Middle-East Journal of Scientific Research 24 (11): 3479-3484, 2016
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2016.3479.3484

Voice Based Application As Medicine Spotter For Visually Impaired

D. Jayashree, K. Afritha Farhath, R. Amruthavarshini and S. Pavithra

Department of Computer Science and Engineering, Rajalakshmi Institute of Technology, Chennai, India

Abstract: This paper proffers an application in order to expedite easy and innate way to find the medicine for the visually impaired people and to take it according to their Doctor's prescription. Visually impaired people need not be dependent and seek others help to find the medicine to be taken. This android application is used to overcome the difficulties they face in this scenario. In this application, a reminder is set which tells the user when to take the medicines, as voice output. The pictures of the medicine strip held in the hand are captured by the inbuilt camera of the mobile. The image is processed and consequently text localization and extraction is done by which the name of the medicine is identified. A spotter section is also consolidated with this application which checks the prescription which has been already uploaded in the user's mobile, compares with the name of the medicine identified and if the medicine has to be taken at that time, then it tells the quantity of medicine to be taken to the user as voice output. On receiving the voice output from the mobile, the user intakes their medicines according to their prescription. It can also be useful for uneducated people who suffer to find which medicine must be taken. Especially elder people who are not educated usually suffer to read their medicine names on their own. This idea would achieve good results in practice.

Key words: Image Processing · Visually impaired · Medicine identification · Text localization and Extraction · SQLite

INTRODUCTION

Visually challenged people and uneducated people face a lot of adverse challenges in their day to day life. Most of the time they are perplexed in a new environment or surrounding due to issues related to accessibility. So, this prevents them from experiencing the world in the same way as others do. Identifying and accessing things is something many of us may take it for granted. But the visually challenged people are curbed by their disability. Especially in a medicine taking scenario, it is difficult for them to find whether they have identified the medicine correctly or not. They will have to seek others help for it. Moreover a mobile application will be easy to use and the hardware needed is very limited.

In this paper we propose an image processing based android mobile application that provides top- to-bottom guidance and assistance to the visually impaired user for taking their medicines. Throughout the process the user is guided using the voice output rather than text. The Android platform has been used to build this app mainly because of its wide popularity and cost effectiveness in the smart phone market. Android platform has Comprehensive libraries for image processing, SQLite for facilitating data storage and good hardware features for video or image capture. Also it is available on mobile phones from various manufacturers, from Sony Ericsson, Motorola and HTC to Samsung. There are 3 basic modules into which the application has been split up as reminder, identifying the medicine by label reading and voice output.

Related Works: Chucai Yi, Student Member, IEEE, 2013 [1] in the work proposed a camera-based assistive text reading framework to help blind persons read text labels and product packaging from hand-held objects in their daily lives. This paper proposes a Gaussian based approach in which initially the object of interest is identified, followed by region of interest identification and performing various image processing operations on the identified image to retrieve the desired text. Here the object from which the text is to be extracted is isolated [3]

Corresponding Author: D. Jayashree, Department of Computer Science and Engineering, Rajalakshmi Institute of Technology, Chennai, India. E-mail: hod.cse@ritchennai.edu.in.

[4] from the background by shaking the object using motion based object detection. The captured series of frames are analyzed to find the foreground object followedby applying mean of the estimated foreground masks. Then the region of interest is found where the text is localized based on edge or textural properties and finally the desired text is extracted [5]. But the above paper has few a imitations because it uses a separate system for processing, a wearable camera and a Bluetooth earpiece which increases the hardware utilization K. Matusiak, Lodz University of Technology, IEEE 2013 [2], in his work he describes main features of software modules developed for Android smart phones that are dedicated for the visually impaired users. The main module can recognize and match scanned objects to a database of objects, e.g. food or medicine containers. The two other modules are capable of detecting major colours and locate direction of the maximum brightness regions in the captured scenes but it merely helps in object recognition by matching objects with database objects, other accessibility or communication issues are not addressed.

There are quite a few applications available in the market as of now. Recognizer developed by Look Tel [16] is a commercial application dedicated for iPhones, that is supposed to recognize an object within the camera field of view that was previously stored in a local database of object's images. Here and Now, an iPhone app that uses the camera of the iPhone to retrieve product information.

There are many screen reader applications such as AI Reader [17] in android platform and Voice Over [18] in IOS which merely gives the voice output for the text on the screen. Scan Life Barcode and QR Reader [19] an application that can read UPC and QR codes. Once a code is scanned, the app reads the embedded string as a QR code. This is certainly useful for people who have a hard time in identifying their medicines.As a tablet strip has no barcode, the existing systems are not efficient.

Proposed Work

Image Processing Module: The user gets a reminder as a voice output which tells when to take the medicines and this reminder is set based on doctor's prescription. The reminder also gives the information whether the medicine should be taken before food or after food. Before the user could take the medicine, it would be good if he could know if the medicine he has in his hand is the correct one or not. In order to facilitate this, the application allows the user to search for the medicine details and retrieve information about the medicine using voice commands. This will help the visually impaired user to get an overall view about the medicine name and course details of the medicine. In this module, initially the search item is received as voice input from the user and the corresponding details of the medicine is retrieved from the database. These results are given to the user by converting it to speech.

Obtaining Product Details: This is the most important and essential module for the visually challenged user to identify and choose the medicine he has to take. The proposed application uses the inbuilt camera of the android device for capturing the medicine image. The captured image is initially processed using text localization algorithms to separate the text from the background. Then text extraction methods are used to extract the name of the medicine or the label. The extracted text is then compared with the prescription in the database, if the extracted name matches with the name in the prescription then the medicine details are given as a voice output is given to the user.

Course Details: In this module, the application has the course details which is used for the identifying medicines and send them all to the user through voice output.

Implemetation: This paper proposes a prototype of the interactive medicine taking guide assistant application for visually challenged.

Image Processing Module: We have used a Samsung Galaxy S3 mini for our experimental work. It has 153600 pixels(320x480) .On capturing the image through the phone we display it in the designated area through programming which gets stored in the RAM. On further processing of the image text [7] label is identified.

Grey Scale Conversion: Initially the object from which the text has to be extracted is identified by means of background separation method and then the image of the desired object has to be converted to a grayscale image in the YUV colour space. The method we are using uses only the Y channel (luminance) of the YUV colour space so that it can be input to other thresholding functions during future processing. Once this is done luminance value thresholding is done in order to spread luminance values throughout the image thus increasing the variation in contrast between the required regions of the image and the rest.

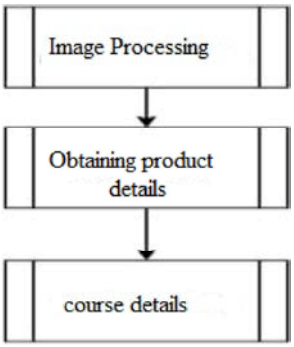

Fig. 1: Process Modules

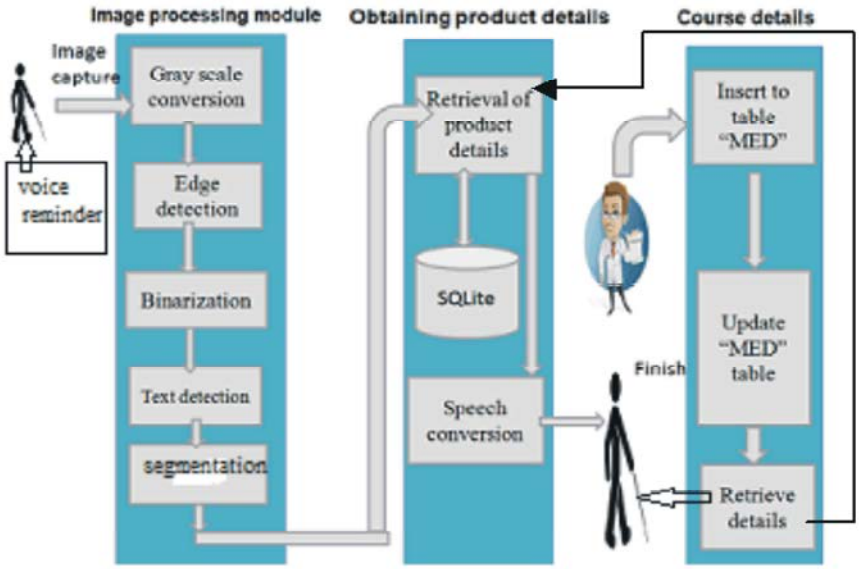

Fig. 2: Architecture Diagram

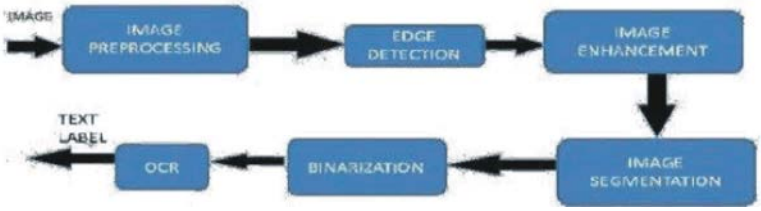

Fig. 3: Image Processing Steps

Fig. 4: Input Image

Fig. 5: Converted Grayscale Image

Edge Detection: Once the image is converted into the YUV colour space the next step is to detect the edges in the image. There are a lot of edge detection algorithms available using various operators Canny, Robert, Prewitt, Laplacion of Gaussian etc. Canny edge detection algorithm finds edges by looking for local maxima of the gradient to detect strong and weak edges, including weak edges in the output if they are connected to strong edges. Laplacian of Gaussian edge detection algorithm finds edges by looking for zero-crossings after filtering I with a Laplacian of Gaussian filter. Prewitt algorithm finds edges using the Prewitt approximation to the derivative. It returns edges at those points where the gradient of I is maximum. Robert's algorithm finds edges using the Roberts approximation to the derivative. It returns edges at those points where the gradient of I is maximum.

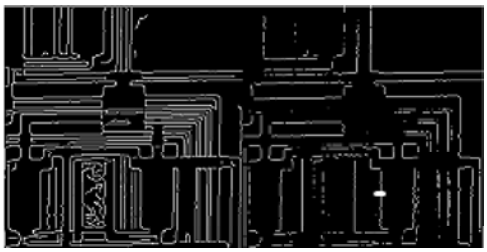

Fig. 6: Canny vs Prewitt [15]

By using two thresholds, the Canny method is less likely than the other methods to be fooled by noise and more likely to detect the weak edges. Thus, Canny edge detection method is used.

Binarizing the Image: The reduced image has to be binarized by erasing out all pixels outside the bounding boxes. (i.e.) All pixels outside the bounding box will be set to white (0) and those within the box will be set to black (255). The next step is to fill the gaps within the bounding box. To do this consider this condition: If one white pixel is surrounded by two black pixels in any direction either horizontal or vertical, then it is also filled with black (255). Finally using both the edge image generated in step 1 and the gap image generated in this step, the text coordinates are refined [10].

Fig. 7: Binarized Image

Text Detection: Once the edge image is got it is analyzed to find out text regions within it. For this a Histogram is computed which will tell how the pixel [8] intensities varies. Texts in an image will usually have a higher contrast than the rest of the contents. So this idea is used here to detect texts. Two thresholds are used to find the local maxima. To find out if a line from the edge image can be a part of the text that is going to be detected two conditions should hold true.

- It should contain a sufficient number of sharp/minimum edges.

P1 (belongs to)H
P1(y) > Minedges

- Difference between the edge pixels in one line to the previous line is bigger than a threshold.

P1(y) – P1(y-1) > MinlineDiff

- Here P1(y) refers to a line along with y-axis from the edge image.
- H refers to the Histogram computed.

In this way, a text region is isolated which will have several texts aligned horizontally with the y-coordinates defined. The X coordinates are then calculated for the left, right, top and bottom of the detected text. This step ends with the creation of bounding boxes for the text regions by Figureuring out the exact co-ordinates.

Image enhancement is done by separating out texts according to their geometric properties [9]. Those characters or parts of text in the text region which do not have their height, width and other geometrical properties within a predefined set of values will be eliminated. This generates a reduced image from the original edge image.

Fig. 8: Text detection

Segmentation of the Image: To do this the text candidates are extracted from the grayscale image [11]. Segmentation concludes by enhancing the contrast on the obtained text image.

Crocin pain relief

Fig. 9: Extracted Text

Database Information Retrieval: Since the scope of the proposed application is limited to a specific person, the information related to the medicines of the user are readily stored in the database. To achieve this, SQLITE LIBRARY of the android framework is used and the details are stored in the database with necessary key constraints. Also details of the medicine availability in a specific location are stored similarly. When the user wants to know the medicine information, the name of medicine obtained in the image processing module is used. The recognized medicine name would be passed as a search parameter to the database function and the resulting information is given to the visually impaired user using voice commands. Similarly other information regarding medicine such as composition and quantity details would be provided using voice commands.

Text To Speech Conversion: A spotter section consolidated with this application uses the text recognized to check whether the name of the medicine in user's hand matches with the prescription which has been already uploaded in the user's mobile, compares with the name of the medicine identified and if the medicine has to be taken at that time, then it tells the quantity of medicine to be taken to the user as voice output. On receiving the voice output from the mobile, user intakes their medicines according to their prescription.

CONCLUSION

In this paper we have suggested an application for the visually impaired and the uneducated people to provide complete assistance while taking medicine through label reading. In order to do this we have presented an approach to detect, localize and extract texts appearing in grayscale or colour images. This is based on employing a colour reduction technique, a method for edge detection and region segmentation and selecting text regions based on their horizontal projection and geometrical properties [14]. The application is implemented on the android platform owing to its ease of use and wide market. All the input and output are given by means of speech in order to address the accessibility issues of the visually impaired.

Future work includes enabling a multiple login option in a same device so that more than one visually impaired person can use the same device. To update the prescriptions into database on their own without doctors help by using a printed prescription.

REFERENCES

1. 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, 2014.
2. Mobile Camera Based Text Detection and Translation Derek Ma Department of Electrical Engineering Stanford University, Qiuhau Lin Department of Electrical Engineering Stanford University, Tong Zhang Department of Mechanical Engineering Stanford University.
3. Extraction of Text in images Malik, R. Dept. of Electr. Eng. & Computer Eng., New Jersey Inst. of Technol., Newark, NJ, USA, 1999.
4. Text Extraction from Document Images Using Edge Information Author(s) Grover, S. Nat. Inst. of Technology., Rourkela, India Arora, K. ; Mitra, S.K, 2009.
5. Morphological text extraction from images Author(s): Hasan, Y.M.Y. Dept. of Electr. Eng., Arizona State Univ., Tempe, AZ, USA Karam, L., 2000.
6. Multiscale Edge-Based Text Extraction from Complex Images Author(s)Xiaoqing Liu Dept. of Electr. & Comput. Eng., Univ. of Western Ontario, London, Ont. Samarabandu, J, 2006.
7. Scene text extraction in natural scene images using hierarchical feature combining and verification Author(s):Kim, K.C. Dept. of Comput. Sci., Yonsei Univ., South Korea, 2004.
8. Liang, J., D. Doermann and H.P. Li, 2005. "Camera-based analysis of text and documents: a survey," International Journal on Document Analysis and Recognition IJDAR), 7: 84-104.

9. Thillou, C. and B. Gosselin, 2007. "Natural Scene Text Understanding, "In Vision Systems: Segmentation and Pattern Recognition, Ed. Croatia: I-Tech education And Publishing, pp: 307-332.
10. Jung, K., 2004. "Text information extraction in images and video: a survey," Pattern Recognition, 37(5): 977-997.
11. Jain, A.K. and B. Yu, 1998. Automatic Text Location in Images and Video Frames, Pattern Recognition, 31: 2055-2076.
12. Agnihotri, L. and N. Dimitrova, 1999. Text Detection for Video Analysis. In Proc. of the International Conference on Multimedia Computing and Systems, Florence, Italy, pp: 109-113.
13. Li, H. and D. Doermann, 1998. Automatic Text Tracking In Digital Videos. In Proc. of IEEE 1998 Workshop on Multimedia Signal Processing, Redondo Beach, California, USA, pp: 21-26.
14. Wu, V., R. Manmatha and E.M. Riseman, 1997. Finding Text in Images In Proc of Second ACM International *Conference on Digital Libraries*, Philadelphia, PA, pp: 2326.
15. http://in.mathworks.com/help/images/ref/edge.html #buo 5g3x-1
16. http://www.looktel.com/
17. https://play.google.com/store/apps/details?id=com .neverla nd.alreader&hl=en
18. http://www.imore.com/how-use-voiceover-iphone-and- ipad
19. http://www.androidtapp.com/scanlife-barcode-qr-reader/