JOURNAL OF COLLEGE OF EDUCATION FOR PURE SCIENCES
Vol. 8, No. 3, September 2018
Website: jceps.utq.edu.iq Email: jceps@eps.utq.edu.iq

Smart Vision System forthe Blind People using Image Segmentation and Edge Detection

Dr.Zainab N.Nemer
drzainabnemer@gmail.com
Dept. of Computer Science, College of Computer Science and InformationTechnology, UniversityofBasrah,Basrah,IRAQ.

Abstract
The proposed system OTSS (Object and Textconversion to Sound System) is designed to describe objects, answer questions, read texts and offer guidance about blind people environment. Thesystem helps the visually impaired people to shopping; reading and navigating within any new environment by recognize the environment objects and tell the blind people what their environment has.The efficiency use of image segmentation makes the algorithm more compatible. The conversion of objects and a text to speech synthesizer will be of great help to blind people.

Keywords: Image Processing, Text Recognition, Blind People, Image Segmentation, Edge Detection, Indoor Navigation, Assistive Technology.

1. Introduction

Smart vision's model depend on the path is tracked, detect the obstacles and cover few meters in front of user by using one or two small camera ,CPU and earphone mounted in dark class[1]. The system depends on image processing and speech recognition to drive the robot automatically on the road stop at zebra–crossingtraffic red light and stop at end point. The input image is translated into digital signal to the laptop. The system could be used as driven assistant. Voice command could also be used forpeople cannotdrive [2].The objects are detected from the video by converting this video to frames and processed them. The algorithm is started with noise filtering, segmenting to enhancement the image and edge detection to detect the obstacles [3].The navigation could be enhanced by detecting the obstacles and finding the location in unfamiliar environment [4].There are two challenges are face the impaired people:first is the obstacle avoidance (local problem), second is the wayfinding(global problem). The impaired user can determined the location and orientation in an indoor environment.The system based on database of the building layout[5].Freely moving without any help, at public and private buildings, and in open places like the streets can be defined themobility. Navigation system is used to facility the blind people mobility and makes it moresafety and easier. The system provides the blind people with necessary informationabout the path [6].Smart phone applications likes bluetooth and wifi are used for transmitting messages in indoornavigation. Remote processing computer analyses the images to compute the orientation of the blind and guide him to safe destination [7].The navigation system can detect path borders, moving obstacles and static obstacle to instruct the blinds to correct their direction on path. Imagesquality is the measurement of images similarities and finds thedifferences between two images. The simplest and widely used method is the Mean Square Error (MSE) [8].All human machine interaction used speech synthesis to make alarms and to give information. Synthetic speech is very useful and important in many applications.Blind people can read and communicate with the aid of speech synthesis. It can be used in many educational purposes with different languages and in telecommunications and multimedia like E-mail messages and spoke mobile messages [9].Unfortunately 161billion are visually impaired and 37billon are bind people. Many modern smart technologies are founded to reduce challenges that are faced them and to assist such human being. Theenvironment can be recognized and extracted by taking images and detected text, and objects using cross correlation algorithm [10]. Speech is the efficient medium for communication.A text to speech synthesizer (TTS) is a computer based system which is scanned text and converts it into a computer format [11].Speech processing is based on speech segmentation. Frame features and frame energy are extracted from input audio. Zero crossing rate and pitch value are

calculated. Background noise is detected [12].The speech synthesis can be build depending on recorded database of alphabet and numbers in the form of wave files [13].

2. TheProposal System

The proposed system OTSS (Object and Text conversion to Sound System) is composed of three stages, as shown in Figure (1). In the first stage, basic objects in the image are extracted and recognized such as,door, window, sofa, chair and so on. In the second stage, any text in the image is segmented from the background. Finally,convert objectsand text to sound to alarm the blind people. The proposed system can process images at different resolution. The system can take data from video also.Operational stages of the system can be summarized as:

1. Capture the image.
2. Image pre-processing (filtering, edgedetectionand segmentation, morphological operations...)
3. Recognize the objects and label them to convert the names of these objects tospeech synthesizer
4. Text is extracted depending on canny filter.
5. Convert object's name and Text tospeech.

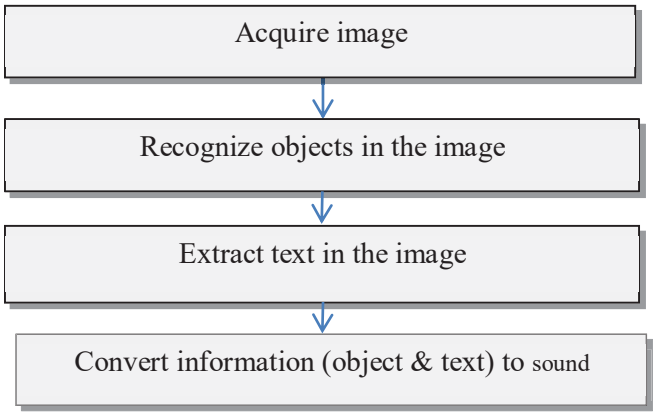


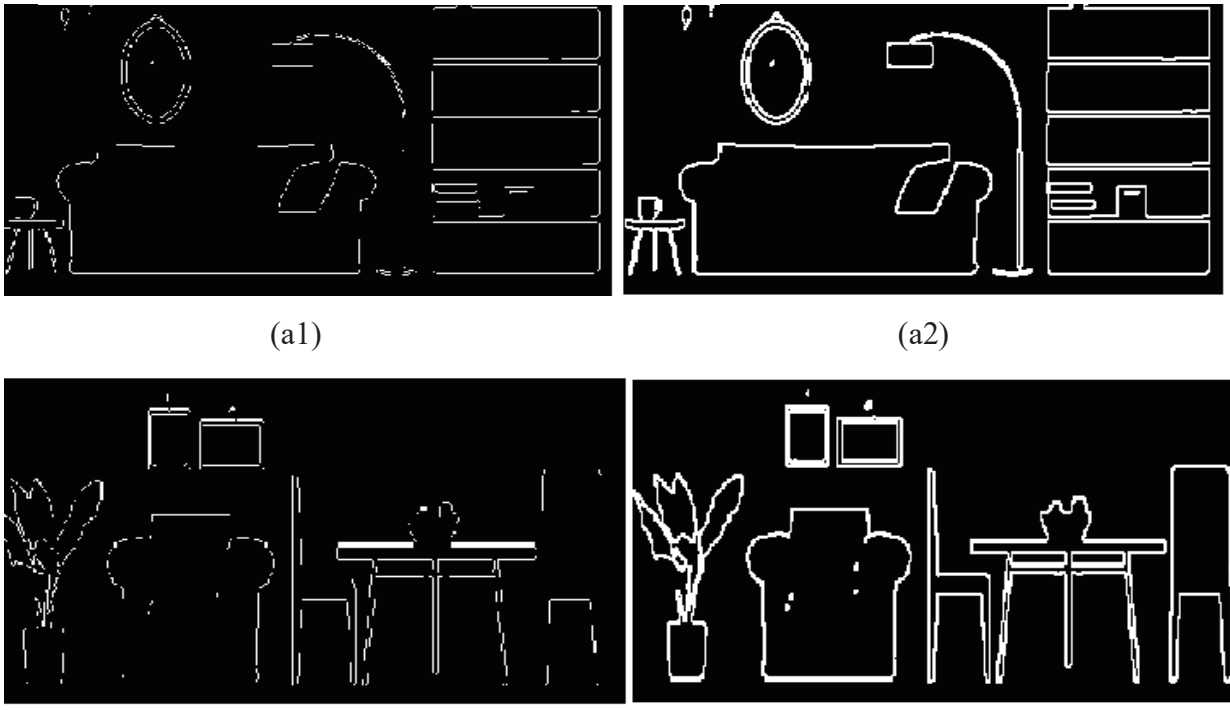
Figure (1): The OTSS stages.

2.1 Objects Recognition

In this stage of the OTSS, the imageprocessing steps have been performed on the input images. These images have taken from any type of camera like phone camera.The images are scanned and converted to binary images. They aredetected and filled and filtered to extract the objects in these images. The selected threshold is very important to separate objects from the back ground. Sobel filter is used to find the objects edges. Filling the holes is implemented to find the exact boundaries of the desired objects and reject any other smallobjects Figure (2) summarized the previous steps.The OTSS depend

on the mean square error (MSE) to estimate the sum of square of the error between two images as in Equation (1) [8][14].

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(x(i,j) - y(i,j)\right)^2 \quad \text{.....(1)} \quad \text{...}$$



(a1) (a2)

(b1) (b2)

Figure (2): two examples of input images, a1, and b1 are the images before objects are recognized.a2, and b2 are the images after processing and detecting objects (categories) in the images.

The extraction phase of the OTSS is summarizedin recognizing of objects in the images. The database is divided into five different categories like sofa, chair, armchair, window and door as in Table (1).The correct category is the category with the lowest error value.After the main category is determined, and to involve all the changes in same object, the input image is compared with the types of that object in the database. The selected category is divided into four branches to consider all the differencesand for best results in the recognition as in Tables (2) and (3). Table (2) shows the MSE value between the input image and the types of sofa in the database. Table (3) shows the MSE value between the input image and the types of window in the database.

Table (1): The MSE error between an input image histogram and the histogram of the proposalcategories (sofa, chair, door, armchair, and window) of the system.

6.0028e+03	6.1035e+03	4.1933e+04	4.1751e+04	3.1219e+04

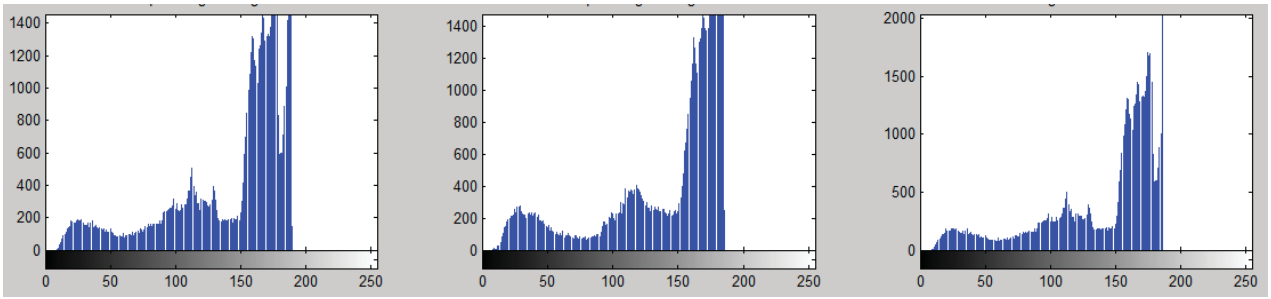
Table (2): The MSE error inside category 1(sofa).If input image is a sofa, it is compared with all types of the sofas in the database.

Cate1.1	Cate1.2	Cate1.3	Cate1.4
6.0028e+03	1.4253e+04	1.0416e+04	1.9846e+04

Table (3): The MSE error inside category 5 (window).If input image is a window, it is compared with all types of the windows in the database.

Cate5.1	Cate5.2	Cate5.3	Cate5.4
4.091484e+02	7.050391e+02	1.88335e+02	1.544219e+02

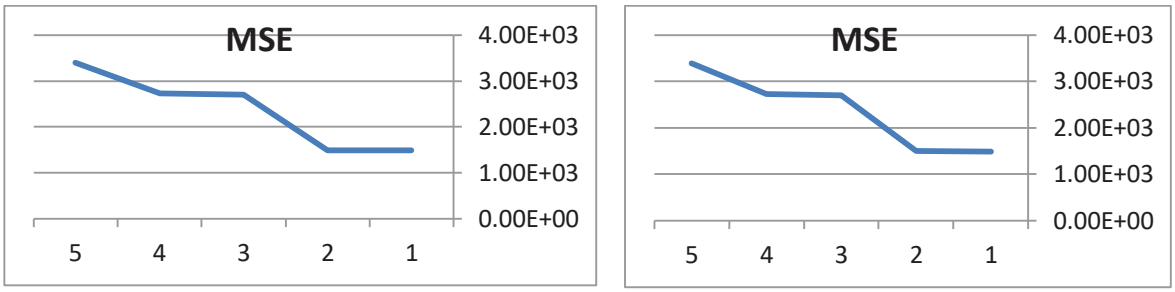
The OTSS used histogram values for the images to found the MSE results as in Figure (3).The system has perfect result compared with those results without histogramvalues as shown in Figure (4).



(a) (b) (c)

Figure (3): (a) input image histogram,(b) template image histogram, (c) matched histogram.

Matching algorithm is used to match each of the extracted objects with objects in the database. After objects recognition process is terminated theobject recognition stage is finished and this will be as preparation for the third stage of OTSS. The input image is compared with the images of the database. The database is rearranged and organized in the manner where each image has a code to simplify the objects reorganization process.



(a) (b)

Figure (4): (a) is a graph shows the error between two images using MSE on two images. (b) is a graph shows the error between two images using Histogram in MSE.

2.2 Text Extraction

The OTSS intend to inform the blind person if there are any phrases in the around environment. The extraction based on several steps as a pre-processing operation to enhance the results. The input images are converted to gray scale to separate text from background. Gray image is converted into black and white pixel for more perfect separation from the background. Filter and smooth the image is very important to discard the noise from them. The process of add or remove pixel is used through morphological operations. Multiplying the resultant image with original image is done to extract the text characters [15].

(a) (b) (c) (d)

Figure (5): (a), (b), (c), and (d) are four examples which are processed by OTSS and texts areextracted.

The proposed system has the ability to recognize the more complicated natural images. It based on detecting the Maximally Stable Extremal Regions (MSER) algorithm to segment the text on the

constant regions. The system uses the canny filter to distinguish the edges. The OTSS can treat the Arabic text as shown in Figure (5) (c). Arabic text could be processed easily in OTSS as connected objects and this is really the exact feature of the Arabic text. The text is segmented into words as preparing for the next stage the speech synthesis.

2.3 Speech synthesis

Similarity the speech to the human voice is used to measure the quality. The speech synthesis part uses the results of the previous parts of OTSS.It depends on image recognition and text extraction to convert the text to speech as shown inFigure (6).

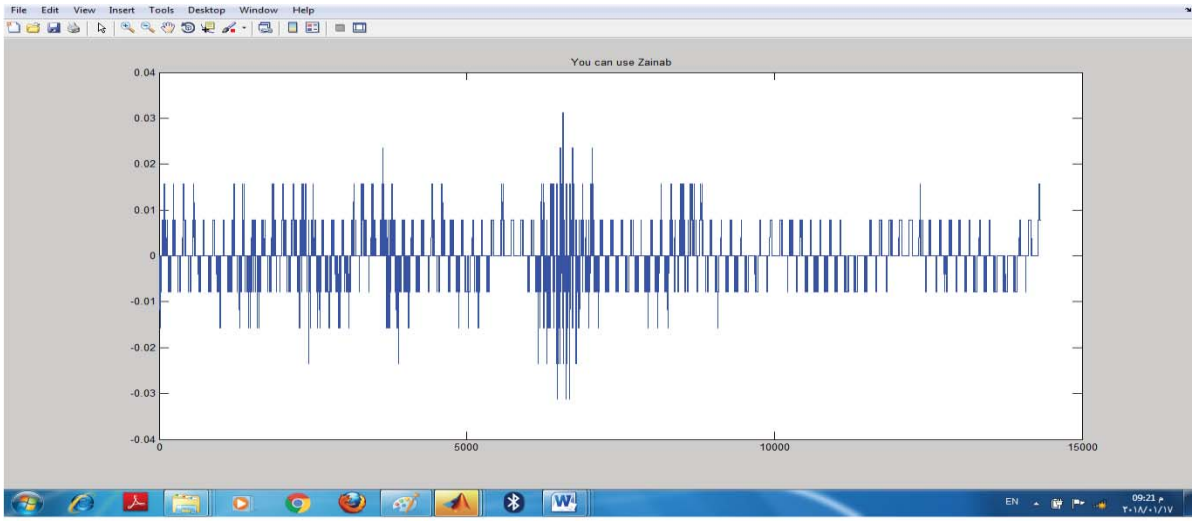


Figure (6): Example of speech synthesis output.

3. Experimental Results

The system depends on matlab to analyze images and recognize objects with low memory size and faster instructions. The SUN database is used in the system, Figure (6). It has 908 scenes categories and 3819 object categories [16]. The OTSS is operated on 40 images.It recognized 70% of images successfully.

Figure (7): samples from the SUN database.

The design of the OTSS's graphical user interface is showed in Figure (8).The system is experimented in a school to teach the children the English language.

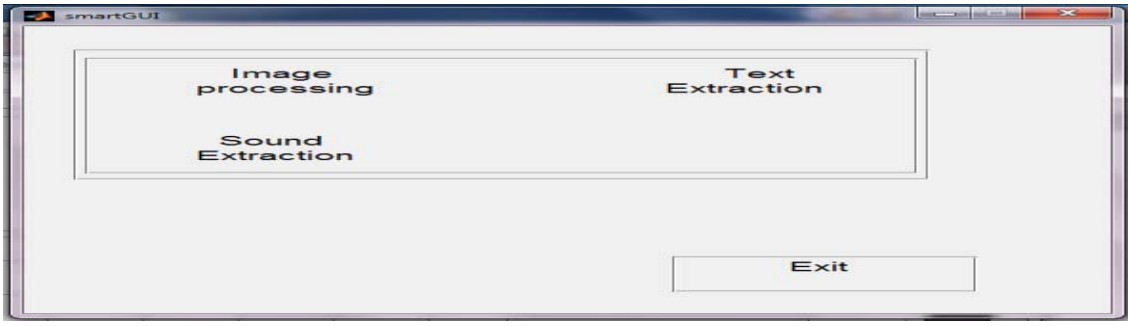


Figure (8): The design of OTSS GU.

4. Conclusions

The system can produce instructions to lead the blind person to desired destination. We proposed system based on smartphones or black glass or any other camera with good resolution. The OTSS could be used by impaired people.

5. Future works

The OTSS could be improved by using the eye to invoke and initiate the objects of the system. It can process the videos instead of images.

6. References

[1] João José, Miguel Farrajota, and João M.F. Rodrigues, " The Smart Vision Local Navigation Aid for Blind and Visually Impaired Persons", International Journal of Digital Content Technology and its Applications, Vol.5, No.5, May 2011.

[2] Shashank," Robot Navigation Using Image Processing And Isolated Word Recognition", International Journal on Computer Science and Engineering (IJCSE), Vol.4, No.05, May 2012.

[3]Sadique Waris, and S.B Somani," Indoor Navigation Approach for the Visually Impaired ",International Journal of Emerging Engineering Research and Technology ,Vol. 3, Issue 7, PP 72-78 , July 2015.

[4] R.R.Bhambare, and Akshay Koul, "

Smart Vision System For Blind ", International Journal of Engineering and Computer Science, Vol. 3 Issue 5, PP. 5790-5795, May 2014.

[5] B.S. Tjan, P.J. Beckmann, "Digital Sign System for Indoor Wayfinding for the Visually Impaired",in Proceedings of the First IEEE Workshop on Computer Vision Applications for the Visually Impaired, 2005.

[6] Edoardo D'Atri, Carlo Maria Medaglia, Emanuele Panizzi," A System to Aid Blind People in The Mobility:A Usability Test and Its Results",Second International Conference on Systems (ICONS'07), 2007.

[7]Vasuki.D, Preethi.T,Sabareeswaran.B,ArunPrasad.M and Mrs.S.Saritha.M.E," Intelligent Navigation System for Blind Persons ",International Conference on Explorations and Innovations in Engineering & Technology (ICEIET), 2016.

[8]C.Sasi Varnan, A.Jagan, Jaspreet Kaur, Divaya Jyoti, D.S.Rao, " Image Quality Assessment Techniques PN Spatial Domain", IJCST, Vol. 2, Issues 3, September 2011.

[9] Mohd Bilal Ganai1, and Er jyoti Arora," Implementation of Text to Speech Conversion Technique ", International Journal of Innovative Research in Computerand Communication Engineering, Vol. 3, Issue 9, September 2015.

[10] PriyankaRathod, and Suvarna Nandyal," Image Acquisition and Text To Speech Conversion for Visually Impaired People ", International Journal of Engineering and Technology, Vol. 6, No.8, August, 2016.

[11] Pooja Chandran, Aravind S, Jisha opinath, and Saranya S S," Design and Implementation of Speech Generation System using MATLAB ",International Journal of Engineering and Innovative Technology (IJEIT), Vol. 4, Issue 6, December 2014.

[12] Mallikarjun Mulgi, Vijaykumar Mantri, and Gayatri. M.," Voice Segmentation without Voice Recognition ", IJACKD journal of research, Vol. 2, ISSUE 1 ,PP.2 2, February 2013.

[13] N. Swetha , K.Anuradha," Text-To-Speech Conversion ",International Journal of Advanced Trends in Computer Science and Engineering, Vol.2 , No.6, pp. 269-278, November 2013.

[14] Kalpana Chaurasia, and Nidhi Shama,"Performance Evaluation and Comparison of Different Noise, apply on PNGimage Format Used in Deconvolution Wiener Filter FFT Algorithm",Evolving Trend in Engineering and Technology,Vol.4, PP. 8-14,Switzerland,2015.

[15] Amit Choksi, and Nihar Desai, "Text Extraction from Natural Scene Images using Prewitt Edge Detection Method ", International Journal of Advanced Research in Computer Science and Software Engineering, Vol.3, Issue 12,December 2013.

[16] https://archive3d.net/?category=3.