Face Detection and Tracking
Using OpenCV

Kruti Goyal
ASET, CSE
Amity University
Noida, India
kaaggoyal@gmail.com

Kartikey Agarwal
ASET, CSE
Amity University
Noida, India
Kta136@gmail.com

Rishi Kumar
ASET, CSE
Amity, University
Noida, India
Rkumar25@amity.edu

Abstract- **An application for tracking and detecting faces in videos and in cameras which can be used for multipurpose activities. The intention of the paper is deep study of face detection using open CV. A tabular comparison is performed in order to understand the algorithms in an easier manner. It talks about various algorithms like Adaboost, Haar cascades. This paper aims to help in understanding the best prerequisites for face detection.**

Keywords---- Adaboost (adaptive boost), open computer vision (OpenCV).

I. INTRODUCTION

Face detection is the most popular area of research in the vision of computer science. It is a computer technology which is being used in a variety of applications that identifies human faces in digital images [1]. The research under this field is expanding in many areas of science such as psychology. Face detection is one of the most talked about in technology. Localization of human faces is considered as the primary and the initial stage in study of face detection. For example in home video surveillance etc. Face localization can be referred to as extraction of facial features using pattern recognition system. Both MATLAB and Open CV can be used for creating such prototypes and systems. In this paper we have carried out our research using Open CV. The Reasons for using open CV have been discussed further in this paper.

II. OPEN CV V/S MATLAB

A. Speed

Since, MATLAB is fabricated from java which in turn is fabricated from C. Therefore, when a code is scripted and run on MATLAB, the computer initializes by interpreting the code and converting it into java and then finally executes the script. Whereas, open CV uses c/c++ library functions. Which directly provides the computer with the machine language code and hence helps in faster execution. Using OpenCV results in more utilization of time and resources in image processing and less in interpreting.

B. Portability

As OpenCV runs on C, therefore any device which runs on C can run OpenCV. It can toil well with Windows, mackintosh or Linux.

C. Cost

MATLAB is much more expensive than OpenCV. MATLAB costs around USD2150 whereas, OpenCV is Free of cost. Even the base MATALAB is expensive as it has commercial, single user License. And OpenCV is a BSD license so it is free of cost.

III. RELATED WORK

Face recognition involves three basic steps which include face detection, face recognition, and face extraction. Any system needs to encapsulate the image and then manage as well as record the vital features to determine the location of the face [2]. For the recognition of the captured image it keeps records of various features such as skin color, skin tone etc.
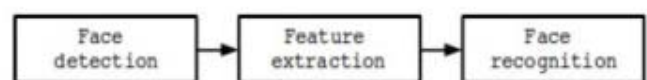

Fig 1. A generic face recognition system

Face recognition takes an image from a video or a camera as input and outputs the identified image topic. Facial features may include regions in the face, variations in the face structure, face cuts and angles that have been formatted and styled. Face extraction includes grabbing of the features from camera. Face detection includes removal of the background and focusing on the foreground eliminating any other elements apart from the face region, however the system still pertains a few drawbacks as it cannot detect the head count which may be present due to overlapping of faces or improper recognition of two faces having similar facial features. Implementation of face recognition using principal component analysis using 4 distance classifiers is proposed in a publication by Hussell Hardy [3].

IV. FINDING FACES

Finding faces is the most integral part in face detection. There are various techniques from which faces could be found. In this paper, we will be comparing the various algorithms used previously by implementing them and analyzing them. Even though face detection is the most important step in the area of image processing still the techniques used for its implementation need to be reworked upon so as to optimize its performance and bring down the hurdles it is cladding. The working of face detection algorithms is majorly based on the accuracy of the face detection, due to this face detection is the cusp point in the entire process of face detection and tracking.

1) FINDING FACES VIA COLOR

A. Images having a definite backgroud .

One procedure is to find images in which we have a monochromatic background containing only grey scale pixels. These images have wavelength of narrow band. While using these images when we eliminate the background from the foreground we get the boundaries of the face. This is by far the easiest technique used for face detection.

B. Images having a coloured background.

The system in which a face is to be detected for colored images is based on two procedural steps

1. By applying a skin filter

Skin filter is applied for skin detection. The skin filter processes and defines the texture of image part that is being masked. After the initial masking the output generated contains discrete areas of human skin. Various morphological techniques such as dilation and erosion are used to develop this kind of filter.

2. By hauling out the features which are being masked.

During this step the various portions of the image which are very dark or very bright are subtracted from the image. These include the area around the eyes, area under the nostrils or the area below the mouth. By subtracting these portions, we get the most relevant area which is covered by skin and effective detection can be done.

The major problem during this process is the variable light sources at the time of face detection by the camera. Now it is not feasible to place all the cameras in the presence of sunlight, some might have been placed at low light source region. These cameras can be used for application only if they are placed according to need and not according to environment.

To solve this problem, we make use of a skin color model proposed by outski and Healey [4]. The input should be in RGB format with caricature intensity values ranging from 0 to 255. The RGB matrices are "zeroed" to supposedly act like a wet blanket desaturation when the image is improved from RGB enlarge space to IRgBy emphasize space. The smallest intensity outlay greater than 10 pixels from entire edge in barring.

C. Images having a complex background

Face detection for faces present in the complex background can be done using MUHULANOBIC metric. It is based on the examination of human faces in a two-dimensional natural scene images. It distinctly makes use of the process of color segmentation of the image that is being taken as the input. This segmentation of the color is being performed by Thresholding the picture in the hue color space. This takes into account the effects of distinctiveness of the color in the human skin when the lighting has been changed in the image. Then the results of the image are accumulated together for any other examination. This is followed by median filtering of the resultant image using a limited number of accumulations of pixels. At the end the difference between the faces and the remaining complex background, a multi-layer perceptron neural network is used with the invariant moments as an input factor. [5]

2) FINDING FACES VIA MOTION

A. Use of Blink Detection Techniques

Blinking is an involuntary task functioned by a human. It is a very hasty process. Some humans might not even detect blinking in the daily procedures but blinking as a process has be proved to detect the presence of a human significantly at any frame of time. Blinking provides a non-periodic time and space signal which is unique to every other person. Therefore, the blinking process can equivalently act as a biometric means of measure to detect the presence. An algorithm is used to make blinking make sense to a computer.

This algorithm includes taking two images of a person simultaneously and subtracting the second image from the first image [6]. This subtraction causes a distinct boundary outside the head and if in one of the images the eyes are blinked and there seems to be a little circled region at the eye portions. To this subtracted image a connected component procedure is pragmatized. For a detection of blinked image there should be a horizontal and vertical bounded regions. These regions signify the horizontal and vertical separation. After this hypothesis, the site of the picture is determined using this center line and the horizontal and vertical bounding regions. The size of these bounding regions help to face. This is one of the most reliable and capable technique to find the size of the face. This can be used in various applications such as face records during a record of criminals in jail, schools with compulsory attendance.

An Efficient security management service, to create software equivalent to and efficient than Picasa which records faces and their identification. To acquire such an arrangement, a mesh of photos is needed to be created. Due to this system robustness against scaling, posture, facial expression and illumination is increased [9].

This algorithm can also be translated to video recording purposes via small variations in the algorithm [7]. In this the moment is signified via calculation of difference between the two videos and if the resultant value is greater than (usage of colors/0) then only the process is taken any further.
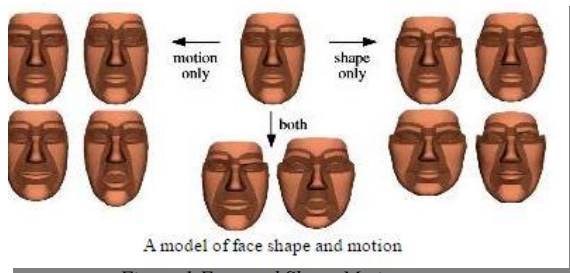

Figure 1 Face and Shape Motions

3) FINDING FACES IN LIMITED AREAS OF PIXELS.

In the images which are unconstrained it is difficult to detect the faces but various methods have now been introduced such as edge detection orientation, haudrof distance, weak classifier cascades.

Edge orientation matching is a technique treated as a template matching procedure. This includes object modelling based on edge orientations. Various templates are created and matched to the image to detect the edges of the face. It takes approximately less than 0.08 seconds.

Haudrof technique makes use of the two distinct points which can be represented as

$G(\square\square\square\square\max(g(\square\square\square\square\square\square\square\square\square\square$
Where $g(\square\square\square\square\max\min|a-b|$.

$G(\square\square\square\square$ is denoted as directed Haudrof distance [8].
Through Edge detection and the modified haudrof distance, after the coarse detection of the portion of the face, the attributes of the face are thus then refined.
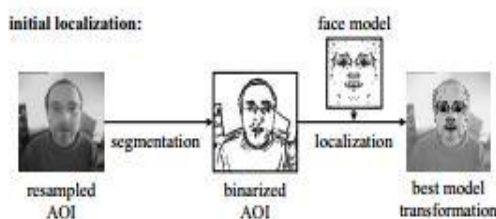

Figure 2.Localization of faces

4) CAMSHIFT ALGORITHM
The camshaft algorithm follows a series of steps performing the face detection:

    a.   Obtain the target
    b.   Generate a histogram and the portion which needs to be initially searched.
    c.   Determine the most populated pixelated region.
    d.   Determine the center position of the entire searched region.
    e.   Move the center of the searched region to the center of the entire region.
    f.   If you need to convert move forward else move back to step d.
    g.   Now, the new search window is established as the center is originated to the past center
    h.   Determine the new area to be processed.
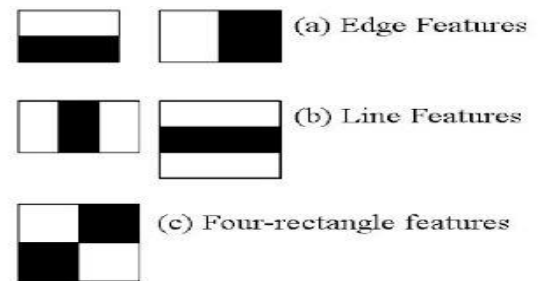   i.   Choose the start of the new window

5) HAAR CASCADES.

Figure 3. Haar cascades Features

Haar cascades makes use of the image subtraction morphological process to detect the face. In this the cascades of different images of the same person is taken and recorded in the database. All the pixels in the influence of white region are subtracted from all the pixels in the influence of black region.

This method of subtraction is performed on each of the image in the cascade but all the images might not give us the best results. Many of the images have a lot of errors. The image with the least error is selected.
The result of all the images are added together and is mentioned as weak classifier. As all the week classifiers are added together to form a strong classifier.
Appling the subtraction process and determining each image error is a very time and space consuming process. Instead of applying it on each of the images the subtraction is applied on images one by one. If the last image is not a useful it is discarded.

This process works with the predefined set of trainers present in open CV but if the developer wants he/she can also create a self-cascade by applying basic code on the compiler.

Many people prefer to prepare a self-cascade for better understanding.

s.no	Algorithm comparisons		
	Technique	Time consuming	Performance
1	Haar Cascades	Initially, it was time consuming, but after the modified version it is not that time consuming.	It is very efficient in performance and is popularly used.
2	Cam Shift Algorithm	It is not very time-consuming	It is very expensive to buy and is not that effective in performance
3	Finding via motion	It is very time-consuming	It is not so reliable as compared to any other algorithm

Figure 4. Calculate time for calculating the Haar cascade results.

Feature	8) face	9) face	10) person	11) person	12) 2 persons left	12) 2 persons right	13) 2 persons left	13) 2 persons right
intensity on-off	0.00	3.85	0.01	0.02	0.07	0.02	0.00	0.01
intensity off-on	6.62	1.74	7.20	3.12	5.67	6.59	6.73	6.36
orientation 0°	1.88	6.27	1.70	0.97	2.09	1.33	1.96	2.07
orientation 45°	1.47	6.50	2.66	2.27	2.58	2.30	2.14	1.70
orientation 90°	4.03	10.65	2.38	2.32	1.79	3.05	1.63	2.41
orientation 135°	3.60	3.49	2.26	3.22	3.21	3.28	2.79	1.52
color green	0.05	1.08	0.70	0.04	0.00	3.93	1.15	0.00
color blue	0.44	1.88	6.96	0.13	0.02	2.31	4.29	12.84
color red	5.25	10.55	3.01	8.09	47.65	0.81	2.95	25.04
color yellow	4.03	2.67	0.00	1.05	14.72	1.92	0.04	0.00
intensity	3.74	2.95	4.17	1.44	2.31	2.67	2.96	2.78
orientation	2.80	7.27	2.32	2.03	2.23	2.23	1.93	2.19
color	2.35	3.28	1.93	3.63	13.62	2.31	2.48	9.49

Figure 5. Calculate time for calculating the camshift results

V. APPLICATIONS OF FACE DETECTION AND TRACKING

Face detection and tracking is used in many of the fields these days. It is collaborated in every section of the trade. While some are effectively and efficiently using it, others still depend the on the old methods.

Following are the most effective and efficient uses of face detection and tracking.

a) One of the most common areas is video surveillances. Police departments all over the world are using face detection as an important and reliable proof to crack a case. Video surveillances help in tracking the faces that are coming in view of the camera. It is widely spread amongst the markets of the local as well as international trade.

b) Another application where it is used is security features. With this increasing change in technology, many companies are in trying to bring in this face Detection in daily lives of human beings by introducing this as a feature in mobile phones and laptops for user recognition, face passwords.

c) In many areas, slowly and steadily face detection is being introduced as means of smart attendance integrated with the payroll systems in the offices particularly. In this, the leave pay is automatically detected from the employee's monthly salary if the face is not detected.

d) In many countries, it is being used as a prevention from illegal withdrawal from the ATM machines ,in this the bakers link the account with a graphical representation of the account holder and while withdrawing the money from the ATM an alarm is automatically beeped if the card is in unauthorized hands which helps the guard of the ATM to nab the imposter, the drawback in this system is that it might be possible that some other family member is withdrawing the money in this case the account may be able to be linked with more than one pictures. This might be added to an additional layer of security to the pins and passwords.

e) Software's like Picasa are used to manage the unlimited number of pictures in one's account. It detects the faces in the still pictures and manages the pictures

via each face.[11] If two members are in the same photo then someone might complain about the redundant data but instead Picasa uses the location of the picture to reach.

f) Companies like Facebook, Instagram use face detection as a part of tagging people which is endorsed in market as an efficient feature which creates some popular demands of its usage within the users of the services.

g) Image and face recognition has brought about a major change in the gaming field. The most popular Xbox games makes use of Kinect which is based on face detection and is used as hardware controllers.

h) Snapchat which is most commonly used amongst the youth makes use of face detection for adding filters to it.

i) Many trading companies such as lens kart and others are using face detection as a technique of attracting more audience. They detect the face and apply frames of each face which gives a rough idea about how the frame looks on them rather than just going to various stores.

VI. RESULT ANALYSIS

In this paper we talked about the various techniques of implementing face detection and tracking .

1. This paper compares all the algorithm in terms of space and time paradigm. It identifies Haar cascades as the most efficient way of face detction.

2. This paper also talks about the difference in face detection via performing it with open cv and performing it with matlab. The advantages and disadvantages of both the methods have been talked about.

3. As we talked about the various algorithms we discovered that the algorithms such as haar cascades , camshift algorithm give a better performance and a much more verified output than detection via motion but as far as time is considered camshift algorithm and detection via motion algorithm is a better option.

4. Haar cascades works gives better accuracy in facial expression [10].

FUTURE SCOPE

Today's advancement is just the beginning in the onset of face detection technology there are many other applications where face detection can be introduced. The various ideas are:

a. It could be introduced in cabs where the drivers drive at night. An algorithm can be stated such that as soon as the driver blinks the eyes for more than 5 seconds or so an alarm might go of in the car that alerts everyone that the driver Is sleepy. [12]This might not be able to diminish the accident rates due to this but might be able to reduce it.

b. In finding the lost kids, adults can come back to home if they are lost but what about the kids which are being lost or being kidnapped to do labor or much worse where people are scared that the small girls are being used of prostitution.by creating a sever of these faces and using of face detection algorithms might help in saving hundreds of lives from the torture.

c. It can be used at various stations be it a metro station or an airport station. if a face detection system is being introduced at the entrances of all the stations and is being compared with the fugitive list of the internationally all the o police departments we might be able to reduce the crime rate.

REFERENCES

[1] G.T. Rado and H. Suhl,International Journal of Soft Computing and Artificial Intelligence(1963), ISSN: 2321-404X, I.S. Jacobs and C.P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, Eds. New York: Academic, , pp. 271-350.

[2] Driver Drowsiness Warning System Using Visual Information for Both Diurnal and Nocturnal Illumination Conditions, . (2010). *EURASIP Journal on Advances in Signal Processin*.

[3] Jalal, m. (2013). Drowsy Driver Warning System Using Image Processing. *INTERNATIONAL JOURNAL OF ENGINEERING DEVELOPMENT AND RESEARCH* .

[4] Robertson, D. J. (2015). Face Averages Enhance User Recognition for Smartphone Security. *PLoS One.*

[5] Gary Bradski, Adrian Kaehler (2008), Learning openCV:computer vision with the opencv Library. "O'Reilly Media, Inc."

[6] Neetu Saini, Sukhwinder Kaur, Hari Singh(2013). A Review: Face Detection Methods And Algorithms, International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 www.ijert.org Vol. 2 Issue 6.

[7] Smita Tripathi,Varsha Sharma.(2011)," Face Detection using Combined Skin Color Detector and Template Matching Method". International Journal of Computer Applications Volume 26– No.7.

[8] Kirby and Sirovich, 1990. Application of KarhunenLoeve procedure for the characterization of human faces. IEEE Trans. pattern analysis and machine intelligence, 12:103-108.

[9] Longin Jan Latecki, Venugopal Rajagopal, Ari Gross, "Image Retrieval and Reversible Illumination Normalization", SPIEIIS&T Internet Imaging VI, vol. 5670, 2005.

[10] Hemant makwana & Taranpreet singh ,"comparison of different algorithm for face reconition"."global Journal of Computer Science And technology graphics & vision. Volume 13 Issue 9 Version 1.0 Year 2013

[11] Duan-Sheng Chen , Zheng-Kai Liu, Aug 2011. "Generalized Haar-like features for Fast Face Detection". Conference on Machine Learning and Cybernetics, 2007. Hong Kong, pp. 2131-2135

[12] P.Viola and M. Jones "Robust real time object detection"Proceedings of International Journal of Computer Vision, pp.137-154.2004.