

B-LIGHT: A Reading aid for the Blind People using OCR and OpenCV

Mallapa D.Gurav¹, Shruti S. Salimath², Shruti B. Hatti², Vijayalaxmi I. Byakod²,
Shivaleela Kanade²

¹ Assistant professor, Department of CSE, KLECET, Chikodi

² Student, Department of CSE, KLECET, Chikodi

ABSTRACT

Optical character recognition (OCR) is the identification of printed characters using photoelectric devices and computer software. It converts images of typed or printed text into machine encoded text from scanned document or from subtitle text superimposed on an image. In this research these images are converted into audio output. OCR is used in machine process such as cognitive computing, machine translation, text to speech, key data and text mining. It is mainly used in the field of research in Character recognition, Artificial intelligence and computer vision. In this research, as the recognition process is done using OCR the character code in text files are processed using Raspberry Pi device on which it recognizes character using tesseract algorithm and python programming and audio output is listened. To use OCR for pattern recognition to perform Document image analysis (DIA) we use information in grid format in virtual digital library's design and construction. This research mainly focuses on the OCR based automatic book reader for the visually impaired using Raspberry PI. Raspberry PI features a Broadcom system on a chip (SOC) which includes ARM compatible CPU and an on chip graphics processing unit GPU. It promotes Python programming as main programming language.

Keywords— document image analysis(DIA), Raspberry PI, audio output, OCR based book reader, python programming.

1. INTRODUCTION

Visually impaired people report numerous difficulties with accessing printed text using existing technology, including problems with alignment, focus, accuracy, mobility and efficiency. We present a smart device that assists the visually impaired which effectively and efficiently reads paper-printed text. The proposed project uses the methodology of a camera based assistive device that can be used by people to read Text document. The framework is on implementing image capturing technique in an embedded system based on Raspberry Pi board. The design is motivated by preliminary studies with visually impaired people, and it is small-scale and mobile, which enables a more manageable operation with little setup. In this project we have proposed a text read out system for the visually challenged. The proposed fully integrated system has a camera as an

input device to feed the printed text document for digitization and the scanned document is processed by a software module the OCR (optical character recognition engine). A methodology is implemented to recognition sequence of characters and the line of reading. As part of the software development the Open CV (Open source Computer Vision) libraries is utilized to do image capture of text, to do the character recognition. Most of the access technology tools built for people with blindness and limited vision are built on the two basic building blocks of OCR software and Text-to-Speech (TTS) engines. Optical character recognition (OCR) is the translation of captured images of printed text into machine encoded text. OCR is a process which associates a symbolic meaning with objects (letters, symbols an number) with the image of a character. It is defined as the process of converting scanned images of machine printed into a computer process able format. Optical Character recognition is also useful for visually impaired people who cannot read Text document, but need to access the content of the Text documents. Optical Character recognition is used to digitize and reproduce texts that have been produced with non computerized system. Digitizing texts also helps reduce storage space. Editing and Reprinting of Text document that were printed on paper are time consuming and labor intensive. It is widely used to convert books and documents into electronic files for use in storage and document analysis. OCR makes it possible to apply techniques such as machine translation, text-to-speech and text mining to the capture / scanned page. The final recognized text document is fed to the output devices depending on the choice of the user. The output device can be a headset connected to the raspberry pi board or a speaker which can spell out the text document aloud. [1] Gives an algorithm for detecting and reading text in natural images for the use of blind and visually impaired subjects walking through city scenes. The overall algorithm has a success rate of over 90% on the test set and the unread text is typically small and distant from the viewer. [2] have proposed a novel scheme for the extraction of textual areas of an image using globally matched wavelet filters. A clustering based technique has been devised for estimating globally matched wavelet filters using a collection of ground truth images.[3] proposes a support vector machine (SVM) is used to analyse the textual properties of texts. The combination of CAMSHIFT and SVM's produces both robust and efficient text detection.[4] tells about the

navigational technologies available to blind individuals to support independent travel, our focus is on blind navigation on large scale.

**Prevalence of blindness (per thousand)
As per the estimate of 2015**

| Types of blindness | Urban | Rural | Total |
|--------------------|-------|-------|-------|
| Total blindness | 4.43 | 5.99 | 5.40 |
| Economic blindness | 11.14 | 15.44 | 13.83 |
| One eye blindness | 8.23 | 7.00 | 7.46 |

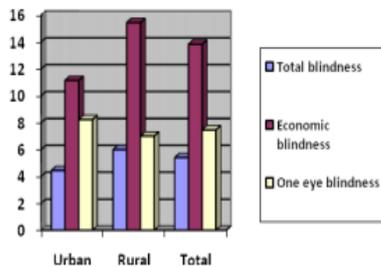


Fig.1 survey of blind people

[5] presents an approach to automatic detection and recognition of signs from natural scenes and its application to sign translation task that further propose a local intensity normalisation method to effectively handle lighting variations followed by a gabor transform to obtain local features.[6] presents a comparative survey among portable/wearable obstacle detection/avoidance systems to inform about the progress in assistive technology for visually impaired people.

2. PROPOSED SYSTEM

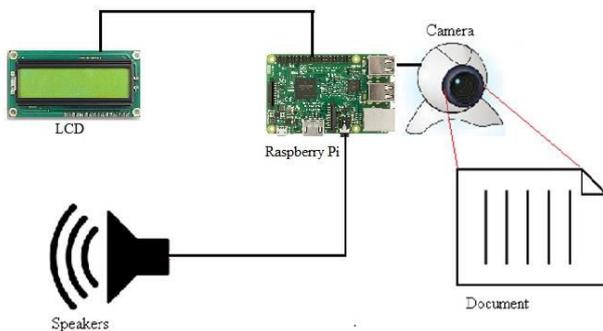


Fig.2 architectural design

To overcome from the problems in the existing system we have developed a project named “B-light A Reading Aid for Blind People using OCR and OpenCV”. The proposed system is to assist blind persons to read text from challenging pattern and background for the purpose of reading document. The main objective of our system is to identify the text in the documents. Firstly the object image is captured by using a webcam which is embedded within Raspberry Pi and is followed by the image processing. To implement an automated system, this would scan a document and read out its content to the user on click of a button the vocal output is produced with the help of speaker which would help the user to readout the text in the document. Our system helps the

blind people to for the purpose of reading without consuming much space.

3. RASPBERRY PI DESCRIPTION

Raspberry pi is a low cost, credit card sized computer that plugs computer monitor and TV and uses standard keyboard and mouse that uses python programming. There are two models of raspberry pi, model A and model B. these two are bit similar with few advance features on model b compared to model a. model b has 512 MB RAM, two USB port whereas model a has 256 mb ram and just a USB port. Besides, model b has Ethernet port while model.



Fig.3 raspberry pi3 model B.

4. ARTIFICIAL INTELLIGENCE

It is the field of computer science which makes the system to behave intelligently by various process of training and cognitive learning. It is the one of the branch of computer science which aims to imitate human vision and form the basis of image acquisition, processing, document understanding and recognition.

A far more streamlined field of Document Recognition and understanding is Optical Character Recognition which attempts to identify a single character from an optically read text image as a part of a word that can be then used to process further information on. The area gains rising significance as more and more information each day needs to store processed and retrieved rather than being keyed in from an already present printed or handwritten source.

4. CHARACTER RECOGNITION

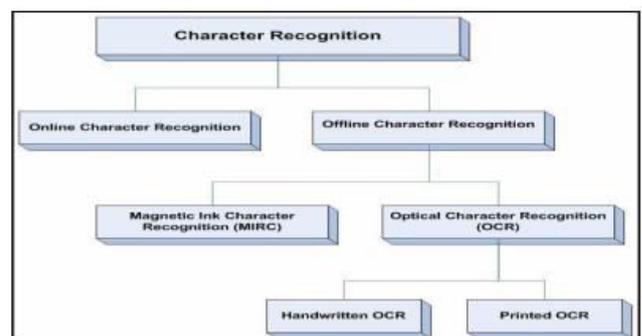


Fig.4 types of character recognition

Character recognition is a sub-field of pattern recognition in which images of characters from a text image are recognized and as a result of recognition respective character codes are returned, these when rendered give the text in the image. The problem of character recognition is the problem of automatic recognition of raster images as being letters, digits or some other symbol and it is like any other problem in computer vision.

5. FLOW OF PROCESS

IMAGE CAPTURING

The first step in which the device is moved over the printed page and the inbuilt camera captures the images of the text. The quality of the image captured will be high so as to have fast and clear recognition due to the high resolution camera

PRE-PROCESSING

Pre-processing stage consists of three steps: Skew Correction, Linearization and Noise removal. The captured image is checked for skewing. There are possibilities of image getting skewed with either left or right orientation. Here the image is first brightened and binarized.

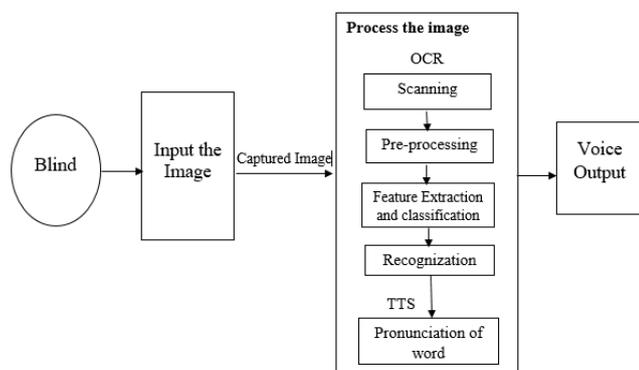


Fig.5 flow of process

The function for skew detection checks for an angle of orientation between ± 15 degrees and if detected then a simple image rotation is carried out till the lines match with the true horizontal axis, which produces a skew corrected image. The noise introduced during capturing or due to poor Quality of page has to be cleared before further processing.

6. CONCLUSION

In this research, we have described a prototype system to read printed text and hand held objects for assisting the blind people. To extract text regions from complex backgrounds, we have proposed a novel text localization algorithm based on models of stroke orientation and edge distributions. The corresponding feature maps estimate the global structural feature of text at every pixel. Block patterns project the proposed feature maps

of an image patch into a feature vector. Adjacent character grouping is performed to calculate candidates of text patches prepared for text classification. An Adaboost learning model is employed to localize text in camera-based images. Off-the-shelf OCR is used to perform word recognition on the localized text regions and transform into audio output for blind users. In this research, the camera acts as input for the paper. As the Raspberry Pi board is powered the camera starts streaming. The streaming data will be displayed on the screen using GUI application. When the object for label reading is placed in front of the camera then the capture button is clicked to provide image to the board. Using Tesseract library the image will be converted into data and the data detected from the image will be shown on the status bar. The obtained data will be pronounced through the ear phones using Flite library

REFERENCES

- [1] Chen X, AL Yuille, Detecting and reading text in natural scenes, in Proc. Computer. Vision Pattern Recognition, 2004;'
- [2] II-366–II-373. 2. Kumar S, R Gupta, et al. Text extraction and document image seg-mentation using matched wavelets and MRF model, IEEE Trans Image Process, August 2007; 16:2117– 2128.
- [3] K Kim, K Jung, et al. Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm, IEEE Trans. Pattern Anal. Mach. Intell, December 2003; 25: 1631– 1639.
- [4] N Giudice, G Legge, Blind navigation and the role of technology, in The Engineering Handbook of Smart Technology for Aging, Disability and Independence, AA Helal, M Mokhtari, B Abdulrazak, Eds. Hoboken, NJ, USA: Wiley, 2008.
- [5] Chen J Y, J Zhang, et al. Automatic detection and recognition of signs from natural scenes, IEEE Trans. Image Process., January 2004 ;13: 87–99.
- [6] D Dakopoulos, NG Bourbakis, Wearable obstacle avoidance electronic travel aids for blind: A survey, IEEE Trans. Syst., Man, Cybern, January 2010; 40: 25–35.