ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

Face Detection for Human Identification in Surveillance

[I]S.G.Kartheeswari, [II]Ms.B.Vijaya Lakshmi (M.E(Ph.D))
[I]Final Year Research Student, [II]Associate Professor
[I,II]Dept. of ECE, K.L.N.College of Information Technology, Madurai, Tamilnadu, India

Abstract

In the video sequence human faces have unlimited orientations and positions so face detection and clustering is very important. In this paper, I have proposed a method to cluster human faces from the video sequence based on Spatio Temporal method.The proposed method is based on three main stages. First I have used a face detector to localize faces in all frames of the video and extract various features of the detections.Face detection is done by using viola jones detector and feature extraction is done by using SURF detector. Optical flow estimation is used to compute the movement between any of two frames.In optical flow estimation pyramidal version of lucas kanade algorithm is used. This estimation is used to compute a dissimilarity matrix such as appearance dissimilarity,time dissimilarity. Clustering tracks the same faces from the video sequence it may have different pose,image orientation or it may occluded by objects. Finally, an optimization method involving clustering of faces.

Keywords

Face Detection, Optical Flow, Clustering, Tracking, Dissimilarity

I. Introduction

Face detection is required as the first step of the automatic face analysis system. This work has been widely investigated in recent years because it overlies many areas of application: face recognition, man-machine interaction systems, visual communication systems, video-surveillance, etc. However,face detection is a challenging task due to variation in illumination, variability in scale, location, orientation and pose,Facial expression, occlusion and lighting conditions also change the overall appearance of face.In this paper face is detected by means of viola-jones algorithm and then feature extraction is done which is used to cluster the faces.Optical flow estimation is calculated between any of two frames which has been extracted from the video.In optical flow method pyramidal version of the Lucas–Kanade algorithm to represent both small and large displacements.From the above result output dissimilarity matrix is calculated and then Clustering is performed based on the feature extraction and output dissimilarity matrix.

II. Literature Survey

Many novel methods have been proposed to resolve Face detection and clustering. For example, the template-matching methods are used for face localization and detection by computing the correlation of an input image to a standard face pattern. The feature invariant approaches are used for feature detection of eyes, mouth, ears, nose, etc. The appearance-based methods are used for face detection with eigenface, neural network , and information theoretical approach. Nevertheless, implementing the methods altogether is still a great challenge. First, the all the faces are vertical and have frontal view, second, they are under almost the same illuminate condition. Face pattern detection discriminates and localizes the face within the identified body parts.Faces and bodies of users are tracked over several temporal scales: short-term (user stays within the field of view), medium-term (user exits/reenters within minutes), and long term (user returns after hours or days). Short-term tracking is performed using simple region position and size correspondences,while medium and long-term tracking are based on statistics of user appearance.Various algorithms are for face detection. We propose a method to cluster-specific object detections of a video sequence, which we applied to face detections. Due to the large number of detections extracted from video, an automatic clustering of face detections is interesting for visual surveillance applications. For archive browsing or for facetagging on videos, it is easier to investigate with an album of faces than with a set of all the detected faces.Our efforts focus on real visual surveillance constraints: cluttered scenes, uncontrolled, and containing multiple small faces.

III. Proposed work

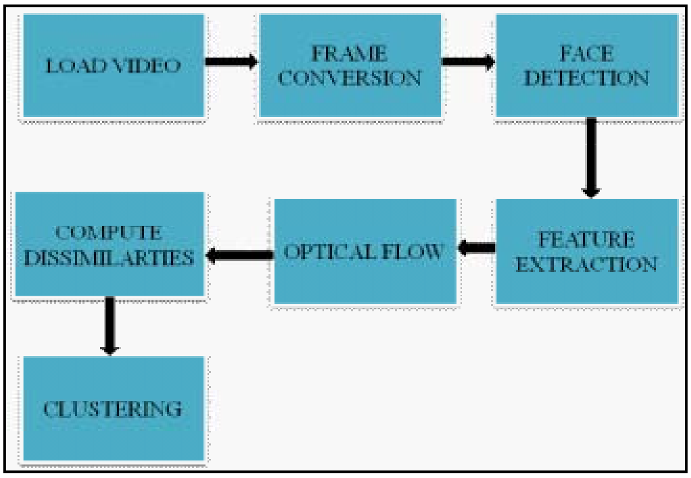

Fig. 1. Block diagram of proposed system

1. Load Video

Load the video using the avi reader method.
Read Audio/Video Interleaved (AVI) file

2. Frame Conversion

In this module, Loaded video is convert into frames.

3. Face Detection

Faces are detected by means of viola-jones algorithm.

Viola-Jones's Detector:

This framework, designed for rapid object detection, is based on the idea of a boosted cascade of weak classifiers [1] but extends the original feature set and provides different boosting variants for learning [2].The cascade learning algorithm is similar to decision tree learning. Essentially, a classifier cascade can be seen as a degenerated decision tree. For each stage in the cascade a separate subclassifier is trained to detect almost all target objects while rejecting a certain fraction of the non-object patterns. The resulting detection rate, D, and the false positive rate, F, of the cascade is given by the combination of each single stage classifier rates:

$$D = \prod_{i=1}^{K} d_i \qquad F = \prod_{i=1}^{K} f_i \tag{1}$$

Under this approach, given a 20 stage detector designed for refusing at each stage 50% of the non-object patterns (target false positive rate) while falsely eliminating only 0.1% of the object patterns (target detection rate), its expected overall detection rate is 0.99920 ≈ 0.98 with a false positive rate of 0.520 ≈ $0.9 * 10^{-6}$. In a single stage classifier one would normally accept false negatives in order to reduce the false positive rate classifiers are combined. The cascaded classifier is composed of stages each containing a strong classifier.

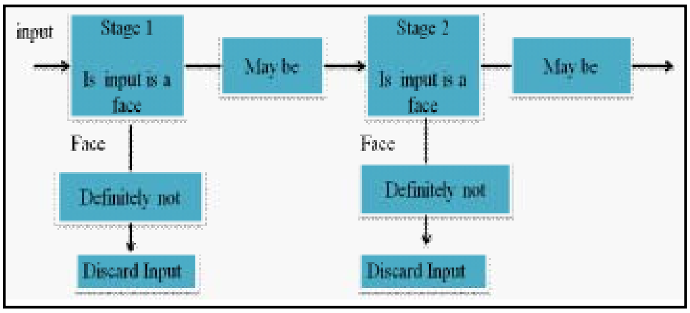

Fig. 2: Cascade classifier

4. Feature Extraction

First we use a face detector to localize faces in all frames of the video and extract various features of the detections. These features are then used to compute a dissimilarity matrix based on appearance and space–time. Here we extract the feature of the video by spatio temporal algorithm.First, we select candidate points based on the SURF detector. Next, we calculate several optical flow features at each point with local temporal units divided in order to consider consecutiveness of motions. We calculate the variation between the each frames.This values will be used as a feature values of video. Initially, from a given image I, SURF constructs the integral image according to

$$I_{\in} = \sum_{m=1}^{x} \sum_{n=1}^{y} I \tag{2}$$

$$D_{SURF}(X;\sigma) = G_{xx}(X;\sigma)G_{yy} - 0.81{G_{xy}}^2(X;\sigma) \tag{3}$$

SURF uses the above approximation, to calculate pixel position at x = (x; y) and scale σ.

5. Optical Flow

There are many methods to compute an optical flow between two frames, one of the main issues is to represent a large scale of displacementsIn this, we used a pyramidal version of the Lucas– Kanade algorithm[4] to represent both small and large displacements. In each frame having a detection, we compute the optical flow from the previous to the current frame and from the current to the next frame, then these two optical flows are averaged. The resulting speed vector of detection is obtained by taking the most representative flow vector in the detection area.

6. Compute Dissimilarities

In this paper we have calculated appearance based dissimilarity and space time based dissimilarity.Finally output dissimilarity is calculated using above two dissimilarity

(i). Appearance Dissimilarity

Detection appearance is represented by an HS-V histogram. This histogram is the concatenation of a 2D HS(Hue Saturation) histogram and a 1D V(Value) histogram of image pixels . where H, S, and V represent hue, saturation, and value of a color, respectively). If the S and V values are large enough for a pixel, the pixel is counted in the HS histogram, or else it is counted in the V histogram.To measure the dissimilarity between two HS-V histograms, we used the Bhattacharyya coefficient.

(ii). Space–time dissimilarity

Space time dissimilarity is almost similar to optical flow estimation which measure the trackelet time difference between the previous frame and in frame.

7. Clustering

Clustering is performed based on feature extraction and output dissimilarity.This method is used to cluster the faces among multiple faces presented in the video.faces may have different position and orientation this method effectively clusters the faces.

IV. Experimental Results

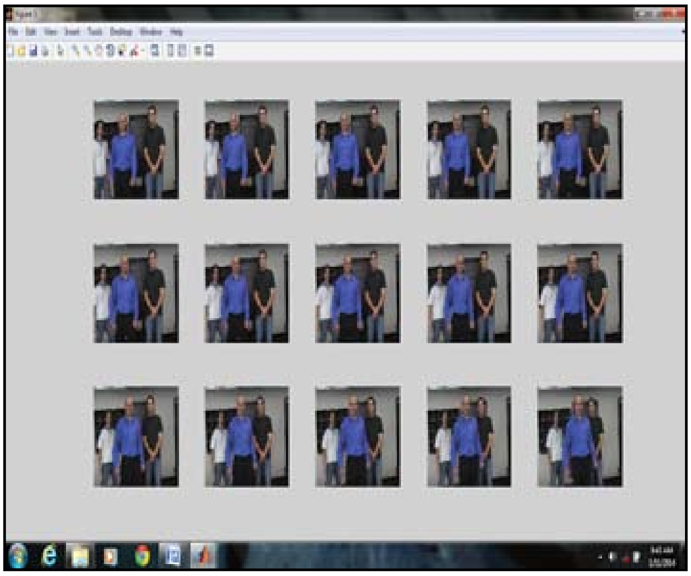

Fig. 3. Extracted frame from video sequence

Figure. 4. Detected faces

ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)

Figure.5.Frames in HSV color space

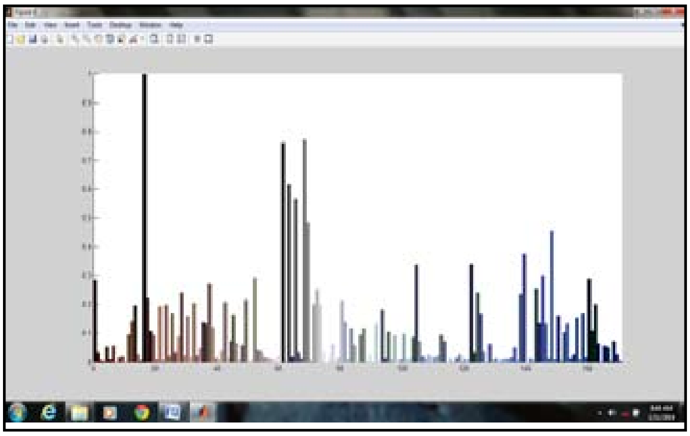

Figure. 6.Color histogram

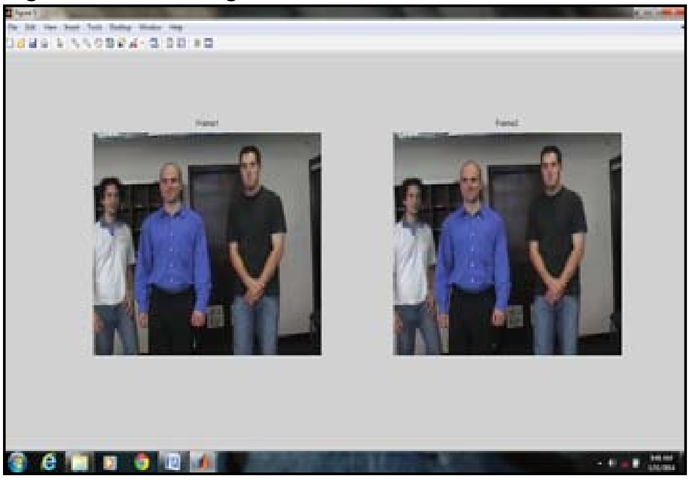

Figure. 7. Two frames for optical flow

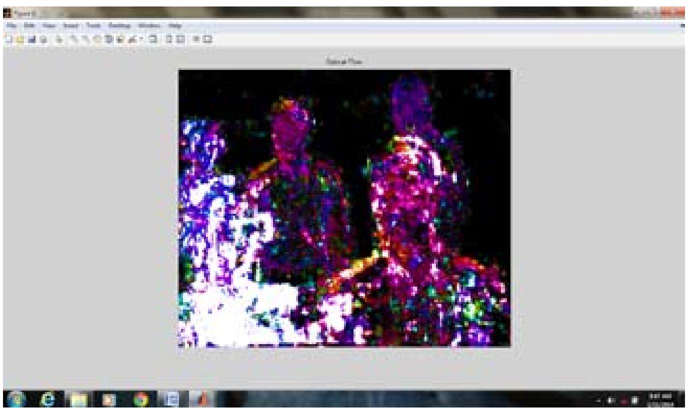

Figure. 8. Optical Flow

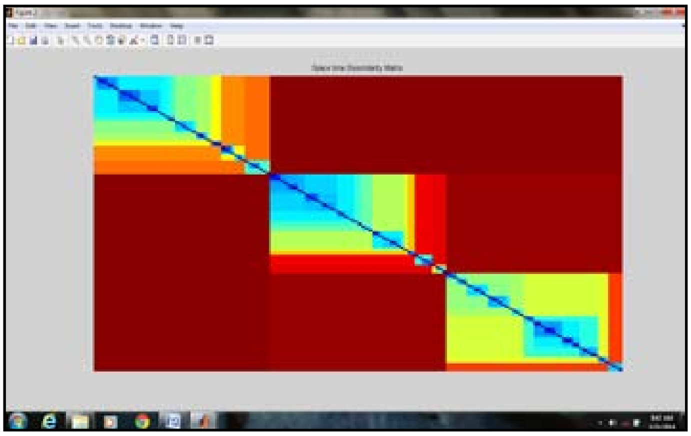

Figure. 9. Space Time dissimilarity

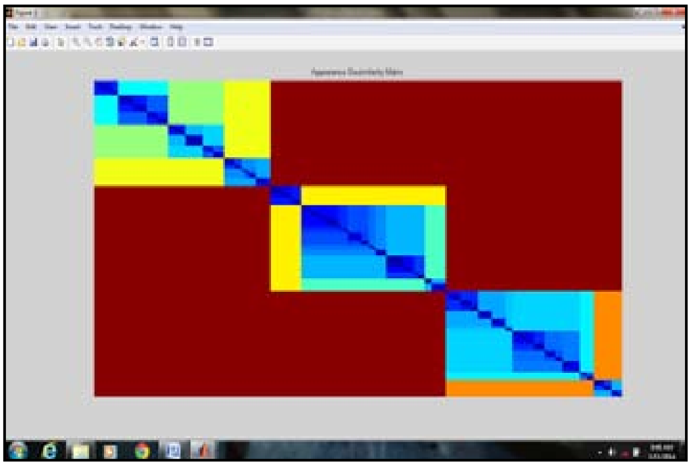

Figure 10. Appearance dissimilarity

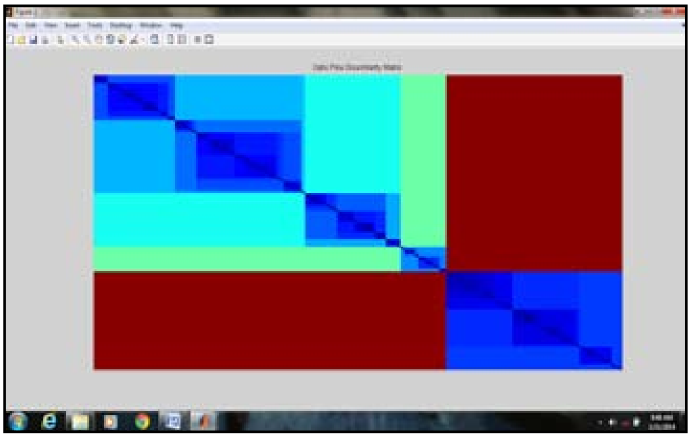

Figure 11. Output dissimilarity matrix

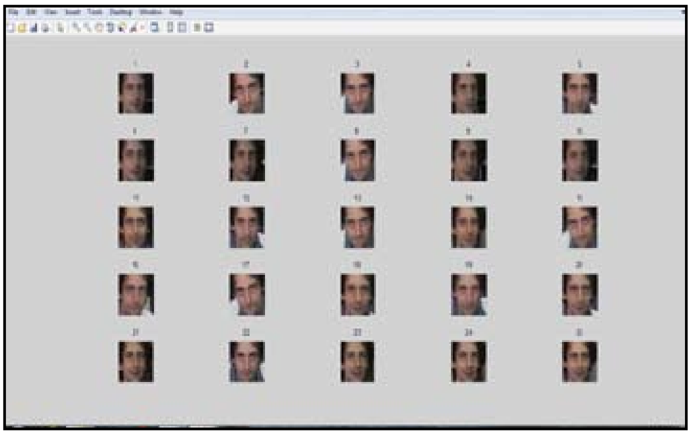

Figure 12.Clustered faces

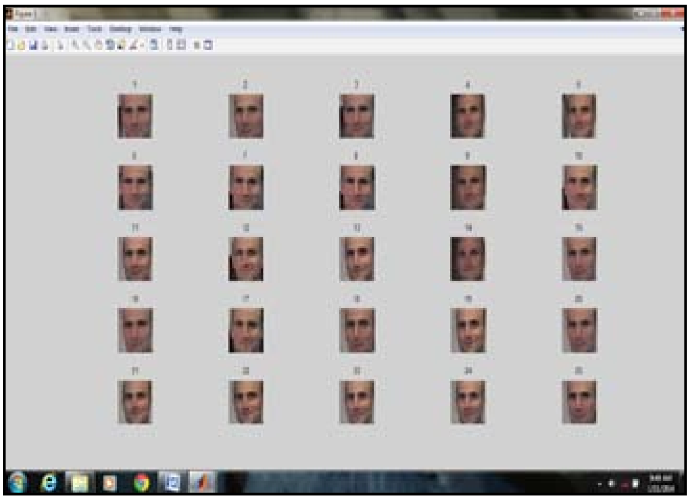

Figure 13. Clustered faces

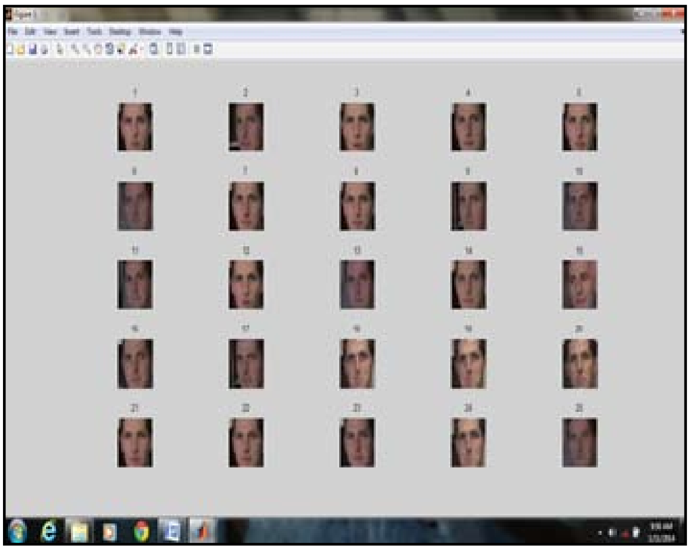

Figure 14. Clustered faces

V. Conclusion

I have proposed a method to cluster faces in video sequences. Face detection is done by using viola-jones detector then feature extraction is done by using SURF detector.Optical flow estimation is calculated by using pyramidal lucas-kanade algorithm.Optical flow estimation,appearance dissimilarity,time dissimilarity is used to compute the output dissimilarity. Feature extraction is used to identify the faces from the video sequence from that result clustering is performed.Faces are efficiently clustered by spatio-temporal clustering method.

References

[1] *Paul Viola, Michael J. Jones, and Daniel Snow.Detecting pedestrians using patterns of motion and appearance..In Proc. of the International Conference on Computer Vision, October 2003.*

[2] *Rainer Lienhart, Alexander Kuranov, and Vadim Pisarevsky,Empirical analysis of detection cascades of boosted classifiers for rapid object detection.In DAGM'03, Magdeburg, Germany, September 2003.*

[3] *M Kim, S Kumar, V Pavlovic, H Rowley, Face tracking and recognition with visual constraints in real-world videos. In IEEE Conference on Computer Vision and Pattern Recognition (Anchorage AK, 23-28 June 2008)*

[4] *J Marzat, Y Dumortier, A Ducrot, Real-time dense and accurate parallel optical flow using cuda.In Proceedings of The 17th InternationalConference in Central Europe on Computer Graphics, Visualization andComputer Vision (WSCG). (WSCG, 2009). [http://www.wscg.eu/]*

[5] *S Dubuisson, J Fabrizio, Optimal recursive clustering of likelihood functions for multiple object tracking,In Elsevier Pattern Recognition Letters , vol.30 (Elsevier, 2009), pp. 606–614. [http://www.elsevier.com]*

[6] *SimeonSchwab,Thierry Chateau, Christophe Blanc and Laurent Trassoudaine, Schwab, A multi-cue spatio-temporal framework for automatic frontal face clustering in video Sequences, EURASIP Journal on Image and Video Processing 2013, 2013:10*

[7] *Z Lu, MA Carreira-Perpin,Constrained spectral clustering through affinity propagation in IEEE International Conference on Computer Vision.(Portland, 23-28 June 2008*

[8] *T Brox, C Bregler, J Malik, Large displacement optical flow. in IEEE Conference on Computer Vision and Pattern Recognition. (Berkeley, 20-25June2009),pp. 41–48*

Ms.B.Vijaya Lakshmi M.E(Ph.D) has received her M.E degree from karunya university and she is working as a associate professor in K.L.N.College of information technology, Madurai, India.

Ms.S.G.Kartheeswari is doing Final year M.E in Digital Communication and Networking in K.L.N College of information technology, Madurai, India.