International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 04, April 2017
www.ijarse.com
IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

ROBUST FACE RECOGNITION FROM MULTI-VIEW VIDEOS

Rinsi.S[1], Manoj.R[2]

[1] *Computer and Information Science, College Of Engineering Munnar, (India)*

[2] *Computer Science ,College Of Engineering Munnar, (India)*

ABSTRACT

The recognition of the persons from videos has numerous applications in Video Surveillances and Computer Vision. Face detection in still images is applied in many of the devices and in many of the applications. Face detection in videos is a recent technique which ensures the identification of person in the video. In a video normally the pose of the persons and the illumination variations were normally present. The main challenge of detecting face images in videos is the pose and the illumination variations and sudden changes in the movement of the object. The videos taken from cameras are taken. The surveillance videos were normally not so clear and there may be some illumination variations and blurness at some places due to rapid movement of the objects. The proposed system analyzes and recognizes the exact face image from the video while the existing systems deals with the recognition of the face images from still images. The proposed system is capable of identifying the face images in the video in a better manner even though there are pose variations illumination variations and blurs occurring due to rapid motions of the subjects in the video. The videos were converted into frames. Preprocessing is applied to the video frames using median filter to remove the unwanted noises from the frames. The face of the person is detected and the face is masked. The face region is masked so that we identify the face images of the particular person more clearly. The gradient value of the image and the histogram values were calculated. These will be helpful in the identification of the face positions continuously. Based on the detected face the position of the person is determined in each frame and the person is tracked continuously. Hog features are extracted and then RBF kernel function is applied along with spherical harmonics function. The Hog features are more reliable and they are useful in the identification of the face image of the particular person.Face images is retrieved from the dataset which is trained based on the distance calculated using Bhattacharya coefficient. The Bhattacharya distance is used for identifying the face image from the Dataset. The results shows that the recognition rate of the proposed system is increased compared to other existing systems. The system is more reliable for the identification of the face images from the videos.

Keywords: ***Face recognition, multi-camera networks, pose variations, spherical harmonics***

I. INTRODUCTION

The identification of persons in the video has numerous applications in video surveillance system. It is easy to identify persons in frontal view cameras by identifying the face. The identification of the persons in other views was difficult. The identification of the persons in different views can be much helpful in identifying the

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

suspicious persons in a video[1]. The process requires some tedious works including a multiple camera network an effective tracking system to track and identify the persons in an effective manner.A facial recognition system is a computer application for automatically identifying or verifying a person from a digital image or a video frame from a video source. One of the ways to do this is by comparing selected facial features from the image and a facial database. It is typically used in security systems and can be compared to other biometrics such as fingerprint or eye iris recognition systems. Face recognition has been one of the most active research topics in computer vision and pattern recognition for more than two decades. Some facial recognition algorithms identify facial features by extracting landmarks, or features, from an image of the subject's face. For example, an algorithm may analyze the relative position, size, and/or shape of the eyes, nose, cheekbones, and jaw. These features are then used to search for other images with matching features. Other algorithms normalize a gallery of face images and then compress the face data, only saving the data in the image that is useful for face recognition. A probe image is then compared with the face data. One of the earliest successful systems is based on template matching techniques applied to a set of salient facial features, providing a sort of compressed face representation. Recognition algorithms can be divided into two main approaches, geometric, which looks at distinguishing features, or photometric, which is a statistical approach that distills an image into values and compares the values with templates to eliminate variances. A facial recognition device is one that views an image or video of a person and compares it to one that is in the database. It does this by comparing structure, shape and proportions of the face; distance between the eyes, nose, mouth and jaw; upper outlines of the eye sockets; the sides of the mouth; location of the nose and eyes; and the area surrounding the check bones. Upon enrolment in a facial recognition program, several pictures are taken of the subject at different angles and with different facial expressions. At time of verification and identification the subject stands in front of the camera for a few seconds, and then the image is compared to those that have been previously recorded. To prevent a subject from using a picture or mask when being scanned in a facial recognition program, some security measures have been put into place. When the user is being scanned, they may be asked to blink, smile or nod their head. Another security feature would be the use of facial thermography to record the heat in the face. The main facial recognition methods are: feature analysis, neural network and automatic face processing. Some facial recognition software algorithms identify faces by extracting features from an image of a subject's face. Other algorithms normalize a gallery of face images and then compress the face data, only saving the data in the image that can be used for facial recognition[2][3].

II. RELATED WORKS

The face detection process is employed in still images. 3D model reconstruction, local feature matching, regularized regressions were employed to identify the face images. The feature based approaches uses SVM classifier to verify the poses. SVM based approaches is used to select the frontal poses from the video frames in video based face recognition [8]. 3D face models based matching process is used to match the face images. In feature based approaches LBP and PCA were more commonly used for face matching. Gaussian Mixture Model, Hidden Markov Models were proposed to find the face regions in the video. The feature matching techniques extracts the feature values from the images using several algorithms that were proposed. Then based

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

on the extracted feature values the face images were recognized from the videos. Two pair of face were got from the user and the similarities between them were measured and the face is recognized [4].

III. PROPOSED SYSTEM

The video is converted into frames. Preprocessing is applied to the video frames using median filter. Then face region in the frame is detected and the face region is masked. Based on the position of the face in the frame the face region is detected in the whole video. Spherical harmonics is applied to the input image. HOG features are extracted and then RBF kernel function is applied. The face images of some subjects were trained. The distance between the train feature and the test image features were calculated. The image that is having the minimum distance is retrieved from the database.

3.1 Preprocessing

The video is first converted into frames. The noises in the frames reduce the quality of the frames. Each frame is considered as images. Inorder to improve the quality of the images we normally employ some filtering operations. Median filter is used for filtering. The median filter considers each pixel in the image in turn and looks at its nearby neighbors to decide whether or not it is representative of its surroundings. The identified noisy pixel is replaced by the median of neighboring pixel values. The median is calculated by first sorting all the pixel values from the surrounding neighborhood into numerical order and then replacing the pixel being considered with the middle pixel value. In preprocessing unnecessary noises in the video were eliminated. An unnecessary noise refers to the unwanted pixels in the frames. Pre-processing methods use a small neighborhood of a pixel in an input image to get a new brightness value in output image. Such pre-processing operations are also called filtration [5]. Local pre-processing methods can be divided into the two groups according to the goal of the processing as smoothing and filtering. Smoothing suppresses noise or other small fluctuations in the image equivalent to the suppression of high frequencies in the frequency domain. Unfortunately, smoothing also blurs all sharp edges that bear important information about the image. Gradient operators are based on local derivatives of the image function. Derivatives are bigger at locations of the image where the image function undergoes rapid changes. The aim of gradient operators is to indicate such locations in the image. Gradient operators suppress low frequencies in the frequency domain (i.e. they act as high-pass filters). Noise is often high frequency in nature; unfortunately, if a gradient operator is applied to an image the noise level increases simultaneously. Filters are used in the preprocessing techniques. The Gaussian smoothing operator is a 2-D convolution operator that is used to `blur' images and remove detail and noise. In this sense it is similar to the mean filter, but it uses a different kernel that represents the shape of a Gaussian (`bell-shaped') hump. The effect of Gaussian smoothing is to blur an image, in a similar fashion to the mean filter. The degree of smoothing is determined by the standard deviation of the Gaussian. (Larger standard deviation Gaussians, of course, require larger convolution kernels in order to be accurately represented.). The Gaussian outputs a `weighted average' of each pixel's neighborhood, with the average weighted more towards the value of the central pixels. This is in contrast to the mean filter's uniformly weighted average. Because of this, a Gaussian provides gentler smoothing and preserves edges better than a similarly sized mean filter.

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

3.2 Face Detection

The face image are detected in the frame. The face region is detected using the vision cascade operator which identifies the face region in the image. It gives the x and y position of the face images. The position of the detected face image is taken and a rectangle is drawn at the particular location [7].

3.3 Face Masking

The face region is then masked based on the detected face region. The region which is detected as face is taken and the particular region is marked separately. The areas other than the region which is detected as a face were blackened and face alone is masked. This will be helpful in the correct recognition of the face images. The vision.CascadeObjectDetector System object comes with several pretrained classifiers for detecting frontal faces, profile faces, noses, upper body, and eyes. However, these classifiers may not be sufficient for a particular application. The Computer Vision System Toolbox cascade object detector can detect object categories whose aspect ratio does not vary significantly.Objects whose aspect ratio remains approximately fixed include faces, stop signs, or cars viewed from one side. The vision.CascadeObjectDetector System

3.4 Tracking

The position of the face regions at each frame is updated each time and the person is tracked all over the video. The position of the face region at each time is identified by analyzing the movement of the person in the consecutive frame. The position is updated each time so that the system is trained to identify the movement of the person in the frame. Each time the position of the rectangle is moved according to the movements identified in the frame.

3.5 Recognition

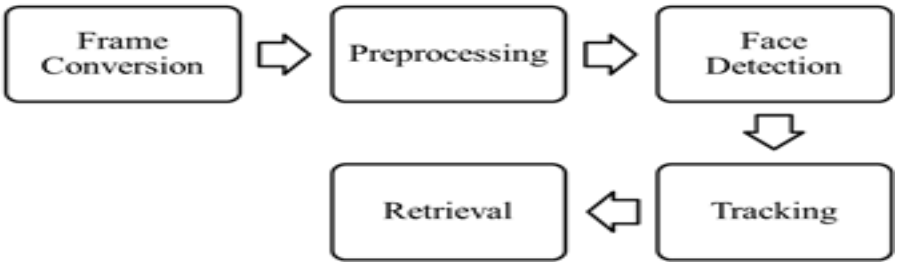
Spherical harmonic pattern is applied to the face images. Then the HOG features are extracted from the image. The extracted features then passed to RBF kernel which projects the feature values based on different views of the subjects. The distance is calculated between the test feature and the train feature. For this Bhattacharya distance is calculated. The image corresponding to the feature having minimum distance is retrieved from the database[5] [6].

3.6 Spherical Harmonics

Spherical harmonics are the angular portion of a set of solutions to Laplace's equation. Represented in a system of spherical coordinates, Laplace's spherical harmonics Y_l^m are a specific set of spherical harmonics that forms an orthogonal system, first introduced by Pierre Simon de Laplace. Spherical harmonics are important in many theoretical and practical applications, particularly in the computation of atomic orbital electron configurations, representation of gravitational fields, geoids, and the magnetic fields of planetary bodies and stars, and characterization of the cosmic microwave background radiation. In 3D computer graphics, spherical harmonics play a special role in a wide variety of topics including indirect lighting (ambient occlusion, global illumination, precomputed radiance transfer, etc.) and recognition of 3D shapes [9] [10].

International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 04, April 2017
www.ijarse.com
IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

IV. SYSTEM ARCHITECTURE



V. EXPERIMENTAL RESULT

VI. CONCLUSION

The face image of the subject is detected exactly. The detection of the face and the retrieval is very useful in many of the processes. There is no need to identify the pose and other such process. The person is identified correctly even though there are pose variations and illumination variations. The performance of the proposed system is high compared to the techniques that are used to identify the face images in multi view videos. The process can be further developed by recognizing the actions of the persons in the video. The features used can be further changed which recognizes the persons in a better way. For feature extraction additional features such as SIFT features SURF features can be used. The process can be developed to identify the reaction and persons or actions performed by the person. Inorder to recognize the action or reaction of the person some classifiers can be used. The classifiers such as Probabilistic Neural Network classifier, multi class adaboost classifiers can be used for this purpose. If the classifiers were trained they can exactly recognizes the action or reaction of the person. Multiclass classifiers have to be used for this purpose. The multiclass classifiers can recognize different actions and reactions of the persons. The supervised classifiers need training of the datasets that we are using to recognize the action. The unsupervised classifiers clusters the features to find the category to which the input feature belongs. The accuracy, sensitivity, specificity of the classifiers can be measured and the confusion matrix and ROC curve can be obtained that represents the performance of the classifiers that we are using. The calculated performance can be compared with some other methods.

REFERENCES

[1] S. Ba and J. M. Odobez, "Probabilistic head pose tracking evaluation in single and multiple camera setups," Multimodal Technol. Perception Humans, vol. 4625, pp. 276–286, Jun. 2008.

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

[2] Q. Cai, A. C. Sankaranarayanan, Q. Zhang, Z. Zhang, and Z. Liu, "Real time head pose tracking from multiple cameras with a generic model," in Proc. CVPR Workshops, Jun. 2010, pp. 25–32.

[3] D. Beymer and T. Poggio, "Face recognition from one example view," in Proc. IEEE Int. Conf. Comput. Vis., Jun. 1995, pp. 500–507.

[4] X. Chai, S. Shan, X. Chen, and W. Gao, "Locally linear regression for pose-invariant face recognition," IEEE Trans. Image Process., vol. 16, no. 7, pp. 1716–1725, Jul. 2007.

[5] H. S. Lee and D. Kim, "Generating frontal view face image for pose invariant face recognition," Pattern Recognit. Lett., vol. 27, pp. 747–754, May 2006.

[6] V. Blanz, T. Grother, P. J. Phillips, and T. Vetter, "Face recognition based on frontal views generated from non-frontal images," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., vol. 1. Jun. 2005, pp. 454–461.

[7] S. Li, X. Liu, X. Chai, H. Zhang, S. Lao, and S. Shan, "Morphable displacement field based image matching for face recognition across pose," in Proc. Eur. Conf. Comput. Vis., Oct. 2012, pp. 102–115.

[8] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3D faces," in Proc. 26th Annu. Conf. Comput. Graph. Interact. Tech., 1999, pp. 187–194.

[9] V. Blanz and T. Vetter, "Face recognition based on fitting a 3D morphable model," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 9, pp. 1063–1074, Sep. 2003.

[10] P. Breuer, K.-I. Kim, W. Kienzle, B. Scholkopf, and V. Blanz, "Automatic 3D face reconstruction from single images or video," in Proc. IEEE Int. Conf. Autom. Face Gesture Recognit., Sep. 2008, pp. 1–8.