

Design & Implementation of an Assistive Device for Vision Impaired People

¹Md. Rakib Ahsan

Dept. of Electrical & Electronic Engineering
American International University-Bangladesh
408/1, Kuratoli, Khilkhet, Dhaka 1229, Bangladesh

²Sheikh Zarif Ahmad

Dept. of Electrical & Electronic Engineering
American International University-Bangladesh
408/1, Kuratoli, Khilkhet, Dhaka 1229, Bangladesh

³Mohammad Shamsul Arefin

Dept. of Electrical & Electronic Engineering
American International University-Bangladesh
408/1, Kuratoli, Khilkhet, Dhaka 1229, Bangladesh

Abstract:- Vision imparity has been a major limitation in the past. Several systems have been developed to allow the visually impaired people to easily interact with the world. The very first one was white stick which is a three-section aluminum cane. By the help of the stick vision impaired people could walk safely outside but there were limitations. After that a sonar device had been invented. A transmitter emits a sound wave which hits an obstacle. The reflected wave is received by a receiver and the interval between the transmitted time and receiving time is fed into a processor. The processor then provides the distance between the device and the obstacle as output, by this way a sonar device works. In this paper, an advance system has described which developed by a system using image processing techniques which may allow the user to communicate and interact with their surroundings with greater ease. Using an open source computer vision library, known as OpenCV, it was possible to create such a system which will act as an assistive device.

Keywords:- Vision imparity, OpenCV, Image processing, Face detection system, Python, NUC kit minicomputer.

I. INTRODUCTION

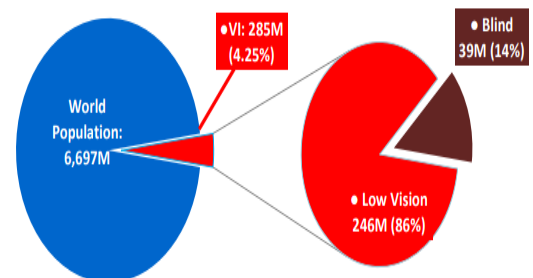
Image processing has been a very useful system for identification and detection of objects from afar using a camera or telescopic vision. There has been a significant increase in the amount of interest in image processing and recognition, full-color image processing, image morphology and computer neural networks in the past five years [1]. The research has started ever since the limitation of the normal camera has been discovered. Most of the advancements were done in the Jet Propulsion Laboratory located in Pasadena, California. The main idea was to gather as much information as possible from a captured image. This has led to the modernization of the system which allowed us to understand more and extract images of distant planets which was impossible using normal cameras. The more it has been developed, the applications of image processing grew wider.

Now it is being used in the assistive device and the working process of the device is successful in identifying the person whose data has been stored in an offline database and after that proper identification of several objects. At last the text is converted to speech mode and played through a headset of the visually impaired people.

II. SECNARIO

According to World Health Organization's (WHO) data in 2010, it have been estimated that there are 285 million individuals who are suffering from visual impairment all over the world. This is shown in the chart below:

Global Estimates: 285 million people are affected by visual impairment*.



80% of Visual Impairment is avoidable
(can be prevented or cured)

Fig 1:- Number of Visual impairments (VI) as studied at 2010 by WHO [2].

Out of these individuals, 39 million are blind and 82% of the blind people are 50 years old or above. Another research data done by experts in Bangladesh states that 750 thousand people suffers from blindness and 250 thousand more are on the verge of losing eyesight. The main causes of blindness is as shown:

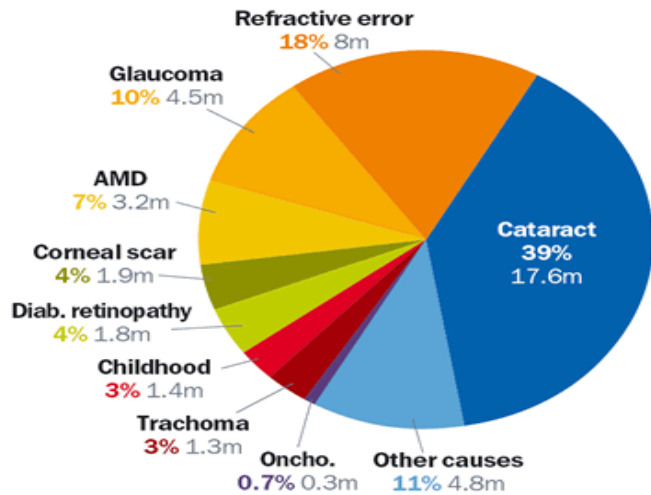


Fig 2:- Major causes of blindness and visual imparity. [3]

According to the chart, it can be seen that cataracts are the main cause of blindness. This problem develops very slowly and turns the lens of the eye blurred and white. Several surgical methods have been developed to counter the problems, however, occasionally it have been observed that the vision could not be regained even after surgery. The “Other causes” involve accidents where the damage is often permanent. Some of the statistic reports state that the problems of visual imparity may increase in future.

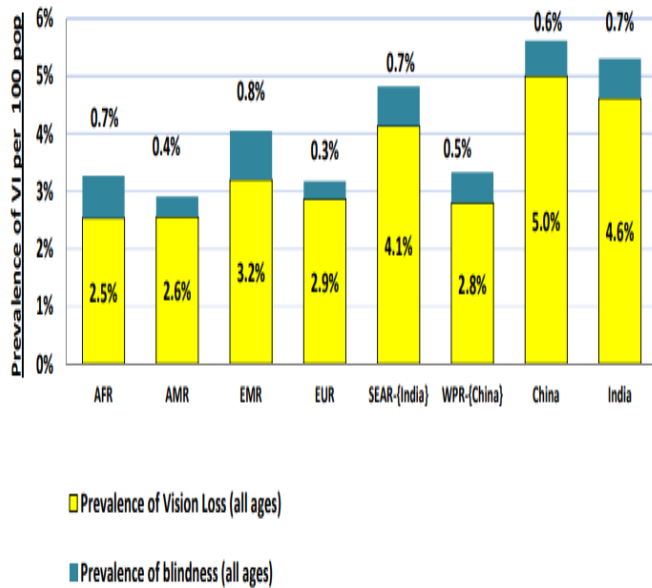


Fig 3:- Distribution of visual imparity around the world. [2]

There are several existing systems that allows a blind person to interact with the environment, however, they have their limitations. For this reason it is important to research on technologies that will allow people with permanent loss of vision to respond with the environment as efficiently as possible.

III. SYSTEM OPERATION

The whole system operates in three main stages. First one is camera which is working as the input .The main working principle of the camera is capturing photos. Secondly, process the input frames by the help of Intel NUC Kit Mini-computer. It does the whole work like image processing, calculation, detection and recognition which shown “Basic procedure for processing an image” part in fig 4. Headset is the last stage of the system to give the outputs. It converts text to speech and delivers to person’s ear. Image processing, Face detection system and text to speech have been described in below:

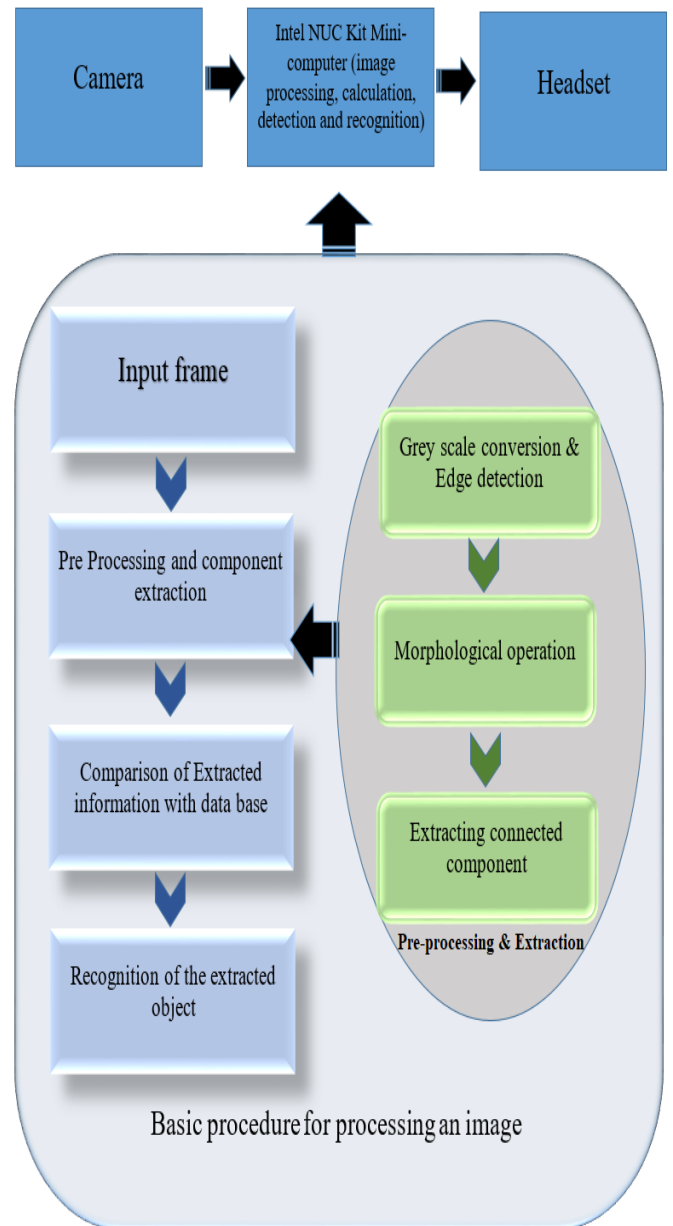


Fig 4:- Block diagram of the overall system

A. Image Processing

First, the data is taken from an input frame and passed through filters which allow faster processing of the image. The image is converted to a greyscale mode which removes color gradients from the frame for simplicity of analysis. This is necessary because, it is observed that the light intensity is a variable factor in almost every day due to change in the atmosphere, for example, clouds may interfere with sunlight or often change the visual appearance of an object. Converting the image to greyscale eliminates such possibility and makes processing much easier. Another important factor for image recognition is the pattern of the borderlines. For this reason, the grey-scaled image is passed through canny edge filter. These modifications done on an image can be termed as “morphological operations” [4]. Several other manipulations, such as contrast, gradient, and smooth edge conversions are also done while refining the image for proper justification. When the desired component has been extracted they are compared with some predefined information, for example, for facial recognition, some predefined face structure stored in the database is the predefined information. Lastly, the user-defined decision is made according to results of processing that took place.

➤ **Canny Edge Detection**

Canny edge detection is an algorithm that detects border or edges of a frame with multiple steps [5]. This algorithm is necessary to remove anomalies that may make it difficult to recognize an image, which is performed after greyscale conversion of an image (pre-processing & extraction part of Fig 4). In still images there is less possibility of noise, however, in real-time mode, there are higher chances of observing interference. Canny edge allows avoiding this errors as shown in the figure below:



Fig 6:- Original image converted to edge detection mode [6]

➤ **Morphological Operations in OpenCV**

Morphological operations are mainly simple manipulations done on an image to improve its details. These operations are done for improvement of the extracted components. The basic morphing techniques involve Erosion, Dilation, Thickening and several other functions. Some of these operations are explained below:

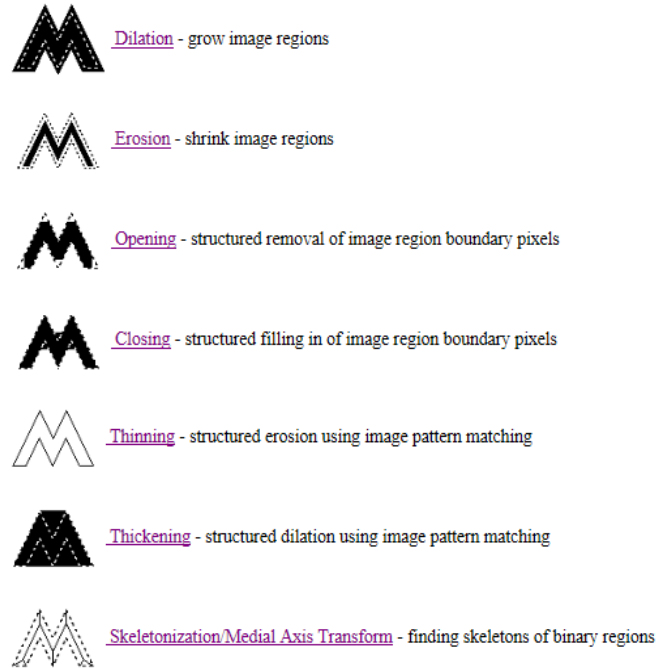


Fig 5:- Basic Morphological operations of OpenCV [7]

B. Face Detection System

Face detection and recognition system mainly define a computer system that is capable of identification of an individual’s facial construct and responds according to a defined algorithm. This system is widely used in surveillance cameras and on reconnaissance satellites to identify any individual which can be seen that the image is directly fed into the system processor. The processor performs several algorithms as mentioned before to convert the image into a binary format (black and white color). After refining the image the processor communicates with the database to collect information related to the input. After that, using some predefined functions, the input is matched with any values stored in the database. Then the processor sends out an appropriate response and displays it in a monitor as output.

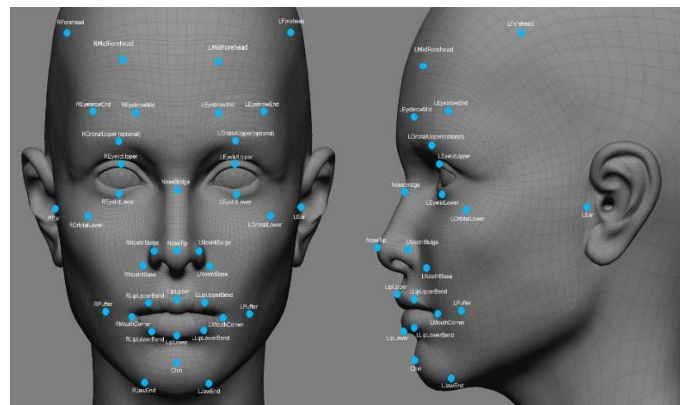


Fig 7:- Some key features of the face that can be used for facial detection and recognition. [8]

➤ *Eigen Faces*

This system defines a method which involves matching faces based on the appearance. Firstly as many samples defined by a user is taken. These images are converted into the low pixel array. Then using all the data a ghost-like image forms which are used as a reference image under a specified identity. During the matching process, the important features that are matched would be the structure of eyes, noses, and lips.

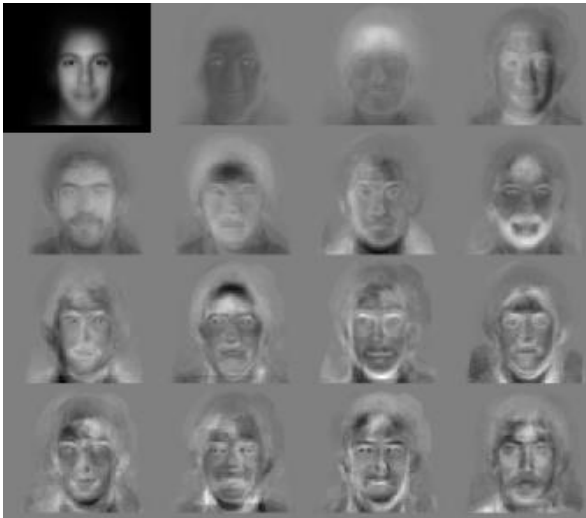


Fig 8:- Eigen faces of a sample image (top left corner) [8]

➤ *Line Edge Mapping*

Line edge mapping is a simple process where the face image is directly into edge detection mode. Then the image is refined such that only the edges are visible as shown in Fig 9.



Fig 9:- Result of converting a face structure into line edge mapping [9]

C. *Text To Speech*

After detection comes to the recognition phase of the system. When a face is detected it is compared with a database as mentioned before. The resulting output is displayed on a monitor with a label that shows the individuals name. To make the computer read this label aloud, there are several systems that are being used nowadays. A noticeable example is the “Narrator” that is a built-in program in Windows operating systems. The user can run this program, which will then read out anything pointed with the cursor. For the project, “win32com.client” has been used in order to convert the text appearing on a screen to speech. This “win32com.client” package contains several modules which can be used in the automation procedures of any system. Using this package the device can be able to read out any text that shall appear on a screen and the Text-to-speech system starts to function.

IV. PROGRAMMING SYSTEM

In this section, the programming components will be discussed. The system will use OpenCV to perform the recognition procedure. OpenCV is an Open Source Software (OSS) that have been developed by Gray Bradsky, a computer vision engineer at Intel. [10] This software was developed in the early 2000s. Initially it was constructed using C++ programming language, now this software supports Python and Java as well. The basic function of this software is:

1. Load information from an image, video or capturing device.
2. Process the image and extract necessary information.
3. Make a decision following the algorithms set by user

The prime benefit of using OpenCV is that it already contains the large and complex algorithms required for image processing. As a result the user needs to code what needs to be done only. The user defines the algorithm for the decision-making procedures.

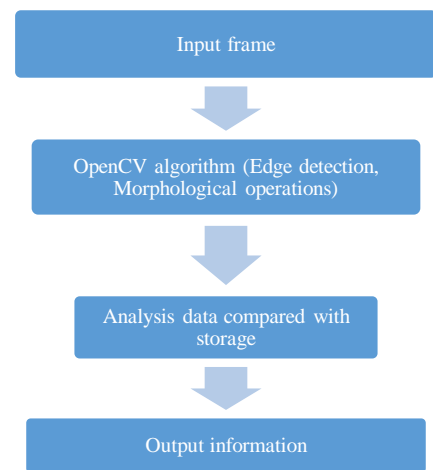


Fig 10:- Operation flow diagram of OpenCV

As mentioned earlier OpenCV supports Python Java language as well. Python is described as a high level all-purpose programming language. The initially Python language was a concept provided by Guido van Rossum in the late 1980 when he was a member of the National Research Institute of Mathematics and Computer Science. The aim of developing python was to develop a programming language as close to English language as possible. Later in 1991, Python 0.9.0 was released. The name “Python” was given after a British TV show named Monty Python. Then in 2000, Python 2.0 was officially released and was developed as an open source project by the members of the National Research Institute of Mathematics and Computer Science. Finally Python 3.0 was released in 2008 which is very similar to Python 2.0 with few extra functions. [11] OpenCV-python is a library designed to work out the problems of computer vision. Numpy is another highly optimized library in OpenCV-python which allows to process an image more easily. [12] It also allows the user to easily connect other libraries such as, SciPy and Matplotlib, which makes the image processing applications more efficient.

V. HARDWARE IMPLEMENTATION

The proposed module has been simulated in OpenCV and python based environment. For equipment usage, we utilized NUC computer kit which has an Intel core i5 7th generation processor, a camera which has a capacity of 1080p video streaming and a power supplier to boot all the device up. The implementation part of the whole system is given below:



Fig 11:- Intel NUC kit and camera powered up and functioning together

VI. RESULTS

At first, we have to give the command to the camera to take some sample data of the face that need to be recognized. And whenever we give instruction to the nuc kit, it will make the camera take input. This input frame will be processed and analyzed for recognition purpose.

A. Person Recognition

In figure 12 a bunch of sample of recognized person has been taken and from those, the system can compare the object to detect.



Fig 12:- Sample images formed while placing an entry in the database

After having the sample data in the library, we need to train the system that this particular sample would be named at this and after training process, OpenCV can recognize the face if it comes in front of the camera. Figure 13 will show the output of a single face has been recognized after gathering the sample of the face and trained the system.

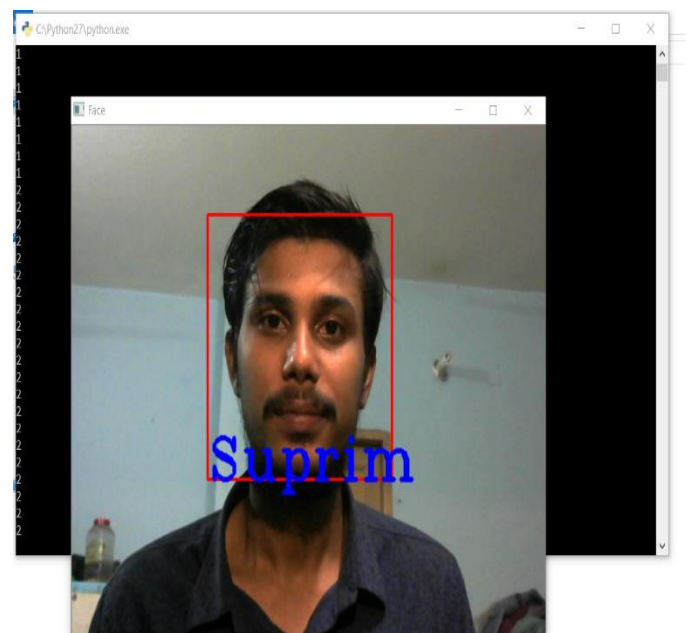


Fig 13:- Recognition of the person whose information was stored in the memory

Like this, OpenCV can also detect multiple faces which has been added in the memory as shown in figure 14.

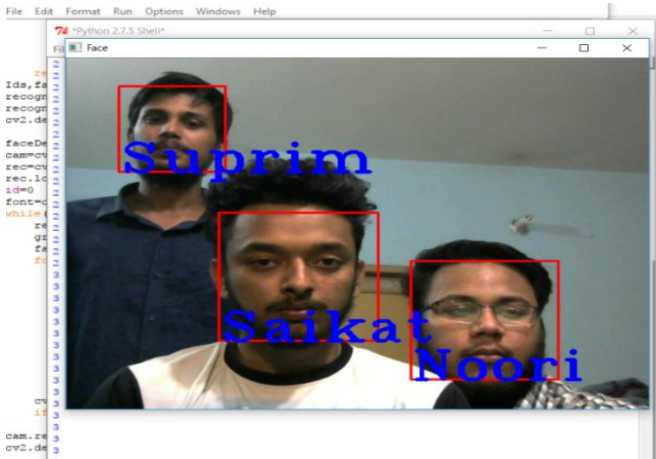


Fig 14:- Recognition of multiple faces

And to have all of the people known the sample directory look like figure 15 given below:



Fig 15:- Appearance of the directory created in the memory unit

Besides, an instruction has been given to the system for the unknown person. If an unknown person crossed by and if he/she has been detected by the camera then it will automatically show an output given in figure 16.



Fig 16:- Presence of unknown identity detected by the device

B. Object Recognition

Besides face recognition, it can also do object recognition. There's a MobileNetSSD_deploy caffemodel which has a bunch of sample in it to recognize pet, vehicle, person, bottle, tv monitor etc. In the following pictures a bunch of detected object has been shown:

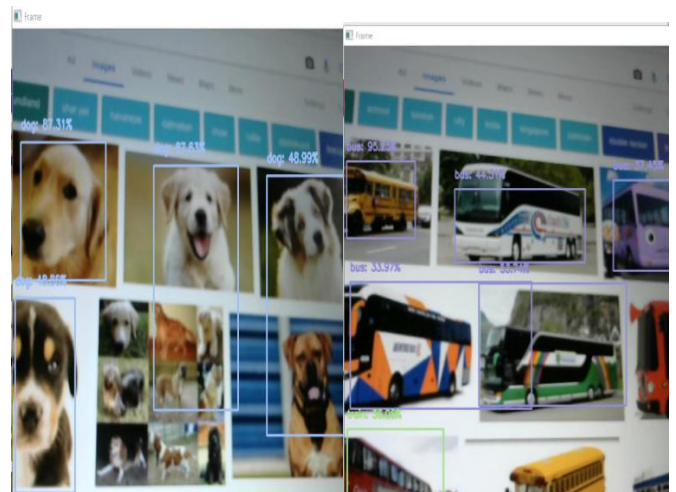


Fig 17:- Identification of non-human identities

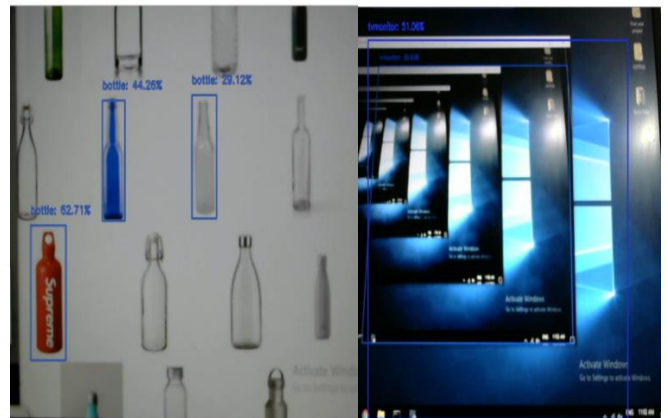


Fig 18:- Identification of objects



Fig 19:- Several others object detections

VII. COST ANALYSIS

Assistive device for the visually impaired people refers to give an identity of an object and subject. For assembling the device several kinds of components have been used. The hardware element used to install all the software elements is the Intel NUC kit minicomputer, which allows to freely use Windows 10 systems efficiently. A camera has been used, which acts as the "eye" of the system. The battery would power the system up so the device can operate remotely. However, there may still be room for improvements.

Equipment name	Quantity	Unit price	Total price(TK)
Intel NUC kit	1	40,000	40000
Camera	1	1800	1800
Battery	1	3000	3000
Head set	1	3000	3000
Bag	1	600	600
Total			48,400

Table 1:- Total cost of the components.

VIII. CONCLUSION

Assistive device for vision impaired individuals intends to give visual assistance to those individuals who are unable to see. This device can help those individuals to identify any object using Mobile Net SSD database. Furthermore, it can recognize the face, which is particularly stored in the database. This device additionally can recognize the unknown appearances as well. Image processing is well-established all through the world by using Matlab, Java and C/C++. We are working with OpenCV, because of the large number of algorithms in its library through which an instructor can spread-out his thought effectively. The motive of our project is to design a simple assistive device for the individuals who are not able to see anything. This device takes the input from the surroundings and gives output as voice, so they can recognize who or what is in front of him. Beside it there is a great future scope of this study. The developments of these systems may lead to advanced applications, such as, the developments of a bionic eye, automated guarding systems, and androids that make decisions by observation of their surroundings.

REFERENCES

- [1]. Uni Assignment Centre, The History of Image Processing information technology, [Visited February 22, 2018] Available at : <https://www.uniassignment.com/essay-samples/information-technology/thehistory-of-image-processing-information-technology-essay.php>
- [2]. https://www.who.int/blindness/data_maps/VIFACTSH-EETGLODAT2010full.pdf?ua=1
- [3]. <https://www.cehjournal.org/article/changing-patterns-in-global-blindness-1988-2008/>
- [4]. Morphological Transformation [Visited March 19, 2018] Available at: https://docs.opencv.org/trunk/d9/d61/tutorial_py_morphological_ops.html
- [5]. Canny Edge Detection [Visited March 20, 2018] Available at: <http://fourier.eng.hmc.edu/e161/lectures/canny/node.html>
- [6]. MathWorks, Edge Detection [Visited March 21, 2018] Available at: <https://www.mathworks.com/disc-overly/edge-detection.html>
- [7]. Morphology [Visited March 22, 2018] Available at: <http://homepages.inf.ed.ac.uk/rbf/HIPR2/morops.htm>
- [8]. The Catholic University of America [Visited March 22, 2018] Available at: <http://www.law.edu/2015-Spring/New-Laws-Needed-to-Brake-Use-of-Facial-Recognition-Technology-Argues-Student-Scholar.cfm>
- [9]. Science & Education Publishing, Face Recognition Using Line edge mapping approach [Visited March 23, 2018] Available at: <http://pubs.sciepub.com/ajjee/1/3/4/index.html>
- [10]. <https://www.oreilly.com/library/view/mastering-opencv-4/9781789533576/2de0893d-e450-417d-b336-a4143799b43d.xhtml>
- [11]. <https://medium.com/@johnwolfe820/a-brief-history-of-python-ca2fa1f2e99e>
- [12]. https://docs.opencv.org/trunk/d0/de3/tutorial_py_intro.html