An Audio Aided Smart Vision System for Visually Impaired

Ramisha Rani K
VIT University, Chennai Campus,
Chennai, Tamilnadu, India.
ramisharanivit@gmail.com

***Abstract*—With help of Image processing foundation, Optical Character Recognition can be handled to distinguish the content in a picture. This would be useful for visually impaired individuals by giving a sound indication to them. This work will be accomplished by utilizing Intel Edison and the application in the framework is composed in python with libraries of OpenCV. At the point when outwardly weakened individuals see the scene through webcam UVC module, which is associated with Intel Edison catch the picture. Captured picture will be handled by yocto module which based open source inside Intel Edison. At the point when the framework process an image, Yocto module would give relating directions to recognize message in a caught picture from the video stream, the framework would call the Bluez module and Alsa sound module to change the guidelines to voice through remote earphones.**

Keywords*— *Intel Edison, Alsa, Debian, OpenCV, Python

1. INTRODUCTION

According to WHOthere are 285 million people estimated to be visually impaired worldwide. Nearly all of folks with poor vision are within the establishing world and over the age of fifty [1]. Individuals with a visual disability are tricky to communicate. Also they have difficulty interacting with the atmosphere, seeing that it can be rough to perceive. Motion can turn out to be confined, leading to having little contact with the surrounding world. The blind or visually impaired people depend mostly on their different senses akin to hearing, touch, and odor with the intention to appreciate their environment. Blindness is most often a difficulty of their living existence. It is tough for them to move out on their own, find toilets, subway stations, eating places and so on. However, there aren't many products that may aid the visually impaired men and women in city life. If the visually impaired persons can "see" the world with the support of alternative gadgets, they are going to attain improved independence and freedom in metropolis existence, and that is precisely why we developed the shrewd glass. The proposed prototype is based on Intel Edison Platform. The Intel Edison [2] is a tiny personal computer offered by Intel as a progress system for wearable instruments. Its dimensions are 35.5x25x3.9mm, with ICs on both sides. The Intel Edison Minibreak board's principal SoC is a 22 nm Intel Atom "Tangier" (Z34XX) that contains two Atom Silvermont cores running at 500 MHz and an Intel Quark core working at 100 MHz (for executing RTOS ViperOS). Integrated board has SoC of 1 GB RAM Intel Edison Mini breakout board has eMMC flash which contains 4 GB.It also integrated with Bluetooth as bluez, Wi-fi and USB controller. small in a form which makes it excellent to build a wearable device.

This utility can robotically become aware of, analyze and recognize textual content in a photograph around the visually impaired and provides corresponding voice hints by means of bluetooth headphones. With the support of this clever reader glass method, the visually impaired may find bus stations, subway stations, eating places, hotels, reading books, etc.

Fig.1. Smart Reader setup

2. IMPLEMENTATION OF SYSTEM DESIGN

The smart glass system design explained as two stages hardware design stage and software design stage. The block diagram of the system is in Fig. 1. The HD camera on the glass catches the video stream around the wearer, and when they click image capture button, the captured image transfers it to the Intel Edison chip. Debian kernel calls the python OpenCV functions to analyze and process the images. When the system extracts text from an image, the system would call the Bluez module and Alsa audio module to transform the information to voice through earphones.

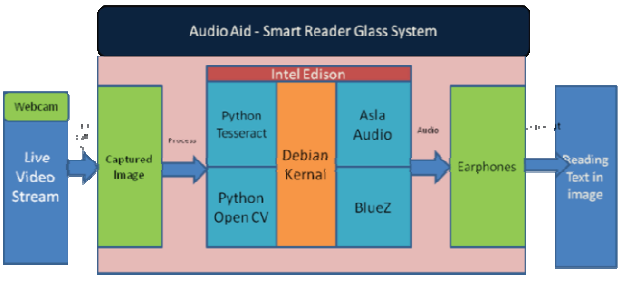

Fig.2. Block Diagram of the System

2.1. Hardware Design

Intel Edison combines a small, low cost, robust, adaptable hardware platform and partner-enabled ecosystem with improved program compatibility and supportive online atmosphere. Fig. 2 indicates the Intel Edison chip. it's particularly small and suitable to construct wearable contraptions. It's high efficiency dual core CPU makes it strong in computing and processing data, so we use this chip to make tricky photo processing. Fig. 4 indicates the prototype of the shrewd glass. Apart from this, a digital camera in front of the glass was fixed. The clever glass is powered with the aid of the 7.4V lithium-Ion battery. The whole device is just like a usual sunglass. It's lightweight, handy to use. The Edison runs the Linux-based procedure which is furnished by way of the Debian project [3]. We installed OpenCV 2.4.9 on Edison.

2.2. Software Design

In software design, programs were written in Python language, and Python IDLE is used for compilation, OpenCV was also use. Images and videos are processed with OpenCV library which has more than 500 algorithms. An algorithm in the Intel Edison is used to acquire the real-time text in a photograph.To extract text from an image stand-alone script was used. It can read all the types of images such as png, jpg,tiff. Tesseract-OCR nearly supports more than 40 languages.It trains each word of different languages. Each language has different training data.

3. Tesseract Algorithm

3.1. Line Discovering

Text region of a processed uniform text size of a page layout is given, Percentile Heightfilter deletes drop-caps and vertically passing letters. After this process approximate text size in a median height, with removed blobs that are smaller than a fraction of the median heightlike noise, punctuation and diacritical marks. The modified test is used to fit a model of parallel, non-over lapping but a sloping line. With a maximum reduced error of incorrect text line by tracking of slope across the page, LMS fit [4] is used to calculate the baselines and the filter-out. The last slope of the line creation process join blobs that overlap by at least half horizontally, putting punctuation marks added with the modified base and associating parts of some fragile characters.

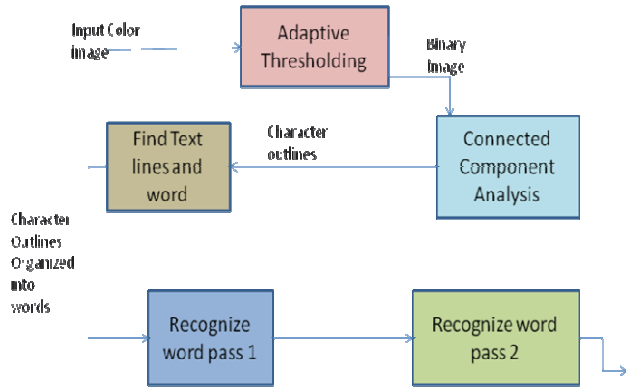

Fig.3. Recognition of text from an image

3.2 Baseline Fitting

When text line processes completes, the baselines are more correctly filtered using a quadratic spline.Quadratic spline was first OCR system to extract text. After that Tesseract is enabled to handle pages with curved baselines[5]. Continuous displacement for original straight baseline is by separating the blobs into groups. The advantage of quadratic spline was stable calculation and disadvantage was discontinuities when various spline segment is required. In earlier stage cubic spline[6] might work better.

3.3 Word Recognition

Word recognition works to correctly identify a word, by segmenting letters. Line finding output was first set of segmentation. The remaining set of segmentation in word recognition fixes only to non-fixed pitch text.

4. Implementation Of Smart Reader

4.1 Smart Reader System

As shown in Fig. 2 smart reader system is designed shown below.

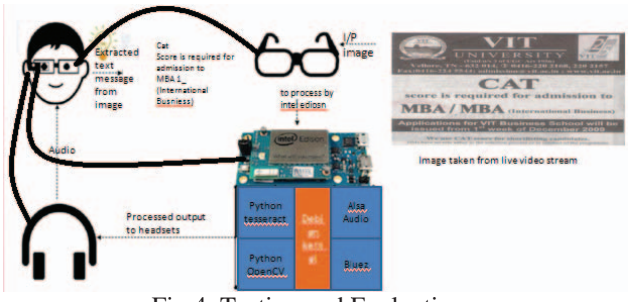

Fig.4. Testing and Evaluation

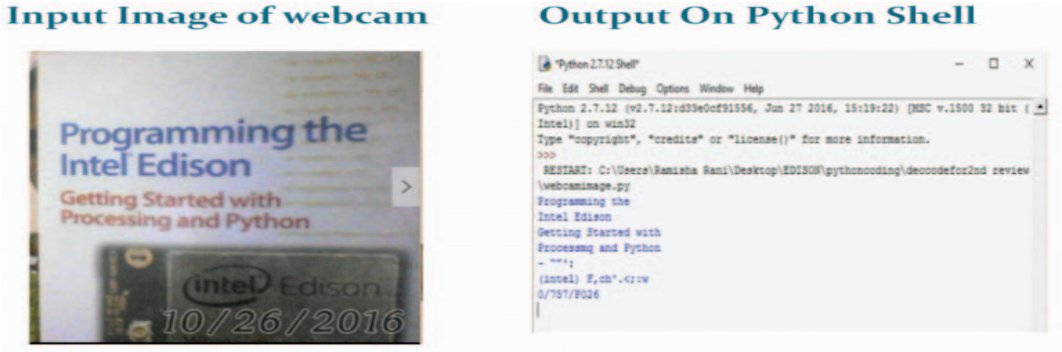

Fig.5. Recognition of text from the camera image with audio

Fig .4 shows a pictorial representation of smart reader in smart glass camera with Intel Edison Processor. Whenever a blind person wants to know information of particular live scene scenario, it will capture the image of the live scene and extract the text content of that captured image.

4.2 Tested with web camera image

To achieve text extraction of live scene image, firstly tested with already taken an image by webcam.

Above Fig.5 represent the output of extracted text from a webcam image.The accurate output of webcam image may vary by a distance of capturing an image.

After extracting text from an image, it will read out all extracted data from an input image. Audio output is heard from a Bluetooth headset because of no enough power supply from Intel Edison Board.

4.3 Tested with live scene captured image

Implementing in real time, a webcam was fixed to Intel Edison board via OTG to capture image and a Bluetooth headset was used to get the audio output of extracted text.

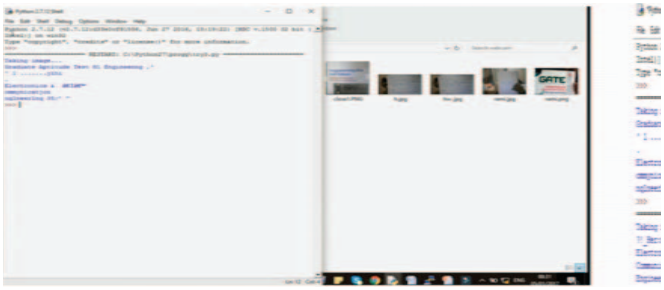

Fig.6. Recognition of text from live video stream

In fig. 6 shows that capturing of the live scene and processing an image to deliver extracted text from an image. To get better result taken image should change to perspective view[7].

4.4 Problems occurred

While researching for this project, following problems occurred.

- Background of image: Normal background of captured image can be easily extracted but a digital background in bus board tough to extract.
- The distance between webcam and target: If visually impaired people click from long distance then accuracy will be low.
- Blurring of image: while clicking if the image becomes blurred then accuracy will be low.
- pronunciation of extracted: Extracted text will read out word by word, in the case of abbreviation in extracted text also it will read like a word not like separate letters for abbreviation

5. Conclusion And Further Work

The paper introduces a prototype approach of lightweight shrewd glass for visually impaired individuals. We established how the intelligent glass was once designed, including hardware design and application design. This approach can become aware of and appreciate the item in actual time. The clever glass can be priceless for the visually impaired men and women of their metropolis existence. And in the quickly future, problems mentioned in the paper will be rectified and come out a perfect clever smart reader.

References

1. Visual impairment and blindness", http://www.who.int/mediacentre/factsheets/fs282/ene/
2. "Intel Edison, one tiny platform, endless possibility," http://www.intel.com/content/www/us/en/do-it-yourself/edison.html.
3. R. Smith, proposed "A Simple and Efficient Skew Detection Algorithm via Text Row Accumulation", IEEE 1995, pp. 1145-1148.

4.S.V. Rice, G. Nagy, T.A. Nartker, Optical Character Recognition: “An Illustrated Guide to the Frontier”, Kluwer Academic Publishers, USA 1999, pp. 57-60.
5. “Edison inteliot module ships,”http://linuxgizmos.com/Edison-iot-module-ships-with-atom-plus-quark-combo-soc/.
6."The Debian Administrator's Handbook",https://debian-handbook.info/browse/stable/sect.role-of-distributions.html#id-1.4.8.4
7. Dr. Gary RostBradski and Adrian Kaehlerpublished,“LearningOpenCV”, 1st Edition, O’Reilly Media, Inc., first edition, 2008.
8. S.V. Rice, G. Nagy, T.A. Nartker, Optical Character Recognition: “An Illustrated Guide to the Frontier”, Kluwer Academic Publishers, USA 1999, pp. 57-60.
9. P.J. Schneider, “An Algorithm for Automatically Fitting Digitized Curves”, in A.S. Glassner, Graphics Gems I, Morgan Kaufmann, 1990, pp. 612-626
10. TrungQuyPhan, PalaiahnakoteShivakumara, ShangxuanTian, and Chew Lim Tan have proposed, “Recognizing Text with Perspective Distortion in Natural Scenes” ICCV,2013.
11. K. Wang, B. Babenko, and S. Belongie proposed “End-to-End Scene Text Recognition” in ICCV 2011.
12. A. Mishra, K. Alahari, and C. V. Jawahar have proposed “Top-Down and Bottom-up Cues for Scene Text Recognition” in CVPR 2012.
13. D. L. Smith, J. Field, and E. Learned-Miller proposed,
“ Enforcing Similarity Constraints with Integer Programming for Better Scene Text Recognition” in CVPR 2011.
14. T. Novikova, O. Barinova, P. Kohli, and V. Lempitsky proposed, “ Large-Lexicon Attribute-Consistent Text Recognition in Natural Images”, in ECCV 2012.
15. S. M. Lucas, A. Panaretos, L. Sosa, A. Tang, S. Wong, and R. Young. ICDAR 2003 “Robust Reading Competitions” in ICDAR 2003.
16. L. Ballan, M. Bertini, A. Del Bimbo, and G. Serra.Proposed “Semantic Annotation of Soccer Videos by Visual Instance Clustering” in Multimedia Tools and Applications, 2010.
17. A. Graves, M. Liwicki, S. Fernandez, R. Bertolami, H. Bunke, and J.Schmidhuber proposed, “A Novel Connectionist System for Unconstrained Handwriting Recognition” IEEE TPAMI, 2009.