

An Automated Scheme For Surveillance Videos Footage And The System Searches



Sakhi Rehman

Department of Computer Systems Engineering,
 University of Engineering and technology (UET) Peshawar, Pakistan.

ABSTRACT

The research presented an automated scheme for surveillance videos footage and the system searches Person of Interest (POI) and retrieves video shots for specific POI.

An automated system has four stages; 1) to detect shot boundaries and divides original video into shots is called shot detection stage, 2) detection and recognition of face stage that detect faces in video footage using based on Haar-like features and using the Principal Component Analysis algorithm (PCA) for extraction of features in order to compute and select the eigenvectors with the major eigenvalues, 3) groups and apply face annotation with similar facial features into the same cluster, and 4) retrieve specific person footage from database user generate query by person image. The proposed scheme is user friendly and simple. It really overcome workload and enhances accuracy to localized person of interest in surveillance videos data. The average recognition accuracy on NRC-IIT [1] face video data set is 83.2 percent.

Key words: About four key words or phrases in alphabetical order, separated by commas.

1. INTRODUCTION

Huge quantity of video footage is being captured and stored. Though, lacking appropriate methods, such video data is becoming more and more problematic to manage[2]. Video summarization and video annotation (labeling of desired person) is vital for retrieval and syntactic understanding. Automatic video summarization and annotation is fetching more and more valuable for improving the accuracy and performance of multimedia information retrieval. However, manually the localization and annotating of Person of Interest in each frames is not possible with quickly increasing amount of video footage, since it not only a human burden and time consuming job, but as well as subject to human errors [3-5].

The scheme is presented a video synopsis and video annotation method for surveillance and other high quality video footage. Synoptic and annotation of surveillance videos

a very meaningful way to retrieve and analysis important information of a specific person. Although the whole process can also done manually, but the problem come when large amount of videos data available so this process can't be done manually in such a huge amount of data. The out of proposed system is a group of similar frames and annotated video clips containing specifics person who are targeted through the whole videos database.

The key contributions of this research are: a) montage or summarize specifics person information from large amount of video data, the scheme also annotating and clustering faces in fast way. b) Face detection and recognition on video clips and extracted important information regarding user generated query. c) The third main and key contributions are reducing human burden, less time consuming and less chances of error.

The structure of this paper as follow: Section II gives the description of previous literature study. Section III presented the overview of the proposed surveillance video summarization and annotation system and its phases. Section IV shows Experimental and obtained result. Finally Section V Discusses research conclusion and future recommendation.

2. PROPOSED AUTOMATED SYSTEM

The proposed scheme processes a surveillance and other videos as an input and finally outputs of a specifics person video footage or annotated all the frames in which the desired person detected and combine all the these frames in which the desired of interest successfully localized and make a mintage of video. The proposed method overview represented in figure 1.

The proposed scheme is consist of the following four stages:

- Shot-Detection Stage
- Face detection & recognition stage
- Video Faces summarization and annotation stage.
- Video montage or retrieval stage

2.1 Shot Detection

Dependency on object and camera movement by frame correlation to detecting the real boundary of shot . Every

single shot is defined as a flow or combination of frames captured by camera in a single space and time continuous action[6]. In the same shot the correlation between frames vital pointer to detect similarity between each frames and when there is visible change between frames we skipped these frames and search for the desired frames [7, 8].

2.2 Face Detection and Recognition

Detection and recognition of human face is very hot and active research area that focus on recognizing human faces within videos data or images. Actually recognition of face is compliment face detection. The process of finding a face in volume of images or video data is called face detection, whereas recognition of face is the method of matching or cross checking the detected face to one of many faces[9, 10].

Image integral

$$ii(x,y)= \sum_{x' \leq x, y' \leq y} i(x',y') \quad \text{Equation (1)}$$

Where original image pixel value is $i(x,y)$ and corresponding image integral value is $ii(x',y')$

Weak AdaBoost classifier

$$h(x,f,p,\theta)= \begin{cases} 1 & pf(x) < p\theta \\ 0 & otherwise \end{cases} \quad \text{Equation (2)}$$

Where polarity indication and direction of the inequality is P , threshold is θ and f denotes the feature value.

Strong Adaboost classifier

$$H(x)= \begin{cases} 1 & \sum_{t=1}^T \alpha_t h_t(x) \geq \gamma_t \\ 0 & otherwise \end{cases} \quad \text{Equation (3)}$$

The variable $\alpha_t = \log \frac{1}{\beta_t}$ and γ_t is selected for correct classification of training samples which are positive.

In this paper we used Principle Component Analysis (PCA) algorithm using Eigen Face Vector, for face recognition.

A set of Eigen faces calculated from the Eigen-vectors of entire scatter matrix St , as explained below:

Eigen Faces

$$St= \sum_{i=1}^n (xi-m)(xi-m)^T \quad \text{Equation (4)}$$

• Face Detection

In surveillance system videos, the key objects are people. For the purpose of identification we can identify interesting object via facial features detection to get better result and measurement about the activity objects. Viola and Jones[11] is used for face detection.

• Feature Extraction

In the proposed scheme, we used the predefined Principle Component Analysis (PCA)[12, 13] algorithm for extraction of features. One of the main advantages of using PCA algorithm is fast and only need less amount of memory. The concept of Principle Component Analysis feature extraction, based on information theory, look for computational model that completely defines a face by extracting the best related information contained in that face. The approaches of Eigen faces is a PCA based technique that is used to define the changes between face images. The objective is to find eigenvalues and eigenvectors, the eigenvalues which has the weight each vector, eigenvector of the covariance matrix of distribution. Hence we choice the most highly weighted eigenvectors and discard the lower to maximize the efficiency and minimize the detected face images training speed time.

• Face Grouping and Annotation

The automated face group or clustering is important part of the proposed scheme, which combines or group faces with similar facial feature into single video shots.

Old-fashioned tagging or tracking technique includes workload proportional to number of videos clips or images to be tracked or tagged. Thus the proposed scheme attempts to track or label groups of similar faces of Person of Interest in fewer operations.

• Up Sampling

The video frames resolution is originally 130 by 180 pixels resulting comparatively lower recognition accuracy. So we initially up sample the video resolution 360 by 480 pixels usinf Gaussian low pass filter and bicubic interpolation. This grant the proposed to detect and recognize human face far better accuracy.

• Retrieval Stage

In this stage we gather and combine the vital information which helps out to find desired person of interest through the video footage. After the face grouping and annotation stage we have a group or more for each entity in the video with addition annotation or tracker name. Finally the user generate query by input desired person of interest face picture and the proposed system find all the relative information to POI in the database. The system retrieves all the videos or frames which is similar to the POI from the database and make a montage of video clip of the desired POI.

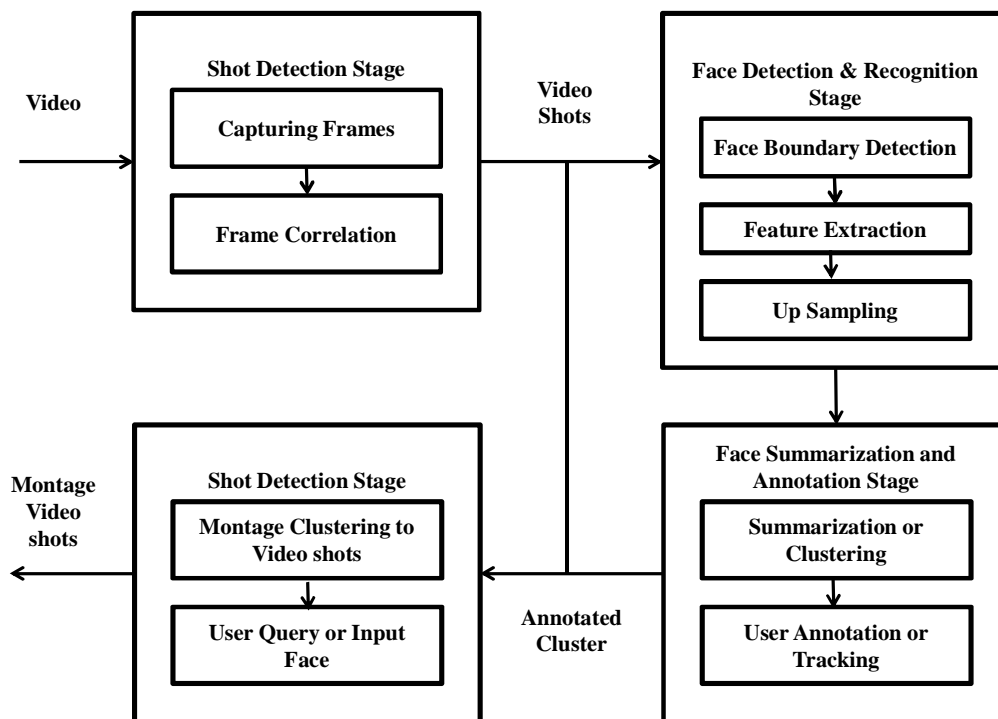


Figure 1: Proposed Automated System overview

3. EXPERIMENTAL RESULT

The proposed scheme was tested for 10 persons each person have two video clips one for training the system and the second video clip for testing purpose. The standard NRC-IIT [1] face database was used for testing the system performance and accuracy. The propose system result compare with the benchmark scheme [1]. Table 2, show some sample result and the total average accuracy is better than previous scheme [1].

The Figure 2 illustrate the accuracy of the proposed system with previous benchmark [1], through confusion matrix and ROI.

Comparison of the proposed scheme against other benchmark [1], we see in Table 2, that the proposed system average accuracy is 7.1 percent better than previous benchmark scheme.

The proposed system have also some limitation and condition. All the experiment was tested on high quality video footage give a good result for facial features detection and facial recognition. In Table 2 if we see the Video number 9, the recognition accuracy rate is less than the other database videos due to the light invariant and poses variation of the person.

The NRC-IIT standard face video data set contain number of videos footage, in Table 2 the video number 2 has the best detection and recognition rate due to very good quality video.

4. CONCLUSIONS AND FUTURE WORK

This research presents an automatic system for efficiently tracking or labeling face in surveillance videos through modest and easy to use interaction. The main and key point of this research work is video summarization or synoptic. The specific tracked or labeled face extracted and makes a montage of similar faces of desired POI. The system can also use for Home video to annotate and search records of home movies. The proposed method really reduces workload of person's identification or saving the time of human. The labeling or tracking detection and recognition accuracy far better than other benchmark as we tested the proposed system on Standard videos database as well as self-made video data set. The average accuracy of proposed and other benchmark on 10 videos is 83.2 and 76.1 respectively.

The recommendation and future, we would like to work on low video footages with pose variation to improve detecting, recognition and tracking or labeling accuracy.

Table 2. Comparison of Proposed Scheme accuracy against other benchmark Scheme [1]

Video ID	Proposed Scheme				Accuracy %
	True Positive	True Negative	False Positive	False Negative	
1	45	7	0	1	84.9
2	173	7	2	2	94.0
3	321	19	0	41	84.3
4	250	3	2	12	93.6
5	112	0	7	20	80.6
6	221	1	6	1	96.5
7	193	22	19	10	79.1
8	252	32	33	30	72.6
9	180	25	17	50	66.2
10	270	14	33	19	80.4
Total	2017	130	119	186	83.2

Gorodnichy [1]

True Positive	True Negative	False Positive	False Negative	Accuracy %
41	11	0	1	77.4
170	10	2	2	92.4
270	19	30	41	75.0
175	32	11	70	60.8
97	2	3	20	79.5
192	15	13	26	78.0
193	22	19	10	79.1
252	32	33	30	72.6
180	25	17	50	66.2
270	14	33	19	80.4
1840	182	161	269	76.1

REFERENCES

1. Gorodnichy, D.O., *Video-based Framework for Face Recognition in Video*, in *The 2nd Canadian Conference on Computer and Robot Vision (CRV'05)*. 2005. p. 330-338.
2. Wei Ren and S. Singh, *An Automated Video Annotation System*, in *Pattern Recognition and Image Analysis*. 2005: Springer-Verlag, Berlin, Heidelberg.
3. Yang SONG, et al., *Semi-Automatic Video Semantic Annotation Based on Active Learning*, in *International Conference on Visual Communications and Image Processing*. 2005: Beijing, China. .

4. Yan, R. and Hauptmann, *Multi-class Active Learning for Video Semantic Feature Extraction*, in *IEEE International Conference on Multimedia and Expo (ICME'04)*. 2004: Taipei, Taiwan.
5. Fangshi Wang, et al., *Automatic Annotation and Retrieval for Videos*, in *PSIVT*. 2006: Hsinchu, Taiwan.
6. Yu Song, Wenhong Wang, and F. Guo, *News Story Segmentation Based on Audio-visual Features Fusion*, in *4th International Conference on Computer Science and Education*. 2009.
7. Ekin, A., *Sports Video Processing for Description, Summarization, and Search*. Chapter 2 Structural and Semantic Video Modeling, 2003. **Ph.D. thesis**.
8. A. Ekin, A. M. Tekalp, and R. Mehrotra, *Automatic Soccer Video Analysis and Summarization*, in *IEEE Trans. on Image Processing*. 2003.<https://doi.org/10.1109/TIP.2003.812758>
9. [wiki.htm. Face Recognition OpenCV](http://www.FaceRecognition.com). 2018; Available from: <http://www.FaceRecognition.com>.
10. Zhao, X., *Using Genetic Algorithm for Classification in Face Recognition*, in *Advances in Swarm Intelligence 2010*. 2010: Springer Berlin, Heidelberg.https://doi.org/10.1007/978-3-642-13495-1_54
11. Viola Paul and M.J. Jones, *Robust Real-time Face Detection*. International journal of computer vision, 2004; p. 137-154.
12. Myoung Soo Park, Jin Hee Na, and Jin Young Choi, *PCA-based Feature Extraction Using Class Information Systems*, in *IEEE International Conference on Man and Cybernetics*. 2005.
13. Cooray, et al., *Facial Feature Extraction and Principal Component Analysis for Face Detection in Color Images*, in *International Conference on Image Analysis and Recognition*. 2004.https://doi.org/10.1007/978-3-540-30126-4_90