ISSN (Online): 2456-0448
International Journal Of Innovative Research In Management, Engineering And Technology
Vol. 1, Issue 4, May 2016

Retrieve Similar Facial Images Using Auto Face Annotation Technique

[1] Mr. Akshay Keshwatkar, [2]Mr. Radhakrishnan k
[1] [2]Computer Science Dept,Velammal Engineering College, Chennai-66,Tamil Nadu,India
[1] akshay.b.keshwatkar@gmail.com, [2]radhakrishnankumarasamy@gmail.com

Abstract: ***In this paper we investigate an auto face annotation technique by mining weak facial images which does not consist any data. These images can be collected from various sources like videos, random images, internet, CCTV, etc. However this is a challenging problem because the images collected have incomplete data, they do not have proper label and are random as well as noisy. Such images are called weak images. To improve the quality of these weak images we build a database by extracting various facial features of these images. Given a weak query image for identification, the auto face annotation technique annotates the query image and obtains resultant values. These values are then matched with the values of images stored in the database to retrieve the top ranked similar facial images. We detect and align the facial image extracted from a weak image with help of Haar cascade classifiers and extract the GIST features from the detected facial image. Different clusters are then formed for improved scalability of these images which significantly boosts up the performance of this technique.***

Keywords ***Gabor filter, cluster, Haar cascade classifiers, Kirsch detector, auto face annotation.***

I. INTRODUCTION

In this paper our main objective is detection and identification of face which can be easily performed at public places, schools and colleges, industries, etc. Nowadays due to availability of resources there is tremendous increase in photo and video sharing, cameras are present at most of the places, as a result of this numerous images are captured. But the disadvantage is that most of these images are not tagged with proper data. It would be a tedious and time consuming task to annotate appropriate data to these images. To overcome this problem we make use of automated image annotation [1], which automatically assigns the images with some data to describe the semantic features of the image. Despite of studying in detail [1, 2 and 3], the existing methods of image annotation give less satisfactory results. In this paper we make use of approach where in, a human face is detected from an image and appropriate data is annotated to it, this process is called auto face annotation [4]. Auto face annotation can be applied to various fields, for example it can be used at railway stations and bus depots, airports, schools and colleges, in military, in news videos in order to detect important people appearing in these videos [5], etc. Conventional annotation methods are usually model based which generally trains data from various facial models using human supervised techniques [6, 7]. Such approach has a few limitations; firstly it is an expensive and time consuming process to collect appropriate human labelled facial images. Secondly it is difficult to organize all the models when new facial images are added; also the process of annotating data is often poor when the number of images are large. To overcome these limitations, in this paper we focus on specific features of face. We make use of Haar cascade classifiers for detection of human face and extract GIST features of the facial image by calculating colour, intensity, edge value. A Gabor filter is used to extract the Gabor value of the face. We then make different clusters of numerous facial images. When a query image is given for identification, the features of the facial image are extracted and the image is annotated with this data. A search is then performed to mine out the top n similar facial images to that of the query image.

II. LITERATURE SURVEY

Facial detection, recognition and verification are the most common problems which are being studied over several years [8, 9]. There is a variety

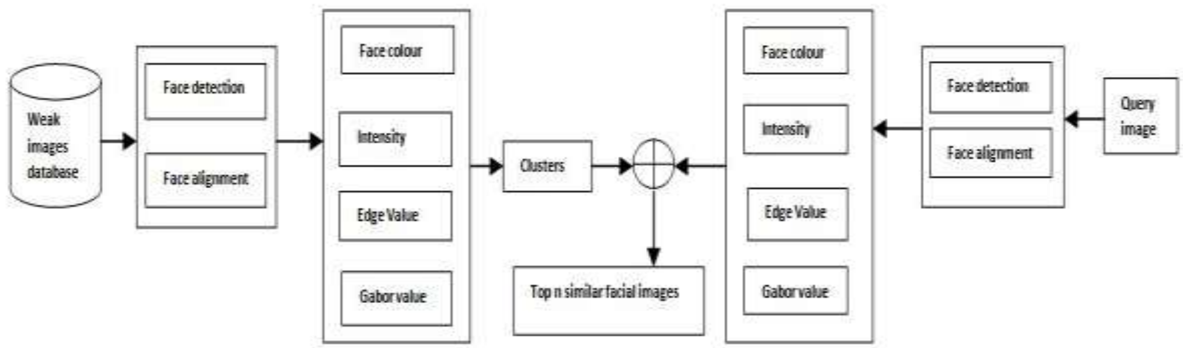

Figure 1: Architecture diagram of auto face annotation framework

of facial image databases for face detection and verification developed by various researchers for example the FERET database [10]. Face annotation in general can be referred as a face recognition and verification problem. For example the face annotation linear discriminate analysis method by Fisher is a Bayesian methodology for face recognition [7] and it uses support vector mechanism for training and predicting human names faces in videos. Attempts of semi supervised face annotation techniques are also practiced. For example kernel fisher discriminate method for facial annotation [11] makes use of unlabeled as well as labelled facial images for classification into models. Conventional facial annotation techniques usually make use of existing object recognition techniques to classify the human labelled facial images into different models [1, 2, and 3]. For example, Russell et al. [12] developed a database consisting of collection of web facial images along with truth table for research in object recognition. Wang et al. [13] made use of efficient hashing techniques for fast retrieval of facial image annotation. Torralba et al. [14] proposed a large scale image repository by suggesting efficient search of images as well as scene matching technique. Unlike the above mentioned work, our work is not model based; we concentrate on specific facial features and pixels by mining out weak facial images and performing auto face annotation.

III. AUTO FACE ANNOTATION FRAMEWORK

Figure 1 shows Auto face annotation framework, initially weak images are collected in weak facial images database. These images undergo facial detection that is detection of face from a weak image, this is performed by using Haar cascading classifiers. After face detection and alignment GIST features are extracted from the facial image. First the colour of the face is extracted this value is given as input for calculating intensity value of the facial image. The value of the image colour is used as input for calculation of intensity by using grayscale. This value is used as input for detection of edges of the face where we make use of Kirsch operator to find edge value of the facial image. The edge value of the facial image is used to find the Gabor value by using Gabor filter. Finally with the help of edge value and Gabor value of the facial image we form clusters of different facial images. Now when a query image is given as an input, the face is detected from the image and aligned, GIST features of the face are calculated to obtain a value. This value of query image is then compared with already calculated facial image values in the clusters to obtain the top n similar facial images thus we retrieve similar facial images as that to the query image. This entire process takes place without any human intervention and so is called auto face annotation.

IV. IMPLEMENTATION AND TESTING

In the initial step weak images are collected in a database, weak images are the images which does not consist any data for example images without any name, date, location, etc. In our application we made use of Microsoft visual studio express 2012 and SQL server for database management. These images can be collected from various sources like news videos, Closed circuit television cameras (CCTV) which are present at various places, random images, World Wide Web (WWW), etc. Then we perform facial feature extraction from these weak images.

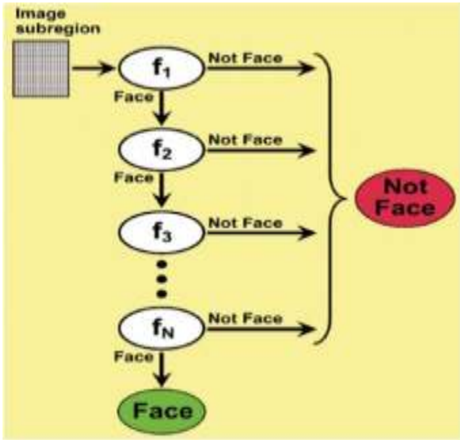

Figure 2: Working of algorithm.

A. Face detection and alignment
For detecting a face from a weak image we make use of Haar cascade classifiers, here we make use of edge features, rectangle and line features. We focus on a property that, the regions of eye points are generally darker than nose and cheeks. In figure 2 it is shown how this algorithm works, instead of using all 6000 features the features are divided into different stages, initial stages contains few features. If the window will fail in first stage of features the remaining features will not be considered on that stage but if the window passes, second stage of features is applied to it and the process is continued. If all the stages are passed then that region is a face. In this way a face is detected from a weak image. Figure 3 shows detection of the face from a weak image and its alignment. When a face is detected, the facial image is cropped such that the distance from eye points is100 pixels only. Then the facial image is aligned in 100 x 100 pixels for easy calculation.

B. GIST feature calculation

For calculation of GIST features of the face we consider Gabor value, edge value, colour and intensity features of the face.
1. Calculation of face colour
Initially the facial image is converted to bytes using 'ImageToByte' function. The colour of the facial image can be represented as either 8 bit or 16 bit. Figure 4 shows the image file representation, it consists of pixel indices and palette. Palette consists of 256 colours. A single byte represents each pixel in the image; this byte is the index, that

Figure 3: Detection of face from weak image.

is location of the colour in the palette. The colour of each pixel of the facial image is calculated and represented in a RGB value. Here we calculate the skin tone using 'System.Drawing' namespace. When RGB value for each and every pixel is calculated, we consider only top 3 most repeated colour values as shown in figure 5, we get the top 3 values as 12, 7, 7 .

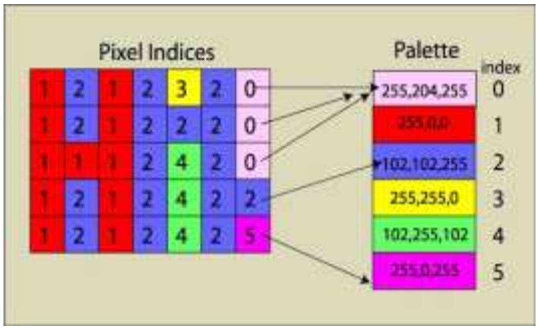

Figure 4: Image file representation.

2. Calculation of intensity

To convert an RGB colour to its grayscale weighted sums are calculated in linear RGB space after removing gamma compression. The gamma expression can be defined as follows
$C_{linear}= C_{srgb}/12.92$ or $C_{linear}= (C_{srgb}+0.055/1.055)^{2.4}$ where C_{srgb} represents any of the primary sRGB that is R_{srgb}, B_{srgb}, G_{srgb} in range of (0, 1). C_{linear} represents the linear intensity values of R, G, and B in the range of (0, 1). The intensity value of each pixel in the facial image is obtained and the average is calculated to get a black and white form of the coloured facial image as shown in figure 5, the calculated intensity of the facial image is 234.

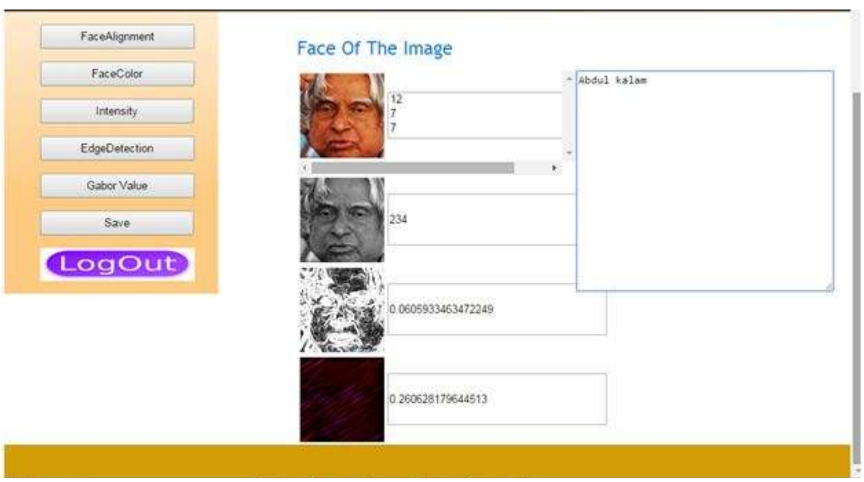

Figure 5: Calculation of face colour, intensity, edge and Gabor value. In the first image the colour feature is extracted, second image is black and white which represents intensity, third image represents the edge value and fourth image represents the Gabor transformed image.

3. Calculation of Gabor value and edge value

Gabor filter views the facial image as a formation of superimposed series of sine waves of different frequencies in all possible directions. Each pixel in Fourier transform describes the intensity of the wave and the frequency of the wave is described by pixel position. Gabor filter is a kind of band pass filter that cuts the Fourier transform to reveal specific information only. Each Fourier pixel has a real and imaginary part and is a complex value. $g(x, y, \lambda, \theta, \psi, \sigma, \gamma)=X$; where X represents a 2D matrix of numbers which contains a convolution mask. The original facial image is filtered with the convolution mask to get another value of the image. This new image is the Gabor filtered image as shown in figure 5, its calculated value is 0.260628179644513.
To calculate the edge value of the facial image we make use of non linear edge detector algorithm called Kirsch operator. It uses the pixel position and direction, to calculate maximum edge strength. It uses a few predetermined directions N, S, E, W, NE, NW, SE, SW, the mathematical description is as follows

$h_{n,m} = \max_{z=1,2,\ldots,8} \sum_{i=-1}^{1} \sum_{j=-1}^{1} g^{(z)}_{ij} . f_{n+i,m+j}$. where z enumerates the compass direction.

$$g^{(1)}=\begin{bmatrix} +5 & +5 & +5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix},\ g(2)=\begin{bmatrix} +5 & +5 & -3 \\ +5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix},$$

$$g(3)=\begin{bmatrix} +5 & -3 & -3 \\ +5 & 0 & -3 \\ +5 & -3 & -3 \end{bmatrix},\ g(4)=\begin{bmatrix} -3 & -3 & -3 \\ +5 & 0 & -3 \\ +5 & +5 & -3 \end{bmatrix}$$

and so on. The Kirsch operator rotates the kernel mask in 45 degree increments and in all 8 directions to find the edge magnitude which is considered as the maximum magnitude in all 8 directions. The calculated edge value of the facial image shown in figure 5 is 0.0605933463472249.

C. Cluster formation

Different facial images will have different edge and Gabor values depending upon the colour and intensity of the images. Depending on edge and Gabor values of the facial images we formed three clusters for sake of simplicity. The range of cluster 1 is ((edgeval>=0.045&&edgeval<=0.065) &&(gaborval>=0.20&&gaborval<=0.50)), cluster 2 range is ((edgeval>=0.061&&edgeval<=0.065)&& (gaborval<=0.19&&gaborval>=0.10)) and cluster 3 range is ((edgeval>=0.056&&edgeval<=0.065)&&(0.20<=gaborval && gaborval<=0.30)) where edgeval represents the edge value and gaborval represents the Gabor value of the facial images. As shown in figure 5, the image lies in cluster 1. Similarly depending on edge and Gabor value other facial images are also divided among these three clusters.

D. Similar face retrieval

When a weak query image is given as input the above mentioned process is repeated, that is initially using Haar cascade classifiers the face is

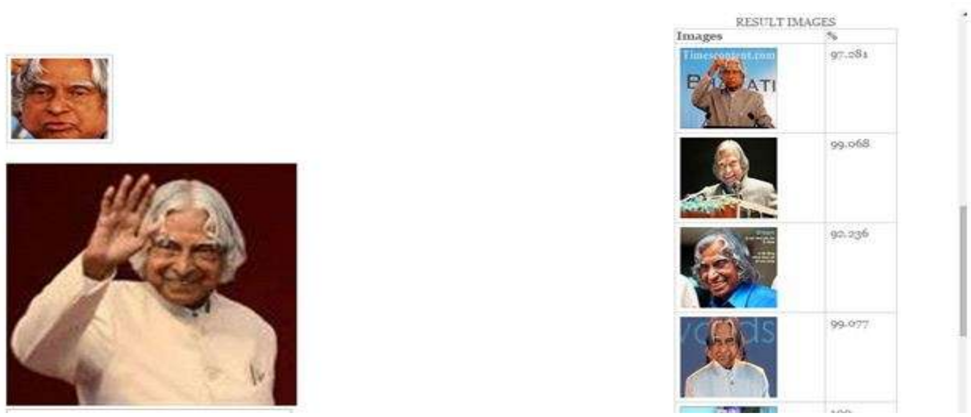

Figure 6: Retrieval of top 5 similar facial images.

detected from weak query image. Considering the eye points the face is cropped. The face is then aligned in 100x100 pixels. Now the GIST features are extracted from the facial image. We obtain the value of face colour, intensity, edge and Gabor transform. We have kept the level of matching as 90% and above so that only those images whose matching percentage is more than 90 are mined and retrieved from the database. The edge and Gabor value of query image is compared among the three clusters and those facial images whose matching percentage is 90 and above are displayed as shown in figure 6 in order to retrieve only top 5 facial images. This entire process takes place without any human intervention and so is called automated face annotation.

V. CONCLUSION AND FUTURE ENHANCEMENT

In this paper we investigated auto face annotation technique where we overcome the problem of face detection and identification. Initial problem was face detection from a weak image; this we overcame with the help of Haar cascade classifiers. Then specific facial features were extracted and clusters were formed for scalability. We retrieved top ranked similar facial images to that of a query image. In future research face identification still needs to be refined in order to get exact match of the query facial image. In future this system must be updated for implementation in smart phones and other hand held devices.

VI. REFERENCES

[1] P. Duygulu, K. Barnard, J. de Freitas, and D. A. Forsyth.Object recognition as machine translation:

Learning a lexicon for a fixed image vocabulary. In ECCV, pages97–112, 2002.

[2] G. Carneiro, A. B. Chan, P. Moreno, and N. Vasconcelos. Supervised learning of semantic classes for image annotation and retrieval. IEEE Tran. PAMI, pages 394–410, 2006.

[3] J. Fan, Y. Gao, and H. Luo. Multi-level annotation of natural scenes using dominant image components and semantic concepts. In ACM Multimedia, pages 540–547, 2004.

[4] Dayong Wang, Steven C.H. Hoi, Ying He. Mining Weakly Labeled Web Facial Images for Search-based Face Annotation. School of Computer Engineering, Nanyang Technological University, Singapore {s090023, chhoi, yhe}@ntu.edu.sg.

[5] S. Satoh, Y. Nakamura, and T. Kanade. Name-it: Naming and detecting faces in news videos. IEEE MultiMedia, 6(1):22–35, 1999.

[6] 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. In CVPR (2), pages 848–854, 2004.

[7] J. Yang and A. G. Hauptmann. Naming every individual in news video monologues. In ACM Multimedia, pages 580–587, New York, NY, USA, 2004.

[8] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Trans. PAMI, 19(7):711–720, 1997.

[9] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. ACM Comput. Surv., 35(4):399–458, 2003.

[10] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss. The feret evaluation methodology for face-recognition algorithms. IEEE Trans. PAMI, 22(10):1090–1104, 2000.

[11] J. Zhu, S. C. H. Hoi, and M. R. Lyu. Face annotation using transductive kernel fisher discriminant. IEEE Transactions on Multimedia, 10(1):86–96, 2008.

[12] B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman. Labelme: A database and web-based tool for image annotation. Int. J. Comput. Vision, 77(1-3):157–173, 2008.

[13] X.-J. Wang, L. Zhang, F. Jing, and W.-Y. Ma. Annosearch: Image auto-annotation by search. In CVPR, pages 1483–1490, 2006.

[14] A. Torralba, Y. Weiss, and R. Fergus. Small codes and largedatabases of images for object recognition. In CVPR, 2008.