ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 4, April 2016**

# Image Clustering Using Face Recognition

Akshay Deo[1], Shubhankar Dongre[2], Vinayak Dhakol[3], Aniket Mandilkar[4], Prof. Jyoti Vetal[5]

Student, Dept. of IT, MIT College of Engineering, Pune, Savitribai Phule Pune University, Pune, India[1,2,3,4]

Prof., Dept. of IT, MIT College of Engineering, Pune, Savitribai Phule Pune University, Pune, India[5]

**ABSTRACT:** Automated human face recognition is mainly divided into two parts one is face detection and second one is recognition of detected faces. The paper proposes a viola Jones algorithm for face detection in color images, with detection of multiple faces in an image. Further the images given in bulk are clustered as – all the images of one particular face into one folder.

**KEYWORDS:** Digital Image Processing, Viola Jones, Cluster, Face Detection

## I. PROJECT OBJECTIVES

1. Using Image processing concept we have handled any type of image.
2. We provide clustering of images.
3. In image processing upload image folder and cluster them as :-all the images of one particular face into one folder. In this way create folders for each face.

## II. THEME OF THE PROJECT / INTRODUCTION

Face is our important focus of attention for conveying identity. Human face detection by computer systems has become a major field of interest. Detecting the multiple faces in a digital image has gained much importance with application in many areas. Face detection is very important approach to recognize the face procedure, head tracking and many more applications. In the presented work, face detection problem in a group photograph is worked out using the viola Jones algorithms. The face recognition is done using combined approach of HAAR Wavelet transform. Face identification from a group photograph is very much required during investigation of a scene from crowd or cluster of faces. The most difficult challenge is to detect faces in clumsier group photos. While face recognition, in general the faces that retrieved from group photo are not giving sufficient information due to poor clarity. This kind of limitation is inherent in the imaging device and circumstances when the photographs are taken. The photographic conditions are uncontrolled when imaging the crowd or at public places in emergency situations. Therefore, poor visibility or incomplete photos are common problem in identification of faces from cluster of faces.

Face recognition is one of the biometrics traits that received great attention of many researchers during the past few decades because of its potential applications in a variety of civil and government regulated domains. It usually involves: initial image normalization, preparing an image for feature extraction by detecting the face in that image, extracting facial features from appearance or facial geometry. In the proposed work, we have done the detection of human face from group photograph using Viola-Jones algorithm. From the detected faces, we have selected one face image from group photograph's detected faces as an test image i.e. cropped the face and have prepared its folder. Similar method has been used for all such face images and multiple folders for their corresponding faces were created.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 4, April 2016**

## III. SOFTWARE AND HARDWARE REQUIREMENTS

**SOFTWARE**

I. APACHE TOMCAT 6.0
II. Eclipse-IDE
III. My-SQL
IV. Web-browsers

**HARDWARE**

I. PENTIUM-5 PROCESSOR
II. RAM-4GB
III. HARD DISK-500GB

## IV. DETAILS OF ALGORITHM AND TESTING TOOLS

1) RGB : -
Any image depends on RGB values. RGB means Red, Green and Blue. The range of RGB value is 0-255.

2) Edge Detection : -
The edge detection is detecting edges of images. That is highlight the edges of images. This is a pre-requisite for using the Voila Jones algorithm.

3) Face Detection: -
In color images we have to find only human faces. Human faces have different types of patches like eye, nose, lips etc.

4) Viola Jones: -
The viola Jones is a face detection algorithm. This algorithm has used haar features of face detection.
The algorithm has four stages:
1. Haar Feature Selection
2. Creating an Integral Image
3. Adaboost Training
4. Cascading Classifiers

- Haar-like features:-

Haar-like features are digital image features used in object recognition. They owe their name to their intuitive similarity with Haar wavelets and were used in the first real-time face detector.
Historically, working with only image intensities (i.e., the RGB pixel values at each and every pixel of image) made the task of feature calculation computationally expensive. Working with an alternate feature set based on Haar wavelets instead of the usual image intensities is discussed. Viola and Jones adapted the idea of using Haar wavelets and developed the so-called Haar-like features. A Haar-like feature considers adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region and calculates the difference between these sums. This difference is then used to categorize subsections of an image. For example, let us say we have an image database with human faces. It is a common observation that among all faces the region of the eyes is darker than the region of the cheeks. Therefore a common haar feature for face detection is a set of two adjacent rectangles that lie above the eye and the cheek region. The position of these rectangles is defined relative to a detection window that acts like a bounding box to the target object (the face in this case).

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

Rectangular Haar-like features

A simple rectangular Haar-like feature can be defined as the difference of the sum of pixels of areas inside the rectangle, which can be at any position and scale within the original image. This modified feature set is called 2-rectangle feature. Viola and Jones also defined 3-rectangle features and 4-rectangle features. The values indicate certain characteristics of a particular area of the image. Each feature type can indicate the existence (or absence) of certain characteristics in the image, such as edges or changes in texture. For example, a 2-rectangle feature can indicate where the border lies between a dark region and a light region.

- Integral image:

An image representation called the integral image evaluates rectangular features in constant time, which gives them a considerable speed advantage over more sophisticated alternative features. Because each feature's rectangular area is always adjacent to at least one other rectangle, it follows that any two-rectangle feature can be computed in six array references, any three-rectangle feature in eight, and any four-rectangle feature in nine.

- Adaboost Training:

The speed with which features may be evaluated does not adequately compensate for their number, however. For example, in a standard 24x24 pixel sub-window, there are a total of $M = 162,336$ possible features, and it would be prohibitively expensive to evaluate them all when testing an image. Thus, the object detection framework employs a variant of the learning algorithm AdaBoost to both select the best features and to train classifiers that use them. This algorithm constructs a “strong” classifier as a linear combination of weighted simple “weak” classifiers.

- Cascading classifiers:

On average only 0.01% of all sub-windows are positive (faces).
Equal computation time is spent on all sub-windows must spend most time only on potentially positive sub-windows.
A simple 2-feature classifier can achieve almost 100% detection rate with 50% FP rate.
That classifier can act as a 1st layer of a series to filter out most negative windows
2nd layer with 10 features can tackle “harder” negative-windows which survived the 1st layer, and so on…
A cascade of gradually more complex classifiers achieves even better detection rates.

5) Matching:-
Detecting human faces and matching them with other faces to find out similarity of images. The pixel values (eg. Pixel intensities) of images are compared for the matching purpose.

6) Clustering: -
The matching photograph is clustered and saved in its respective folder.

## V. Result Analysis

The results are taken for two different cases. First when there is a single face in the image. Second when there are mutiple faces in the image. Image clustering is made possible for both the scenarios. Further the lighting conditions, picture quality also have an impact on the result.

## VI. Conclusion

In Face recognition system, it is challenging to detect the face from group photograph & recognize the faces at different poses. In the proposed work, we have done the detection of human face from group photograph using Viola-Jones

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 4, April 2016**

algorithm. The photographic conditions are uncontrolled when imaging the crowd or at public places in emergency situations. Therefore, poor visibility or incomplete photos are common problem in identification of faces from cluster of faces.

## VII. PROJECT OUTCOME

In image processing there are different types of operations in image. We have different type of photos- family, friends, trip, etc. If you want to sort it then you have to check all the photos one by one. This process is too long. To avoid this problem we implement the system using viola Jones algorithm. This is for detecting the faces. Main motto of project is - In image processing upload image folder and cluster them as :-all the images of one particular face into one folder. In this way create folders for each face.

## REFERENCES

[1] S. Ganesh Bashyal and K. Venayagamoorthy, "Recognition of facial expressions using Gabor wavelets and learning vector quantization", Engineering Applications of Artificial Intelligence, vol. 21, no. 7, (2008) October, pp. 1056-1064.
[2] T. Ahonen, "Face description with local binary patterns: Application to face recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 12, (2006), pp. 2037-2041.
[3] H. Moon, "Computational and performance aspects of PCA-based face-recognition algorithms", Perception, vol. 30, no. 3, (2001), pp. 303-321.
[4] D. L. Swets and J. Weng, "Using discriminant eigenfeatures for image retrieval", IEEE Trans. Pattern Anal. Mach. Intell., vol. 18, no. 8, (1996), pp. 831-836.
[5] L. Juwei, "Face recognition using LDA-based algorithms", IEEE Transactions on Neural Networks, vol. 14, no. 1, (2003), pp. 195-200.
[6] I. Dagher and R. Nachar, "Face recognition using IPCA-ICA algorithm", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 6, (2006), pp. 996-1000.
[7] Z. Baochang, "Local derivative pattern versus local binary pattern: face recognition with high order local pattern descriptor", IEEE Transaction on Image Processing, vol. 19, no. 2, (2010), pp. 827-832.
[8] K. Fukunaga, "Introduction to statistical pattern recognition", Second ed., Academic Press.
[9] L. Wiskott, J. M. Fellous, N. Krüger and C. von der Malsburg, "Facerecognition by elastic bunch graph matching", IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 7, (1997), pp. 775-779.
[10] D. Lowe, "Distince image features from scale-invariant key-points", Int. Journal of Computer Vision, vol. 60, no. 2, (2004), pp. 91-110.
[11] ORL face database, website: http://www.cam-orl.co.uk/facedatabase.html, AT&T Laboratories Cambridge.
[12] UMIST face database, http://images.ee.umist.ac.uk/danny/database.html, Daniel Graham.
[13] Yalefacedatabase, www1.cs.columbia.edu/~belhumeur/pub/images/yalefaces/,ColubmbiaUniversity.
[14] Bioid face database, website: www.bioid.com/downloads/facedb/.
[15] M. Aly, "Face Recognition using SIFT Features", Technical Report, Caltech, USA, (2006).
[16] L. Lenc and P. Král, "Novel Matching Methods for Automatic Face Recognition using SIFT", AIAI, (2012), pp. 254-263.