IDENTIFY THE TEXT AND PRODUCT LABEL FROM HAND-HELD OBJECTS WITH AUDIO OUTPUT FOR BLIND PERSONS

Ms Komal Mohan Kalbhor[1], Mr. Kale S.D.[2]

[1]PG Scholar, [2]Proffessor, Departement of Electronics and Telecommunication,

SVPM College of Engineering,Malegaon(BK),Maharashtra,India.

.

ABSTRACT

For blind person new approach a identify text name and product labels and product packaging from hand-held objects in their daily lives. To separate the object from cluttered backgrounds or other surrounding objects in the camera view, first developed an efficient and effective motion based method to define a region of interest (ROI) in the video basking the user to shake the object. The work consists of three stages. First is image capturing –by using a mini camera, the text which the user want to read will get captured as an image and have to transfer to the image processing platform. Second is text recognition –by using pre-processing and feature extraction method extract various feature of text image and given to ANN for training. Third is speech output - the filtered text will be shared into system to get an audio speech output. This work will be useful for blind persons in their daily life. The entire process is done with the help of MATLAB software.

*Keywords**: **Hand-held objects, Image capturing, Text, Speech output, Matlab software.**

I. INTRODUCTION

Reading is necessary in today's society. Text is everywhere in the form of printing reports, receipts, bank statements, restaurant menus, product packages, instructions on medicine bottles etc. And while optical aids, video magnifiers, and screen readers can help blind persons and those with less vision to access documents, there are few devices that can provide better access to common hand-held objects such as product packages and objects printed with text such as prescription medication bottles.

This paper presents a prototype system of assistive text reading. There are three main parts included scene capture, data processing, and audio output. The scene capture component collects scenes containing objects of interest in the form of images or video. The data processing component is used object-of-interest detection and text localization to obtain image regions containing text, and text recognition. The recognized text codes is informed by audio output component to the blind user. A Bluetooth earpiece with minimicrophone is employed for speech output. The main contributions of this paper are:

1) a novel motion-based algorithm are used to solve the aiming problem for blind users by their simply shaking the object of interest for a brief period.

2) a novel algorithm of automatic text localization to obtain text regions from complex background and multiple text patterns.

3) a portable camera-based assistive framework to aid blind persons reading text from hand-held objects.

II. RELATED WORK

A. Detecting and reading text in natural scenes

X. Chen and A. L. Yuille[2] ,proposes an algorithm for detecting and reading text in natural images. The algorithm is meant for use by blind and visually impaired subjects walking through city scenes. This text includes stereotypical forms – such as street signs, hospital signs, and bus numbers as well as more variable forms such as shop signs, house numbers, and billboards. This paper selects this feature set guided by the principle of informative feature . calculate joint probability distributions of these feature responses on and off text, so weak classifiers can be obtained as log-likelihood ratio tests.

B. Automatic detection and recognition of signs from natural scenes

This paper presents an method to automatic detection and recognition of signs from natural scenes, and its application to a sign translation task. The proposed approach embeds multiresolution and multistate edge detection, adaptive searching, colour analysis, and affine rectification in a hierarchical framework for sign detection, with different emphases at each phase to handle the text in different sizes, orientations, colour distributions and backgrounds by Xilin Chen, Jie Yang, Jing Zhang, Alex Waibel[3].

C. Wearable Obstacle Avoidance Electronic Travel Aids for Blind: A Survey

Dimitrios Dakopoulos and Nikolaos G. Bourbakis, Fellow [4]presents a comparative survey among portable/wearable obstacle detection/avoidance systems (a subcategory of ETAs) in an effort to inform the research community and users about the capabilities of these systems and about the progress in assistive technology for visually impaired people. The survey is depend on various features and performance parameters of the systems that classify them in categories, giving qualitative–quantitative measures. Finally, it offers a ranking, which will serve only as a reference point and not as a critique on these systems.

D. Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm

The above paper presents by Kwang In Kim, Keechul Jung, and Jin Hyung Kim[5] a novel texture-based method for detecting texts in images. A support vector machine (SVM) is used to analyse the textural properties of texts. External texture feature extraction module is not used, but rather the intensities of the raw pixels that make up the textural pattern are fed directly to the SVM, which works well even in high-dimensional spaces. Next, text regions are identified by applying a continuously adaptive mean shift algorithm (CAMSHIFT) to the results of the texture analysis.

E. Text Detection in Natural Images Based on MultiScale Edge Detetion and Classification

Long Ma, Chunheng Wang, Baihua Xiao[6] propose, a robust method for text

detection in color scene image. The algorithm is depend on edge detection and connected-component. multi-scale edge detection is achieved by Canny operator and an adaptive thresholding binary method. the filtered edges are classified by the classifier trained by SVM combing HOG,LBP and several statistical features, including mean, standard deviation, energy, entropy, inertia, local homogeneity and correlation. k-

means clustering algorithm and the binary gradient image are used to filter the candidate regions and re-detect the regions around the candidate text candidates. Finally, the texts are relocated accurately by projection analysis.
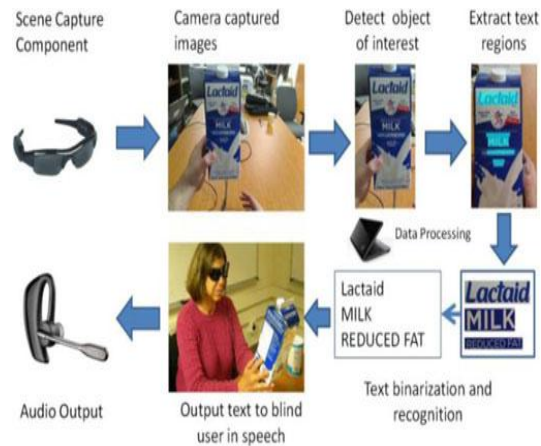
III. SYSTEM ARCHITECTURE

Fig. 1. Flowchart of the proposed framework to read text from hand-held objects for blind users.

This paper presents a prototype system of assistive text reading. As illustrated in Fig. 3, the system framework consists of three functional components: scene capture, data processing, and audio output.

Scene capture

The scene capture component collects scenes containing objects of interest in the form of images or video. In our prototype, it corresponds to a camera attached to a pair of sunglasses.

Data processing component

The data processing component is used for deploying our proposed algorithms, including.

1. **Object-of-interest detection**:-To selectively extract the image of the object held by the blind person from the cluttered background or other neutral objects in the camera view.

2. **Text localization: -** To obtain image regions containing text, and text recognition to transform image-based text information into readable codes.

Audio output component

The audio output component is to inform the blind person of recognized text codes. A Bluetooth earpiece with mini microphone is used for speech output. This simple hardware configuration ensures the portability of the assistive text reading system.
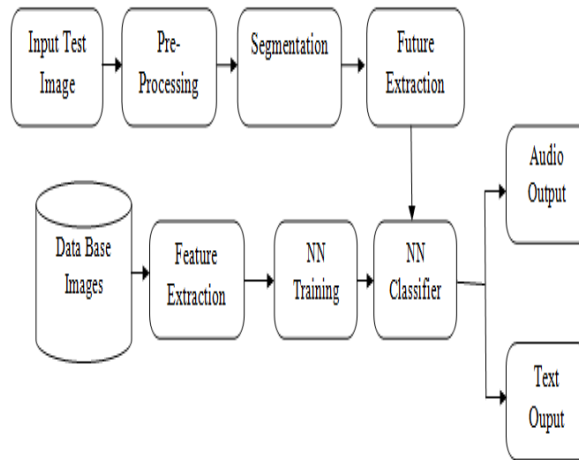
IV. SYSTEM BLOCK DIAGRAM

Fig.2 Block diagram of system.

V. 1INPUT TEXT IMAGE:-

A) Document images :-

 Document images can have text and graphics. This type of images are generated by scanners or camera phones, which acquire printed documents, historical documents, handwritten documents, books, etc.

 B) Scene images :-

 Scene images contain the text, such as the advertising boards, banners, which is captured by the cameras; therefore scene text appears with the background part of the scene. These types of images are very challenging to detect and recognize, because the backgrounds are complex .

Image should be taken by camera, Internet, Database.

5.1 Image Pre-Processing

The next stage is image pre-processing module. Image pre-processing relates to the preparation of an

image for later analysis and use. Images captured by a camera or a similar technique are not necessarily in a form that can be used by image analysis routines. Some may need improvement to reduce noise; other may need to simplified, enhanced, altered, segmented, filtered, etc [4]. The first step in the preprocessing block is to transform the color image into a gray scale image and this result to noisy gray scale image. In the next step, filtering is used in order to cancel the presented noise. Then, edge detection algorithm is applied for obtaining edge of the noiseless gray scale image. Image pre-processing module is consist of following operations-

(i) Gray scale image

(ii) Noise removal

GRAY SCALE IMAGE:

In this proposed system hand image is captured through digital camera so the original image is colored image. For digital image processing it is necessary first colored hand image convert in to gray-scale image. Each pixel has single sample which has intensity information. Now color image is converted in to gray scale image with

noise because there is some noise present in the input colored image due to dust and atmospheric conditions. This noise removal is therefore essential for the system.

NOISE REMOVAL:

The next step in image pre-processing is noise removal. It is necessary to remove the noise from the image because it may produce difference between the actual palm and captured image. This causes the variation in data base feature and measured feature and also affected the accuracy of the system. Edge detection is difficult in noisy image. Noise and the edges contain high- frequency content. Basically the noise produced in the image is due to device using for capturing image, atmosphere condition or surrounding. There are many methods to remove the noise in Matlab. In this proposed system the noise is removed by Lowpass filter. So before extracting features from the image, it is very important to remove the noise from the image. Noisy images are used by operator larger in scope. So less accurate localization of the detected edges are available

5.2 SEGMENTATION:-

The purpose of segmentation is to simplify and change the representation of an image into something that is more meaningful and easier to analyse. Image segmentation is typically used to locate objects and boundaries in images. Simply, 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.

Morphological operation and histogram equalization technique are used.

5.3 EDGE DETECTION:-

Detect edges in the given images various edge detection method is available. Canny operator is chosen because it can detect most edge even in worse condition.

5.4 FEATURE EXTRACTION:-

If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input.

Gray lever co-occurence matrix (GLCM) is used to calculate the special dependence of gray level in an image. In GLCM the number of rows & column are exactly equal to the number of gray levels in the image. Co-occurrence matrixes are constructed in four special orientations(0,45,90 &135 degree). Another matrix is constructed as the average of precending matrixes. Let the co-occurrence matrix be Pij & the size of the matrix is N*N. Each element(i,j) represent the frequency by which pixel with gray level i is specially related to pixel with gray level j. Cosstruction of GLCM from gray scale image is illustrated in fig.2
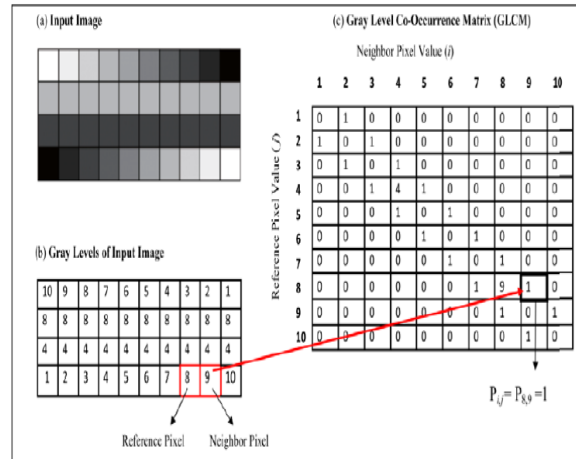
Fig.3. Construction from grey-scale image.

Texture feature are calculated by GLCM contrast ,correlation, Dissimalirity, Energy, Entropy, Mean as shown in table below.

Sl.No	GLCM Feature	Formula
1.	Contrast	$\sum_{i,j=0}^{N-1} P_{i,j}\,(i-j)^2$
2.	Correlation	$\sum_{i,j=0}^{N-1} P_{i,j}\left[\dfrac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}}\right]$
3.	Dissimilarity	$\sum_{i,j=0}^{N-1} P_{i,j}\,|i-j|$
4.	Energy	$\sum_{i,j=0}^{N-1} P_{i,j}^2$
5.	Entropy	$\sum_{i,j=0}^{N-1} P_{i,j}\left(-\ln P_{i,j}\right)$
6.	Homogeneity	$\sum_{i,j=0}^{N-1} \dfrac{P_{i,j}}{1+(i-j)^2}$
7.	Mean	$\mu_i = \sum_{i,j=0}^{N-1} i\left(P_{i,j}\right)\ ,\ \ \mu_j = \sum_{i,j=0}^{N-1} j\left(P_{i,j}\right)$

Table1. Features of GLCM.

5.5 AUDIO OUTPUT:-

The recognized text codes are recorded in script files. Then,we employ the Microsoft Speech Software Development Kit to load these files and display the audio output of text information

ADVANTAGES

* Extract text information easily from complex backgrounds.
* Propose a text localization algorithm that combines rule based layout analysis and learning-based text classifier training.

- System is portable.
- Easy to operate.
- Accuracy is more than Existing system.

APPLICATIONS

- For blind people in mall, public place etc.
- For Illiterate people.

VI. RESULTS:

The input image captured by digital camera is a colored image. Before features are extracted from an image, it may be useful to pre-process the image to reduce irrelevant information or noise and to enhance the image properties that will make feature measurement easier and reliable. There may be random noise that is generated due to different factors such as dirt, dust particles, etc. It can cause significant degradation in the feature extraction process which in turn may lead to higher error rates in the classification process. This noise removal is therefore essential for the system.
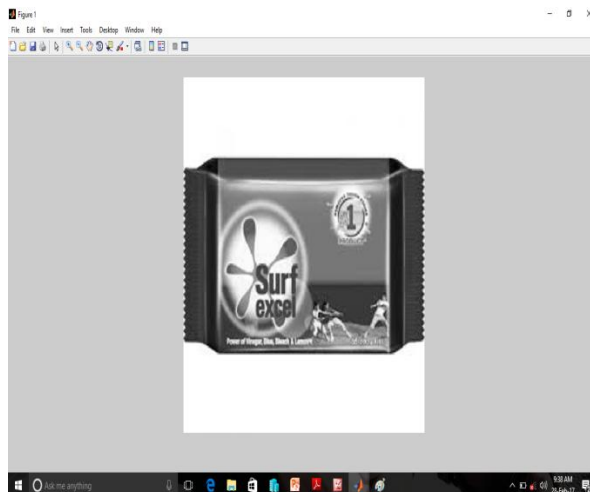

Fig.4 Input original image.
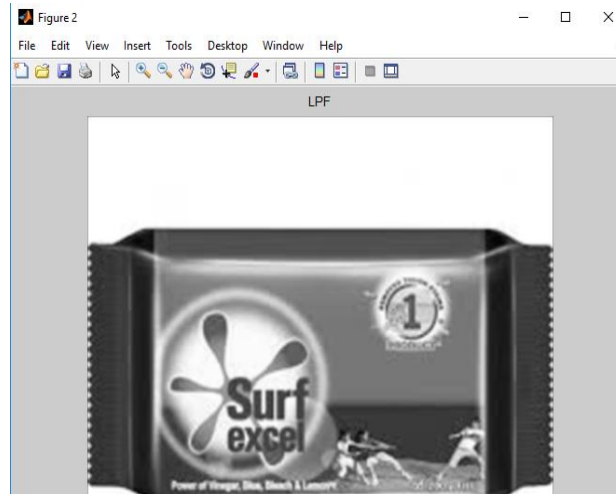

Fig.5 Gray-scale converted image.

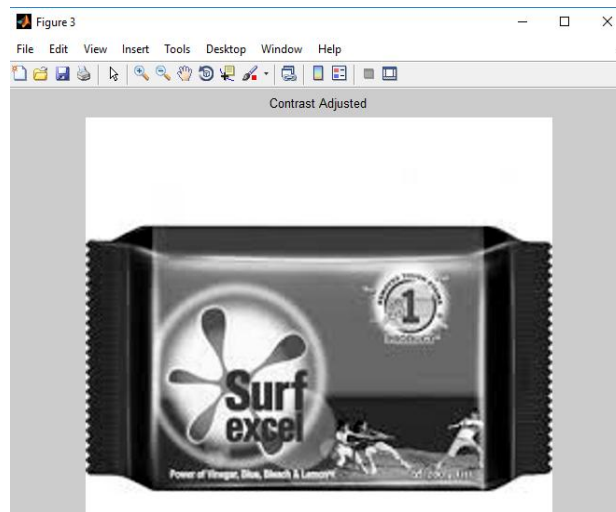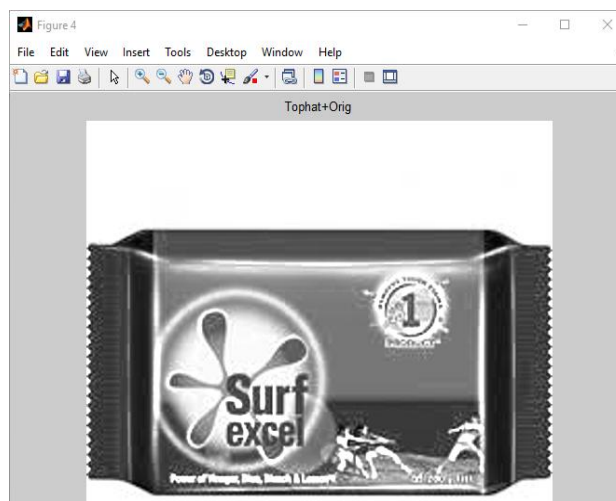**Fig.6 output of low-pass filter.**

Fig.7 Contrast adjustment.

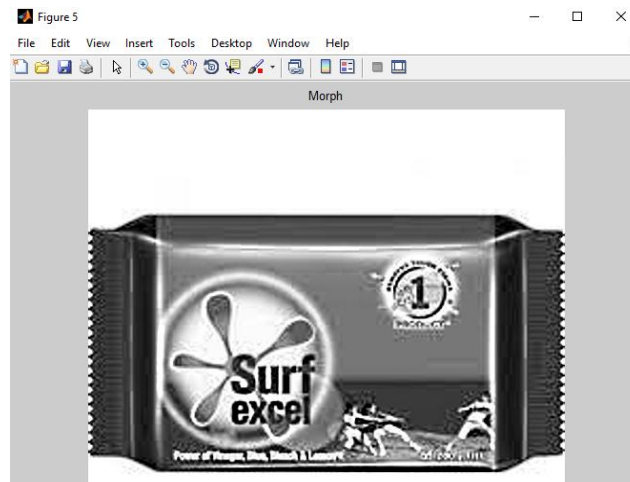**Fig.8 Output of Morphological top-hat operation.**

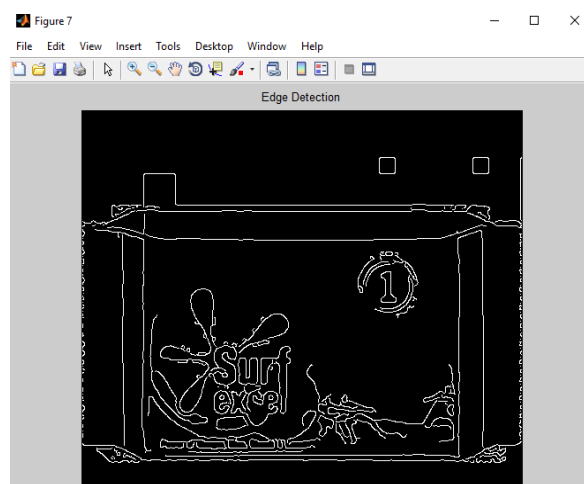**Fig.9 Output of Morphological Bottom-hat operation.**

Fig.10 Output of edge detection.

VII. CONCLUSION

This paper focuses a prototype system to read printed text on hand-held objects for assisting blind persons In order to solve the common aiming problem for blind users, This method can effectively distinguish the object of interest from background or other objects in the camera view. In the proposed system Canny edge detection algorithm is used which will recognize the input image by detecting the edges of objects in the image. GLCM algorithm is used for extract various feature of input image.It is capable of handling the different input images and translates them into text and speech. The proposed system is trained on predefined dataset. The future work we are trying to work in advanced technology such as video conferencing, and try to make android application.

REFERENCES

[1   Chucai Yi, Yingli Tian, Aries Arditi,'' Portable Camera-Based Assistive Text and Product Label Reading From Hand-Held Objects for Blind Persons", IEEE/ASME TRANSACTIONS ON MECHATRONICS, vol. 19, NO. 3, June 2014

[2]X.   Chen and A. L. Yuille, "Detecting and reading text in natural scenes," in Proc. Compute. Vision Pattern Recognise. 2004, vol. 2, pp. II-366–II-373

[3]   Xilin Chen, Jie Yang, Jing Zhang, and Alex Waibel, "Automatic Detection and Recognition of Signs from Natural Scenes," *IEEE TRANSACTIONS ON IMAGE PROCESSING, vol. 13, no. 1, January 2004*

[4]   K. Kim, K. Jung, and J. Kim, "Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm," IEEE Trans. Pattern Anal. Mach. Intel., vol. 25, no. 12,pp. 1631–1639, Dec. 2003.

[5]   M. A. Faisal, Z. Aung, J. R. Williams, and A. Sanchez, "Data-stream based intrusion detection system for advanced metering infrastructure in smart grid: A feasibility study," *IEEE Syst. J.*, vol. 9, no. 1, pp. 1–14, Jan. 2014.

[5]   L. Ma, C. Wang, and B. Xiao, "Text detection in natural images based on multi-scale edge detection and classification," in Proc. Int. Congr. Image Signal Process., 2010, vol. 4, pp. 1961–1965.

[6]   C. Yi and Y. Tian, "Text detection in natural scene images by stroke gabor words," in Proc. Int. Conf. Document Anal. Recognit., 2011, pp. 177–181.

[7]   World Health Organization. (2009). 10 facts about blindness and visual impairment [Online]. Available: www.who.int/features/factfiles/ blindness/blindness_facts/en/index.html

Advance Data Reports from the National Health Interview Survey (2008).[Online]. Available: http://www.cdc.gov/nchs/nhis/nhis_ad.html.

BIOGRAPHIES:

Ms KOMAL MOHAN KALBHOR

Pursing Master Engineering(M.E.E&TC), From SVPM College of Engineering, Malegaon(BK),Maharashtra,India

Santosh D.Kale:- currently working as a Assistant Professor at college of engineering, Malegaon (Bk), Baramati. He received B.E.Degree in Electronics & Telecommunication in 2001, from North Maharashtra University of Jalgaon, Maharashtra, India.he received M.Tech Degree in (Electronics Instrumentation) in Electronics & Telecommunication, from college of Engineering,Pune(COEP), India,he guided several UG & PG projects.his research area includes signal and image processing.

E-mail-santosh007_kale@rediffmail.com.