International Journal of Computational Intelligence Research
ISSN 0973-1873 Volume 13, Number 6 (2017), pp. 1483-1492

http://www.ripublication.com

Detecting Face in Videos and Naming Celebrity in Videos

[1]Dr. Pramod Patil

Department of Computer Engg, DYPIET, PIMPRI, PUNE
Savitribai Phule Pune University, India

[2] Ms. Tejaswini B Patil

Department of Computer Engg, DYPIET, PIMPRI, PUNE
Savitribai Phule Pune University, India

Abstract

The Internet contain large amount of data of different things including text data, video information and images data. Huge amount of data is uploaded by user now a day due to which information or the data present over internet is growing day by day. Most of the time information uploaded by user is not related to label which they provide. Because of that it gets very difficult to get the correct information while searching for a particular data or video. Most of the time the search operation is related to label assign by user or by the index present over internet. Due to this loose labeling result provided by search operation is not related to information for which user looking for. Web video which shows up with the names is new and viral approach in numerous applications. The exclusive verification of celebrity face descriptions are provided which makes a person's name rundown of obvious detection of well-known people, get a prior understanding of images and contrast emphasize vector with each person's name associated with the image. The unmatched faces are discarded and the faces with the matched name are displayed. An identification of the images is perform on the basis of intra-demonstrate investigation, when a image of that persons face is matched then the recognition is complete with very high correctness in spite of big variations in

look, illumination, and expressions Datasets of well-known people have turn out to be accessible, an attempt to recognize celebrities in the news has also occurred.

Index Terms— Mask-name, feature vectors, face recognition, face image instance, celebrity face images, non-face images, intra-model, biometric models, celebrity videos.

1. INTRODUCTION

The process of tagging individuality i.e., person's name or topic on individual photos is known as face footnote or person's name classification [3]. From the recent development in internet large amount of digital data is available through news regarding data media issue online and user sided content provided online like YouTube, Face book, Twitter etc [2]. As fast growth in film industries large amount of data is present over internet. Because of large variation in faces expression it becomes difficult to identify a celebrity in web video. Most of the name provided to the video is user uploaded name which may relate or may not relate to that video. Due to all this it gets very difficult to search people related video over a large data present on internet.

Relating the name of celeb is done by relating the face with the given name. Such type of problems are occurs in field of information videos [1], [2], cinema [3] and television sequence [4]. Concentration on the wealthy place of time coded data which contain speech record and subtitles. As a consequence, directly expand these loom towards Web video area is not easy. Deployment of wealthy context data for mask identification is also deliberate in the area of individual album compilation [5]–[8], by means of timestamps, geo tags, individual information lists and public networks. However, such conditions cannot be straightly applied for field unrestricted videos, since the lack of context cues and earlier knowledge such as people relationships for complexity formulation.

To eliminate this kind of problems of face naming of celebrities over web video a method based on conditional random field (CRF) [9], [10] is proposed. This technique is based on three main approaches, out of which first is the face to name approach. In this mode name related to face is given on the basis of prior knowledge from image domain. Second method is the face to face in which name is assign on the basis of background context, visual similarity, temporal does not connectivity, spatial overlapping etc. and the third mode is the Name to name relationship, which consider dual facade of famous person by leveraging common network build based on the co-occurrence information between superstars. From above mentioned three types the first two are used to label the name to the celebrity on web video. This paper contain two major tasks give a video to v1 “within-video” tagging which builds a chart with

the names and countenances in the film as vertices. Based upon the face-to-face and face-to-name associations, boundaries are recognized between the vertices for deduction of face labels using CRF.

2. LITERATURE SURVEY

Table 1: Literature Survey

Sr. No.	Reference Paper Name	Algorithms Used	Disadvantage	Advantage
1	Collaborative face recognition for improved face annotation in personal photo collections shared on online social networks	Eigenfeature regularization and extraction (ERE), marginal Fisher analysis (MFA), kernel direct discriminate analysis (KDDA), Bayesian , Fisher linear discriminate analysis (FLDA), and PCA. Viola-Jones face detection algorithm	1. Individuals are only part of the constructed SGM when they were manually tagged in the personal photo collections. 2. This approach may not be optimal when manually tagged photos do not clearly represent the social relations between different members of the OSN.	1. Improve the accuracy of face annotation. 2. Low Computational cost. 3. Suitable for decentralized system. 4. Low attention on OSN as suitable FR engine is belonging to OSN group.
2	Mining weakly labeled web facial images for search-based face annotation.	Unsupervised face alignment technique, GIST algorithm, locality sensitive hashing (LSH),	1. Not so good to label face on web for Lesley known person. 2. Duplicate name can be a practical issue in real-life scenarios.	1. This technique enhances the label quality and proposed URL technique. 2. Achieve promising result under verity of settings.
3	A conditional random field approach for face identification in broadcast	CRF Conditional Random Field, OCR technique and Named Entity Recognition technique, Automatic Speech	1. Not able to recognize face on open set data. 2. Use of automatic speech reorganization technique is correct always to label the	1. Improves the face labeling by using audio modality. 2. Uses joint CRF clustering of face track and speaker

	news using overlaid text.	Recognition	name in video.	segment.
4	Retrieval-based face annotation by weak label regularized local coordinate coding.	Content-based image retrieval techniques, Search-Based General image annotation techniques, Unsupervised Face Alignment technique, Locality-Sensitive Hashing technique.	1. The experiment result based on top n result. Hence if number is reduced it gives poor performance. 2. When extra noise labels are added to the retrieve images the annotation performance is decreases.	1. Maintains comparable results. 2. Better annotation performance validates effectiveness of proposed algorithm for retrieval based face annotation. 3. For small value of t gives significant performance.

3. BACKDROP

Different type of methods and systems are proposed on which most of the methods are related to Web images and video such as news, series and movies video. Many face detection and labeling methods are proposed for multi face detection and assigning name to them. [1] Proposed a improved graph dependent method in such a methods a point associated to faces and corners corresponds highly matched faces for labeling name in video related to people. For the naming purpose they first assign a name to the detected caption and then searching for the text set on the graphical data set to assign name to that caption.

Many methods are projected to handles the situation related to unverified face-name relationship. Such methods are not able to handles large scale real time data that is they are suffers from computational expensive problem. Also these types of methods are hard to assign null decisions. To overcome from all this problem [8] proposed a novel structure called FACD (stands for Face-name Association via Commute Distance). For all this purpose they firstly used CD based approach and after that they used anchor-based commute distance (ACD) method to rapid up the process.

To solve the problem related to mask name relationship and broadcast a new loom which associates the effect of two Conditional Random Field (CRF) models are proposed. First is the grouping of manifold cues together with as major assistance the use of recognition sources and recurring local face visual background (LFB) singing the role of a named less characteristic. A second CRF is used for combined recognition of the person group to get better recognition presentation. In this faces are

identified by blinking among a grouping step and a identification step of those group. All steps are given by a committed CRF [7].

To get over from the current index framing technique based on the complex distance measure between two face set a new method is proposed to encode the bag of faces into a single sparse representation. In this technique every sack consists of bare symbol, using the non-zero entries in the bare symbol as distinct code words. A upturned index is generated with millions of faces extracted from videos and can allow scalable recovery over large-scale database [9].

4. ALGORITHM

Contained by capture face labeling from video

Input: The sets of faces and names in a video

Output: Frames or Images get from video with labeling

1. Initialize frame rate=0
2. Dataset contain labels and images
3. Take a video as a input
4. for (File file : listOfFiles)
5. if (Img!= null)
6. get the frames with label
7. if (getImage==DatasetImage)
8. faceMatch()
9. while (image. next())
10. else facial image not found
11. return facial Image

5. MATHEMATICAL MODULE

S={ U={u1,u2,…un }, V , Fv , Img , FD , D , URL , FR , $M_{I,}$ $Top_{Img,}$ CFN}

Where,

S=System.

U= {u1, u2….un} =Set of Users.

V=Video. Fv=Frames or Images get from video.

FR=Face Recgnation.

FD=Face Detection.

FD ∈ Fv---------- (1)

D=Dataset.

D= {URL, Images….etc} ---------- (2)

$$f(\mathrm{Top_{Img}}) = \mathrm{Sim}(\mathrm{D_{I,}} \in \mathrm{Fv}) - - - - - - - - - - - (3)$$

D_I=Dataset Images. M_I=$\mathrm{Sim_{Results}} \int_0^n (\mathrm{Top_I}) - - - - - - - - - - (4)$

$M_I = \mathrm{Measurity\ Of\ Images.}$

$\mathrm{Top_I}$=Top Images.

$\mathrm{Sim_{Results}} = \mathrm{Similaer\ Results.}$

CFN=$M_I \in D \,||\, FD \in Fv$ ---------- (5)

CFN=Celebrity Face Name.

6. PROJECT APPROCH

Put forward an efficient unsupervised label refinement (ULR) loom for cleansing the tags of web facial metaphors using machine wisdom procedure. Give a Discriminative Random Fields (DRF), structure for the categorization of usual image regions by including area spatial craving in the tags as well as the experiential data. The projected method gives the restricted discriminative model and permits smoothing the consideration of experiential data provide the tags generally used in Markov Random Field (MRF) structure. The constraints used by DRF models are well-read using penalized utmost pseudo-likelihood technique. In addition an image processing algorithm is provided with the MAP interface for binary classification allow by the DRF model. The presentation of the replica is confirmed on the basis of artificial as well as the real-world images. Along with this also propose another approach for contextual feature for tagging images, in which every pixel is assign with one of the restricted set of tags. The facial appearance is associated on the basis of probabilistic structure which unites the output of numerous mechanisms. Several concentrated on the image-label chart, while some concentrated exclusively on patterns inside the tagged field. Mechanism also varies in level as some are based on ll declaration pattern at the same time other on courser and more worldwide structure. An advance side of contrastive divergence algorithm is used to study from tagged image records. User applied method on two actual dataset and compare it to a differentiator and a Markov random field.

A Dynamic Conditional Random Field representation is used to unite the related information for object fragmentation in picture series. Spatial and sequential reliance is united inside segmentation method are unfired using a active probabilistic structure depends on the conditional random field (CRF). A well-organized estimated filtering method is resulting for the DCRF model to repetitively approximate the fragmentation from the past of video frames. The fragmentation technique deployed both intensity and movement cues, and it unites active data and spatial interface of the experiential information.

SYSTEM ARCHITECTURE

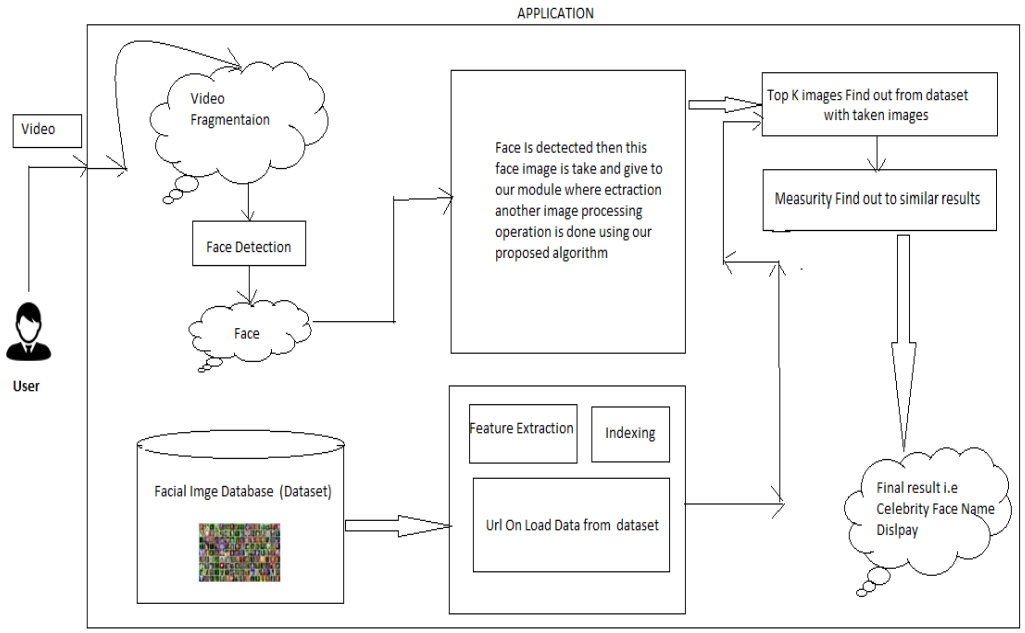

Fig 1. System Architecture

The above outline demonstrate the system architecture of the project. Scheme conisist of video input, different data set of images with there label of celebraty face name and final is the result of detected face with correct label. The folw of the system is as follow, first a video has to upload by the user contaning celebraty. Then this video is divided into number of frames that is the fragmentation process. After the fragmentation process the frames are given for the face detection. A face detection algorithm is aaplied to detect wether the frame contan face or not. The detected face is then match with given data set. If the face is matches with the images present in the data set then a lablel is assign to that face and display as face detected with name label at output.

The above process is explain in mathematicla term as according to some constants as well as some function:

UP_{video} , D, FD(x), LB(x), F, L.

$F = split(UP_{video}) \in F$ ----------(1)

Where, F = Frame

UP_{video} = Uploaded Video.

D = Dataset.

L = Labels.

$FD(x) = \int_0^n F + FD_{VG}$ ---------- (2)

FD=Face Detection frames viola gones.

$LB = \sum_{k=0}^{n} FD(x) + D + L$ ---------- (3)

Where FD(x) is a face detection function and LB(x) is used to assign the labeling into frame.

A. DATAFILE ASSESMENT

The dataset comprise of picture URLs for 202792 appearances. The marks of the appearances are consequently produced by the calculation, with high exactness [14]. To encourage downloading the pictures, which give various URLs to the close copies of every face? Likewise, the pictures and facial components (LBP) are additionally accommodated representation and benchmarking purposes. Because of copyright reasons, Creator doesn't give the first web pictures.

Creator has used different real world data set to get the face label for the corresponding face in the given video stream. First user has to choose the dataset file from the given data set and then upload the video. From the video stream first it segmented the video in number of frame and then detects the faces in given video set. After the face is detected the corresponding face is matches with the given dataset which contain number of images and if match found then it assign the label related to that image.

B. OUTCOME ASSESMENT

Below figures (b) shows the result of the proposed system. Provide a video as input to the system. This video is fragmented i.e. divided into number of images depending on the fragmentation time provided to system. Fig (a) shows one of the frame which give after the fragmentation is applied on the video frame.

Fig (a)

After getting the number of frames from the fragmentation process a Viola–Jones algorithm is applied on each frame to detect the face present in the images. Fig (b) shows celebrity face detected in frame with square box on the face.

Fig (b)

In the third stage the face detected frame with detect face is compare with the different faces present in the dataset. A dataset is consisting of number of different celebrity faces with their name label. If the face is get match with the provided data set the on the top matched result label is assign to that face. Fig (c) shows frame with face label frame.

Fig (c)

7. CONCLUSION:

This approach have developed and studied face labeling algorithm for celebrity features labeling in network video. The initial stem is based on the chronological alignment of the evidence of fans; sub-text and video structure specified labels are limply credited to the matching body in the clip. Related to face labeling of celebrity in video the proposed approach is based on CRF model can smoothly encodes F2F and F2N relationships and also permitting null category by considering uncertainty labeling. The proposed scheme results a good effect than the previous face labeling method.

REFERENCES:

[1] M. Guillaumin, T. Mensink, J. Verbeek, and C. Schmid, “Automatic face naming with caption-based supervision,” in Proc. IEEE Comput.Vis. Pattern Recog. Jun. 2008, pp. 1–8.

[2] Y. F. Zhang, C. S. Xu, H. Q. Lu, and Y. M. Huang, “Character identification in feature-length films using global face-name matching,” IEEE Trans. Multimedia, vol. 11, no. 7, pp. 1276–1288, Nov. 2009.

[3] J. Choi, W. De Neve, K. N. Plataniotis, and Y. M. Ro, “Collaborative face recognition for improved face annotation in personal photo collections shared

on online social networks," IEEE Trans. Multimedia, vol. 13, no. 1, pp. 14–28, Feb. 2011.

[4] D. Y. Wang, S. Hoi, Y. He, and J. K. Zhu, "Mining weakly labeled web facial images for search-based face annotation," IEEE Trans. Knowl. Data Eng., vol. 26, no. 1, pp. 166–179, Jan. 2014.

[5] J. Yang and A. G. Hauptmann, "Naming every individual in news video monologues," in Proc. ACM Int. Conf.Multimedia, 2004, pp. 580–587.

[6] Z. Chen, C.-W. Ngo, W. Zhang, J. Cao, and Y.-G. Jiang, "Name-face association in web videos: A large-scale dataset, baselines, and open issues," J. Comput. Sci. Technol., vol. 29, no. 5, pp. 785–798, 2014.

[7] G. Paul, K. Elie, M. Sylvain, O. Marc, and D. Paul, "A conditional random field approach for face identification in broadcast news using overlaid text," in Proc. IEEE Int. Conf. Image Process., Oct. 2014, pp. 318–322.

[8] J. Bu et al., "Unsupervised face-name association via commute distance," in Proc. ACM Int. Conf. Multimedia, 2012, pp. 219–228.

[9] B. C. Chen et al., "Scalable face track retrieval in video archives using bag-of-faces sparse representation," IEEE Trans. Circuits Syst. Video Technol., submitted for publication.

[10] D. Y. Wang, S. C. Hoi, Y. He, J. K. Zhu, T. Mei, and J. B. Luo, "Retrieval-based face annotation by weak label regularized local coordinate coding," IEEE Trans. Pattern Anal. Mach. Intell., vol. 36, no. 3, pp. 550–563, Mar. 2014.

[11] K. Barnard, P. Duygulu, D. Forsyth, N. de Freitas, D. Blei, and M. Jordan. Matching words and pictures. Journal of Machine Learning Research, 3:1107–1135, 2003.

[12] T. Berg, A. Berg, J. Edwards, M. Maire, R. White, Y. Teh, E. Miller, and D. Foryth, "Names and faces in the news," in Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, 2004, vol. 2, pp. 848–854.

[13] Z. Liu and Y.Wang, "Major cast detection in video using both speaker and face information," IEEE Trans. Multimedia, vol. 9, no. 1, pp. 89–101, 2007.

[14] A. Gursel and S. Sen, "Improving search in social networks by agent based mining," in Proc. Int. Joint Conf. Artificial Intelligence, 2009.