MOTION BASED OBJECT DETECTION METHOD FOR VISUALLY IMPARIED PEOPLE

Article · December 2018

4 authors, including:

Saranya .K
Sri Ramakrishna Institute of Technology

MOTION BASED OBJECT DETECTION METHOD FOR VISUALLY IMPARIED PEOPLE

R N Devendra Kumar[1], A.Praveena[2], N.V. Shibu[3], K.Saranya[4], D. Betteena Sheryl Fernando[5]
Assistant Professor/CSE [1][2][3][4][5],
Sri Ramakrishna Institute of Technology [1][3][4][5], Jansons Institute Technology[2], Coimbatore

ABSTRACT

Object recognition by the blind people often requires a manual intervention more which is more difficult to perform. In this proposed design, automated prediction of objects has been implemented to enhance the visually impaired people communication. In order to solve the common aiming problem for blind users, a motion-based method to detect the object of interest is projected, while the blind user simply shakes the object for a couple of seconds. This method can effectively distinguish the object of interest from background or other objects in the camera view. Localized text regions are transformed into audio output for blind users. The audio output component is to inform the blind user of recognized text codes. In order to solve the common problem for blind users, we proposed a motion based technique to detect the object of interest, while the blind user simply shakes the object for few seconds. This method can effectively distinguish the object of interest from background or other objects in the camera vision. To extract text regions from complex backgrounds, we have proposed a "Haar Cascade Classifier Algorithm" text localization algorithm the corresponding feature maps estimate the global structural feature of text at every pixel

Key words: Blindness, assistive devices, text reading, hand-held objects, text region localization, stroke orientation, distribution of edge pixels, OCR.

I. INTRODUCTION

If you're a person with normal vision and you have to probably spend most of the days staring at screens — smartphones, tablets, and computers are our window to the world, and we interact with them primarily using our eyes. So at first glance, it seems like people who can't see would be totally cut off from these technologies. The truth is that you don't want to use devices like smartphones. Thanks to some simple accessibility tweaks and software, blind people can navigate websites and use apps just like the rest of us.

A product needs to be relevant because the users must have an immediate use for it. A product needs to be functionally able to do what it is supposed to, and do it with a good quality. Visually impaired people are made to use several different products in their day-to-day life. Our process of identification of products with ease helps them to recognize products in a swifter manner. This project is developed to make blind people life easy. This is very beneficial in case of finding out the description of product to the blind people and thus helping them in deciding to purchase a product or not especially which are packaged. In order to use this system, all the user needs to do is capture the image on the phone in the mobile phone which then resolves the voice of blind people. The image to voice conversion are used in this application really benefits blind and visually impaired people and thus making their work of identifying products easy.

A person who is visually blind has a decreased ability to see, even with corrective lenses, that adversely affects his visual access or interferes with processing visual information. The visual challenges an individual may have range from not being able to see a newspaper print to not being able to read at all. Other challenges may include not being able to find a friend or person in a room until he or she is standing within arm's reach or until she identifies herself. A simple and easy product also can't be able to identify by them without the ease of touch. The person needs assistant or another person's help to identify a product or a person. In the ideal world, all team members have equal ability, equal interest in the problem, and worked equally hard. In the real world that may not happen because of errors and major bugs in the code.

The offline version of Google now described in the research paper is feature-complete. Aside from a slimmed-down dictionary and the loss of some accuracy, it is capable of supporting all the features of the online algorithms. Identifying a product which has voice customization and error detection, all running on the limitations of the phone's processor. It may be a little way off yet but offline voice recognition looks likely to appear in Android at some point in the future, making Google Now a more versatile assistant that is helpful outdoors as well as in the office.

II. RELATED WORKS

Chucai Yi, YingLi Tian ,Ariess Arditi 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. A motion based method to define a region of interest (ROI) in the video by asking the user to shake the object .This method extracts moving object region by a mixture of Gaussians based background subtraction method. In the extracted ROI, text localization and recognition are conducted to acquire text information.

Ananth Noorithaya, Kishore Kumar M , Dr. Sreedevi stated that the visuallyimpaired are at a considerable disadvantage because they often lack the information for avoiding obstacles and hazards in their path. They have very little information on self- velocity, objects, direction which is essential for travel. Previously developed navigation systems use costly equipment which is often not affordable by the common blind community. The navigation systems available are heavy and very complicated to operate. This research has been aimed at design and development of a smart and intelligent cane which helps in navigation for the visually impaired people. The navigator system designed will detect an object or obstacle using ultrasonic sensors and gives audio instructions for guidance.

T.Rubesh Kumar, C.Purnima 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. We first propose an efficient and effective motion based method to define a region of interest (ROI) in the video by asking the user to shake the object. To automatically localize the text regions from the object ROI, we propose a novel text localization algorithm by learning gradient features of stroke orientations and distributions of edge pixels in an Ada boost model.

The proposed system by Rajalaskhmi P, Deepanraj S, Arun Prasath M. and Dinesh Kumar S which is a portable camera based visual assistance prototype for bind people to identify currency notes and also helps them to read printable texts from the handheld objects. To read printable texts, an efficient algorithm that combines an Optical Character Recognition (OCR) with Hierarchical optimization is used. In Pattern Recognition OCR every character is localized and separated then the resulting character image is sent to a pre-processor to reduce noise and to perform normalization.

A camera based assistive text reading framework by Vasanthi.G and Ramesh Babu.Y is to help blind persons read text labels and product packaging from handheld object in their daily resides is proposed. To isolate the object from cluttered backgrounds or other surroundings objects in the camera view, we propose an efficient and effective motion based method to define a region of interest (ROI) in the video by asking the user to shake the object. In the extracted ROI, text localization and recognition are conducted to acquire text information. To automatically localize the text regions from the object ROI, we propose a novel text localization algorithm by learning gradient features of stroke orientations and distributions of edge pixels in an Adaboost model.

Interaction with standard databases are possible only if we know about the standard SQL queries. This paper done by S.Nareshkumar, N.Mariappan, K.Thirumoorthy mainly focuses on interacting with the DBMS with speech. Automatic speech recognition is becoming famous now a days and it is widely used in many applications. Here users can interact with the database with their voice for retrieving details from it. Hence it is not necessary that user must have a prior knowledge about the SQL queries, they could retrieve details with their knowledge. The main purpose of this paper is that novice users who have no knowledge about the SQL queries can use it for retrieving the details from the database.

Priyanka Patil, Sonali Solat, Shital Hake Prof.S.T.KHOT says that nowadays printed text appears everywhere like product names, restaurant menus, instructions on bottles, signed boards etc. Thus blind people need some assistance to read this text .This paper presents a camera-based product information reader to help blind persons to read information of the products. Camera acts as main vision in detecting the label image of the product then image is processed internally and separates label from image by using MATLAB and finally identifies the product name and identified product information is pronounced through the optical character recognition (OCR).

III. OBJECT RECOGNITION FRAMEWORK FOR VISUALLY IMPAIRED PEOPLE

Speech is one among the way of exchanging information among people. Automatic speech recognition ASR is becoming famous nowadays. Many speech recognition systems can recognize lots of words. ASR has a lot of applications I many aspects of our daily life. A number of portable reading assistants have been designed specifically for the visually impaired. In this system we use a camera based object identification process for identifying the products but has problems recognizing colored text or text on a coloured background. We overcome this by using object detection algorithm for identifying the products easily without moving the objects or placing the objects correctly.

3.1 BLOCK DIAGRAM

This is a system design for product identification to vocalize. In this dataset which is stored in the database through an admin based database system that will store the product and compare the result with the given datasets through output and it's fully offline which does not require the features of internet. Image processing. The result is given as an audio output. If image quality and representation are not up to the mark there might be more chances to lose the business or customers. In the Fig 1 the process is done by capturing the image with a device and the image is Thresholded (colour) and Dilationed with Erosion. Then it is detected through edge pixels.

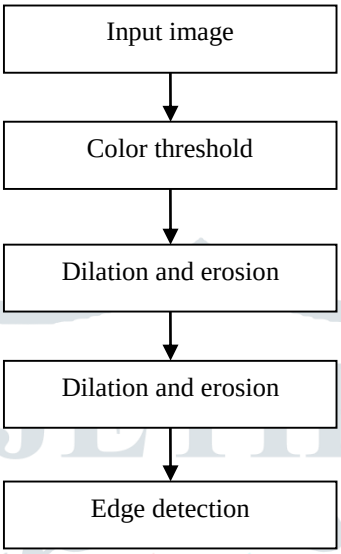

Fig 1 System design for image processing

3.2. COLOUR THRESHOLD BASED FEATURE EXTRACTION

Module used to remove parts of the image that fall within a specified colour range. From grayscale image, thresholding can be used to create binary images. Thresholding can be done in six methods.

- Histogram shape based
- Cluster based
- Entropy based
- Object attribute based
- Spatial based
- Local method based

In this system we are going to use the Local method based Threshold. In local method it adapts the threshold value on each pixel to the local image characteristics.

Dilation: Adding pixels to the boundaries of object in an image.

Erosion: Removing pixels on object boundaries.

Edge Detection: It is used for image segmentation and data extraction in areas such as image processing, computer vision and machine vision. It is used for finding the boundaries of objects within images.

3.3 TEXT ANLAYSIS PROCESS

Once the captured image is being processed with the above process the image is next mapped with dataset in the database. If the image in the dataset is matched then the description given is converted to text and the text with its description is converted into a voice output by the following process.

When the image is being processed then the following steps has to be done for generation of waveform or voice which is the reliable output for this system. Fig 2 explains how a text is converted into a waveform and generates an output for the given pre-processed text.

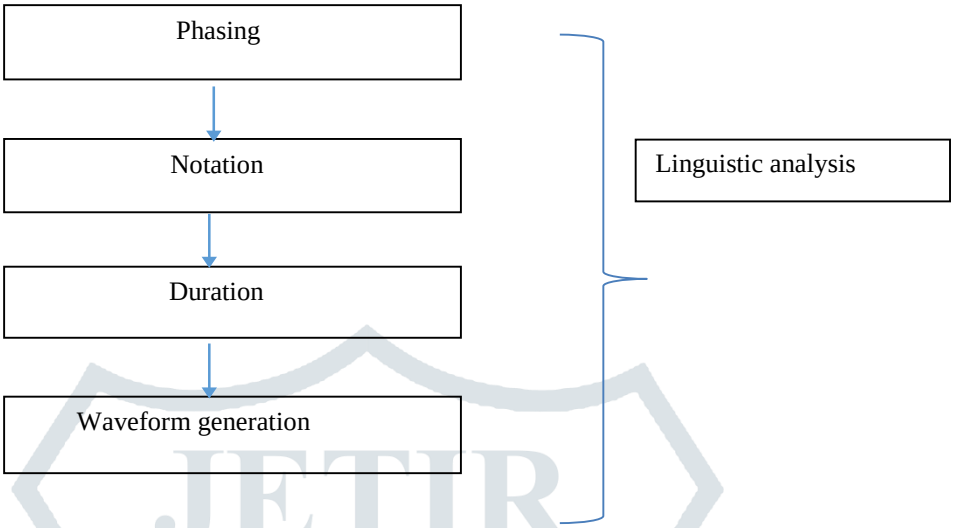

Fig 2. Text-to-speech conversion

It is a text-to-speech (TTS) system which converts normal language text to speech where other system renders symbolic linguistic representation like phonetic transcriptions into speech. Synthesized speech can be created by concatenating pieces of recorded speech that are stored in the database. Once the linguistic analysis is done the text will be converted into speech or voice. With a click of a button or the touch of a finger, TTS can take words on a computer or other digital device and convert them into audio.

The advantages of the proposed system are listed as follows:

- ODA detection and text localization algorithms are independently evaluated.
- Employed with a camera at a reasonable wide angle.
- User captured dataset generates low quality edge maps and text boundaries which result in better identification of images.
- User can adjust speech rate volume and tone according to their preferences.

3.4. PROPOSED SYSTEM OVERVIEW

This system runs on Java environment Eclipse and works on both Eclipse and Android Studio. The Main Activity of the program runs and executes the Text to Speech Recognition Process while Home Activity process the Image Recognition. Once Both the codes are given in this activity , Only the Home Activity run as an Android Application while Main Activity runs in the Background. After the whole process is done the Main screen or the UI layer of the Application will have two users.

1) User
2) Admin

Admin captures the product and will add the Images to the Data set .Then the user captures the product and the system identifies the Product Name with its Specification and the output is given to user as an Audio output.

3.5. IMPLEMENTATION STEPS

STEP 1: Install eclipse IDE and the workspace is given as parent folder, where the code is given.
STEP 2: The layout of the design is first created which shows the working model. This code executes image mapping process by mapping the edge pixels of the image captured.
STEP 3: The image which is mapped with the images same as that of in the dataset is then converted to text format.
STEP 4: The converted text matches the data in the Meta data which is retrieved and matched.
STEP 5: The text which is matched in the Meta data will read the description of the product.
STEP 6: The description product is given as an audio output to the user.

IV. RESULTS AND DISCUSSION

Any person can access data or add a data set by using the Admin login. On clicking the Admin login, it will ask for a Product capture along with its specification which will be sent as an audio output while executing the Application.

4.1 LOGIN PAGE

The login page shows the information of the user side and admin side where the datasets are being stored from the admin person. The required image is captured in different positions and therefore stored with added specifications.

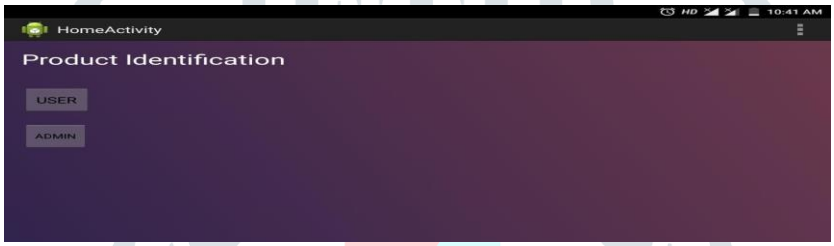

Fig 3. Login Screen

4.2 ADMIN DATASET ENTRY

The admin enters the dataset by capturing the product which is previewed in the same screen and then the following specification of the product is added.

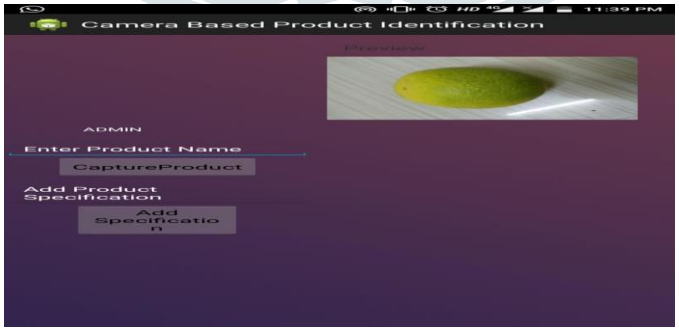

Fig 4 Dataset Entry

4.3. BLIND USER

Blind user can be able to capture the required image from the person's hand and then process with the given requirements.

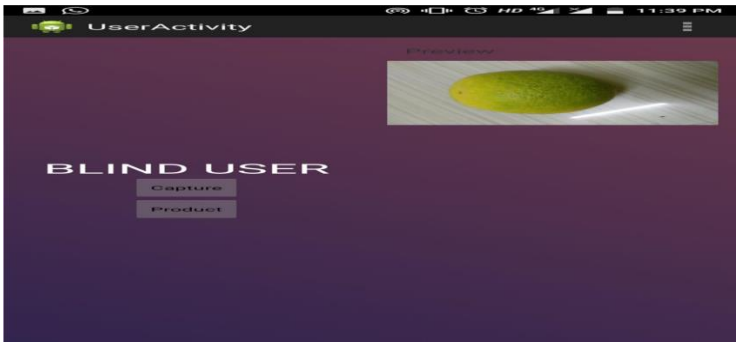

Fig 5. Blind user interface

4.4. PRODUCT IDENTIFICATION

The product is made to compare within the given android code so that the captured image is being processed and then the information about the image is given as voice output. The identified image is compared with the dataset entry to give the required output. Comparison is made using Object Detection algorithm and the result is given in percentage. The accuracy can be calculated with the given percentage but displays if the image is not captured properly (NaN).

V. CONCLUSION

This system that we designed here helps the life of Visually Impaired Person and this system is mostly offline and there is no need of Internet Connection. In order to solve the common aiming problem for blind users, a motion-based method to detect the object of interest is projected, while the blind user simply shakes the object for a couple of seconds. This method can effectively distinguish the object of interest from background or other objects in the camera view. Localized text regions are transformed into audio output for blind users. The audio output component is to inform the blind user of recognized text codes. In order to solve the common problem for blind users, we proposed a motion based technique to detect the object of interest, while the blind user simply shakes the object for few seconds. This method can effectively distinguish the object of interest from background or other objects in the camera vision. To extract text regions from complex backgrounds, we have proposed a "Haar Cascade Classifier Algorithm" text localization algorithm the corresponding feature maps estimate the global structural feature of text at every pixel. Our future work will extend our localization algorithm to process images with complex backgrounds and to design more robust block patterns for image feature extraction. We will also extend our algorithm to handle non horizontal photo or images. Furthermore, we will address the significant human interface issues associated with identifying the products by blind users.

REFERENCES

1. Chucai Yi, YingLi Tian ,Ariess Arditi," Portable Camera-based Assistive Text and Product Label Reading from Hand-held Objects for Blind Person" IEEE 2013

2. Ananth Noorithaya, Kishore Kumar M , Dr. Sreedevi. A "Voice Assisted Navigation System for the Blind" 2014

3. T.Rubesh Kumar, C.Purnima "Assistive System for Product Label Detection with Voice Output For Blind Users" 2014.

4. Rajalaskhmi P, Deepanraj S, Arun Prasath M. and Dinesh Kumar S "Portable camera based Visual Assistance for Blind People" 2015.

5. Vasanthi.G and Ramesh Babu.Y "Vision Based Assistive System for Label Detection with Voice Output" 2014.

6. S.Nareshkumar, N.Mariappan, K.Thirumoorthy "Database Interaction Using Automatic Speech Recognition" 2014.

7. Priyanka Patil, Sonali Solat, Shital Hake Prof.S.T.KHOT "A Voice Based Product Identification for Blind Persons" 2015.

8. Akshay Bal, "RFID Based Identification System" 2009

9. Noyal James, Dennise Koshy Sam, Roncy Annie Cherian, "Smart Shopping Facilitator for Blind People"2013
10. Rajkumar.N, Anand M.G, Bharathi Raja.N , "Portable Camera Based Product Label Reading for Blind People" 2013

11. V.Ramya, Lakshmiraja, B.Palaniappan, "Voice Assisted Embedded Navigation System for the Visually Impaired" 2014

12. Rui Jiang, Qian Lin, Shuhui Qu, "Real Time Visual Recognition with Results Converted to 3D Audio" 2010

13. Sarvesh Athawale, Mohd.Javed Ali, Tejal Birajdar, Deepak Patil, Prof.Saurabh Saoji "Object Deduction in a Smartphone for Visually Impaired"