International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 04 Issue: 06 | June -2017 www.irjet.net p-ISSN: 2395-0072

Face Annotation Using Co-Relation Based Matching For Improving Image Mining in Videos

Vijay W. Pakhale[1], Mayank Bhargav[2]

[1]Student, Patel College of Science and Technology, RGTU, Bhopal
[2]Assistant Professor, Patel College of Science and Technology, RGTU, Bhopal

---

Abstract - In today's world, Sharing of information online between friends have boosted the requirement of better Image retrieval techniques. In online social networking, sharing of photos and videos in the form of information has increased the need of Face Annotation. The face annotation means something particular about the Facial features. Face Annotation is a bit complex in the videos. Since face annotation provides the special or the to the point information about the targeted part of face or the facial feature, it got the special attention of the programmers.

In this paper, the problems of the face recognition techniques have been detected and solved them. This is done using the features correlation based matching techniques. We observed that the proposed system is quite effective compared to the others.

Key Words*: *Image Retrieval, Content-Based Image Retrieval (CBIR), Unsupervised Label Refinement (ULR), Search Based Face Annotation (SBFA), Viola Jones Algorithm.

1.INTRODUCTION

From the popularity of cameras and social media tools, images and videos gain much more interest of the users. We can capture videos and images at any time with digital camera and post it onto the internet using some social networking systems. Also these images contain some meta-information like GPS co-ordinates, date-time or quality of image, which are may be automatically put by our digital camera [1]. From these images, a large number of images are human facial images on internet/social networks. Many people like to tag their facial images by their names or some text. But many people does not tag images properly which leads to inaccurate information about that face. These images are weakly labeled and their meta-information is known as weakly labeled data [4]. Numbers of techniques are used in networking for refining the weakly labeled data, for getting correct information.

1.1 Biometric Systems:

A person can be generally identified by his face in social environment. In the same way in machine environment a person is identified by the Biometric features. Every person has some unique biometric features and these features are detected by the biometric system [2][11]. These features are compared with other person's biometric features to decide the uniqueness of the first person. In same way, the face recognition systems perform their function to detect the faces among other faces. The advantages of these features are uniqueness and can't be forgotten. The examples of the biometric functions are fingerprints, eye retinal features etc. facial features can also be stated as the biometrics.

1.2 Unified Label Refinement:

For weakly labeled facial images collected from internet, an ULR is an effective method to fix the errors in annotations. It can effectively correct images those are incomplete and full of noise. By using ULR in algorithms we can improve scalability, efficiency and performance of face detection systems [4][13].

But ULR also has some limitations according to face annotation. First, it consumes time and collecting a huge amount of human-named training facial images is expensive. Next, it is normally hard to generalize the models when new training data or new persons are added, in which an intensive retraining process is usually required [2][6]. Frequently the annotation/recognition performance will be poor if the number of persons/classes is very huge.

1.3 Content-Based Image Retrieval:

Textual information about images can be easily searched using existing technology, but this requires humans to manually describe each image in the database. Initial CBIR systems were developed to search databases based on image color, texture, and shape properties. An image is retrieved in CBIR system by adopting several techniques simultaneously such as Integrating Pixel Cluster Indexing, histogram intersection and discrete wavelet transform methods [3]. It can compare content using image distance measures and by some query techniques like semantic retrieval, relevance feedback, iterative machine learning, etc.

There are some limitations according to CBIR techniques. The learning of image similarity, the interaction with users, the essentiality for databases, the problem of evaluation, the semantic gap with image features, and the understanding of images are degrade this technique [7][8].

2. LITERATURE SURVEY

In this section we discuss about previous surveys which are related to this paper. We found many previous papers and by examining those we describe our method and how it is better than those others. Some of these are given below:

A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "*Content-Based Image Retrieval at the End of the Early Years*," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 12,pp. 1349-1380, Dec. 2000:
It uses image retrieval systems for image processing and retrieves texture, colour and local geometry of the images [3]. It considers some features of image retrieval, like salient points, shape and object features, their structural combinations and global and accumulative features. For each of that features it can review the similarity of pictures and objects presents in that images. So that, it provides interaction by considering close connection of types and feedbacks means for those images. Also author presents field's driving force, computer vision heritage and influence of computer vision for concluding their system.

T.L. Berg, A.C. Berg, J. Edwards, M. Maire, R. White, Y.W. Teh, E.G. Learned-Miller, and D.A. Forsyth, "*Names and Faces in the News*," Proc. IEEE CS Conf. Computer Vision and Pattern Recognition (CVPR), pp. 848-854, 2004:
By applying appropriate discriminate coordinates, we can cluster these face images with associated set of names, breaks the ambiguities between labeling and identify faces which are incorrectly labeled, from set of images [5]. After that it use merging procedure to identify variants of name applied to same individual from news pictures. It can also used as process that removes and correct noisy unsupervised data from any network.

X.-J. Wang, L. Zhang, F. Jing, and W.-Y. Ma, "*Anno Search: Image Auto-Annotation by Search*," Proc. IEEE CS Conf. Computer Vision and Pattern Recognition (CVPR), pp. 1483-1490, 2006:
In their proposed system at least one keyword is required which is accurate for enabling text-based search, then to retrieve visually similar images by content-based search. After that, from titles and surrounding texts and URLs of these images, annotations are mined. It ensures high efficiency with visual features which are mapped by some hash codes [8]. The search process is speed-up by these hash codes. They choose real web images for experimental results.

Z. Cao, Q. Yin, X. Tang, and J. Sun, "*Face Recognition with Learning-Based Descriptor*" IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 2707-2714, 2010: They present Unsupervised Learning techniques as encoder for invariance of training images. For getting compact face descriptor they apply PCA technique and simple normalization. The result of that technique produces high discriminative and easy to extract learning-based descriptor [13]. They also propose pose-adaptive matching methods for handling large pose variations of matching face pair by using pose-specific classifiers for some pose combinations. They also achieved 84.45% recognition rate on the different datasets like Labelled Face in Wild Benchmark.

2.1 Limitations of Existing Systems:

We have many existing systems which resemble our work, but all these existing systems have more or less limitations in their fields. These limitations cause poor result or inconveniency as we get output from any existing application which uses these existing concepts. Some of these limitations from existing systems are considered here:

2.2 Content Based Image Retrieval:

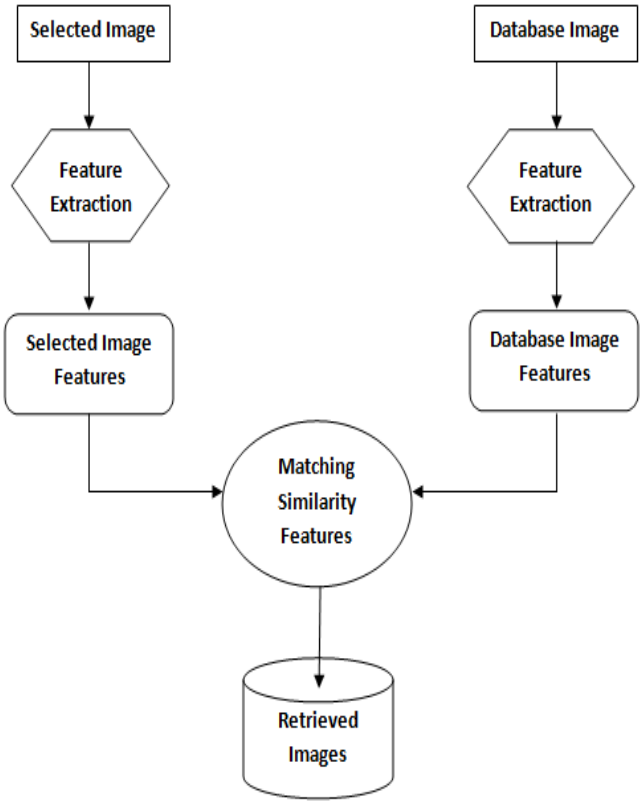

Fig. 1: Algorithm for Content-Based Image Retrieval Systems

i. In this technique, the tagging of the image is done efficiently. But the metadata is to be entered manually for the creation of database and tagging [11].
ii. The Systems using this technique is not efficiently good enough to have a better interaction with users.
iii. The complexity of statistical evaluation gets increased [10].
iv. The manual tagging of the image features creates the semantic errors.
v. The accuracy of the system is compromised

2.3 Search-Based Face Annotation:

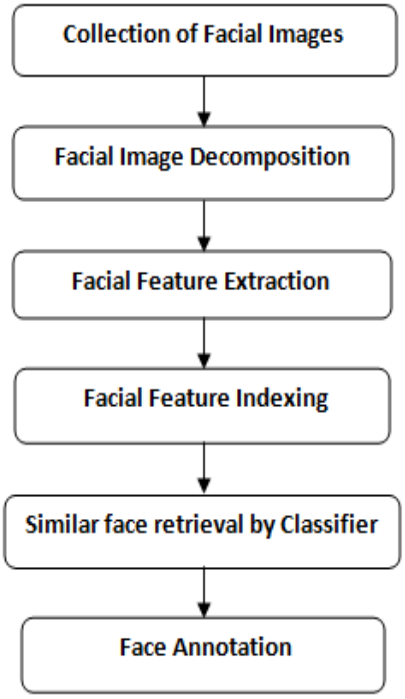

Fig. 2: Algorithm for Search-Based Face Annotation Systems

i. In this technique, mostly the annotation is performed on the images which are Famous like celebrities. So it's generalization of the models is difficult when new data or new persons are considered [9].
ii. Due to the long process of matching, the technique consumes the time
iii. It is an expensive task to collect the huge amount of face features data primarily required here.
iv. Classifier defined classes have the inverse effect on the performance of the systems. The number of classes increases, the performance decreases [12].

3. PROPOSED SOLUTION

Here we define our algorithm for face annotation which is accurate and less time consuming with respect to other systems. In our proposed solution first we get our system raw data from online sources in the form of videos then we divide these videos in several images. After that by running our algorithm we retrieve those images which are best images for providing the resource data to the system accurately. After selecting images we perform feature extraction operations for collecting the features of the faces which are present in those images. These features can be compared with other facial features present in our database. These matched facial features can further define the exact person present in a particular image. After that our system can automatically tagged that images with those data through which another system can also gets the accurate information.

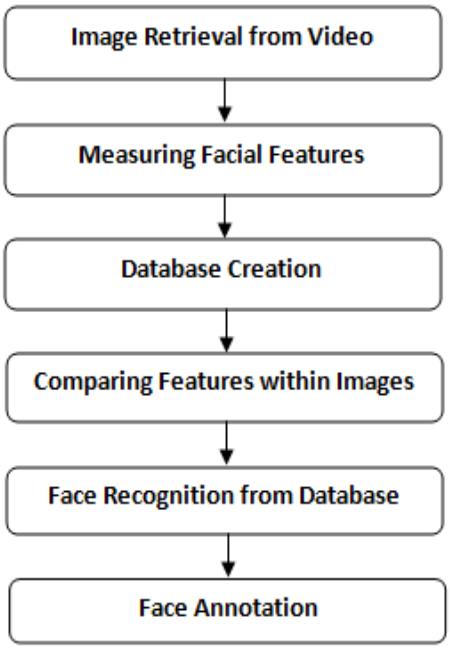

Fig. 3: Proposed Algorithm for Face Annotation in Video

Image Retrieval from Video: In our proposed system, the videos which are uploaded by the users are observed. The images are retrieved from them using frames separation. The images which are not proper are rejected. The proper ones are considered for further process.
Measuring Facial Features: The next step is measurement of the facial features of the proper images chosen. In this proposed solution we measure several different facial features of the image. Like,

i. Width of an eye
ii. Height of an eye
iii. Width of nose
iv. Height of nose
v. Width of mouth
vi. Height of mouth
vii. Distance between two eyes
viii. Distance between left eye to nose
ix. Distance between right eye to nose
x. Distance between left eye to mouth
xi. Distance between right eye to mouth
xii. Distance between nose to mouth

To measure these features, we have to consider only the face region of every frame of the image. For this very purpose, grey scaling of the image is performed which helps in the detection of these features.

1. *Database Creation:* Every image which has undergone the measurement of features is stored in the database. This database is specially formed to assign the feature statistics for the particular image. The Database is also accessed while comparison process. The statistical information stored in database is reused in face recognition process.
2. *Comparing Features within Images:* By observing the images we get the individual facial features. The features of the images stored in the database are compared with these observed facial features. The relative statistics are considered for matching or the non-matching of the images. For example, the distance between the eyes of one image is compared with the distance between the eyes of the other image.
3. *Face Recognition from Database:* In the face recognition step, the comparison is checked within the system for the maximum matching of the respective features. Through this, the system knows the differences of the face present in the images. When the observed matching ratio is high, the possibility of those two faces to be similar gets higher. In our proposed method for face recognition the accuracy of the matching ratio is higher than the other algorithms.
4. *Face Annotation:* This above mentioned whole process is the backbone for the face annotation. The statistical information we get through this entire process is the note or the explanation added to the image. This fulfils the necessary conditions of face annotation.

4. RESULT ANALYSIS

To analyze the results of the proposed technique, the process is to be carried out is two main parts. These two parts are divided on the basis of functionality performed in respective parts of the system. The second part uses the information which is generated by first part.

To create the database, the image or video is input in the system. If the input is video, the proper usable images are retrieved from the video by dividing it in the frames of images. Face detection process is carried out on the image using Viola-Jones algorithm. The grey scale of the detected face also called the encoded face is obtained. This grey scale image helps in providing the proper locations of the facial features like eyes, nose etc. Once the facial features are extracted from the image, the computational analysis of the features is begun. The twelve parameters mentioned above are measured. Identification is provided to the frames or the image. The measurements tagged with image are stored in the database.

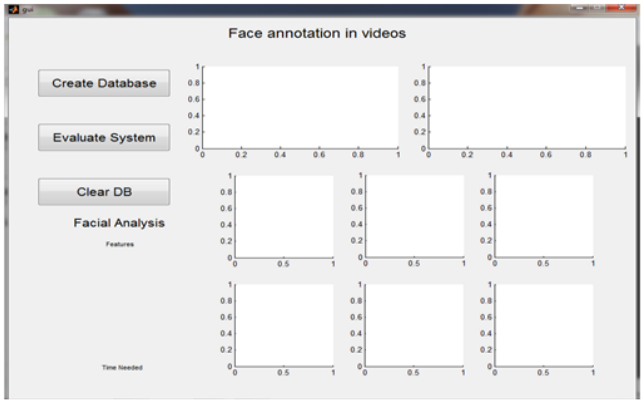

Fig. 4: GUI of our Proposed System

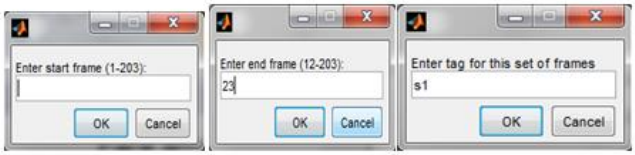

Fig. 5: Framing of videos and tagging of retrieved images

Next part of the process is to evaluate the system. The image or video to be verified against the database created earlier. The video is again divided in the frames and exposed to Viola-Jones algorithm for face detection. Whereas the image is directly fed into system. The process of grey scaling and facial features extraction is carried out as above. The computation of the features is done. Now the related parameters are compared with the parameters of the images saved in the database. The images from database with highest matching relative parameters are shown as output.

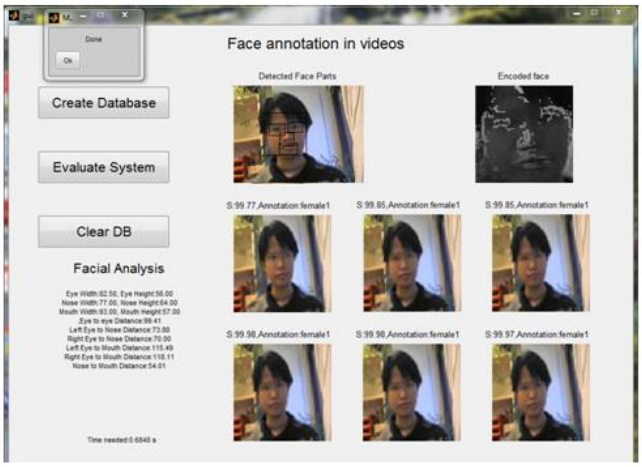

Fig. 6: Final Output of Our Proposed System

In our System, we created the database using both videos and images of 4 different persons. The total number of the images fetched by system in the database was 22. Then for the evaluation of the system we chose one video of a person with 20 frames. After the evaluation performed by the

system, it gave 6 most matched images from database with the percentage accuracies as 99.77%, 99.85%, 99.85%, 99.98%, 99.98% and 99.97% respectively. This resulted in the average percentage of accuracy as 99.90% for our proposed system. The delay factor in this process for comparison given by our system is 0.6848 seconds for 22 frames. This gives the average delay of 0.03113 seconds, which is less than the existing systems.

5. CONCLUSION

Through our proposed technique, we have proved that the face annotation technique by measurement of face structural features is quite adequate. The accuracy of the system has been increased. The performance factor has been increased with respect to time. Due to improved accuracy and decrease in the delay, the reliability of our system increases. The systematic database used in our system helped to achieve the desired performance. In this technique of face annotation, the drawbacks of some existing face annotation systems have been eliminated to some extent.

This research indicates a new direction for the findings of asymmetric systems to eliminate the dependence on the passwords and smart cards etc. This face annotation technique may be a useful tool for the improvement of identity recognition parameter in online social networking.

6. REFERENCES

1. S. Satoh, Y. Nakamura, and T. Kanade, "*Name-It: Naming and Detecting Faces in News Videos*", IEEE MultiMedia, vol. 6, no. 1, pp. 22-35, Jan.-Mar. 1999.
2. Jianke Zhu, Steven C.H. Hoi and Micheal R. Lyu, "*Face Annotation Using Transductive Kernel Fisher Discriminant*", IEEE Trans. Multimedia, vol. 10, no. 1, pp. 86-96, Jan. 2008
3. A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "*Content-Based Image Retrieval at the End of the Early Years*," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 22, no. 12,pp. 1349-1380, Dec. 2000
4. D.-D. Le and S. Satoh, "*Unsupervised Face Annotation by Mining the Web*," Proc. IEEE Eighth Int'l Conf. Data Mining (ICDM), pp. 383-392, 2008
5. T.L. Berg, A.C. Berg, J. Edwards, M. Maire, R. White, Y.W. Teh, E.G. Learned-Miller, and D.A. Forsyth, "*Names and Faces in the News*," Proc. IEEE CS Conf. Computer Vision and Pattern Recognition (CVPR), pp. 848-854, 2004
6. Ms. Malashree Patil, Prof. Jyoti Neginal, "*A Robust Face Annotation Method by Mining Facial Images"*, IJARCET, May 2015.
7. Dayong Wang, Steven C.H. Hoi, Ying He and Jianke Zhu, "*Mining Weakly Labeled Web Facial Images for Search-Based Face Annotation*", IEEE, Jan 2014.
8. X.-J. Wang, L. Zhang, F. Jing, and W.-Y. Ma, "*Anno Search: Image Auto-Annotation by Search*," Proc. IEEE CS Conf. Computer Vision and Pattern Recognition (CVPR), pp. 1483-1490, 2006
9. Steven C.H. Hoi, Dayong Wang, Yeu Cheng, Elmer Weijie Lin, Jianke Zhu, Ying He, Chunyan Miao, "*FANS: Face Annotation by Searching Large-scale Web Facial Images*", IW3C2, May 2013.
10. M. Kan, S. Shan, H. Zhang, S. Lao and X. Chen, "*Multi-View Discriminate Analysis*", in proc, 12th Eur. Conf. Comput. Vis., vol. 7572. Oct. 2012, pp. 808-821.
11. G. Guo and X. Wang, "Kinship *Measurement on Salient Facial Features*", IEEE Trans. Instrum. Meas., vol 61, no. 8, pp. 2322-2325, Aug. 2012.
12. Y. Yang, Y. Yang, Z. Huang, H.T. Shen, and F. Nie, "*Tag Localization with Spatial Correlations and Joint Group Sparsity*," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 881-888, 2011.
13. Z. Cao, Q. Yin, X. Tang, and J. Sun, "*Face Recognition with Learning-Based Descriptor*" IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 2707-2714, 2010