MPGI National Multi Conference 2012 (MPGINMC-2012) 7-8 April, 2012 **"Recent Trends in Computing"**
Proceedings published by International Journal of Computer Applications® (IJCA)ISSN: 0975 - 8887

A Novel Algorithm for Automatic Human Face Retrieval in Video

Z. G. Sheikh
Dept. of Computer Science
SGB Amravati University , Amravati

V.M. Thakare
Dept. of Computer Science
SGB Amravati University , Amravati

S. S. Sherekar
Dept. of Computer Science
SGB Amravati University , Amravati

ABSTRACT
Huge amount of video data is being generated every day, with enormous growth of security and surveillance system. It is immensely challengeable for researcher to search and retrieve accurate human face of interest from video with utmost speed. The proposed work is stimulated from the same concern. It would be the future demand for searching, browsing, and retrieving human face of interest from video database for several applications.

This paper proposes the novel algorithm for human face retrieval from video database based on holistic approach. The Viola and Jones frontal face detector detect the face region. The next stage is face extraction which have input for grouping individual faces. The individual group of faces has converted into single normalized mean face using PCA. The final face group contains single face for each person occurred in video. After the pre-processing of normalized faces, recognition is performed on the basis of query face image.

General Terms
Computer Vision, pattern recognition, video processing

Keywords
Face detection, face recognition, tracking

1. INTRODUCTION
Human face detection and recognition from video database is very intuitive to computer and human [1], still various challenges to computer for face detection described in [2] like pose, scale, illumination, expression etc. Work is in progress to overcome the problems in real time applications. The objective of this paper is to build the methodology for retrieve the human face from the video database using holistic approach, under the assumption that face is independent of the above problems i.e. explore the human face retrieval with (near) frontal face with some variation in face pose and facial expressions. Based on the proposed work, users can easily acquire the information that, interested human face image is available in video or not. If it is available, then retrieve the corresponding key frames from video with face model. This model is not only bring a new browsing and searching experience, but also provide an alternative of video summarization.

This paper presents a novel algorithm for human face retrieval from video. Videos have different categories like feature-length films, news videos, and surveillance and family videos. A human face image works as query by detecting and extracting features. The selected features are matched with key frames. The multiple face detection from video is worked with scene and key frame extraction methods for efficient detection.

This framework involves three steps for human face retrieval, in first step uses Viola- Jones detector for face image detection. Second step for conversion of crop face image into face group called face extraction, the faces of similar group are identified and the identical face image made using mean of faces. The face images goes for normalization and pre-processing. A third and last step is to recognize the query face image from video sequences using chamfer distance.

The paper is organized as follows. Section II described the related work, the proposed model appears in Section III, and Section IV concludes with future work.

2. RELATED WORK
Image retrieval based applications on visual content such as QBIC, Netra, VisualSeek, WebSeek, Virage, VideoQ, MARS are available for use. The limitations and challenges were discussed in [3] related with image/video searching and retrieving closely associated with CBIR. Google is also working on object matching in video [4] with text retrieval approach using SIFT descriptor for view invariant. Soft biometrics is applied [5] for facial marks identification on FERET database using Active Appearance model (AAM) for improving face matching and retrieval. Whereas, Josef Sivic *et al.*[6] takes efforts to find all occurrence of a particular person in shot with changes in scale, pose and partially occlude using Gaussian mixture modal in RGB colour space. New people detected in video stream [7] by Viola and Jones face detector and a kernel based regressor face tracking.

While the fusion of face and naming approaches were used for retrieval from videos. [8] Proposed a readily available texture source, the film script, which contain character name in front of their spoken lines. Yi-Fan Zhang et al [9], applied global matching between names and clustered face tracks with association network. Towards person Google [10], is the combination of face detection and speaker segmentation for multimodal person retrieval using statistical normalization PCA. Similarly, [11] shows the framework for retrieve faces in the TV show video frame sequence. In[12], we presents a frame for face retrieval using KLT tracker.

Whereas, O. Arandjelovic and A. Zisserman [13] are uses face image as a query to retrieve particular characters. Affine warping and illumination correcting were utilized to alleviate the effect of pose and illumination variations. Whereas, [14] is proposed kernel-based SVM for visual feature retrieving actors in films. To overcome the problems in content based image retrieval, Pablo Navarrete *et al.*[15] is projected interactive face retrieval system using self–organizing maps. An integration of statistical and structural information [16] for the local feature constructed from coefficient of quantized block transforms. DCT features were used in [17] for face image retrieval based on centered position of two eyes. Chon Fong Wong *et al.* [18] is employed Adaboost based face

detection and Lifting Wavelet Transform (LFWT) for feature extraction for in video sequences. The [19] utilized intelligent fast-forwards to jump video to the next scene containing that face, affine covariant region tracker for face region tracking.

The efforts which are more significant to propose approach such as, Mark Everingham and A. Zisserman uses combination of generative and discriminative head models for identifying individuals in video [20],[21] 3-D ellipsoid approximation, [22] coarse 3-D model with multiple texture for character identification in situation comedies or feature-length films.

3. PROPOSED METHODOLOGY

The problem is a combination of detection, tracking and recognition from image databases. However, the video has more generic and specific problems such as quality of video, frame speed, motion, color, etc. The aim of this method is to present an approach for human face retrieval from video databases. In other words, browsing or searching the specific character, person or face image of interest from available offline video databases.

Face detector from still images, real time face detectors are available under the controlled situation or databases. The uncontrolled condition involves background, scene change, lighting condition, size and quality of frames [2]. However, the face detection and processing on large scale video needs to improvement.

This model involves three stages (figure 1) for human face retrieval, in first stage uses Viola- Jones detector for face image detection. Second step for conversion of crop face image into face group called face extraction, the faces of similar group are identified and the identical face image made using mean of faces. The face images goes for normalization and pre-processing. A third and last step is to recognize the query face image from video sequences using chamfer distance.

3.1 Holistic Approach

The holistic approach generally refers to methods those use the entire face image for face identification or varification. Basically it includes methods like Principal Component Analysis (PCA) or Eigenfaces, Linear Discriminate Analysis (LDA) or Fisherface.

3.2 The Viola-Jones Face Detector

Three main components that make it possible to build a successful face detector [23] which can run in real time are namely, integral image, classifier learning with AdaBoost, and the attentional cascade structure.

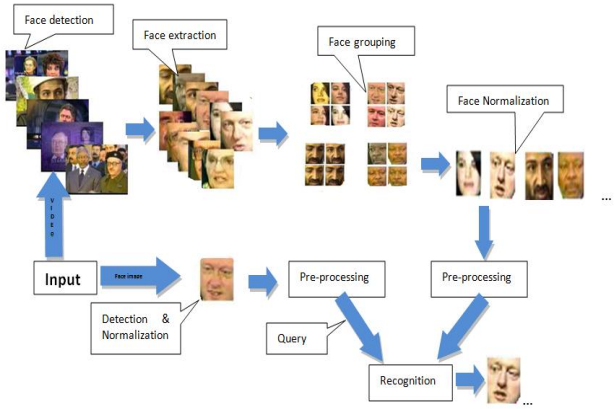

Figure 1. Structure of proposed human face retrieval

Figure 1 shows multistage architecture for human face retrieval from video database. It involves face detection using Viola and Jones frontal face detector using Adaboost implemented on video sequences. The next stage is face extraction which have input for grouping individual faces. The individual group of faces has converted into single normalized mean face using PCA. The final face group contains single face for each person occurred in video. After the pre-processing of normalized faces, recognition performed on the basis of query face image using chamfer distance [22].

Here, Viola and Jones face detector using Adaboost [23] has been implemented for large scale video databases. Which detect multiple nearly frontal faces from each video frames. The problem reading large scale video databases solved by managing buffer i.e. discarding the earlier frames and replaced new after fixed number of frames.

The detected faces has rectangle with different size in each frame. The faces were extracted from video frame and store in face group. The every detected faces were extracted and stored. The grouping of faces performed using simple matching of the closet and minimum variation in the detected face from two neighboring frames. If such face find in the next frame then it includes in face group of specific face, otherwise, make new face group. We treated that group as training data for future processing, the restrict size of faces in individual group not person face group. The individual group of faces has normalized using simple mean and average using PCA or SVM with fixed image size. As a result we get a mean face for each group.

For the next stage requires per-processing for on normalized faces. We can add a filter which discards the frame having negative outcomes and preformed better. For the speed up processing and minimized the computational cost, resized the face images into fixed sized 35 x 35 pixels.

The face recognition performed using [22] as input query normalized face image on normalized face group. If input face is recognized then display the result with tracking specific face in video and stop other operations except detection. Otherwise, performed till recognition or last frame, and discard the face group and create new if finding new face. The other recent face recognition techniques [25-26] can be also works with some changes.

4. ALGORITHM AND FLOW CHART

Input: face image, Video

Output: tracks of query face image in video

Step 1: face detected from each frame (f1,f2,…fn)

Step 2: face extraction for each frame

Count k;

If face detected=new or k<5 then extract face region and add to group z = {fi}

Else go to step 1

Step 3: face grouped from extracted faces by simple closest match between the bounding box positions of the face candidates in current and previous frames, where k=5, k++. If k=5 then go to step 4 else step 2

Step 4: normalized the face by average face from face track from each face group

Step5: average face image group convert in to gray scale image and resizing.

Step6: applying face recognition

If face recognized then go to step 1 and track the particular face region as a result of query.

Else discard the face image group until new face detected

If new face detected go to step 2.

Step 7: Stop: if terminate request or no.s of frames is empty in video.

Figure 2 . Algorithm for automatic human face retrieval

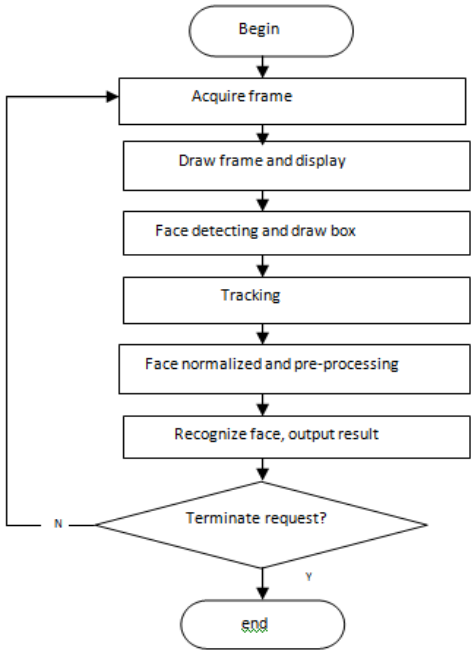

Figure 3. Flow chart for system design

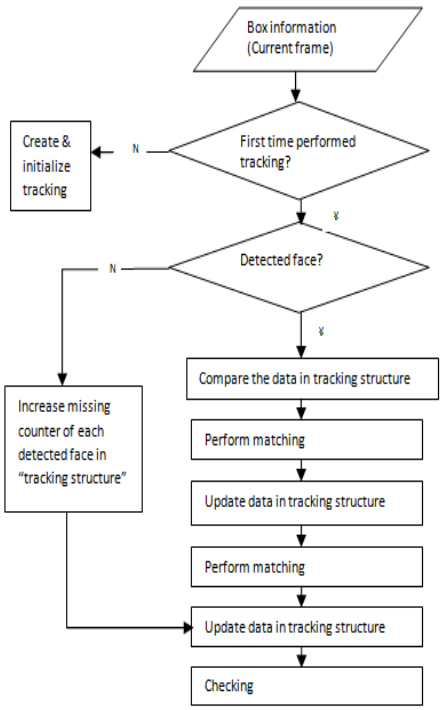

Figure 4. Flow chart for face tracking

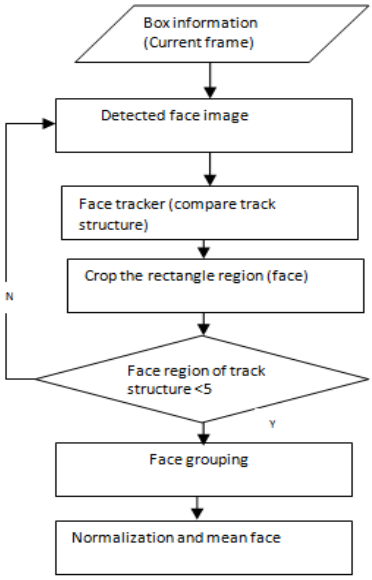

Figure 5. Flow chart for pre-processing

5. RESULT ANALYSIS AND DISCUSSION

The proposed model based on holistic approached with face detector using cascaded AdaBoost. In [20],the task is to detect the frames contains each character, and identify the image position and pose of the face correctly. Pose of the ground truth faces in the video covers poses of around +/-60 azimuth, +/-30 elevation and +/-45 in-plane rotation. Faces vary in

scale from 15 to 200 pixels. By using face detector in place of color segment would increases the rate of accuracy and speed.

Algorithm [21] tested on 1,500 key-frames taken one per second from the episode 'A Touch of class' of the BBC sitcom 'Fawlty Towers', detecting three main characters. The pose variation exceeding +/-30 about 3-D axis and correctly identify the character. The correct identification requires both detection and recognition of character, in contract to face detection or recognition. The Viola and Jones face detector [24] has been tested with 93% detection precision in video sequences. However, by applying the detector it would be achieve more proficiency than existing work. Whereas, [22] provides recall level of 50% the precision is around 80% for all characters in video sequence and by applying detector the speed of detection increases.

The part-based approach [12] which has been undergoes many stages and not considering the possibility of detection under the variation of pose and facial expression. The presented methodology works for invariant pose and facial expression with more precision rate.

6. CONCLUSION

The proposed algorithm works for holistic approach with the objective of human face retrieval from video sequences. The whole face is the input to the video sequences. The chamfer distance is match in every frame sequence of the video. The more accurate and prominent face detector and recognition approaches increases the rate of precision and speed of retrieval.

It is new experience to search and browsing the human face image of interest from video.

7. FUTUTE SCOPE

The algorithm may be improved by adding the scope for filter and enhancing techniques for the detected faces. E.g. reducing noise, changing format, discarding negative images, enhanced the detected face images and also the actual implementation needs to work on real life situation.

The implementation work has been considered the standard NIST and HONDA video databases.

8. REFERENCES

[1] Rama Chellappa, Pawan Sinha, and P. Jonathon Phillips, *Face recognition by computers and humans*, International Journals of IEEE Computer Society, 2010

[2] Ming-Hsuan Yang, David J. Kriegman and Narendra Ahuja, *Detecting Faces in Images: A Survey*, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 24, NO. 1, pp-36-58, JANUARY 2002

[3] Nida Aslam et al, *Limitation and challenges: image/video search & retrieval*, International workshop on Multimedia information retrieval , Netherlands , 2009

[4] Josef Sivic and Andrew Zisserman, *Video google: A Text Approach to object Matching in videos*, Proceedings of the British Machine Vision Conference, 2006

[5] Unsang Park and Anil K. Jain , *Face matching and retrieval using soft biometrics*, International journal IEEE Transaction on Information forensics and security(TIFS) , 2010

[6] Josef Sivic et al., *Finding people in repeated shots of the same scene*, proceeding of the British Machine Vision Conference, 2006

[7] Nicholas Apostoloff et al., *Who are you? – Real time person identification*, Proceedings of the British Machine Vision Conference, 2007

[8] Mark Everingham et al., *Hello! My name is… Buffy" –Automatic Naming of character in TV video*, Proceedings of the British Machine Vision Conference, 2006

[9] Yi-Fan Zhang, Changsheng Xu et al, *Character Identification in feature length films using global face –name matching*, IEEE transaction on multimedia, Vol. 11 No. 7, Nov 2009

[10] Lutz Goldmann , Amjad Samour and Thomas Sikora, *Towards Person Google: Multimodal Person Search and Retrieval* , Technical report, 2008

[11] Yina Han et al, Speaker retrieval for TV show video by associating audio speaker recognition result to visual faces, Proceedings of 2nd International conference the K-Space, 2008

[12] Zafar G. Sheikh, Dr. V. M. Thakare and Dr. S. S. Sherekar, *Towards Retrieval of Human Face from Video Database: A novel framework*, International Journal of Information Systems and Communication, ISSN: 0976-8742 , E-ISSN: 0976-8750, Vol. 3, Issue 1, 2012

[13] Ognjen Arandjelovie et al, *Automatic face recognition for film character retrieval in feature-length films*, In proceeding of IEEE Conference on Computer Vision and Pattern Recognition, San Diego , 2005

[14] Shuji Zhao et al, Actor *retrieval system based on kernels on bags of bags*, 16th European signal processing conference, (EURASIP), Switzerland , Aug 25-29, 2008

[15] Navarrete, P. and Ruiz-Del-Solar, *Interactive face retrieval using self-organizing maps* , Proceedings of the 2002 International Joint Conference on Neural Networks , IJCNN '02 , 687 - 691, 2002

[16] Daidi Zhong and Irek Defee , *Face retrieval based on robust local features and statistical structure learning approach* , EURASIP Journal on Advances in Signal Processing , Volume 2008, Article ID 631297, 12 pages

[17] Aamer S.S. Mohamed et al, *An efficient face image retrieval through DCT features*, In proc. of international conference, 2006

[18] Chon Fong Wong et al, Face image retrieval in video sequence using lifting wavelets transform feature extraction, Proceedings of the Ninth IEEE International Symposium, 2005

[19] Josef Sivic et al, Person *spotting: video shot retrieval for face sets*, International Journal of computer vision 2006 Springer Science + Business Media, 2006

[20] Mark Everingham et al., *Identifying individuals in video by combining generative and discriminative head models*, Proceedings of the International Conference on Computer Vision, 2005

[21] Mark Everingham et al., *Automated visual identification of character in situation comedies*,

Proceedings of the British Machine Vision Conference, 2004

[22] Mark Everingham and Andrew Zisserman, *Automated person identification in video*, Proceedings of the International Conference on Image and Video Retrieval, 2004

[23] P. Viola and M. Jones, *Robust Real-Time Face Detection*, International Journal of Computer Vision 57(2), pp-137–154, 2004

[24] P. Viola and M. Jones, Rapid object detection using boosted cascade of simple features, In Proc. of CVPR, pp-1-13, 2001

[25] Ran He, Wei-Shi Zhengand Bao-Gang Hu, *Maximum Correntropy Criterion for Robust Face Recognition*, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 33, NO. 8,PP. 1561-1576, AUGUST 2011

[26] Stefano Berretti, Alberto Del Bimbo, and Pietro Pala, *3D Face Recognition Using Isogeodesic Stripes*, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 32, NO. 12, PP. 2162-2177, DECEMBER 2010