a much more complex problem to solve. The most important factors behind this being, its ability to provide numerous images of the same individual varying in pose, expression, illumination etc. and thus also adding to the nuisance factors encountered while recognizing faces in videos.

Face recognition using video presents various challenges and opportunities. It focuses on achieving accurate face recognition results in significantly degraded viewing conditions. Numerous performance evaluation techniques have demonstrated that algorithms that perform well in controlled environments, may not work as efficiently in surveillance contexts. To address these issues face recognition algorithms have been developed that utilize the abundant information provided by videos.

Utilizing the extra information available in a videois a major challenge. Also, there is variation in illumination, pose,resolution, and facial expressions in the different video sequences of the same subject. Thus these

variations make designing an effective video-based face recognition algorithm a challenging problem.

Needless to say, in spite of all the work that has been done for face recognition, while the algorithms may perform well for still face images, the performance for the very same algorithms may degrade for face recognition in videos as it is a difficult task and does not give accuracy. Thus, to improve accuracy, human faces are categorized basedon the classification using skin color detection, gender classification and eyeglasses classification. In fact, individuals are able to process a face in a variety of ways using characteristics, such as gender, ethnicity, and age.

Face recognition is a process of recognizing individuals using various facial features. In particular, recognizing human gender is important since it is an important factor. In addition, performance of many applications can be boosted with the help of successful gender classification. In a similar manner, categorizing faces based on eyeglasses and their complexion can go a long way in reducing the complexity of recognizing individuals in a video.

2. LITERATURE SURVEY

As explained earlier, although several face recognition algorithms may perform well in controlled environment, the very same algorithms tend to suffer when presented with surveillance videos. Some of the most challenging nuisance factors include pose and illumination, but recent studies show that expression changes and occlusion too may have a significant impact in uncontrolled environment. Needless to say most algorithms perform significantly better on high resolution videos. As known, face recognition in videos suffers greatly due to variation in appearance of the same individual's images. Therefore, while taking advantage of the abundant video data, much of the research of face recognition from videos focuses on handling the aforementioned nuisance factors.

One of the main concerns with videos is the resolution of the video. An algorithm called Iterative Back Projection algorithm[4] is used to convert low resolution image to high resolution image. According to B. Gunturk, this algorithm proves to be an effective approach since real time face videos are of low quality face images. Iterative Back Projection algorithm projects similar images one behind another in order to enhance details and thus produce a high resolution image. However this method comes with the drawback of high computational cost associated with aligning images.

Another method that has been implemented by K. Lee, uses probabilistic appearance manifold[6] for video based face recognition. An appearance manifold is built in this approach. Using piecewise linear subspaces and the changing aspects among them that are embodied in the transition matrix,which is learned from an image sequence is used to approximate it further. An appearance manifold is created for a person's face which depicts the appearance of how the face may look from several angles. This helps in identifying a face in a video when a person's face may not be turned in the direction of the camera. Thus, this method can efficiently handle large motions. It helps to approximate large non-linear appearance manifold for large 2D and 3D rotations as well as partial occlusions. This method has a few limitations as it tends to degrade in performance when subjected to changes in facial expression. Another factor that poses to be a hurdle in its efficiency is large changes in the illumination of images.

Gender Classification is the much discussed topic since two decades, in spite of which a gap exists between the desired and the actual results. Pose, illumination, changes in expression are the major factors that contribute to this gap. Gender classification is the process of identifying genderusing facial images. According to recent studies, to determine a better approach for gender classification using face images, performance of two

feature extraction algorithms have been compared by the authors[3]. One of the methods was Local binary pattern (LBP) and the other used for comparison was Histogram of oriented gradient (HOG). To detect face from an image, Haar Cascade Classifier is used. In order to normalize illumination, Histogram equalization normalization technique has been used. The classifier used for gender classification is Support vector machine (SVM). Gender classification system architecture has been implemented by the authors using OpenCv 2.4.2. Database used for the experiment is Indian face database (IFD). Experimental results conducted on Indian face database show that HOG is more efficient approach that can be used gender classification since gender recognition rate is improved by it up to 95.56%.Finally, in order to classify whether the image that of a male or female, Support Vector Machine (SVM). It was observed that LBP is less accurate for feature extraction as compared to HOG, making HOG more accurate. A classifier needs to be fast enough to handle large datasets. One such classifier is Haar Cascade classifier. In spite of changes in illumination, HOG proves to be an accurate method for gender classification.

3. PROPOSED METHOD

In this paper, the proposed system categorizes face images based on

- Gender
- Eyeglasses
- Complexion.

3.1 SYSTEM ARCHITECTURE

Working of the system with respect to architectural component is explained below:

Haar Cascade Classifier for Face Detection

Haar classifier algorithm is used for face detection. In order to pass the face to the next stage in the detection process, Haar Cascade classifier is used. A sub-window is a cropped rectangular image consisting of only the face from the original image. Faces may be of several sizes due to which the sub-window needs to be scaled. A face candidate is denoted to each section after the algorithm scans the whole image. Not all sections may be valid face candidates. Thus, in order to eliminate non-face candidates, a cascade of stages is used.Each stage consists of several features called as Haar features.Haar feature classifiers are used to classify each of these Haar features. Several features are grouped together and then applied to a stage of classifiers thus forming a cascade of classifiers. The output generated by this cascade of classifiers is further compared with a threshold value which determines if the section is a valid face candidate. If not, the section is discarded; else the next stage of classifiers is applied. This comparator is known as the stage comparator. The face candidate is concluded to be a valid face, upon passing all the stages of classifiers.

a) Haar Feature Classifier :

The value of a feature is calculated by the Haar Feature classifier using the rectangular section. Weight of each rectangular section is multiplied by its area and sum of results is calculated. Although multiple Haar features together compose a stage, the number of features in each stage is not fixed. It depends on the parameters of the individual in the training stage.

b) Haar Features:

Two or three rectangles together form a Haar feature.

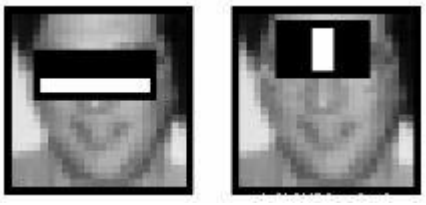

Fig -1: Haar Features

Based on the appearance, the rectangles are used to define a particular feature. As shown

in the Fig. 1 above, the dark region of the eyes forms a black rectangle, and the bridge between eyes is denoted with a white rectangle to show that the eyes on both sides form a darker region.

Support Vector Machine

The SVM is a learning algorithm that is used to minimize the classification errors for unseen patterns by finding optimal hyper plane. Not all data is linearly separable. Such data is mapped on to a high dimensional feature space where the input can be separated using a hyper plane. The equation $W^Tx+b$is used to denote the best hyper plane for separating the non-separable data. Ideally, the separation should be such that the distance between the two classes is maximized. SVM aims at finding such a hyper plane for the given set of training samples.A model is built by the algorithm, which helps to classify a test data into one class out of two. The computational cost is very high since a large amount of data is required by SVM to identify an optimal decision boundary. SVM has been used in the proposed system to classify face images based on gender.

Description of system architecture:

- **Frame Extraction :**
  User uploads a video on the application, application then extracts frames from the video. These frames are saved on local machine.
- **Face Detection :**
  Apply the Haar Cascade Classifier for the face detection in images.
- **Pre-Processing on images :**
  Once we get the faces apply the pre-processing on images like noise removal, normalization etc.

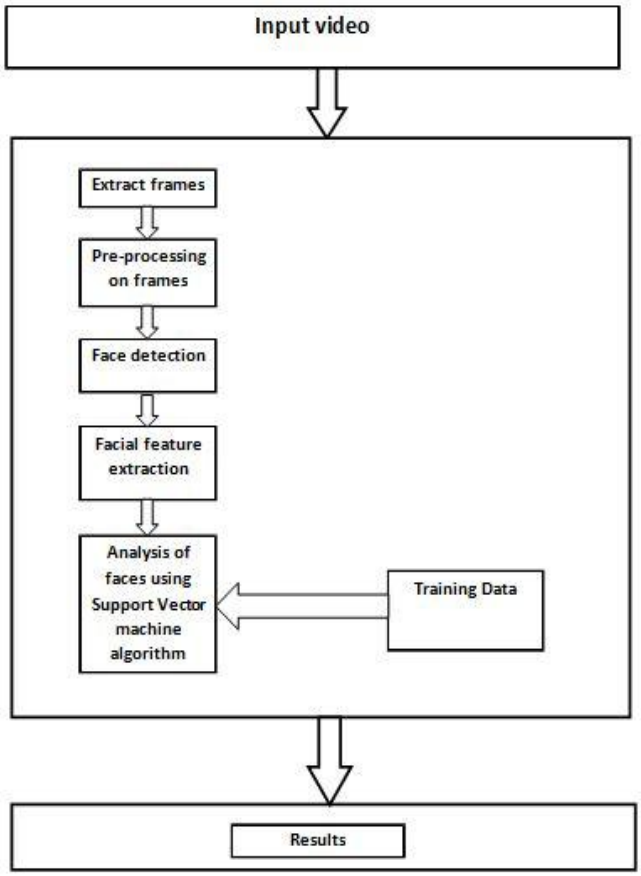

Fig-2: System Architecture

a. **RGB to Gray Scale Image :**
   Convert the image into Gray scale by taking the average of the each pixel's RGB values.

b. **Image Normalization :**
   To avoid distraction or lassitude from the images normalization of the images is necessary. It is a process used to change the range of values of pixel intensity.

c. **Noise Removal :**
   During the image acquisition process noise removal is used to remove errors, so that the true intensities of the pixels are reflected.

- **Facial Feature Extraction :**
  Several features of the face such as eyes, nose, lipare extracted in the form of points.
- **Analyzethe Faces :**

Using Support Vector Classification algorithm analysis on the facial features and makes the classification gender wise.

4. CONCLUSION

The system will enable efficient preprocessing of face images and present a group of face images categorized on the basis of facial features. Support Vector Machine is used to categorize faces.

In future work, real time face recognition can be applied in various fields. Also, parallel processing technique can be implemented which will further reduce the processing time.

ACKNOWLEDGEMENT

We would like to thank our project guide Prof. SanjeevkumarAngadi and Head of the Department Prof. Shweta Kale for their valuable guidance, support, time and suggestions. We would also like to thank them for providing us all the required facilities that helped in completing the paper. We would also like to thank all the staff members of the Department of Information Technology of RMD Sinhgad School of Engineering, Warje. We would also like to thank the institute for providing all the required facilities.

REFERENCES

[1] Y. C. Chen, V. M. Patel, S. Shekhar, R. Chellappaand P. J. Phillips, "Video-based face recognition via joint sparse representation," in *Proc.10th IEEE Int. Conf. Workshops Automatic Face and Gesture Recognition*, Apr. 2013, pp. 1–8.

[2] Md. Hafizur Rahman, Suman Chowdhury, Md. Abul Bashar, "An Automatic Face Detection and Gender Classification from Color Images using Support Vector Machine",Journal of Emerging Trends in Computing and Information Sciences, Vol. 4, No. 1 Jan 2013

[3] Vandna Singh, Dr. Vinod Shokeen, Mr. Bhupendra Singh, "Comparison Of Feature Extraction Algorithms For Gender Classification From Face Images", International Journal of Engineering Research & Technology (IJERT), Vol. 2 Issue 5, May – 2013

[4] J. R. Barr, K. W. Bowyer, P. J. Flynn, and S. Biswas, "Face recognition from video: A review," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 26, no. 5, p. 1266002, 2012.

[5] V. M. Patel, T. Wu, S. Biswas, P. J. Phillips, and R. Chellappa, "Dictionary-based face recognition under variable lighting and pose," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 3, pp. 954–965, Jun. 2012.

[6] K. Lee, J. Ho, M. Yang, and D. Kriegman, "Video-based face recognition using probabilistic appearance manifolds," in *Proc. IEEE Int. Conf.CVPR*, Jun. 2003, pp. 313–320.

[7] Himanshu S. Bhatt, Richa Singh, MayankVatsa, "On recognizing faces in videos using clustering-based re-ranking and fusion", *IEEE Transactions on information forensics and security, vol. 9, no. 7, July 2014*

[8] Michael R Blaha, James R Rumbaugh, Object-Oriented Modeling and Designwith UML,2nd Edition. This edition of Object Oriented Modeling and Design with UML offers a clearexplanation of basics of object-oriented thinking via UML models which has been used extensively for system design.

[9] http://docs.opencv.org/3.1.0/d7/d8b/tutorial_py_face_detection.html . This site was used to study Haar Cascade classifier in detail.