

## Vision substitution using audio-tactual sensation

Dinesh R. Damodar<sup>\*</sup>, Umang V. Suthar and Hemang D. Solanki

Department of Biomedical Engineering, Gujarat Technological University, Ahmedabad, Gujarat, India

©2018 ACCENTS

### Abstract

*In the contemporary era, technologies available for blind individuals can assist them in narrow and short range while they lack amenities to feel, observe and identify long range objects. Hence they are still in trouble of conventional method using blind stick, laser cane, vibrators etc. The proposed technique implement camera to determine shape and size of object, pose of person by extracting its dimensions using object detection and pose estimation method, converting visual details in auditory and feed in auditory output simultaneously with vibration matrix for accurate determining of object with its location for assisting in terms of far sight vision. In short developing a neural network based artificial intelligence unit which can continue adding up significant information of new vectors, images and speak out to the user whenever he/she need. It gives accurate and precise control to become more and more independent every day with practice and guidance.*

### Keywords

*Vision substitution, Vision augmentation, Neural networks, Artificial intelligence, Image recognition, Pose estimation, Electro-oculography.*

### 1.Introduction

Vision is integral part of Human body plays pivotal role in vision, reacting and balancing of entire body. Encountering loss of vision begin with refractory errors consequences in certain disorders like retinal detachment, glaucoma, cataract, tumors which might have detrimental impact on individual if not cure on time. Probably they need spectrum if only suffering with partial vision impairment but ophthalmic surgeries are needed where individual is entirely blind and unable to percept anything. Traditional innovations propose use of braille, tactile substitution, ultrasonic cane, laser cane to guide blind people to estimate about obstacle encounter in their pathway considerably classified as object comes in short sights or nearby circumstances and has poor or negligible performance for visualize far-sight scenario. It only alerts about obstacles while exempting details of that shape, size is in appearance.

It is strongly desired by blind individual to feel scenario as it is rather than just approximate alerts of obstacle. Previous technologies have been become very popular but only in case of short- sighted object. In case of predicting details of far sight object they have no choices, they need to combat and compromise with it.

After taking the reviews of few blind individual we realize and found kind of problems faced by them and thus promote us to substitute vision of individual with auditory and vibrational matrix to provide more detailed information to assist blind one. The main objective and ambition behind employing this method is to assist blind user with more privileges to let him became more familiar with physical environment and collaborate with it.

### 2.Research methodology

Audio-visual substitution method is explained by the following block-diagram (*Figure 1*). The purpose is to design it in aesthetic design by considering from aspect of blind individual. This method deals by processing and extracting details of real time graphics and converting this details in to auditory and vibration to let blind one know about relevant image with its spot prediction.

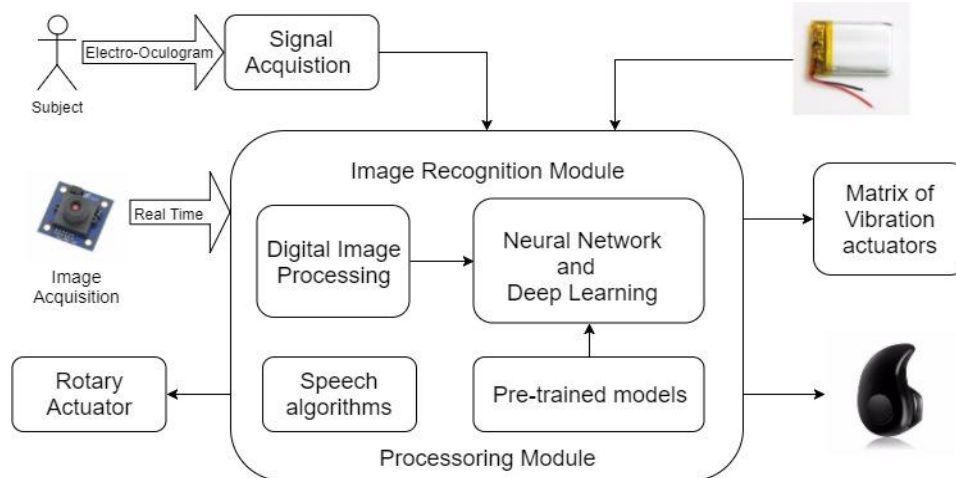
1) Real time graphics are capture with the help of image sensors. Acquiring and processing real time image to extract features and identify object and relevant information.

2) Controller has been employed to gather real time graphics since it has been fascinated with inbuilt RAM, memory, easy accessibility of required peripheral and environment to create algorithm of image identification, speech and pattern of vibration to provide it as output.

<sup>\*</sup>Author for correspondence

3) Electro-Oculogram signals are captured by placing electrodes on the frontal lobe of the head, driven by signal acquisition circuit comprises of high gain instrumentation amplifier having appropriate common mode rejection ratio, filter configurations to minimize

noise and maintain signal to noise ratio. This signal are purposed to control the direction of camera lens [1].



**Figure 1** Block diagram of audio-visual substitution method

4) Visual output of camera need to bifurcate in terms of pixel so only limited view can be processed in present time to avoid escalated useless information interference. Camera has settled with limited perception using controller to acquire real time images purposes to reduce interference of overloaded object in descriptive manner. Rotation of camera is driven by EOG consequences of blind person. Rotary actuator setup is requires to let blind people concentrate on particular direction without moving his head similar to normal individual.

- Even though individual is blind and his sight has not working, human muscle of eyes can get contract and relax based on action of user.
- Consequently if user is willing to see left-hand side, his eyes will contract and generate pattern of EOG signals which can be acquired with EOG acquisition circuit and fed to controller for further actions of comparing and computing with pre-measured reference thresholds.
- In this way, horizontal and vertical movement of the eye can be detected using electrooculogram to ensure about direction which blind user is trying to concentrate

5) Image recognition module consistently driven by real time graphics which has undergone through a convoluted system of image processing, neural network and deep learning.

- Converting real time data captured by camera to digital data so arithmetic and logical operation can be done to convert type of image, execute image segmentation and edge detection method to create classified data of image in terms of cluster so that can get compared with information associated with pre-trained module.
- Developing models with data sets of pre-assumed information which has become more frequent involves car, traffic signal, mobile phones, remote, image pattern and vector information of shape. Similarly to identify the pose of human approaching non-parametric representation recognized as Part Affinity Fields (PAFs) to learn associate body parts of individual in real time has become possible [2].
- Recognizing image even with already available data of image information and statistics is difficult because during acquiring every time same image may appear with alters in some degree. Many of times background of the image alters according to place and occasions as well due to motion which need to subtract from image using frame differencing method or traditional background subtraction method [3].
- Introducing neural network and deep learning let controller make self-computation for predictions by making each artificial neural layer. Different protocols of layers discriminate different information of data percept in real time like

- brightness, resolution, type, contrast etcetera. Therefore appearance of image out of pre-trained model will be computed to determine and assign with appropriate notations and category [4].
- 6) Since image is 2D representation of data hence occupies large source of information which need to compute spontaneous according with acquisition, platform supporting high performance numerical calculations required to achieve optimum efficiency in real time.
  - 7) Matched image vector consequences in output which is feed to further speech algorithms module to convert assigned output in auditory at the same time stimulating two matrix of vibrator. Each matrix of vibrator has designed to give sensation about spot of object. Two matrixes are utilized to classify and clarify about accurate horizontal and vertical geographical view.
  - 8) Optical character recognition employs the same method as extracting featured information from camera so its output can be fed to the auditory canal through headset to fulfill the purpose of read. Since it occupies another set of memory in controller to employ standalone OCR it consume



**Figure 2** Training module

This array includes some of the pictures that we had included in our pre-processing model, from certain possible angles.

- 2) Post-processing: It employs a method of K-means clustering which is unsupervised method which let identify and assign unlabeled data in particular activity when it is not matching with pre-processing contents. So if any new angle view introduction will get forced by k-means clustering so particular object can be compared and assigned with group of data in autonomous manner.

### 3.Result and discussion

This image has taken from stationary mode as to elaborate about output, and fed to controller which has been trained with this all object as pre-processing and able to identify. Once controller get identify of this all object, its start speaking labels of all this in to the auditory of blind simultaneously with vibrational sensation on matrix in according manner. Initiating

more power, hence server based OCR can be used to keep performance steady and stick to main goal of ambition to collaborate with physical environment [5].

- 9) Battery used to eliminate the need of wired power supply so this module can work standalone and as a portable unit.

#### 2.1 Training data module

Train controller to identify and classify objects can be divided in two distinct sects.

- 1) Pre-processing: It includes utilization of large variety of pictorial data with major possible alteration in its resultant view that can be encounter from various angles and view. To exemplify, we have pre-fed images of cell phone to the controller, but even change in small degree of angle it will not get matched with pre-fed images consequently it will be unable to identify the images. It is shown in *Figure 2*.

with upper right corner, it will continue to speak its name and at the same time giving its estimation of degree of angle of spotting's in the form of vibrational sense on matrix. *Figure 3(a)* shows the normal things comes in daily usage. *Figure 3(b)* shows the object detection with its accuracy. *Figure 4(a)* shows the natural scenario to elaborate blind. *Figure 4(b)* shows the human detection with its accuracy.

All this identification of materialistic world has facilitate with audio output using small headset, so whenever any material will get identify it will be fed to speech converter embedded with algorithms and then to the auditory canal of blind individual. It will assist in imaging all the scenario and deal with that. *Figure 4(c)* gives illustration about posture estimation of all persons are in range of camera revealing their activity of playing cricket with their heat maps and vector maps . It has further use in identifying activity that has going on nearby to exemplify, if any person is

offer his hand willing to handshake with blind person, Real time expression and posture of that person is capture by camera and fed to raspberry pi and as a consequences it gives auditory in the ear of blind one.

A support of batteries gives amenities to make it portable and carry whenever required to desired place.



Figure 3(a) Normal things comes in daily usage



Figure 3(b) Object detection with its accuracy



Figure 4(a) Natural scenario to elaborate blind



Figure 4(b) Human detection with its accuracy

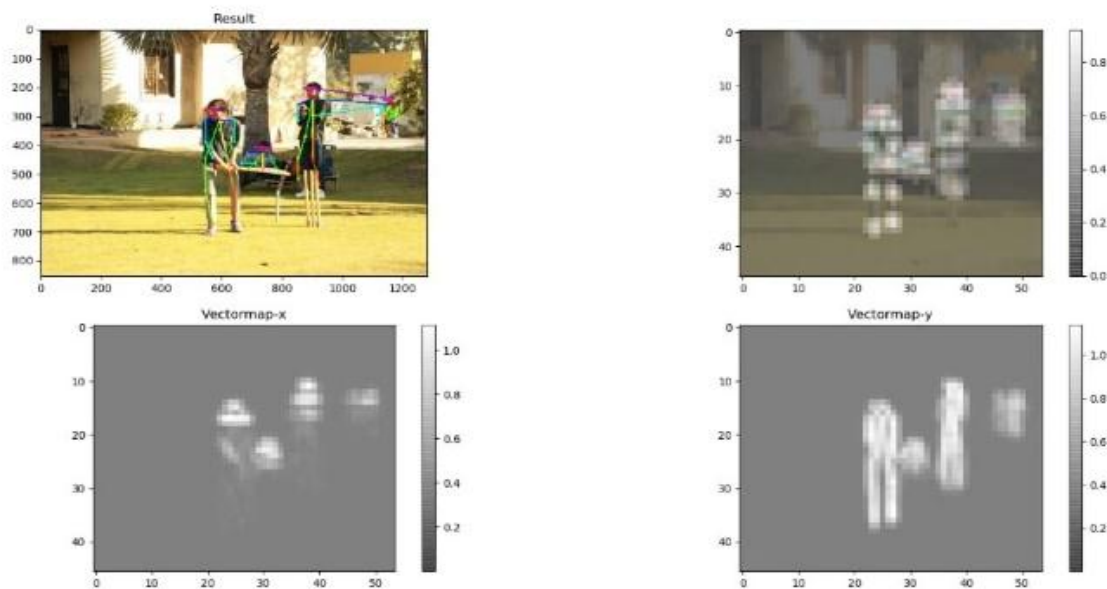


Figure 4 (c) Illustrations about currency estimates of all individuals are in the limit of the camera



#### 4. Conclusion and future work

Vision substitution is novel way to let blind individual relate with real and actual world to navigate far sight path, imagine objects and pose estimation of human to give relevant expression according to. It can be achieve by using capturing real time visuals and extract its sub contents in terms of object, pose identification technique to identify and fed it to auditory by converting in audio/speech simultaneously on two matrix of vibrator placed on shoulder of blind individual to estimate about spot location of object.

Future work reveals that implementing complex algorithms to give access in native languages of blind user in which user is habitat to collaborate so he didn't need to go through learning stages of new language "preferably English" which has implemented in this technique.

#### Acknowledgment

We are thankful to Arpit D. Shrimankar, student at government engineering college, Gandhinagar has contributed his tremendous skill and knowledge in image processing and neural networks to achieve some of results.

#### Conflicts of interest

The authors have no conflicts of interest to declare.

#### References

- [1] Raman SK, Jayaram JK, Murugan S, Saha A, Kavitha RK. Design of a robust method to acquire EOG signals using Bio-medical signal processing. In online international conference on green engineering and technologies 2016 (pp. 1-4). IEEE.
- [2] Cao Z, Simon T, Wei SE, Sheikh Y. Realtime multi-person 2d pose estimation using part affinity fields. arXiv Preprint arXiv:1611.08050. 2016:7291-9.
- [3] Kartika I, Mohamed SS. Frame differencing with post-processing techniques for moving object detection in outdoor environment. In international colloquium on signal processing and its applications 2011 (pp. 172-6). IEEE.

- [4] Chen Q, Song Z, Dong J, Huang Z, Hua Y, Yan S. Contextualizing object detection and classification. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2015; 37(1):13-27.
- [5] Mantoro T, Sobri AM, Usino W. Optical character recognition (ocr) performance in server-based mobile environment. In international conference on advanced computer science applications and technologies 2013 (pp. 423-8). IEEE.



**Mr. Dinesh R. Damodar** currently pursuing his skills in research as a independent researcher and self-employed. He has pursued his bachelors in Biomedical engineering from Gujarat Technological University particularly from Government Engineering College. He is aspiring towards healthcare industry innovation. Major area of interest are Rehabilitaion Instrumentaton, Neuroscience, Biomechanis, Biomedical Instrumenation, Artificial Intelligence, Robotics, Machine learning.  
Email: damodardinesh109@gmail.com



**Mr. Umang V. Suthar** currently working as researcher and has just pursued his Bachelors of Engineering in Biomedical from L.D.College of Engineering, Ahmedabad, Gujarat. He completed his Diploma in Biomedical Engineering from Government Polytechnic, Ahmedabad from Gujarat Technological University.



**Mr. Hemang D. Solanki** is Founder and CEO of HG wellness, handling activities from main office ahemedabad. He has pursued his graduation from Gujarat technological university and has experience of 3 months as an Officer Biomedical in reputed hospital Zydus.