Review On Portable Camera Based Assistive Text and Product Label Reading From Hand Held Object by Using Android App For Visually Impaired People.

[1]Monika Bhosale, [2]Tejshri Thombare , [3]Shruti Kant, [4]Ashwini Marathe,[5]Prof.P.M.Mane
Department of Computer Engineering Zeal College of Engineering & Research, University of Pune.

ABSTRACT

We propose a camera-based Android app. This app helps the visually impaired people to browse the text on different products. In this system camera captures the particular text on product. Multiple techniques are applied on the text. Such as Optical Character Recognition that supply the scanning and recognition of text and a few have the voice output. From a grayscale image, thresholding are often accustomed produce binary pictures i.e. image with solely black or white colors, Filtering are often used to cut back the noise of image, Next image segmentation technique is employed to perform the method of partitioning a textual image into multiple parts .The goal of segmentation is to modify and/or amendment the illustration of a picture into one thing that's a lot of significant and easier to recognized. Image scaling is that the method of resizing a textual image. Next technique employed in this project is template matching. Temples matching is a way in digital image process for locating tiny components of a picture that match a template image. Also template extraction are often employed in producing as a vicinity of internal control, some way to navigate a mobile golem or as some way to notice edges in images then finally voice output are going to be generated then visually impaired people will simply listen the text on it explicit object.

Keywords : ***Assistive devices, sightless, distribution of edge pixels, hand-held objects, optical character recognition (OCR), stringer orientation, text reading, and text region localization***

1. INTRODUCTION

The 314 million visually impaired folks worldwide, forty five million square measure blind handicap that concerning sightlessness. Even during a developed country just like the us, the 2008 National Health Interview Survey (NHIS) reportable that AN calculable twenty five.2 million adult Americans (over 8%) square measure blind. The human baby generation ages square measure quickly increasing. all over written text square measure obtainable as an example receipts, bank statements, building menus, room handouts, report, product packages, directions on medication bottles, room handouts etc. While screen readers, optical aids and video magnifiers will facilitate blind users and people with low vision to access documents. The few devices offer sensible access to common objects like product textual labels and product written with text as an example prescription medication bottles. The visually handicapped person to scan written labels and merchandise packages can enhance freelance living and foster economic and social self-reliance. Today, there square measure already some systems for transportable use however they can't handle product labeling. The transportable Universal Product Code readers designed to assist blind people identify totally different product in an in depth product info will change user for visually handicapped person to access data regarding these product through speech and Braille. during this system an enormous limitation is that it's terribly exhausting for blind users to seek out the position of the Universal Product Code and to properly purpose the Universal Product Code reader at the Universal Product Code however Some reading-assistive systems like pen scanners could be used in these and similar things. This technique integrates OCR computer code use by perform for scanning and recognition of text and a few have integrated voice output. The OCR is optical character recognition. These systems square measure typically designed and perform best with document pictures with appropriate backgrounds, commonplace fonts, a tiny low vary of font sizes, and well-organized characters. This technique used instead of business product boxes with multiple decorative patterns. The OCR computer code cannot directly handle scene pictures with complicated backgrounds and therefore the document to be scan should be nearly flat, placed on a transparent, dark surface and contain principally text.

Even though variety of reading assistants square measure designed specifically for the blind folks, to our knowledge, there's no existing scanning assistant will read text from the varieties of difficult patterns and

Even though variety of reading assistants square measure designed specifically for the blind folks, to our knowledge, there's no existing scanning assistant will read text from the varieties of difficult patterns and backgrounds that found on many everyday business product. Like text data will obtainable in varied scales, fonts, colors, and orientations. camera captured images. (b) Localized text regions that show in fig mark blue (c) Text region cropped from image. (d) Text codes recognized by OCR (optical character recognition). Text at the top-right corner of bottom image is shown during a increased callout. Mobile accurately reads black print text on a white background however there some issues to recognizing colored text or text on a colored background. It cannot scan text with complicated backgrounds as a result of they can't simply detected the text from background Mobile accurately reads black print text on a white background however there some issues to

recognizing colored text or text on a colored background. It cannot scan text with complicated backgrounds as a result of they can't simply detected the text from background.

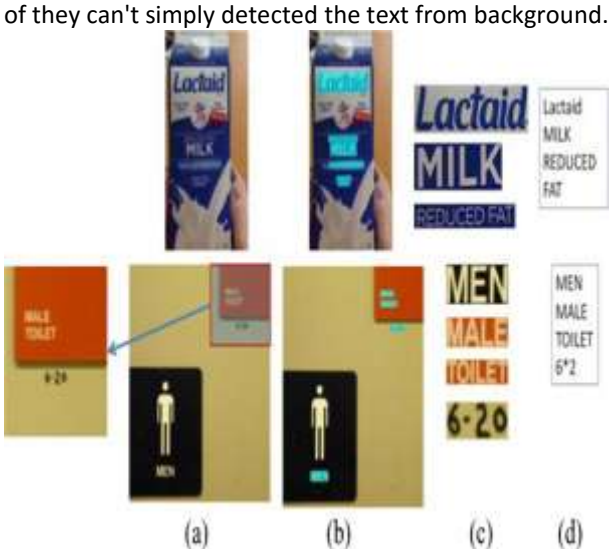

Fig:1

Mobile accurately reads black print text on a white background however there some issues to recognizing colored text or text on a colored background. It cannot scan text with complicated backgrounds as a result of they can't simply detected the text from background. Text printed on cylinders with crooked or incomplete pictures as an example as soup cans or medication bottles. These systems require a blind user to manually localize areas of interest and text regions on the objects. As shown in Fig. 1, such text data will appear in multiple scales, fonts, colors. to help blind persons to scan text from hand-held objects, we have planned of a camera-based helpful text reading frame work to track the item of interest at intervals the camera read and extract print text data from the item. the prevailing rule will effectively handle complicated background and multiple patterns and extract text data from each hand-held objects and assemblage as shown in Fig. 2.The hand-held object extract camera image and we tend to develop a motion-based technique to obtain region of interest (ROI) of the item. During this ROI perform the text recognition technique. The localization of text regions in scene pictures they're divide in 2 classes: rule-based and learning –based. Rule primarily based rule to used component level image process to extract the text data from predefined text layout like character size, ratio ,edge density ,character structure ,color uniformity of text and learning primarily based rule square measure used model text structure and extract representative text options to make text Classifiers.

2. EXISTING SYSTEM

In these days society, there square measure already a couple of systems that square measure moveable use for blind persons. For instance, moveable code readers. The blind those who want to access data regarding these product, the moveable code reader helps the blind individuals to spot totally different product in an intensive product information. However properly position the code reader at the code is incredibly arduous task for blind users. Jointly there some system that uses camera-based helpful text reading framework.

Fig 2

Foremost object of interest inside the camera read is track so written text data from the item is extracted by this framework. During this framework, the item of interest is positioned at the middle of camera's read. This object ought to be positioned specified, it ought to be seem within the camera read. These wide angles of this camera accommodate the users with solely particular/approximate aim. The system extract hand-held object from the camera image. To get the region of interest (ROI) of the item this framework uses the motion primarily based technique. Then text recognition is performed just for that region of interest (ROI).

Sometimes multiple scales, fonts and colors square measure won't to write the text characters. Conjointly immense quantity noise is contained within the captured pictures that contain text instead of industrial product boxes with multiple ornamental patterns the pictures with straightforward background, commonplace fonts, tiny vary of font sizes and well organized character square measure used for these systems provide distinguishing source code text. If Times Roman is not available, try the font named Computer Modern Roman. On a Macintosh, use the font named Times. Right margins should be justified, not ragged.

3. PROPOSED SYTEMS

In existing system camera capture the text on image and so process is finished in portable computer and voice output is created however it's not possible for visually handicapped person to hold the portable computer every time. To overcome the matter of existing system the new system is introduced that's transportable camera primarily based golem app to helps the visually handicapped person.

We projected golem app this app facilitate the visually handicapped person to browse the text on explicit object. In this system camera capture the actual text on image and multiple techniques square measure applied on it text and finally voice outputs are generated and so visually handicapped person will simply listen the text on it explicit object.

3.1 RGB To Gray scale Conversion

• **Steps**

– Traverse through whole input image set.

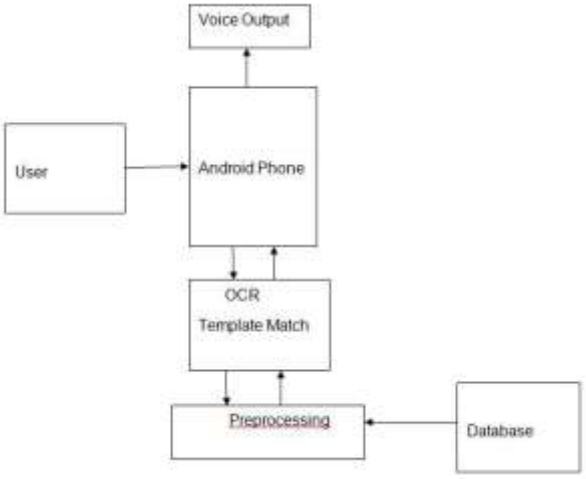

Fig 3: Architecture

– Read individual pixel color value (24-bit).
– Separate the color value into each part, R, G and B 8- bit pixel values.
– Calculate the grayscale constituent (8-bit) for given R,Gand B values using a conversion formula.
– formulate a 24-bit pixel value from 8-bit grayscale value.
– Store the new value at same place in output image.
Educe 8-bit R, G and B values from
24-bit Color Value
b = pix & 0xff;
g = (pix >> 8) & 0xff;
r = (pix >> 16) & 0xff;

3.2 FILTERING:-

In signal processing, it is often desirable to be able to perform some kind of noise reduction on an image or signal. The **median filter** is a nonlinear digital filtering technique, often used to remove noise. Such noise reduction is a typical pre-processing step to improve the results of later processing (for example, edge detection on an image). Median filtering is very widely used in digital image processing because, under certain conditions, it preserves edges while removing noise

3.3 THRESHOLDING:-

• Steps / Algorithm
– Traverse through entire input image array.
– Read individual pixel color value (24-bit) and convert it into grayscale.
– Calculate the binary output pixel value (black or white) based on current threshold.
– Store the new value at same location in output image.
Thresholding Logic:

```
gs = (r+g+b) / 3; // greyscale
if(gs < th) {
pix = 0; // pure black
}else {
pix = 0xFFFFFF; // pure white
}
```

3.4 THINNING:

The Stentiford Algorithm can be stated as following

1. Find a pixel location (i, j) where the pixels in the image match those in template T1. With this template all pixels along the top of the image are removed moving from left to right and from top to bottom.
2. If the central pixel is not an endpoint, and has connectivity number = 1, then mark this pixel for deletion.
Endpoint pixel: A pixel is considered an endpoint if it is connected to just one other pixel. That is, if a black pixel has only one black neighbour out of the eight possible neighbours.
3. Repeat steps 1 and 2 for all pixel locations matching T1.
4. Repeat steps 1-3 for the rest of the templates: T2, T3, and T4.T2 will match pixels on the left side of the object, moving from bottom to top and from left to right. T3 will select pixels along the bottom of the image and move from right to left and from bottom to top. T4 locates pixels on the right side of the object, moving from top to bottom and right to left.
5. Set to white the pixels marked for deletion.

3.5 SEGMENTATION

In computer vision, image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyses.
Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. More precisely, image segmentation is the process of assigning a label to every pixel in an image such that pixels with the same label share certain characteristics. The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image.

3.6 TEMPLATE MATCHING

Template matchingis a technique in digital image processing for finding small parts of an image which match a template image. It can be used in manufacturing as a part of quality control, a way to navigate a mobile robot, or as a way to detect edges in images.

3.6.1 APPROACH

Template matching can be subdivided between two approaches: feature-based and template-based matching. The feature-based approach uses the features of the search and template image, such as edges or corners, as the primary match-measuring metrics to find the best matching location of the template in the source image. The template-based, or global, approach uses the entire template, with generally a sum-comparing metric (using SAD, SSD, cross-correlation, etc.) that determines the best location by testing all or a sample of the viable test locations within the search image that the template image may match up to.

3.7 SERIALIZATION

In computer science, in the context of data storage and transmission, **serialization** is the process of converting a data structure or object state into a format that can be stored (for example, in a file or memory buffer, or transmitted across a network connection link) and "resurrected" later in

the same or another computer environment.[1] When the resulting series of bits is reread according to the serialization format, it can be used to create a semantically identical clone of the original object. For many complex objects, such as those that make extensive use of references, this process is not straightforward. Serialization of object oriented objects does not include any of their associated methods with which they were previously inextricably linked. This process of serializing an object is also called **deflating** or **marshalling** an object.[2] The opposite operation, extracting a data structure from a series of bytes, is **deserialization** (which is also called **inflating** or **Unmarshalling**).

4. IMPLIMENTATION:-

This is the first step after installing the app. Then we have to click on the Ok button, on page there is Crop, Image processing, Original symbol pack and OCR after all this we are start with app

Main Activity:

Fig 4

This activity is open after successfully enter In this first we have to click on Crop button and Select the image and crop the particular text. After that select the image from the gallery for further processing. After selecting image click on process image to apply thinning on that text.

After that thinning will be applied on that image.

After thinning for that image select the process image option for segmentation.Segmentation will be applied on that image that will separate and create rectangle for each character.

For generation of template and matching of template we need to condition the alphabets and symbols and make a set of character.we will train a character A from that text.

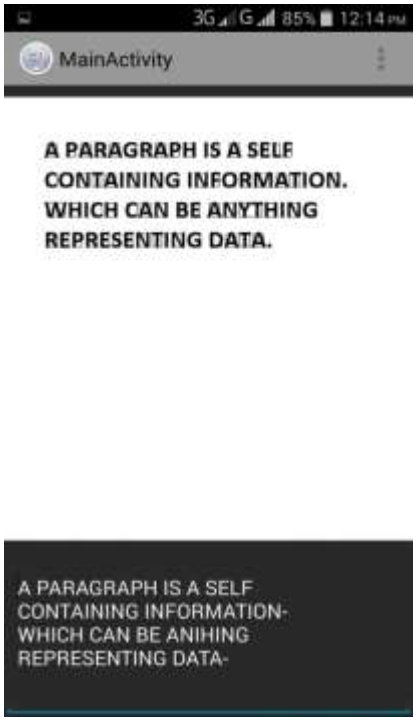

Fig 5

Output

voice output will be generated for that text.

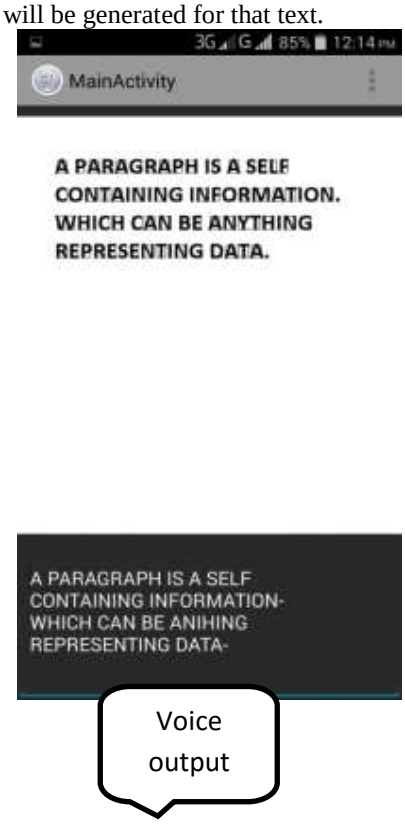

Fig 6

It will show that the voice output will be generated which will heard by the visually impaired people.

4.1Mathematical model:-

_ Input: Text Image
_ Output: Audio Output

_ 1.Set Theory- Let G be the global set
G=I, Op, A, Ds, U
I-Set of Input images

I=i1,i2,i3,. . . . . .in
Op-Set of output Images
OP=op1,op2,op3,. . . . . ,opn
A-Set of Algorithm
A=a1,a2,a3,. . . . . .ak
Ds-set of dataset symbols
Ds=ds1,ds2,ds3. . . . . .dsk
U-ste of users
U=u1,u2,u3,. . . . . .un.

_ 2.Morphism- B:tmap(B)-Capture image(Input Image);
(R,G,B)- RGB Separation(Bitmap);
(Gs)-Grayscale(R,G,B);
(Br)-Blur(Gs);
(Th)- Thresholding(Br);
(Selection Pixel)-Thinning (Th);
(Symbols)- Segmentation (Thinning Bi);
(Symbol Matrix)-Template generation(Symbols) ;
String op-Template Matching(Symbol Matrix);
Voce op-Voice Declaration(op);

Thus we had analyzed feasibility assessment using NP Hard, NP-Complete

5. CONCLUSION

In this paper we've represented a system that scan written text on hand-held object for aiding visually handicapped person.

6. REFERENCES

[1] Portable Camera-Based Product Label Reading For Blind People ,Rajkumar N,Anand M.G,Barathiraja N Department Of Electrical And Electronics Engineering,The Kavery Engineering College, Mecheri, Tamil Nadu, India.

[2] Portable Camera-Based Assistive Text and Product Label Reading From Hand-Held Objects for Blind Persons.Chucai Yi, Student Member, IEEE, Yingli Tian, Senior Member, IEEE, and Aries Arditi

[3] Text Detection in Natural Scene Images by Stroke Gabor Words,Chucai Yi Dept. of Computer Science,The Graduate Center, City Univ. of New York,New York, U.S.A.,e-mail: CYi@gc.cuny.edu. Yingli Tian Dept. of Electrical Engineering,The City College and Graduate Center, City Univ. of New York, New York, U.S.A.,e-mail: ytian@ccny.cuny.edu

[4] Automatic Detection and Recognition of Signs From Natural Scenes,Xilin Chen, Member, IEEE, Jie Yang, Member, IEEE, Jing Zhang, and Alex Waibel, Member, IEEE.

[5] Text Extraction and Document Image Segmentation Using Matched Wavelets and MRF Model Sunil Kumar, Rajat Gupta, Nitin Khanna, Student Member, IEEE, Santanu Chaudhury, and Shiv Dutt Joshi.

[6] Context-based Indoor Object Detection as an Aid to Blind Persons Accessing Unfamiliar Environments. Xiaodong Yang, Yingli Tian Chucai Yi Aries Arditi Dept. of Electrical Engineering, The graduate Center Arlene R Gordon Research Institute, the city college of New York. The City University of New York Lighthouse International NY 10031, USA NY 10016, USA New York,NY 10222,USA {xyang02,ytian}@ccny.cuny.edu cyi@gc.cuny.edu aarditi@lighthouse.org

[7] Experiments with a New Boosting Algorithm.,Yoav Freund and Robert E. Schapire, AT&T Laboratories 600 Mountain Avenue, Murray Hill, NJ 07974-0636 , yoav, schapire@research.att.com

[8] Assistive Text Reading from Complex Background for Blind Persons,Chucai Yi
Dept. of Computer Science,The Graduate Center, City Univ. of New York New York, U.S.A. e-mail: CYi@gc.cuny.edu,Yingli Tian Dept. of Electrical Engineering.The City College and Graduate Center, City Univ. of New York, New York, U.S.A.