Camera Based Assistive Analog reader model For visually handicap person using pattern Recognition technique

Dr. Vilas Nandwadekar
Director
Sinhgad Institute of Management

Prof. Sapankumar B Singh,
Assistant Professor,
Dr D Y Patil School of MCA

Abstract:

This analysis paper is predicated on helpful text reading model to help the visually visually handicapped person, the person having week eye sites. In general, the amount of image are going to be captured and prosed by making single image. the quality size of the exhausting material (News Paper) are going to be outlined and a block wise text reading or image capturing occur. when this method investigator has tried to use text detection rule to extract text from advanced background when this method the hand-held tool acknowledge the pinnacle line and browse them loud mistreatment speech synthesis code. investigator has discovered only a few individuals area unit operating during this space and this can be one in all the distinctive researches wherever it's social importance at the big scale.

Keyword: OCR, Text Extraction, Segmentation, Speech Synthesis.

I. Introduction:

Of the 725 million population of word [1] 246 million individuals area unit visually handicap and thirty-nine million individuals area unit blind as survey conducted by the national Health survey [2]. concerning sixty five exploit all folks that area unit visually impaired area unit aged fifty and older, whereas this people contains concerning twenty exploit the world's population. With AN increasing old population in several countries, additional individuals are going to be in danger of impairment as a result of chronic eye diseases and ageing processes [2]. Even in developing country like Asian country out of 127 million individuals [1]about three.5 million across the country with 30000 new cases being else every year[3].India shoulders the biggest burden of world cecity. This range is increasing speedily because the soul generation ages. Recent developments in pc vision, digital cameras, and moveable pcs create it possible to help these people by developing camera-based merchandise that mix computer vision

technology with alternative existing industrial merchandise such optical character recognition (OCR) systems. Reading is clearly essential in today's society. written text is everyplace within the sort of reports, receipts, bank statements, building menus, room handouts, product packages, directions on medication bottles, etc. And whereas optical aids, video magnifiers, and screen readers will facilitate blind users and people with low vision to access documents, there area unit few devices that may give smart access to common hand-held objects like product packages, and objects written with text like prescription medication bottles. the flexibility of individuals UN agency area unit blind or have vital visual impairments to browse written labels and products packages can enhance freelance living and foster economic and social independency.

II. Literature Review

Today, there are already a few systems that have some promise for portable use, but they cannot handle News Paper Reading. 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 [22/13] through speech and braille. 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 with multiple decorative patterns, Font Size, Font Color. Most state-of-the-art OCR software cannot directly handle images with complex backgrounds. Eye of the Beholder[4]use a client-server architecture: a custom application on a mobile phone allows the user to snap a photo of an object and automatically send it to a server using an HTTP request over the GPRS network; a server-side script invokes the OCR engine, which extracts the text from the image. The server sends the extracted text back to the phone, where it is displayed and enunciated using a speech-synthesis engine. Limitation of phone-based text recognition system for visually-impaired users are,1. The image contains out-of-focus text. For example,

if the subject is too close, the image is too blurred for any text to be recognizable. 2. The image contains very small text. This

occurs when the text itself is very small (such as the ingredients list on grocery items) or when standing further away from the text to fix the out-of-focus problem described above. 3. The image contains angled text. It is difficult even for the sighted users to align themselves perfectly perpendicular to the text. 4. The image has low contrast. This can result if the text and background have similar colors. 5. The environment has poor lighting conditions. Poorly-lit photos are common because non-sighted people have no awareness of ambient light levels. 6. Camera phones lack the memory and processing power needed for OCR. OCR engines require a large number of processing cycles and rely heavily on floating point operations. 7. Leveraging streaming video is impractical. Video enables more sophisticated recognition techniques, but it is currently difficult to stream video through cellular data networks due to latency and bandwidth limitations. 8. The visually-impaired user needs to locate text in the environment and point the camera-phone in the relevant direction.

A number of portable reading assistants have been designed specifically for the visually impaired [5], [6], [7], [8], [9], [10], [11]. KReader Mobile runs on a cell phone and allows the user to read mail, receipts, fliers, and many other documents [5]. 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. Furthermore, KReader 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, text printed on cylinders with warped or incomplete image. Furthermore, these systems require a blind user to manually localize areas of interest and text regions on the objects in most cases.

III. System design and architecture

The proposed module can be divided in to three part first part is image capture second is text extraction and third is text to voice conversion using TTS. As given in below fig.1 First process is to capture the image using the camera then captured image is transfer to text extraction system which consist OCR system. OCR include text preprocessing, text segmentation, Character recognition and post processing. Text preprocessing is nothing but the removal of noise from captured image, in text segmentation region text are binaraized and segmented in text and non text region. In

character recognition individual character are isolated and normalized in order to better understanding. Post processing is grouping the character to form meaning full meaning. Extracted text is then transfer to speech synthesis system where extracted text are converted into speech.

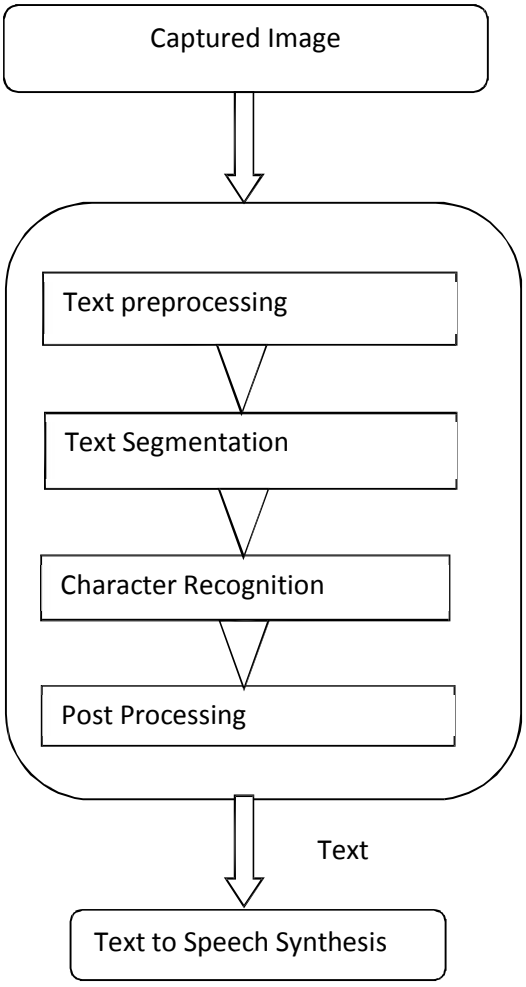

Fig: 1 System design and architecture

IV] Text Extraction Method

The text extraction method include classification of alphabet, numeric character or other character. It include following stage.

4.1. Text Pre Processing:-

Noise Removal- Linear low pass filtering is basic method of noise removal. But the disadvantage of this method is it bluer the quality of image. Therefore this method damage the readability of text from image. Another method is anisotropic filtering method with this method we can remove the drawback of linear filtering method by approximately weighing the filtering coefficient. Anisotropic filtering is a method of enhancing the image quality of textures on surfaces of computer graphics that are at oblique viewing angles with respect to the camera where the projection of the texture appears to be non-orthogonal. Like bilinear and trilinear filtering, anisotropic filtering eliminates aliasing effects, but improves on these other techniques by reducing blur and preserving detail at extreme viewing angles. It is based on monotonically decreasing function. Let

f: I → R be monotone decreasing with range an interval . Then f is continuous on (a, b). Suppose for contradiction that f has a jump at x0. Then at least one of the intervals

$(y_0^-, f(x_0))$ or $(f(x_0), y_0^+)$. Must be nonempty. Pick one and call it J and note

that J $\subset (y_0^-, y_0^+) \subset (f(a), f(b))$ so the image of f is not an interval.

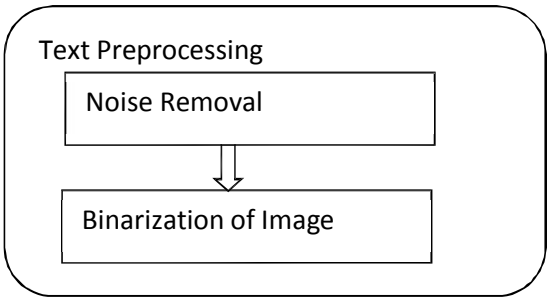

Fig 2: Text Preprocessing

Binarization of Image :-Binarazation is technique by which we can convert the grayscale picture in binary form. Most common way to select a entry point and after that all the intensity below the entry point and all the value above the entry point either 1 to 0. this value represent either black or white.

4.2 . Text Segmentation

This paper proposes two stage of text segmentation. In first level document contain both text and graphics. First of all we have to separate text from graphic by using connected component analysis. Second stage is segmentation based on locating paragraph, line and column.

We are using connected component analysis [12, 13] to segment document into picture or text by using histogram analysis. Connected-component labeling (alternatively connected-component analysis, blob extraction, region labeling, blob discovery, or region extraction)

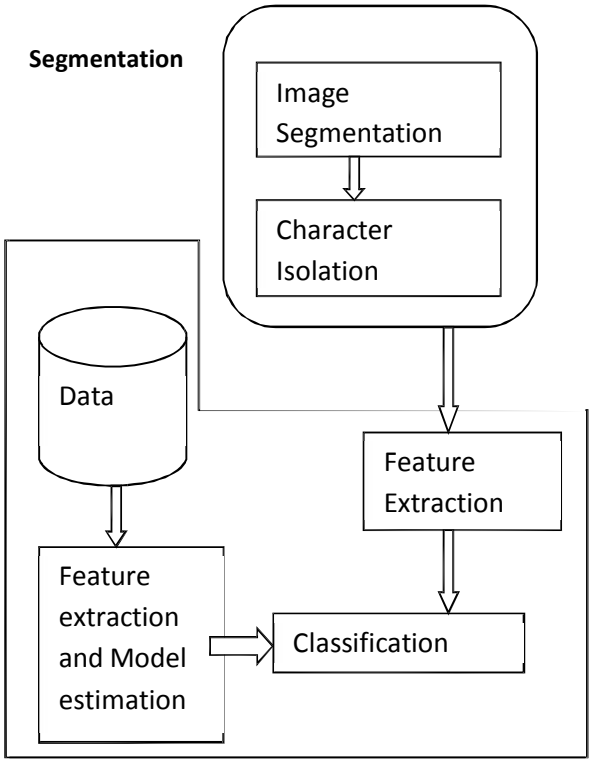

Fig 3: Segmentation

is an algorithmic application of graph theory, where subsets of connected components are uniquely labeled based on a given heuristic.

A connected component may be a character or a picture. Labeling is done for connected component 0 for non background and 1 for background. In this way we have make the group of similar component as a binary image.

A number of statistical formulas have been proposed to segment the text and non text region of binary image. we are using Mean length horizontal black run(MLBR) and white black transaction count per unit(MTC)

MLBR=BRL/TC

MTC=TC/W

MLBR and MTC both are use for classification of text and non text region.

V] Character Recognition: In character recognition is consist of matching of extracted character with store training data. In training data various image are given as a input to OCR . Ten the features of given image are extracted and store as a training data.

Speech Synthesis: Next stage is speech synthesis. In speech synthesis the extracted text are passed as input to speech synthesis engine. Speech synthesis use contaminative synthesis to procedure sound.

VI Conclusion

In this paper we have presented a method of image capture and extraction of text and conversion of extracted text to audible voice to prepare an assistive device which will be helpful for the visually handicap and person with low vision. It work well for the text having in different size and different color. Anisotropic diffusion reduce the additive noise efficiently which is mostly found in news paper without increasing blurring. Post processing error removal can be added to increase the efficiency of system.

Reference:

[1] http://www.worldometers.info/worldpopulation/ date 11/09/2014
[2] http://www.who.int/mediacentre/factsheets/fs282/en/ date 11/09/2014
[3] http://www.netradaan.com/blind-statistics-india.php
[4] Tudor Dumitraş, Matthew Lee, Pablo Quinones, Asim Smailagic, Dan Siewiorek and Priya Narasimhan "Eye of the Beholder: Phone-Based Text-Recognition for the Visually-Impaired"
[5] K. Kim, K. Jung, and J. Kim, “Texture-based approach for text detection in images using SVM and continuously adaptive mean shift algorithm,” In IEEE Trans. on PAMI,Vol.25, No.12, pp.1631-1639,2003
[6] Kreader Mobile user Guide http://www.knfbReading.com dated 12/12/2014
[7] J. Coughlan Computer vision without sight, ACM volume-55 number-1page96-104,2012
[8] N. Nikolaou and N. Papamarkos, “Color Reduction for Complex Document Images,”In International Journal of Imaging Systems and Technology, Vol.19, pp.14-26, 2009.
[9] T. Phan, P. Shivakumara and C. L. Tan, “A Laplacian Method for VideoText Detection,” InICDAR, pp.66-70, 2009
[10] P. Shivakumara, T. Phan, and C. L. Tan, “A gradient difference based technique for video text detection,”In ICDAR, pp.66-70, 2009.
[11] S.Shoval, J.Borenstein, and Y.Koren. Auditory Guidance with the Navbelt-a Computerized

Travel for the Blind. In IEEE Trans.on Systems,Man, and Cybernetics, Part C,Vol. 28,No. 3, 1998.

[12] J.Park, J.Limtext segmentation in color image using tensor voting,image and vision computing 671-685,2007