International Journal of Technical Research and Applications e-ISSN: 2320-8163,
www.ijtra.com Special Issue 40 (KCCEMSR) (March 2016), PP. 60-64

INNOVATIVE PRODUCT LABEL READING SYSTEM FOR BLIND PEOPLE

Aditya G. Tambe[1], Siddhesh D. Suryawanshi[2], Akshay A. Palande[3], Pradeep A. Patel[4]
Department of Electronics & Telecommunication
K.C. College of Engineering and Management Studies and Research
Thane, Maharashtra, India

Abstract— **Disability of visual text reading has a huge impact on the quality of life for visually disabled people. Although there have been several devices designed for helping visually disabled to see objects. The project is to detect the text from the camera and announce those words to guide the blind people for product assistance like shopping and safe path movement. A camera-based assistive text reading framework is implemented to help the blind persons to read the text labels and product packaging from hand-held objects in their casual lives. In the deduced region, localization of text and recognition are conducted to acquire the text information. To automatically focalize the text regions from the object ROI, a text localization algorithm by learning gradient features of edge pixels. The localized text characters in the text regions are then binarized and recognized by optical character recognition in software. These processes are done with the help of**

MATLAB. This text is then converted into speech.

Index terms: **Assistive devices, hand-held objects, Optical Character Recognition (OCR), Text To Speech (TTS)**

I. INTRODUCTION

There are over 314 million visually challenged people worldwide , a major part of this population are still blind even in developed countries like united states, the national health interview survey conducted in 2008 reported that over 85 of the adult Americans lack the ability to see. In recent times development in computer vision, digital cameras and portable computers help to aid these individuals by developing camera based products that integrate computer vision technology with already existing products such as optical character Recognition (OCR). Reading is one of the basic necessity today, everything around us are in the form of reports, receipts, bank statements, product packages, restaurant menus etc. Contain printed text on it although optical aids and video magnifiers and screen readers help blind users and those with lower vision help to facilitate text reading. There are few devices that can render better access to common hand-held objects much as product packages and objects printed with text. Formulating devices which are even more portable and sophisticated can promote independent living and foster economic and social self-dependency. There are already some portable systems in use that cannot handle product labeling for example, bar code readers help identify various products in the extensive product database to enable blind users to access information about products through Braille and speech . But there was difficulty in finding him possession of the barcode so pen scanners might be employed in the cases. These system combine OCR software for the purpose of scanning and text recognition and integrated voice output. But most OCR software proved insufficient to handle images with complicated background. Reader mobile runs on cell phone and reads document which are really flat with the dark surface and mostly contain text. Accurately read black print on a white background but cannot read text with complex background and text on uneven surface. There are no existing reading assistance that can read from all kinds of challenging background and patterns

Our proposed algorithm that effectively handle likely all the drawbacks of existing models and extract text from hand-held objects and nearby sources which ever captured through the camera.

II. EXSISTING SYSTEMS

Today, there are already a few systems that have some promise for portable use, but they cannot handle product labeling. For example, portable bar code readers designed to help blind people identify different products in an extensive product database can enable users who are blind to access information about these products through Speech and Braille. But a big limitation is that it is very hard for blind users to find the position of the bar code and to correctly point the barcode reader at the bar code. Some reading assistive systems such as pen scanners might be employed in these and similar situations. Such systems Integrate OCR software to offer the function of scanning and recognition of text and some have integrated voice output. However, these systems are generally designed for and perform best with document images with simple backgrounds, standard fonts, a small range of font sizes, and well-organized characters rather than commercial

International Journal of Technical Research and Applications e-ISSN: 2320-8163,
www.ijtra.com Special Issue 40 (KCCEMSR) (March 2016), PP. 60-64

product boxes with multiple decorative patterns. Most state-of-the-art OCR software cannot directly handle scene images with complex backgrounds. A number of portable reading assistants have been designed specifically for the visually impaired "K-Reader

Mobile" runs on a cell phone and allows the user to read mail, receipts, fliers, and many other documents. However, the document to be read must be nearly flat, placed on a clear, dark surface (i.e., a non-cluttered background), and contain mostly text. In addition, "K-Reader Mobile" accurately reads black print on a white background, but has problems recognizing colored text or text on a colored background. It cannot read text with complex backgrounds. Furthermore, these systems require a blind user to manually localize areas of interest and text regions on the objects in most cases. Even though a number of reading assistants have been designed specifically for the visually impaired, to our knowledge, no existing reading assistant can read text from the kinds of challenging patterns and backgrounds found on many everyday commercial products. Such as text information can appear in various scales, fonts, colors, and orientations

Fig.1 Printed text from hand-held objects with Multiple colours, complex backgrounds, or non flat surfaces.

III. PROPOSED SYSTEMS

To overcome the problems defined in problem definitions and also to assist blind persons to read text from those kinds of challenging patterns and backgrounds found on many everyday commercial products of Hand-held objects, then have to conceived of a camera-based assistive text reading framework to track the object of interest within the camera view and extract print text information from the object. Proposed algorithm used in this system can effectively handle complex background and multiple patterns, and extract text information from both hand-held objects and nearby signage. To overcome the problem in assistive reading systems for blind persons, in existing system very challenging for users to position the object of interest within the center of the camera's view. As of now, there are still no acceptable solutions. This problem approached in stages. The hand-held object should be appears in the camera view, this thesis use a camera with sufficiently wide angle to accommodate users with only approximate aim. This may often result in other text objects appearing in the camera's view (for example, while shopping at a Supermarket). To extract the hand-held object from the method to obtain a region of interest (ROI) of the object. Then, perform text recognition only that ROI.

It is a challenging problem to automatically localize objects and text ROIs from captured images with complex backgrounds, because text in captured images is most likely surrounded by various background outlier "noise," and text characters usually appear in multiple scales, fonts, and colors. For the text orientations, this thesis assumes that text strings in scene images keep approximately horizontal alignment. Many algorithms have been developed for localization of text regions in scene images. We can divide them into two categories: Rule-Based and Learning-Based.

In solving the task at hand, to extract text information from complex backgrounds with multiple and variable text patterns, here propose a text localization algorithm that combines rule-based layout analysis and learning-based text classifier training, which define novel feature maps based on stroke orientations and edge distributions. These, in turn, generate representative and discriminative text features to distinguish text characters from background outliers.

IV. FRAMEWORK AND ALGORITHM

The system framework consists of three functional components they are:
Scene Capture by Camera

Image Acuisition & Processing Audio Output

The scene capture component collects scenes containing objects of interest in the form of images or video, it corresponds to a camera attached to a pair of sunglasses. The live video is captured by using web command it can be done using OPENCV libraries. The image format from the webcam is in RGB24 format. The frames from the video is segregated and undergone to the preprocessing.

V. TEXT SYNTHESIS

Flowchart of Optical Character Recognition

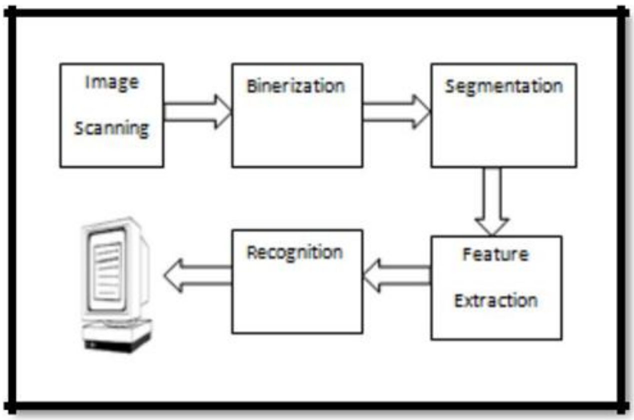


Fig.2 Flowchart of Optical Character Recognition

Algorithm for Optical Character Recognition

1) Scanning

OCR system uses Android mobile camera. Camera captures image of document. This is nothing but the process of scanning. In short we can say that scanning makes original document as digital image. Generally, original documents are made up of the black coloured text print on the white coloured background. Scanning comes with thresholding which makes the digital image as gray scale image. Thresholding is the process which converts multi level image into bi-level image i.e. black and white image. Fixed threshold level is defined in thresholding. If the gray levels are below the threshold level, identified as black. Whereas if gray level is above the threshold level, identified as white. This results in saving memory space and computational efforts.

2. Segmentation

The process of locating regions of printed or handwritten text is segmentation. Segmentation differs text from figures and graphics. When segmentation is applied to text, it isolates characters or words. The mostly occurred problem in segmentation is: it causes confusion between text and graphics in case of joined and split characters. Usually, splits and joints in the characters causes due to scanning.

If document is dark photocopy or if it scanned at low threshold, joints in characters will occur. And splits in characters will occur if document is light photocopy or scanned at high threshold. OCR system also gets confused during segmentation when characters are connected to graphics.

3) Pre-processing

As we seen above, some noise may occurred during scanning process. This results in poor recognition of characters. This usually occurred problem is overcomed by preprocessing. It consists of smoothing and normalization. In smoothing, certain rules are applied to the contents of image with the help of filling and thinning techniques. Normalization is responsible to handle uniform size, slant and rotation of characters.

4) Feature Extraction

It extracts the features of symbols. Features are the characteristics. In this, symbols are characterized and unimportant attributes are left out. The feature extraction technique does not match concrete character patterns, but rather makes note of abstract features present in a character such as intersections, open spaces, lines, etc. Tesseract algorithm is used to implement feature extraction. Feature extraction is concerned with the representation of the symbols. The character image is mapped to a higher level by extracting special characteristics of the image in the feature extraction phase.

5) Recognition

OCR system works with Tesseract algorithm which recognizes characters. Tesseract identifies characters in foreground pixels, called as blobs, and then it finds lines. Word by word recognition of characters is done throughout the lines. Recognition involves converting these images to character streams representing letters of recognized words. In short, recognition extracts text from images of documents.

VI. TEXT TO SPEECH SYNTHESIS

Speech is the vocalization form of human communication. Speech communication is more effective medium than text communication medium in many real world applications.

Speech synthesis is the artificial production of human speech. A system used for this purpose is called a speech synthesizer, and can be implemented in software or hardware. Synthesized speech can be created by concatenating pieces of recorded speech that are stored in a database. The quality of a speech synthesizer is judged by its similarity to the human voice, and by its ability to be understood. A Text-To-Speech (TTS) synthesizer is a computer-based system that should be able to read any text aloud. The block diagram given below explains the same. A text-to-speech system (or "engine") is composed of two parts: a front-end and a back-end. The front-end has two major tasks.

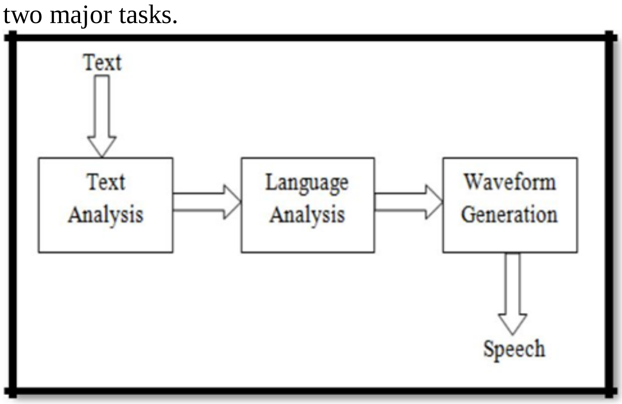


Fig.3 Text to Speech Synthesis

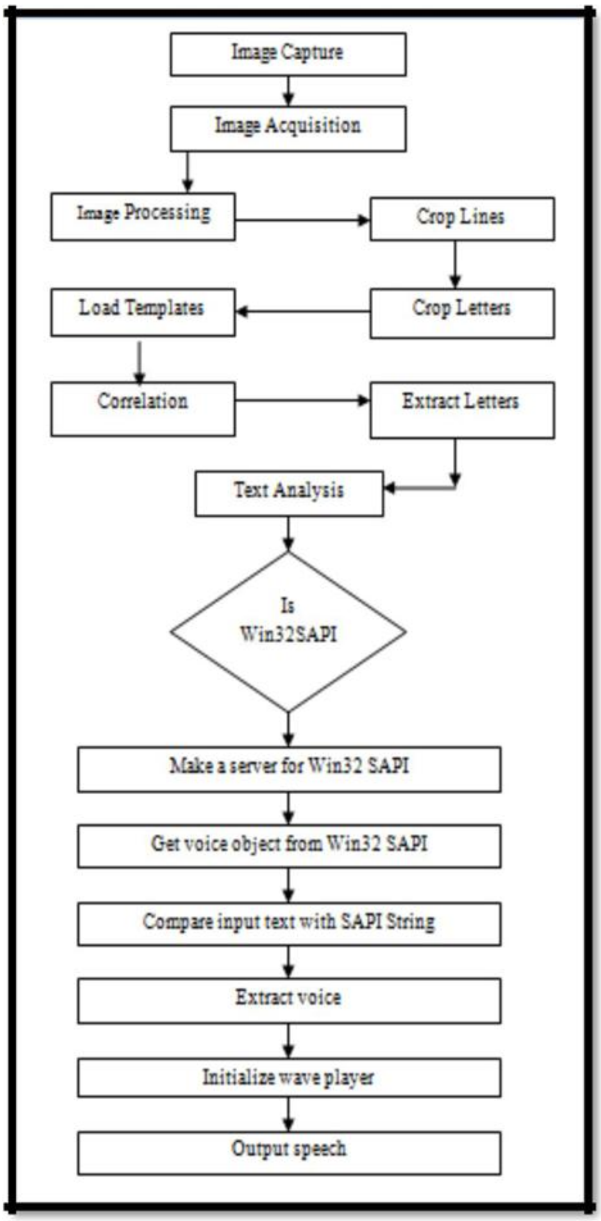
First, it converts raw text containing symbols like numbers and abbreviations into the equivalent of written-out words. This process is often called text normalization, pre-processing, or tokenization. The front-end then assigns phonetic transcriptions to each word, and divides and marks the text into prosodic units, like phrases, clauses, and sentences. The process of assigning phonetic transcriptions to words is called text-to-phoneme conversion. The back-end often referred to as the synthesizer—then converts the symbolic linguistic representation into sound.

VII. . SYSTEM IMPLEMENTATION

System architecture

The system consists of a portable camera, a computing device and a speaker or headphone. Images can be capture using the camera. For better results we can use a camera with zooming and auto focus capability. OCR based speech synthesis system applications require a high processing speed computer system to perform specified task. It possible to do with 100MHz and 16M RAM, but for fast processing (large dictionaries, complex recognition schemes, or high sample rates), we should shoot for minimum of a 400MHz and 128M RAM. Because of the processing required, most software packages list the minimum requirements. It requires an operating system and sound must be installed in PC. System applications required a good quality speaker to produce a good quality of sound.

Process flowchart



VIII. RESULT

Text reading system has two main parts: image to text conversion and text to voice conversion. Image into text and then that text into speech is converted using MatLab. For image to text conversion firstly image is converted into gray image then black and white image and then it is converted into text by using MatLab. Microsoft Win 32 speech application program interface library has been used to produce speech information available for computer in MatLab. This library allows selecting the voice and audio device one would like to use. We can select the voices from the list and can change the pace and volume, which can be listened by installing wave player.

IX. APPLICATIONS

[1] We can convert image into text and text into speech. In short we can convert image into speech without any physical contact.

[2] We can Use the system for blind person as well as illiterate person to read the text from the image.

X. CONCLUSION

This paper is an effort to suggest an approach for image to speech conversion using optical character recognition and text to speech technology.

The application developed is user friendly, cost effective. In this project, we have described a prototype system to read printed text on hand-held objects for assisting blind persons. This method can effectively distinguish the object of interest from background or other objects in the camera view. 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. Word recognition on the localized text regions and transform into audio output for blind users.

REFERENCES

[1] C. Yi and Y. Tian, "Assistive text reading from complex background for blind persons," in Proc. Int. Workshop Camera-Based Document Anal. Recognit., 2011, vol. LNCS-7139, pp. 15–28

[2] X. Chen, J. Yang, J. Zhang, and A. Waibel, "Automatic detection and recognition of signs from natural scenes," IEEE Trans. Image

Process.,vol. 13, no. 1, pp. 87–99, Jan. 2004..

[3] S. Shoval, J. Borenstein, and Y. Koren, "Auditory guidance with the Nav-belt: A computerized travel for the blind," IEEE Trans. Syst., Man,

Cybern. C. Appl. Rev., vol. 28, no. 3, pp. 459–467, Aug. 1998.

[4] J. Zhang and R. Kasturi, "Extraction of Text Objects in Video Documents: Recent Progress," In IAPR Workshop on Document

Analysis System, 2008.

[5]B. Epshtein, E. Ofek and Y. Wexler, "Detecting text in natural scenes with stroke width transform," In CVPR, pp. 2963-2970, 2010.

[6] L. Ma, C. Wang, B. Xiao, "Text detection in natural images based on multi-scale edge detection and classification," In the Int. Congress on

Image and Signal Processing (CISP), 2010.