PUBLICATION OF Little Lion Scientific R&D, Islamabad PAKISTAN
Journal of Theoretical and Applied Information Technology
30th June 2011. Vol. 28 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

CHARACTER IDENTIFICATION IN VIDEOS USING IRIS RECOGNITION

RAMYA G [1], VIDHYALAKSHMI J [2], Dr.A..UMAMAKESWARI [3]

[1, 2] M.Tech - CSE, School Of Computing, SASTRA University, Tamilnadu, India
[3] Associate Dean, School Of Computing, SASTRA University, Tamilnadu, India

ABSTRACT

In Character based browsing which is predominant in films, the identification of speaking faces plays a vital role. After identifying the speaking face, clustering of those faces is a difficult task. The difficulty is in extracting the exact speaking faces and clustering them. In the existing techniques, lip movement of a character is taken as Region Of Interest (ROI) for identifying the faces in the films. For clustering those faces, the k-means clustering technique is used. By using the above technique, retrieval of speaking faces cannot produce an accurate result. So, in this paper, lip movement as ROI and IRIS recognition for clustering of speaking faces are used. This paper proposes a mechanism to generate precise results by comparing with the existing techniques.

Keywords: *Speaking face identification, ROI, lip alignment, IRIS, PCA*

1. INTRODUCTION

With the prosperous development in the film industry, more number of movies is being released. There are 'n' number of characters involved in all the movies. In recent years, the creators are more interested in adding subtitles for a corresponding video. Generally subtitles are added to make the video interactive. In order to add subtitle, it is necessary to identify the characters in the video. Characters are the heart of the movie. Their interactions make movie as a small society. Character identification in the movie poses a challenge to the individual who retrieves the information from it.

In the existing technique, the speaking faces are identified. This can be done by identifying lip movement of the character in the movie. For adding subtitle in the movie, the identification of speaking faces alone is not enough. This is followed by clustering of speaking faces. Clustering of speaking faces is more important. Clustering has been done by comparing the pixel of the characters which may not be accurate. Same faces may not be clustered due to the variation in the poses or illumination of the character. This results in chaos.

The objective of the paper is to identify the characters and label them with their names. In this paper, the characters have been identified by taking lip movement as ROI. After identifying the speaking faces, the clustering can be done by IRIS techniques. Based on this paper, it is easy for the viewer to retrieve the character of their interest. Since every human will have a unique pupil size, clustering will produce exact results.

2. RELATED WORK

The detection of any moving object for example lip movement in case of speaking faces involves verification of the lips and its movement in image sequences. Even in the case of a simple video sequence it is very challenging. [1][2].

Name-it [3] is the first method which identifies the speaking faces in the video. They identified the character which frequently co-occurs with the particular character. Neural network methodology was used for extracting faces from the video. The neural network helps to predict the front view faces with different sizes, and thus the faces are extracted. They used the Eigen vector technique to find similarity between the faces.

PUBLICATION OF Little Lion Scientific R&D, Islamabad PAKISTAN
Journal of Theoretical and Applied Information Technology
30th June 2011. Vol. 28 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Character identification in feature length film using global face name matching [4] detects the speaking face from the video and labels those faces with their names. They used lip movement or alignment as their ROI and changes in the movement represents the speaking faces. The k-Means clustering technique is used to cluster the speaking faces. They compared the (face) pixel value of the corresponding faces.

A Unified framework for object retrieval and Mining [5] proposes a framework to retrieve all the objects involved in the video. It uses novel shot cut detection and shot detection to extract the objects from the video. It also proposes the index which represents the occurrence of all the objects involved in the video.

Finding X person: Correlating Names with visual appearances retrieve the particular person in the video. It uses the Eigen face technique to find the faces in the video and face detector to store the faces.

Video Segmentation [7] partitions the video into the frames or shots [2]. They proposed the fast video segmenting algorithm. It uses the baseline mode and shadow cancellation mode and adaptive threshold method for partitioning the video.

Multiple Instances for labeling faces in the Broadcasting News Video [8] labeled the important faces by the way they appear in the news. They used the multiple instances learning for identifying and naming in the videos.

3. APPROACH PROPOSED

Generally, in films more number of characters are involved. Each and every character has its own uniqueness. If more than one character appears in the scene, it represents the interactions between them. Hence in this approach, the aim is to retrieve the accurate speaking faces. The exact faces can be obtained by using certain unique features in the human.

Generally the unique feature in human is fingerprint and pupil (IRIS). In order to retrieve the exact speaking faces IRIS can be considered as a unique feature. Even though the pixel value of face is same for more than one character, IRIS can be used to differentiate those faces.

PHASES

This approach includes the following phases:

a. Video Segmentation
b. Face Recognition
c. Identification of speaking faces
d. Clustering of speaking faces

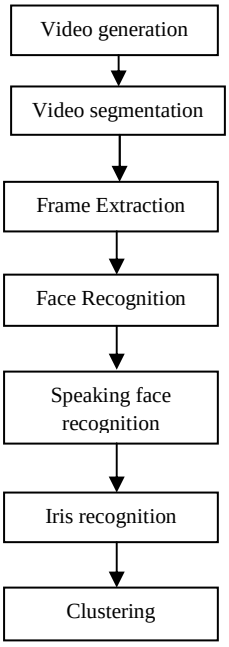

Figure 1. Proposed Framework

A. VIDEO SEGMENTATION

Video Segmentation is the process of dividing the video into number of shots in order to analyze the video in different perspectives. The video can be divided into 'n' number of frames. Each and every frame will be of same size. This helps to get a clear idea about the objects which is present in that video. By segmenting the video into frames, it is easy to identify the number of characters in the video. This could be first step in the retrieval of the speaking faces. Change detection method is one of the video segmentation algorithms which is used here. The frame difference can be calculated by setting a threshold value. This helps to identify the change detection.

B. FACE RECOGNITION

Before identifying the speaking faces, first the faces which are present in that video are identified. Face recognition algorithm can be used. Even though more number of characters are involved in the video, is easy to extract only human faces from the video by using the face recognition algorithm it. Faces can be identified by using PCA (Principal Component Analysis) [9] and stored.

PUBLICATION OF Little Lion Scientific R&D, Islamabad PAKISTAN
Journal of Theoretical and Applied Information Technology
30th June 2011. Vol. 28 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

C. IDENTIFICATION OF THE SPEAKING FACES

After extracting all the faces from the video the next step is to identify the speaking faces. The speaking faces are identified from the lip movement. Here lip movement is taken as ROI. The SIFT algorithm identifies the changes in the alignment of the lip. The changes in the ROI detect whether somebody is speaking or not. The difference can be calculated by using Normalized Sum of Absolute (NSAD).

4. CLUSTERING OF FACES

In existing technique, the (face) pixel value of the person is taken for clustering of face. For clustering k-means algorithm is used. The pixel values of same person may vary in the different poses or gestures. This leads to inaccurate results. In any situation, the value that never changes is the IRIS of a human. So IRIS is taken as a ROI for clustering the speaking faces.

IRIS RECOGNITION

Figure 1 shows the human eye [10]. The pupil is the darkest area in the human eye. The IRIS should be separated from the pupil. This can be calculated by using the following formula [9]:

$$C_1 = \arg\min (\Sigma_y I(x.y))$$

$$C_2 = \arg\min (\Sigma_x I(x.y))$$

In this I (x,y) is the intensity of the iris at that point [9]. LLE (Locally Linear Embedding), which is a high dimensional data embedding, can be used for comparing the values from the faces. An Earth Mover's Distance (EMD) based technique can be used to track the faces.

5. CONCLUSION

In this paper, the accurate speaking face is identified from the video by using IRIS as a unique factor. In existing techniques, lip movement of a person is used to identify the speaking person and the pixel values of those persons are compared to form a cluster. In order to avoid the inaccurate results IRIS can be used to retrieve the exact speaking faces and to cluster them. It can be used in the character based browsing. This can generate the accurate results when identical twins appear in the video. Clustering of speaking faces is difficult when a speaker wears spectacles. This work can be extended by identifying the hidden objects by removing the surface. Movie trailers can be generated by taking particular character as their query.

REFERENCES:

[1]. A. Umamakeswari and A. Rajaraman, "Object Based Video Analysis, Interpretation and Tracking", *Journal of Computer Science* 3(10): 2007, pp.818-822.

[2]. A. Umamakeswari and A. Rajaraman, "Interpretation and manipulation of video images", *BES Review*, Vol.13(2),April-June 2007, pp.24-28.

[3]. S.Satoh and T.Kanade, "Name-it: association of face and name in video", *Proc. of IEEE International Conference on Computer Vision and Pattern Recognition*, 1997, pp. 368-373.

[4]. Yi-Fan Zhang, Changsheng Xu, Hanqing Lu, Yeh-Min Huang, "Character Identification in Feature-Length Films Using Global Face-Name Matching", *IEEE Transactions on Multimedia*, Vol.11, November 2009.

[5]. Arasanathan Anjulan and Nishan Canagarajah, "A Unified Framework for the Object Retrieval and Mining", *IEEE Transactions on Circuits and Systems for Video technology*, Vol.19, January 2009.

[6]. Yasira Beevi C P and Dr.S.Natarajan, "An Efficient Video Segmentation algorithm with Real time adaptive threshold technique", *International Journal of Signal Processing*, December 2009.

[7]. J. Yang, R.Yan and A.G.Hauptmann, "Multiple instance learning for labeling faces in broadcasting news video", *Proc. of the 13th Annu. ACM International Conference in Multimedia*, 2005, pp 31-40.

[8]. W. Zhao, R. Chelappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Compu. Surv.*, Vol. 35, No. 4, 2003, pp. 399-458.

[9]. http://www.google.co.in/iris-recognition-Biometric-passport.jpg.

[10]. A.Basit, M.V.Javed, M.A.Anjum, "Efficient IRIS Identification method for human identification", *World of Science, Engineering and Technology*, 2005, pp.26-28.

PUBLICATION OF Little Lion Scientific R&D, Islamabad PAKISTAN
Journal of Theoretical and Applied Information Technology
30th June 2011. Vol. 28 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

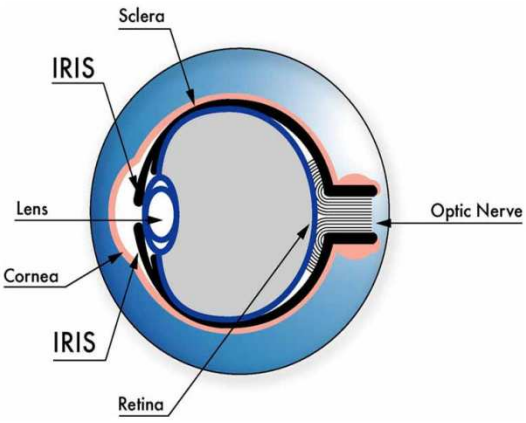

Figure 2. Human Eye