Image Processing & Feature Detection

Project report submitted

in partial fulfillment of the requirement for the degree of

Bachelor of Technology (B.Tech.)

By

Beauty Singh (4166454026)

Saloni Mehta (4166454023)

Vaishali Pathak (4166454018)

Under the supervision of

Dr. Amita Kapoor

SHAHEED RAJGURU COLLEGE OF APPLIED SCIENCES FOR WOMEN

Department of Electronics

SHAHEED RAJGURU COLLEGE OF APPLIED SCIENCES FOR WOMEN

University of Delhi

Acknowledgement

We begin our thankfulness by appreciating god, for everything. Our sincere gratitude to our college Principal Dr. Payal Mago, (Shaheed Rajguru College of Applied Sciences for Women, University of Delhi) for providing us an opportunity to undertake such an interesting project.

We would also like to thank all the teachers of Electronics department (Dr. Venika Gupta, Mrs Preeti Singhal, Mrs Monika Tyagi, Dr Neha Katyal, Ms Sonia, Mrs Surbhi, and Mr. Deepak Jaiswal) for their knowledge and encouragement. Also thanking our lab staff for being there, whenever we needed them.

We are very thankful to our mentor, Dr. Amita Kapoor, for her continuous support and guidance in our B.Tech study and research, for her patience, immense knowledge and providing us an excellent environment for doing the work. Mam, we cannot thank you enough!

Lastly, we would like to thank our friends and families for their immense love and moral support.

Certificate of Originality

SHAHEED RAJGURU COLLEGE
OF APPLIED SCIENCES
FOR WOMEN

This is to certify the project entitled "**Image Processing and Feature Detection**" is a bonafide work carried out by the following students, under the guidance of Dr. Amita Kapoor:

Beauty Singh - 4166454026

Saloni Mehta - 4166454023

Vaishali Pathak - 4166454018

The work has been undertaken taking care of engineering ethics. It contains no material previously published or written by any other person or material which has been accepted for the award of any other degree or diploma of any university, except where due acknowledgement has been made in the text.

Signature of Supervisor …….

Name with Designation …….

Date:

TABLE OF CONTENTS

ABSTRACT

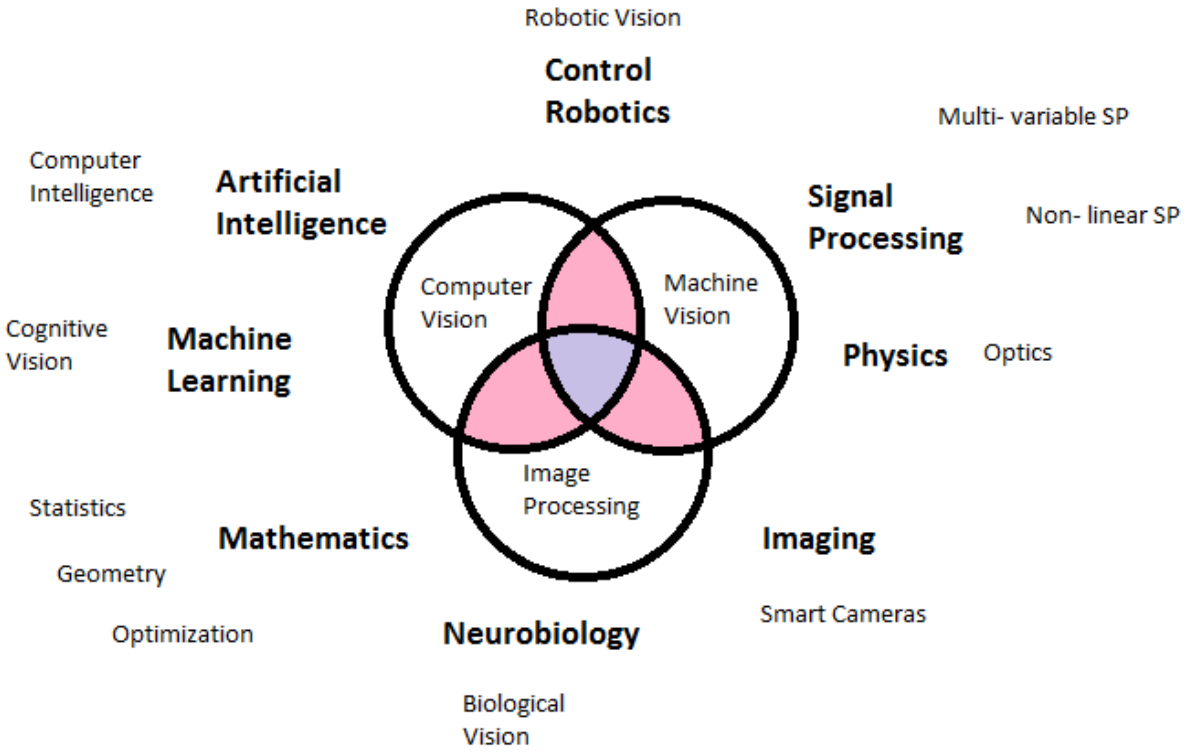
Image Processing and feature detection are the two most important areas of research in Computer Vision. In our one year project we successfully implemented various techniques and coding schemes that deal with Image Processing and Feature Detection of both still images and live video. We started with converting raw images to usable forms like grayscale, plots. We performed various mathematical operations on the image transforms like blurring (Gaussian Transform), sharpening etc. We learned to perform weighted addition of images. On both still images and live video features like face, eyes, smile, upper/lower body and full body have been detected. In this project, we used Haar Classifier to detect human face and eyes. We also employed the use of support vector machines to successfully identify human upper/lower and full body. The work can be extended to pedestrian detection in case of self-driving cars.

INTRODUCTION

An image may be defined as a two-dimensional function *f(x,y)* where *x* and *y* are plane(spatial) coordinates, and amplitude of '*f*' at any pair of coordinates (*x,y*) is called the Intensity or Gray level of the image at that point. When *x,y*, and the amplitude values of '*f*' are all finite, discrete quantities, then the image is called Digital Image. The manipulation of digital image using mathematical procedures with the help of a digital computer is called *Digital Image Processing*.

The digital image is composed of a finite number of elements, each of which has a particular location and values. These elements are referred as picture elements, image elements, pels and pixels.

A branch of artificial intelligence and image processing concerned with processing of images from the real world is computer vision. Computer vision typically requires a combination of low level image processing to enhance the image quality (e.g. remove noise, increase contrast) and higher level pattern recognition and image understanding to recognize features present in the image. The following chart is relation between different fields [1].



Overview of Image processing, Computer Vision and Machine Learning

Face Recognition is one of the most representative and classic application in Computer Vision. Face Detection is the most important step of face recognition. Not only is the face recognition, face detection also a first step in Human Computer Interaction (HCI) systems. Unlike traditional HCI device, keyboard, mouse and display, it provides more effective methods to increase user experience with computer used. Hence, it speeds up human's work.

In biometric approaches, like fingerprints and iris, human faces are unique objects which are widely used in security systems. Many types of personal authentication systems have been developed related to these approaches which take advantage of unique and special characteristics.

Human faces are non-rigid objects and appear in different scale, pose, angle and facial expressions. Human faces always have variation (for example, glasses). In addition, the images have different brightness, contrast. These results are the challenges of the face detection.

The system would be developed by cross- platform components, i.e. the face detection system can be operated in windows platform, Mac system or Linux- related system.

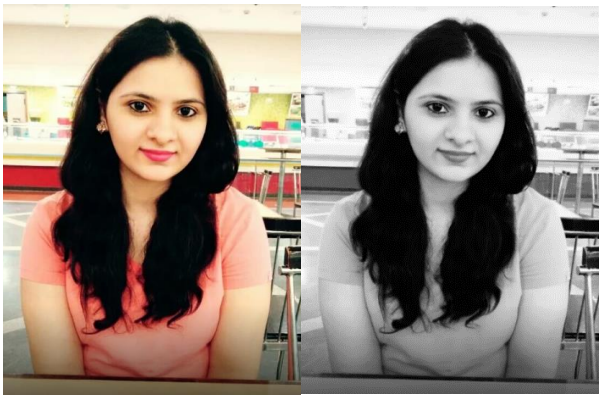
In computer vision and Image processing, the concept of feature detection refers to methods that aim at computing abstractions of image information and making local decisions at every image point whether there is an image feature of a given type at that point or not. Feature detection, description and matching are essential components of various computer vision applications; thus they have received a considerable attention in the last decades. The feature detectors and extractors must have certain properties keeping in mind that the importance of these properties depends on the actual application settings and compromises need to be made. The following properties are important for utilizing a feature detector in computer vision applications:

- Robustness, the feature detection algorithm should be able to detect the same feature locations independent of scaling, rotation, shifting, photometric deformations, compression artifacts, and noise.
- Repeatability, the feature detection algorithm should be able to detect the same features of the same scene or object repeatedly under variety of viewing conditions.
- Accuracy, the feature detection algorithm should accurately localize the image features (same pixel locations), especially for image matching tasks, where precise correspondences are needed to estimate the epipolar geometry.
- Generality, the feature detection algorithm should be able to detect features that can be used in different applications.
- Efficiency, the feature detection algorithm should be able to detect features in new images quickly to support real-time applications.
- Quantity, the feature detection algorithm should be able to detect all or most of the features in the image. Where, the density of detected features should reflect the information content of the image for providing a compact image representation.

Image Processing:

- **Image transforms:**

  Grayscale: Grayscale is a range of shades of gray without apparent color. Intermediate shades of gray are represented by equal brightness levels of the three primary colors (red, green and blue). The reason for differentiating such images from any other sort of color image is that less information needs to be provided for each pixel. In the case of transmitted light (for example, the image on a computer display), the brightness levels of the red (R), green (G) and blue (B) components are each represented as a number from decimal 0 to 255. For every pixel in a RGB grayscale image, R = G = B.



Histogram: A histogram is a graph or plot, which gives and overall idea about the intensity distribution of an image. It is a plot with pixel values (ranging from 0 to 255) in X-axis and corresponding number of pixels in the image on Y-axis.

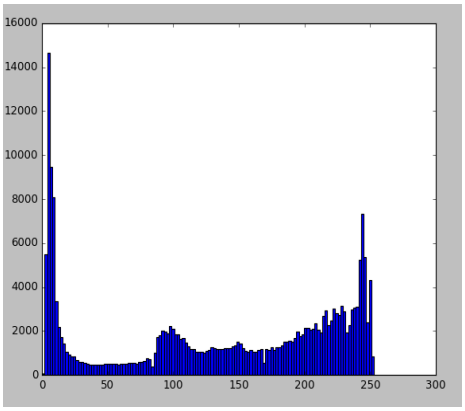
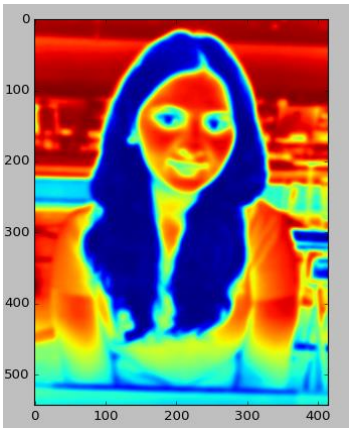


Image Blurring: It is used to define an image scale to work in, for interpolation, for computing interest points etc. An example of image Gaussian blurring is shown



- **Image contours:** Contours are simply curves joining all the continuous points (along the boundary), having same color or intensity. The contour is a useful tool for shape analysis and object detection and recognition. In OpenCV, finding contours can be related to finding white object from black background.

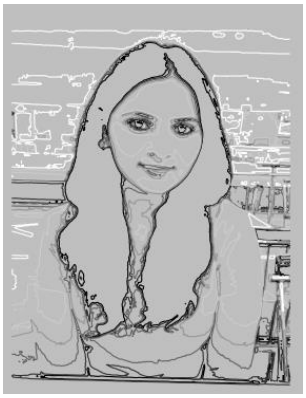


Image blending: This is an image addition, but different weights are given to images so that it gives a feeling of blending or transparency. Images are added as per the equation below:

$$g(x) = (1 - \alpha)f1(x) + \alpha f2(x)$$

where α = 0-1 (weightage)

Original Image 1 Original Image 2 Blended Image

Here, two images have been taken to blend them together. The first image is given a weight of 0.8 and second image is given 0.5.

FEATURE DETECTION USING HAAR CASCADE CLASSIFIERS:

Object Detection using Haar feature-based cascade classifiers is an effective object detection method. It is a machine learning based approach where a cascade function is trained from a lot of positive and negative images. It is then used to detect objects in other images.

Introduction of Haar-like features

A One such method would be the detection of objects from images using features or specific structures of the object in question. However, there was a problem. Working with only image intensities, meaning the RGB pixel values in every single pixel in the image, made feature calculation rather computationally expensive and therefore slow on most platforms. This problem was addressed by the so-called Haar-like features, developed by Viola and Jones on the basis of the proposal by Papageorgiou in 1998. A Haar-like feature considers neighboring 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. An example of this would be the detection of human faces. Commonly, the areas around the eyes are darker than the areas on the cheeks. One example of a Haar-like feature for face detection is therefore a set of two neighboring rectangular areas above the eye and cheek regions [2].

Cascade Classifier

The cascade classifier consists of a list of stages, where each stage consists of a list of weak learners. The system detects objects in question by moving a window over the image. Each stage of the classifier labels the specific region defined by the current location of the window as either positive or negative – positive meaning that an object was found or negative means that the specified object was not found in the image. If the labeling yields a negative result, then the classification of this specific region is hereby complete and the location of the window is moved to the next location. If the labeling gives a positive result, then the region moves of to the next stage of classification. The classifier yields a final verdict of positive, when all the stages, including the last one, yield a result, saying that the object is found in the image. 3 A true positive means that the object in question is indeed in the image and the classifier labels it as such – a positive result. A false positive means that the labeling process falsely determines that the object is located in the image, although it is not. A false negative occurs when the classifier is unable to detect the actual object from the image and a true negative means that a non-object was correctly classifier as not being the object in question. In order to work well, each stage of the cascade must have a low false negative rate, because if the actual object is classified as a non-object, then the classification of that branch stops, with no way to correct the mistake made.

However, each stage can have a relatively high false positive rate, because even if the n^{th} stage classifies the non-object as actually being the object, then this mistake can be fixed in n+1-th and subsequent stages of the classifier [3].

SOFTWARE AND TECHNIQUES USED

We chose Python as the programming language. It was conceived by Guido van Rossum in 1989 [4]. Also, python is one of the three "official languages" in Google which means that applications in Google were deployed in this language. Here are the points that make python selected:

1. Free. Python is open source. People are allowed to use it in business or commercial without any charge.

2. Easy to read. Syntax in Python is clear and readable. Beginners can easily read and handle Python coding very well.

3. High portability. Python can work on different platforms, because Python is written in portable ANSI C.

4. Reusability. Python has easily reused modules and packages. People can develop their own library and reuse it later.

5. Object- Oriented Programming. Unlike scripting languages, Python is designed to be object-oriented. OOP means you can implement it using the concepts of inheritance and polymorphism.

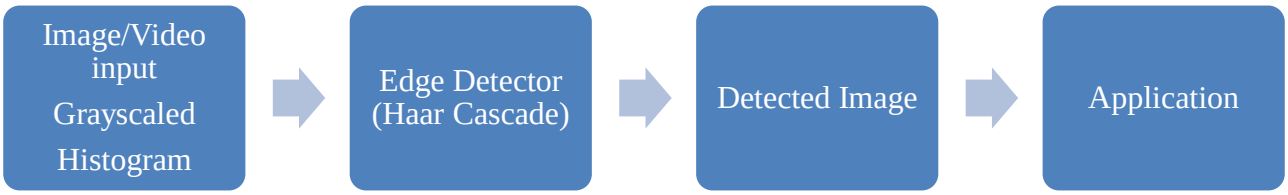
Another tool that we would be using is **Open Computer Vision Library (OpenCV)**, which has at least 500 algorithms, documentation and sample code for real time computer vision. It is originally developed by Intel and launched in 1999. It is free for commercial and research uses. OpenCV library is cross- platform which means it can be executed on Windows, Mac OS x or Linux. The library is mainly written in C, which makes it easier to transfer into specific platforms. Example applications of OpenCV include Human- Computer Interaction, Object Identification, segmentation and Recognition.

Technical Details of the Software used:

- Python 3.5.3
- Anaconda 3.5.3
- Numpy 13.0rc1
- Scipy 0.19.0
- Matplotlib 2.0.2
- OpenCV 3.2.0

The Python Imaging Library (PIL) provides general image handling and lots of useful basic image operations like resizing, cropping, rotating, colour conversion and much more. With PIL, one can read images from most formats and write to the most common ones.

PyCharm 2016.3.2 provides smart code completion, code inspections, on-the-fly error highlighting and quick-fixes, along with automated code refactoring and rich navigation capabilities.



Haar-cascade Detection in OpenCV

OpenCV comes with a trainer as well as detector. Here we are using a detector. OpenCV already contains many pre-trained classifiers for face, eyes, smile etc. Those XML files are stored in OpenCV data.

For face and eye detector, first it is needed to load the required XML classifiers. Then load the input image (or video) in grayscale mode.

Next we find the faces in the image. If faces are found, it returns the positions of detected faces as rectangle(x,y,w,h). Once we get these locations, we can create a Region of Interest (ROI) for the face and apply eye detection on this ROI.

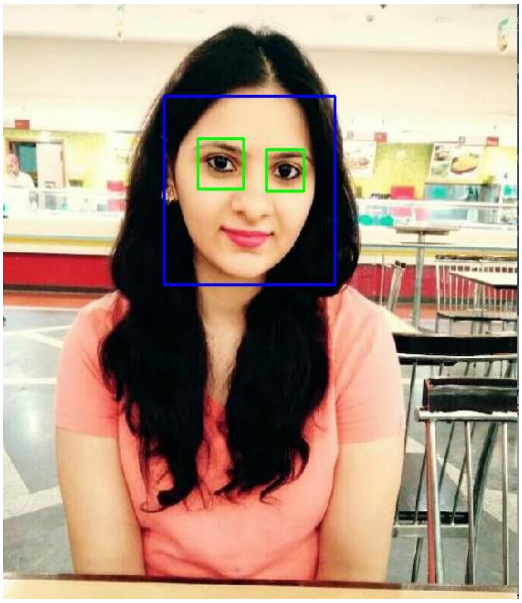
Support Vector Machines

In our project, SVM in Full Body detection is used. Originally, support vector machines (SVM) was a technique for building an optimal binary (2-class) classifier. Later the technique was extended to regression and clustering problems. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall.

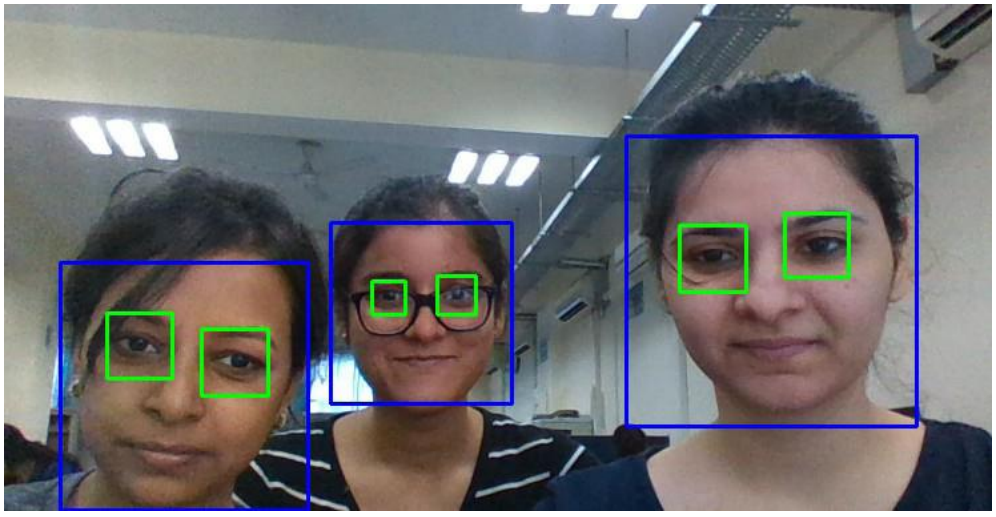
OBSERVATIONS AND RESULTS

Face and Eye Detection

- Face and Eye Detection in an Image



- Face and Eye Detection in a Video

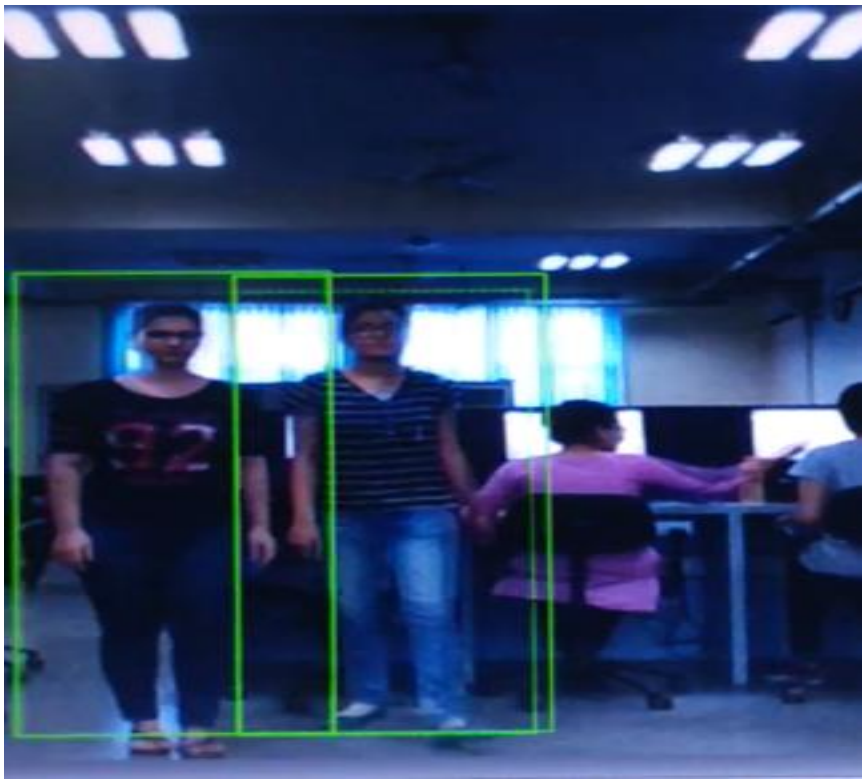


Body Detection

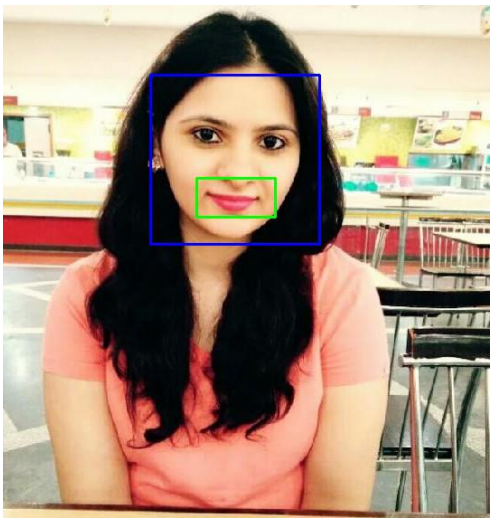
- Upper Body Detection in an Image

- Lower Body Detection in an Image

Full Body Detection in a Video

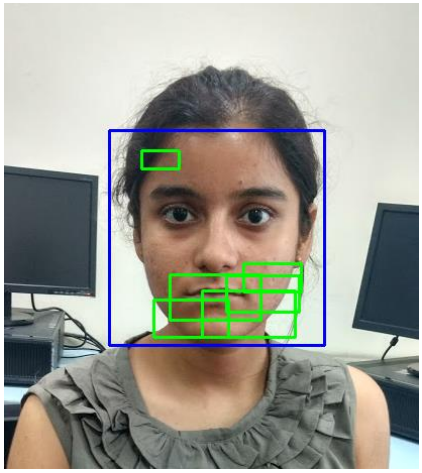


Smile Detection in an Image



Background errors

As shown by the testing process, the classifiers trained can produce errors – either false positives or negatives. Some of them are shown below.



There is a considerable diversity among different sample images, with parameters like spatial position, shape, color, intensity and motion. Many uncontrollable factors such as lighting, weather, or intruders cause unexpected errors, and may affect the efficiency of detection. The system can be improved by various methods like:

- Training the haar cascade for detecting human facial features, such as the mouth, eyes etc.
- Another way to overcome this is by exploring the advantages of the fast saliency-based visual attention to select the image regions which may contain the most relevant parts of video frames i.e., rapidly finding the attractive face regions. This method uses fast spatiotemporal saliency, combined with skin color prefiltering to select specific regions,

and then apply the face detection technique with skin color-post-filtering.

FUTURE WORK

Eye movement Detection

Probably the most common use, as far as consumers are concerned, has been in cameras and smartphones. The aim of this detection has been to help the photographer improve their photographs, by telling them when their subjects have blinked. The blink detection technology focuses on the eyes of people in the photograph (they can often work with up to twenty faces) and whenever a pair of eyes are occluded there will either be a message displayed on the LCD screen telling the photographer to delay taking their photograph, or the more advanced cameras are smart enough to simply snap the photo at a moment when all eyes are open.

Nowadays, people spend more time in front of electronic screens like computers, laptops, TV screens, mobile phones or tablets which cause eye blink frequency to decrease. Each blink spreads the tears on the eye cornea to moisture and disinfects the eye. Reduced blink rate causes eye redness and dryness also known as Dry Eye, which belongs to the major symptoms of the Computer Vision Syndrome. Our future goal is to design eye blink detector which can be used in dry eye prevention system.

Applications in healthcare

Feature detection system can be used as an assisting system for paralyzed people. We can capture and study the eye movement that is blink detection and used to control household devices. The system can be used to control two appliances simultaneously where one eye is used to control one device like bulb and other eye's blink is used to control second device like alarm bell.

Hence, in future, we can present a method based on image processing techniques for detecting human eye blinks and generating inter-eye-blink intervals. We can efficiently track eyes of a person from video image sequences; and propose a new algorithm to analyse the eyelid's states. The analyzed states are then further used for generating an inter-eye-blink interval graph, which can be used to study different eye-related behaviour analyses, e.g., fatigue test for drivers, sleep driving, physical-eye related diseases and lie-detecting process.

It can also be used for Rapid eye movement (REM) analysis of people suffering with sleep disorders.

Face Recognition

Facial Recognition based Password Authentication system can be generally used for a system where security is a major concern for the user. Today's institutions are facing major security issues; consequently, they need several specially trained personnel to attain the desired security. These personnel, as human beings, make mistakes that might affect the level of security. A proposed solution to the aforementioned matter is a Face Recognition Security System, which can detect intruders to restricted or high-security areas, and help in minimizing human error. This system is composed of two parts: hardware part and software part. The hardware part consists of a camera,

while the software part consists of face-detection and face-recognition algorithms software. When a person enters to the zone in question, a series of snapshots are taken by the camera and sent to the software to be analyzed and compared with an existing database of trusted people. An alarm goes off if the user is not recognized.

Facial recognition is widely used because of its benefits. The advantages of facial recognition are that it is not intrusive, can be done from a faraway distance even without the person being aware that he/she is being scanned . Such thing is needed in banks or government offices for example, and this is what makes facial recognition systems better than other biometric techniques in that they can be used for surveillance purposes like searching for wanted criminals, suspected terrorists, or missing children.

Pedestrian Detection

The presence of cameras in our surroundings grows every day, allowing object detection, and more specifically pedestrian detection, to be used in a broad amount of applications. Possible applications include surveillance applications like traffic safety, blurring pedestrians for privacy issues and human-robot interaction.

An Intelligent Pedestrian Detector (IPD) provides real-time information to the Traffic Signal regarding the number of pedestrians waiting to cross, detected via the IPD, as they approach the crossing and they enter the detection area. The Traffic Signal extends the pedestrian green phase based on how many people are waiting to cross or on the number of still crossing pedestrians.

It has an obvious extension to automotive applications due to the potential for improving safety systems. It can also be applied to the concept of “Self driving Cars”

CONCLUSION

In this project, continued for a period of one year, we have worked on Feature Detection methods used in computer vision. Different image processing techniques were studied and successfully implemented. Using Haar cascade filters and SVMs the eyes, face and full body have been detected. The efficiency of outcome depends greatly upon the cascade filter algorithm used. This could be used in various applications like the full body detection technique can be used in pedestrian detection, eye and face detection can be used in iris recognition, authentication systems. The developed project is a good starting point for expanding image processing in a safety and security system and also in helping the paralyzed people by using various eye related behavior analysis.

BIBLIOGRAPHY

[1] Computer Vision – Wikipedia, free encyclopedia,

https://en.wikipedia.org/wiki/Computer_vision

[2] V. Jones, "Rapid object detection using a boosted cascade of simple features," Computer Vision and Pattern Recognition, 2001.

[3] T. M. Inc., "Train a Cascade Object Detector," [Online]. Available: http://www.mathworks.se/help/ vision/ug/train-a-cascadeobject-detector.html#btugex8. [Accessed Nov 2014].

[4] Face Detection – Wikipedia, free encyclopedia,

https://en.wikipedia.org/wiki/Face_detection

[5] Manisha M. Kasar, Debnath Bhattacharyya and Tai-hoon Kim, "Face Recognition Using Neural Network: A Review", International Journal of Security and Its Applications, vol. 10, no. 3 (2016), pp.81-100.

[6] Henry A. Rowley, Shumeet Baluja and Takeo Kanade, "Neural Network- Based Face Detection", Computer Vision and Pattern Recognition, 1996.

[7] M.Nandini, P.Bhargavi, G.Raja Sekhar, "Face Recognition Using Neural Networks",International Journal of Scientific and Research Publications, volume 3, Issue 3, March 2013.

[8] Thai Hoang Le, "Applying Artificial Neural Networks for Face Recognition", Advances in Artificial Neural Systems Volume 2011, Article ID 673016, 16 pages.

[9] Michel Owayjan, Amer Dergham, Gerges Haber, Nidal Fakih, Ahmad Hamoush, Elie Abdo, "Face Recognition Security System" American University of Science and Technology (AUST), Departments of Computer and Communications Engineering and Computer Science