ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X

International Journal of Advanced Research in Biology, Engineering, Science and Technology (IJARBEST)
Vol. 1, Special Issue 5, May 2013

Face Recognition From Real Time Videos-An Optimal Approach

BINU CHANDRAN R
Department of computer science
Mohandas College of Engineering
Anad, Nedumangad
binuadarsh@gmail.com

Prasanth.R.S
Department of computer science
Mohandas College of Engineering
Anad, Nedumangad
prasanthrs@outlook.com

Abstract—**Face label propagation tags people in news videos with their respective names. If unknown persons appear in the news video, we can eventually identify most of them by watching only the video. For this we will identify the faces appearing in news video using eigen values and Open C V Implementations that use HAAR detection method and tag each faces with their respective names through proper training. Face Label Propagation extracts face-name correspondences not only from pre-recorded news videos but also from live videos. Several Information sources available from news videos like transcripts, video captions can also be in cooperated optionally. To associate names and faces together, the proposed system takes full advantage of several advanced image processing techniques and natural language processing techniques like face sequence extraction and similarity evaluation from videos.**

Keywords- **eigen values, Open C V Implementation, HAAR Detection method, transcripts, face sequence extraction, similarity evaluation.**

I. INTRODUCTION

Today multimedia might be defined as the seamless digital integration of text, graphics animation, still images, audio and motion video in a way that provides individual users with high levels of control and interaction. To handle these large pool of information we must identify efficient ways to access its contents. Image processing and natural language processing play an important role in accessing these contents. Accessing the contents from a video is such a tremendous work compared to image or audio content extraction..A video will be containing sequence of frames captured over time and hence accessing the video content will be involving the process of frame grabbing.. The goal is to associate a face appearing in a video, whether live or recorded, with its respective name once trained.

In general visual identification refers to checking whether two images depicts the same object from a certain class and in visual identification face recognition is a more challenging problem due to the possible variations in the appearance. In this method we propose two methods for face detection over a given representation space. The first method uses Eigen values, a fast method, which helps to functionally operate on lots of faces in very little time Considering its imperfections, when faces are viewed at different levels of light or angles, we propose the second method of face detection, that is, Open CV Implementation that use HAAR Detection method which detects faces by passing the integral image through a number of classifiers.

II. RELATED WORKS

A semi automatic method is proposed in [10] to name face images in family albums. It starts with manually naming faces in some photos and then propagates the labeled names to other photos based on image similarity and facial similarity under certain temporal constraints. Though it offers a step beyond face recognition by including image similarity this work still deals with a finite identity matching problem.

The Name-It[1] system uses a similar technique to associate names and faces in news videos based on the co-occurrence between the detected faces and names extracted from the transcript. A face image is labeled with the name that has the largest temporal overlap with a group of images containing faces similar to the given one. The face images corresponding to a given name can be found in a similar way. Though theoretically sound, the robustness of this method is affected by the unreliable eigen face-based face similarity in the heterogeneous settings of news videos, as well as the temporal misalignments between names and faces. No serious performance evaluation has been reported on this work. A similar work [10] labels face images in online news articles with names extracted from news captions by exploring co-occurrence between the clustered faces and the names, which obtains good results. However, the fact that a person's face images in news videos are of much lower quality and homogeneity than those in textual news articles (where many are "standard" frontal faces) makes face clustering unlikely to be effective in our task.

Named Faces system [11] built a database of named faces that allows users to query a person's name by submitting an image containing the target face. The database is populated with the faces from news videos whose names are overlaid on video frames. These names are first recognized using video optical character recognition (video OCR), and the inaccurate recognition results are corrected using a dictionary of names by fuzzy match. Unfortunately, this

ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X

International Journal of Advanced Research in Biology, Engineering, Science and Technology (IJARBEST)

Vol. 1, Special Issue 5, May 2013

work does not label the large number of people whose names appear not in overlaid text but only in the closed-captions. Since these people's names are not visible from the video frame, labeling their names is of greater research importance, which is also the motivation of this paper.

III. LABEL PROPAGATION

In this paper name labeling can be applied for news videos which helps improving video content identification and search tasks. We are using a strategy of face similarity in identifying the anchor thus avoiding the ambiguity when two or more persons appear in the news for discussion. This system can also be used for live camera recording where persons appearing in front of the camera will be labeled with his/her respective name once trained.

In this paper a mixed strategy of eigen vector and Open C V Implementation using HAAR Detection method is used for face identification in videos. This mixed strategy avoids the ambiguity occurring in the case of multiple anchors. If a group of faces are similar, they may have the same name. The level of similarity between these faces will decide the confidence that they share the same name. Generally, the labeled faces can contribute to the naming of an unlabeled face by their names with the confidence estimated by their similarities with this unlabeled face. Moreover, the unlabeled faces can also affect the labels of each other by their similarity.

A. *Face spotting in videos*

The processing starts with the spotting or identification of faces in the given video. This can be done mainly through two methods. First method uses identification using Eigen vector and second using Open C V Implementation which uses HAAR Detection method.

1. *Eigen Vector Method*

   Eigen faces are a set of eigenvectors used in the computer vision problem of human face recognition. The approach of using eigen faces for recognition are used in face classification..These eigenvectors are derived from the covariance matrix of the probability distribution of the high-dimensional vector space of possible faces of human beings. Facial recognition was the source of motivation behind the creation of eigen faces. For this use, eigen faces have advantages over other techniques available, such as the system's speed and efficiency. Using eigen faces is very fast, and able to functionally operate on lots of faces in very little time. Unfortunately, this type of facial recognition does have a drawback to consider: trouble recognizing faces when they are viewed with different levels of light or angles. For the system to work well, the faces need to be seen from a frontal view under similar lighting. Due to these demerits we depend on one more face detection method.

2. *Open C V Implementation using HAAR Detection method*

   Haar wavelets are single wavelength square waves (one high interval and one low interval). In two dimensions, a square wave is a pair of adjacent rectangles - one light and one dark.

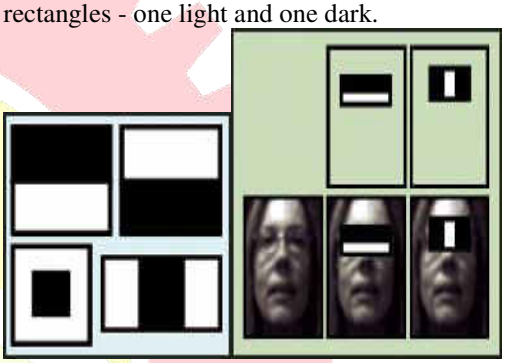

Figure 1 Examples of the Haar features used in Open CV

The actual rectangle combinations used for visual object detection are not true Haar wavelets. Instead, they contain rectangle combinations better suited to visual recognition tasks. Because of that difference, these features are called Haar features or Haar like features, rather than Haar wavelets. Figure 1 shows the features that OpenCV uses.

The presence of a Haar feature is determined by subtracting the average dark-region pixel value from the average light-region pixel value. If the difference is above a threshold (set during learning), that feature is said to be present.

To determine the presence or absence of hundreds of Haar features at every image location and at several scales efficiently, a technique called an Integral Image is used. In general,"integrating" means adding small units together. In this case, the small units are pixel values. The integral value for each pixel is the sum of all the pixels above it and to its left. Starting at the top left and traversing to the right and down, the entire image can be integrated with a few integer operations per pixel.

B. *Face Tracking*

Manually labeling a face appearing in the video each time is a tremendous time consuming task. These faces arise from just a few hundred"tracks" of a particular character each in a single shot. Applying a face tracking method reduces the amount of data to process and allows us to select the best

ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X

International Journal of Advanced Research in Biology, Engineering, Science and Technology (IJARBEST)

Vol. 1, Special Issue 5, May 2013

faces in each face track to help improve the quality of face comparison and classification. Each frame may contain faces using which is to be tracked using HAAR method using Open C V. If there is a face in a frame then we have to track the same face for the next frame using eigen vectors. If it doesn't happen then that information has to be notified. If the next frame contains the same face then it has to be tracked for label propagation. The movement of the detected faces is tracked in this module. Depending on the analysis of these movements we can choose the required name for the respective face.

C. *Anchor Detection*

This module has got more significance when we apply face labeling for news videos where anchors need to be tagged separately. Here we are a mixed strategy of scene similarity and face similarity for identifying anchors. When scene similarity is used , we will be identifying the anchors by their studio background. So any person appearing in the video with studio background will be tagged as the anchor. But ambiguity occurs when multiple anchors are there in the news video for discussion since all these anchors will be having studio background. To avoid this ambiguity we will be using face similarity too which not only checks the background scene but also adapting scene similarity technique but a mixed strategy of scene similarity and face similarity is used so as to avoid the ambiguity that occur when more than one anchor appear in the news video for discussion.

D.Face- Name Pair Selection

We first annotate a small number of name face seed pairs by randomly selecting several faces for each name mentioned. We select faces from the ground truth data. We describe techniques for retrieving face candidates by a given name. For this we need to find out the co occurrences of all face candidates with a given name. Then we will display the names or faces that correspond to the largest co occurrences. The temporal co occurrence similarity of the face track ft and

the name n can be defined as follows, where $(t_{ft}^{start}, t_{ft}^{end})$ be the start time and end time of the face track ft.

$$TS(ft,n) = \begin{cases} 1 & \text{if} \quad t_{ft}^{start} \le t_n \le t_{ft}^{end}, \\ \alpha & \text{otherwise} \end{cases}$$

$$\alpha = exp\left(\frac{((t_{ft}^{start} + t_{ft}^{end})/2 - t_{off} - t_n)^2}{((t_{ft}^{end} - t_{ft}^{start})/2)^2}\right)$$

D. Fcae Label Propagation

Along with labeled faces we need to name unlabeled faces also. When unlabeled faces come they will get matched with the labeled faces already stored. If any match occurs then these unlabeled faces will be given the respective name pair of the already labeled face. If no match is found it will go for a new labeling. Labeled faces contribute to the naming of an unlabeled face by their names with a confidence estimated by their similarities to this unlabeled face. Similarity co-occurrence is the key factor in labeling the unlabeled faces.

IV.CONCLUSION

We have implemented a face-naming method that learns from labeled and unlabeled examples rely on iterative label propagation in a graph of connected faces or name face pairs. A mixed strategy of scene similarity and HAAR Detection method is used here. We designed, implemented and evaluated a method for naming faces in video broadcasts that learns from labeled and unlabeled examples using a mixed strategy of face similarity and HAAR Detection method. Eigen metric values are used for comparing faces that uses the distribution of face similarities in a video broadcast. We used very few labeled data points (about 10-11 % of the detected faces). This yields better results than a Support Vector Machine classifier trained on the same labeled data.

V.FUTURE ENHANCEMENTS

In our future work we plan to focus on action recognition by jointly processing textual and visual data. that is, the actions performed by the person can be identified whether the person is talking, smiling, walking etc. For this we plan to in cooperate several advanced image processing techniques like covariance matching and a bag of 3D points.

REFERENCES

[1] S. Satoh and T. Kanade, "Name-It: Association of Face and Name in Video," in *Proc. of Computer Vision and Pattern Recognition*, IEEE Computer Society

[2] T. Zhang and F.J. Oles, "A Probability Analysis on the Value of Unlabeled Data for Classification Problems," Proc. 17th Int'l Conf.Machine Learning, Morgan Kaufmann, 2000, pp. 1191-1198.

[3]. X. Zhu and Z. Ghahramani, Learning from Labeled and Unlabeled Data with Label Propagation, tech. rep., Carnegie Mellon Univ., 2002.

[4]. T. Berg et al., "Names and Faces in the News,"

Proc. IEEE Conf. Computer Vision and Pattern Recognition,vol. 2, IEEE CS Press, 2004, pp. 848-8

[5] J. Yang and A. Waibel, "Tracking Human Faces in Real Time,"Tech. Rep. CMU-CS-95-210, School ofComputer Science, Carnegie Mellon University,Pittsburgh, 1995.

[6] M. Everingham, J. Sivic, and A. Zisserman, "Hello! My name is. . . Buffy—Automatic Naming of Characters In TV Video," Proc. British Machin VisionConferenceBritish Machine Vision Assoc., 2006.

[7]. M. Desmet, R. Fransens, and L. Van Gool, "AGeneralized EM Approach for 3D Model Based Face Recognition Under Occlusions," Proc. IEEE Conf. Computer Vision and Pattern Recognition,vol. 2, IEEE CS Press, 2006, pp. 1423-1430.

[8]. J. Shi and C. Tomasi, "Good Features to Track,"Proc. Conf. Computer Vision and Pattern Recognition,IEEE CS Press, 1994, pp. 593-600.

[9]. L. D'Anna et al., "A Multi-Stage Approach for Anchor Shot Detection," Structural, Syntactic, and Statistical Pattern Recognition, vol. 4109, Springer,2006, pp. 773_782.

[10]. J. Yang and A.G. Hauptmann, "Naming Every Individual in News Video Monologues," Proc. ACM Multimedia, ACM Press, 2004, pp. 580-587.

[11] Houghton, R. Named Faces: Putting Names to Faces. In *IEEE Intelligent Systems Magazine*, 14(5): 45-50, 1999.