Seminar Report

on

FINGER READER

Submitted in partial fulfillment of the requirements for the award of degree of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING

by

SAMA NISCHALA

(18WH5A0506)

Department of Computer Science and Engineering

BVRIT HYDERABAD College of Engineering for Women

(NBA Accredited EEE, ECE, CSE, IT B.Tech Courses,

Accredited by NAAC with 'A' Grade)

(Approved by AICTE, New Delhi and Affiliated to JNTUH, Hyderabad)

Rajeev Gandhi Nagar, Nizampet Road, Bachupally, HYDERABAD – 500 090

Feb, 2021

VISHNU
UNIVERSAL LEARNING

BVRIT HYDERABAD College of Engineering for Women

(NAAC and NBA Accredited EEE, ECE, CSE, IT B.Tech Courses)

(Approved by AICTE, New Delhi and Affiliated to JNTUH, Hyderabad)

Rajeev Gandhi Nagar, Nizampet Road, Bachupally, HYDERABAD – 500 090

Department of Computer Science and Engineering

CERTIFICATE

This is to certify that the Technical Report titled "**FINGER READER"** is being submitted by SAMA NISCHALA**(18WH5A0506)** of **IV B.Tech Computer Science and Engineering, (Academic year 2020-2021)** during the **I semester** in partial fulfillment for the award of B.Tech degree in Computer Science and Engineering, to Jawaharlal Technological University, Hyderabad.

The results embodied in the Report have not been submitted to any other University or Institute for the award of any degree or diploma.

Dr. Nara Sreekanth

Associate Professor

Seminar Guide

Department of Computer Science and Engineering

Dr. Srinivasa Reddy Konda

Head of the Department

Department of Computer Science and engineering

DECLARATION

I hereby declare that the work presented in this report entitled **"FINGER READER"** submitted towards completion of Seminar in IV Year of B. Tech of CSE at **BVRIT HYDERABAD College of Engineering for Women,** Hyderabad is an authentic record of our original work carried out under the guidance of**, Department of CSE.**

Ms. S. NISCHALA

(18WH5A0506)

ACKNOLWLEDGMENT

Firstly I would like to express my immerse gratitude towards **BVRIT HYDERABAD College of Engineering for Women**, which created a great platform to attain profound technical skills in the field of Computer Science, thereby fulfilling my most cherished goal.

I would like to extend my sincere thanks and gratitude to **Dr. K. V. N. Sunitha**, Principal, BVRIT HYDERABAD College of Engineering for Women for her constant encouragement in carrying my Technical Seminar successfully.

My sincere thanks and gratitude to **Dr. Srinivasa Reddy Konda**, Head of the Department, Computer Science and Engineering, for his inspiring guidance in carrying out the technical Seminar presentation,

I would like to record my sincere thankfulness to the technical seminar coordinator **Dr. Nara Sreekanth,** Associate Professor, Computer Science and Engineering for his valuable guidance and skillful management.

Ms. S.NISCHALA

(18WH5A0506)

ABSTRACT

Accessing printed text in a mobile context is a major challenge for the blind. In the present world visually impaired people use Braille lipi language to read the books, but those books are available only in particular places like library etc. It has become very difficult for the visually impaired peoples to come out into the real world.

Braille cannot be used in every place like in the mobiles, supermarket, and newspapers. Due to this it has become difficult for the visually impaired people to read the text. To overcome this difficulties faced by the visually impaired, finger reader with speech assistance is used to read any printed text from the mobiles, newspapers etc. we are using the assistive technology in this device to provide the assistance for the visually impaired people.

The technology currently used in the market is having problem like focusing, accuracy, mobility and efficiency. Hence here we want to propose a device that will solve all the problems. A camera will be mounted on the device which will fit on the finger of the reader.

FingerReader that assists blind users with reading printed text and also we introduce a novel computer vision algorithm for local sequential text scanning that enables reading single lines, blocks of text or skimming the text with complementary, multimodal feedback. This system is implemented in a small finger worm form factor that enables a more manageable eyes-free operation with trivial setup.

LIST OF FIGURES

INDEX

INTRODUCTION

According to the estimates from World Health Organization (WHO) about 285 million people are visually impaired worldwide: 39 million are blind and 246 million have low vision (severe or moderate visual impairment).

Braille is the standard tactile reading tool used by the blind for printed text, and the system is credited for boosting literacy levels among the visually impaired. However, many books and the materials are still not available in Braille.

Dr. Pattie Maes, founder and director of the MIT Media Lab's Fluid Interfaces research group, who developed FingerReader.

Accessing text has become very difficult for the visually impaired (VI) people in many scenarios, such as reading text on the go and accessing text in less than ideal conditions(i.e. low lighting, columned text, unique page orientations, etc.). As per the survey with the VI users revealed that available technology such as screen reader, desktop scanner, smart phone application and eBook reader are commonly underutilized due to slow processing speeds or poor accuracy.

To overcome these problems faced by the visually impaired people we can make use of the wearable device access the text easily and accurately. We can use this device just by wearing this to our pointer finger and pointing to the text which we want to read. By making use of this device visually impaired people can easily read the newspaper, the text messages from the mobile without anybody help.

HISTORY

DR. Pattie Maes, founder and director of the MIT Media Lab’s Fluid Interfaces research group, who developed FingerReader. In March 2015, the Massachusetts Institute of Technology (MIT) announced that researchers in its Media Lab had developed a prototype of a reading device that is worn on the finger.

Many people in the accessibility community were very excited by this prospect. Unlike other common OCR (optical character recognition) apps that first scan and then process the page, the MIT device, dubbed the Finger Reader, reads text in real time.

The concept for the device was developed by Roy Shilkrot, an MIT graduate student in Media Arts and Sciences. He and Media Lab postdoc Jochen Huber are lead authors on a paper describing the FingerReader.

Additional co-authors were Pattie Maes, the Alexander W. Dreyfoos Professor in Media Arts and Sciences at MIT; Suranga Nanayakkara, an assistant professor of engineering product development at the Singapore University of Technology and Design, who was a postdoc and later a visiting professor in Maes' lab; and Meng Ee Wong of Nanyang Technological University in Singapore.

Dr. Huber presented the paper in April at the Association for Computing Machinery Computer-Human Interface conference. Shilkrot graciously agreed to be interviewed for this article. He explained, "I came up with the idea about two years ago.

I thought it would be very interesting to think about reading because I know accessing print material is not a solved problem for people with a visual impairment. Because of the tactile sensitivity of the finger and the directionality of the finger when you point at something, it just made sense to think about reading and using the finger."

ABOUT FINGER READER

FINGER READER: A WEARABLE READING DEVICE

FingerReader is an index-finger wearable device that supports the blind in reading printed text by scanning with the finger and hearing the words as synthesized speech(Fig 1.c). Our work features hardware and software that include video processing algorithms and multiple output modalities, including tactile and auditory channels.

The design of the FingerReader is a continuation of our work on finger wearable devices for seamless interaction, and inspired by the focus session. Exploring the design concepts with blind users revealed the need to have a small, portable device that supports free movements, requires minimal setup and utilizes real-time, distinctive multimodal response.

The finger-worn design keeps the camera in a fixed distance from the text and utilizes the inherent finger's sense of touch when scanning text on the surface. Additionally, the device provides a simple interface for users as it has no buttons, and affords to easily identify the side with camera lens for proper orientation.

Figure1: Finger Reader Device

Figure 1.1: Finger Reader Prototypes.

PURPOSE

The main purpose of this paper is to make visually impaired people to read the text content without anybody help. This device can be used at any time, in any place and it can read the content from any device or from the books.

TECHNOLOGIES USED

1.) HARDWARE DETAILS:

❖ The finger Reader hardware was designed using:

➢ Multimodal feedback via vibration motors.

➢ A new dual-material case design.

➢ A high-resolution mini video camera.

❖ Here the Vibration motors embedded on the top and bottom of the ring can be used to provide feedback and dual-material case design is used to improve flexibility.

2.) SOFTWARE DETAILS:

- To accompany the hardware, a software stack that includes:
  - A text extraction algorithm.
  - Hardware control driver.
  - Integration layer with Tesseract OCR.
  - File Text-to-Speech (TTS).
- We start with image binarization and selective contour extraction. Thereafter we look for text lines by fitting lines to triplets of pruned contours; we then prune for lines with feasible slopes. We look for supporting contours to the candidate lines based on distance from the line and then eliminate duplicates using a 2D histogram of slope and intercept. Words with high confidence are retained and tracked as the user scans the line. For tracking we use template matching, utilizing image patches of the words, which we accumulate with each frame. We record the motion of the user to predict where the word patches might appear next in order to use a smaller search region. When the user veers from the scan line, we trigger a tactile and auditory feedback. When the system cannot find more word blocks along the line we trigger an event to let users know they reached the end of the printed line. New high-confidence words incur an event and invoke the TTS engine to utter the word aloud. When skimming, users hear one or two words that are currently under their finger and can decide whether to keep reading or move to another area.

COMPONENTS

- TTL Serial JPEG Camera.
- 16 MHz AVR PROCESSOR.
- Trigger button
- 3.7V polymer Lithium –ion battery.
- 3.3V regulator.
- A push button switch(on/off).

IMPLEMENTATION:

- Here all these components are packaged into a wearable ABS nylon casing.

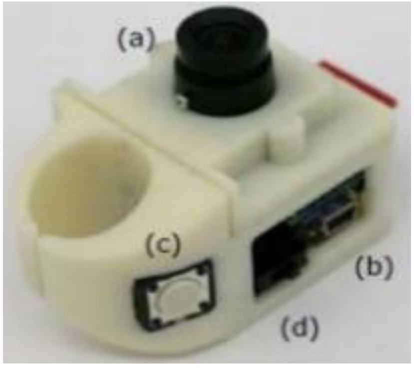

Figure 2: components are packaged into a wearable ABS nylon casing.

❖ The Figure was 3D printed with ABS nylon, and the electronics fit completely inside the skeletal compartment. (a)- VGA mini camera,

(b)- Mini-USB port for recharging and reprogramming

(c)- Trigger button

(d)- On/Off switch

WORKING:

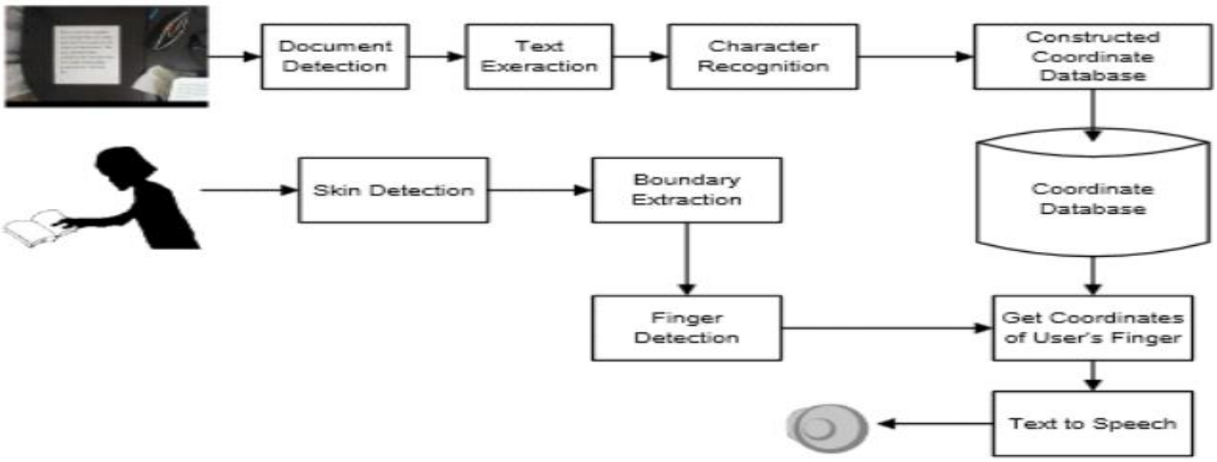

Figure 3: Block diagram for proposed Finger Reading System.

A finger-worn device with an embedded camera, a computation element and an earpiece for information loopback. Autonomous and wireless, and includes a single button to initiate the interaction. Speech processing engine and computer vision engine were implemented on a mobile phone. User would single click the pushbutton switch on the side if the ring using his thumb. A snapshot is taken from the camera and the image is transferred via Bluetooth to phone. Mobile phone then analyzes the images using our computer vision engine. Then the android application uses a TTS module to readout the information through a headset.

A **novel-tracking based algorithm** extracts text locally and sequentially, rather than in whole text blocks and pages like many existing devices use. If the user veers away from a text line, tactile feedback is provided by two vibration motors embedded in the 3D printed case. Auditory cues alert the user at the beginning and the end of reading passages.

Figure 4: The Finger Reader warns the readers when they move away from the line.

ALGORITHM AND SOFTWARE:

We develop a software stack that includes a sequential text reading algorithm, hardware control driver, integration layer with Tesseract OCR and File TTS, currently in a standalone PC applications.

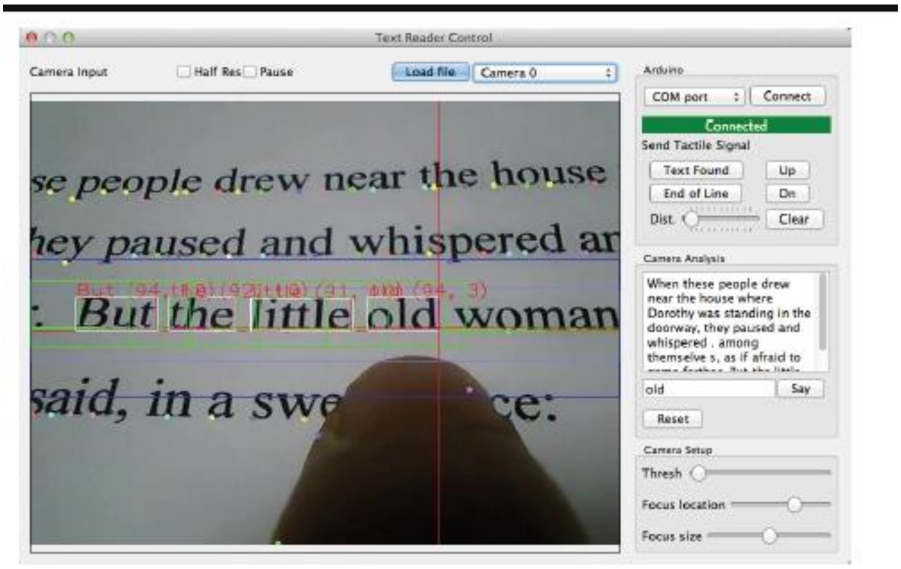

Figure 5: Recognizing the text through image processing

VISION ALGORITHM OVERVIEW:

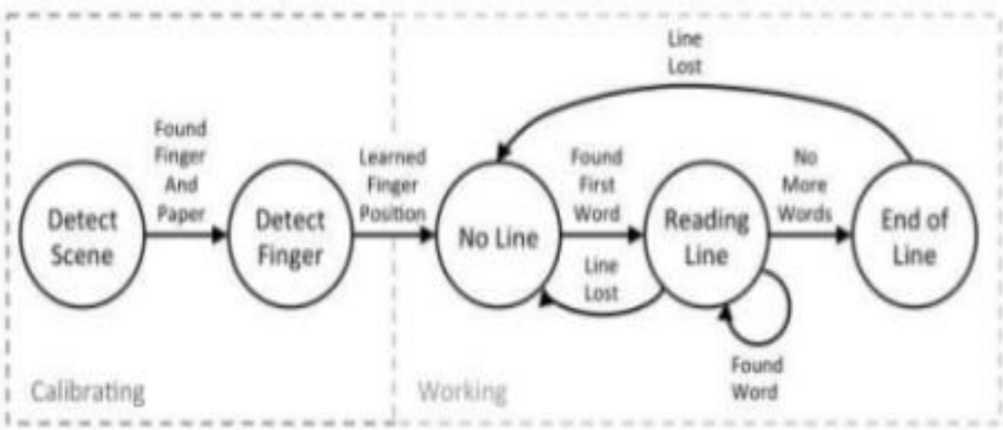

Figure 6: Sequential text reading algorithm state machine.

The sequential text reading algorithm is comprised of a number of sub-algorithms concatenate in a state-machine, to accommodate for a continuous operation by a blind person. The first two states(Detect scene and learn finger) are used for calibration for the higher level text extraction and tracking work states(No line, Line Found and End of line).

Each state delivers timely audio cues to the users to inform them of the process. All states and their underlying algorithms are detailed in the following sections. The operation begins with detecting if the camera indeed is looking at a close-up view of a finger touching a contrasting paper, which is what the system expect in a atypical operations.

Once achieving a stable view, the system looks to locate the fingertip as a cursor for finding charcaters, words and lines. The next three states deal with finding and maintain the working line and reading words.

For Finding a line, the first line or otherwise, a user may scan the page(in No Line mode) until receiving an audio cue that text has been found. While a text line is maintained, the system will stay in the Line Found state, until the user advanced to the end of the line or the line is lost(by moving too far up or down from the line or away from the paper.).

SYSTEM OVERVIEW

LINE EXTRACTION:

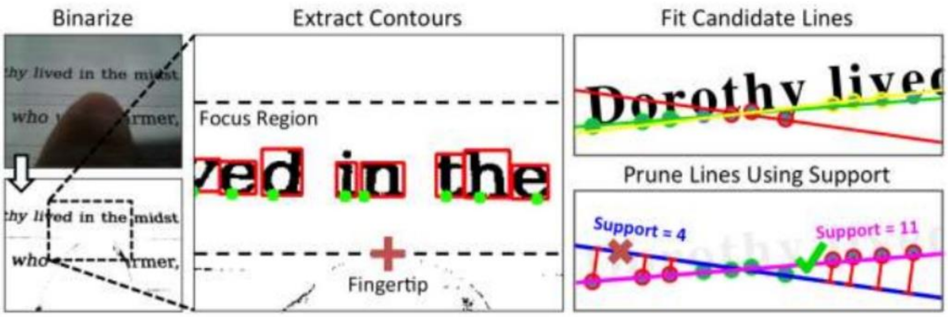

Figure 7: Text line Extraction Process.

Within the focus region, we start with local adaptive image binarization (using a shifting window and themean intensity value) and selective contour extraction based on contour area, with thresholds for typical character size to remove outliers. We pick the bottom point of each contour as the baseline point, allowing some letters, such as y", "g" or j" whose bottom point is below the baseline, to create artifacts that will later pruned out.

Thereafter we look for candidate lines by fitting line equations to triplets of baseline points; we then keep lines with feasible slopes and discard those that do not make sense. We further prune by looking for supporting baseline points to the candidate lines based on distance from the line. Then we eliminate duplicate candidates using a 2D histogram of slope and intercept that converges similar lines together.

Lastly, we recount the corroborating baseline points, refine the line equations based on their supporting points and pick the highest scoring line as the detected text line. When ranking the resulting lines, additionally, we consider their distance from the centre of the focus region to help cope with small line spacing, when more than one line is in the focus region.

WORD EXTRACTION:

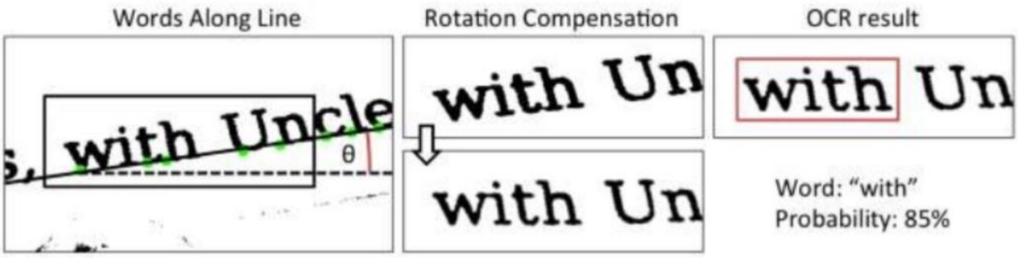

Figure 8 : Word Extraction Process

Word extraction is performed by the Tesseract OCR engine on image blocks from the detected text line. Since we focus on small and centric image blocks, the effects of homography between the image and the paper planes, and lens distortion (which is prominent in the outskirts of the image) are negligent. However, we do compensate for the rotational component caused by users twisting their finger with respect to the line, which is modeled by the equation of the detected line. The OCR engine is instructed to only extract a single word, and it returns: the word, the bounding rectangle, and the detection confidence. Words with high confidence are retained, uttered out loud to the user, and further tracked using their bounding rectangle as described in the next section.

FILE TEXT-TO-SPEECH (TTS):

It is a screen reader application developed by Google for its Android operating system. It converts normal language **text** into **speech**. It is a type of assistive technology that reads digital **text** aloud. It's sometimes called "read aloud" technology. 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.

Figure 9: Text-to-Speech (TTS).

Optical Character Recognition (OCR):

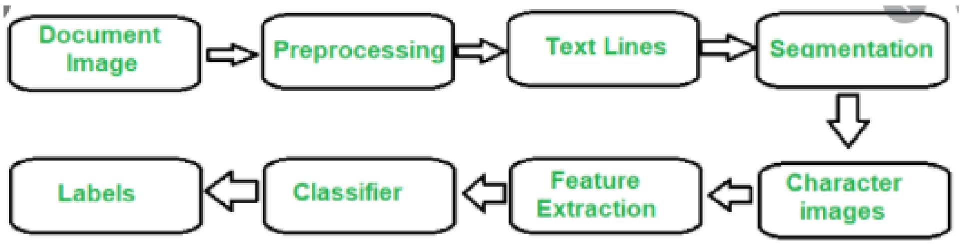

Figure 10: Flow diagram of OCR

The concept of optical character recognition is used in this device. OCR is mechanical or electronical conversion of typed, handwritten or printed text into machine-encoded text. It widely accepts data from any sort of document. It is a common method of digitizing printed texts so that it can be electronically edited, searched, stored more compactly, displayed on-line, and used in machine processes such as text to speech, machine translation, key data and text mining. OCR is a field of research in artificial intelligence, pattern recognition and in computer vision.

WORD TRACKING AND SIGNALING:

Whenever a new word is recognized it is added to a pool of words to track along with its initial bounding rectangle. For tracking we use template matching, utilizing image patches of the words and an L2 norm matching score. Every successful tracking, marked by a low matching score and a feasible tracking velocity (i.e. it corresponds with the predicted finger velocity for that frame), contributes to the bank of patches for that word as well as to the prediction of finger velocity for the next tracking cycle.

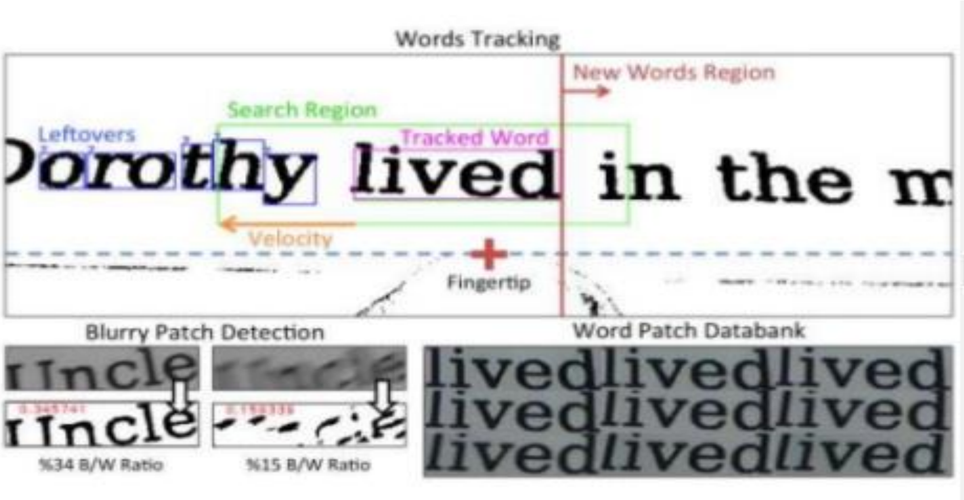

Figure 11: Word Tracking Process.

To maintain an efficient tracking, we do not search the entire frame but constrain the search region around the last position of the word while considering the predicted movement speed. We also look out for blurry patches, caused by rapid movement and the camera‟s rolling shutter, by binarizing the patch and counting the number of black vs. white pixels. A ratio of less than 25% black is considered a bad patch to be discarded. If a word was not tracked properly for a set number of frames we deem as “lost”, and remove it from the pool.

SCENE AND FINGER DETECTION:

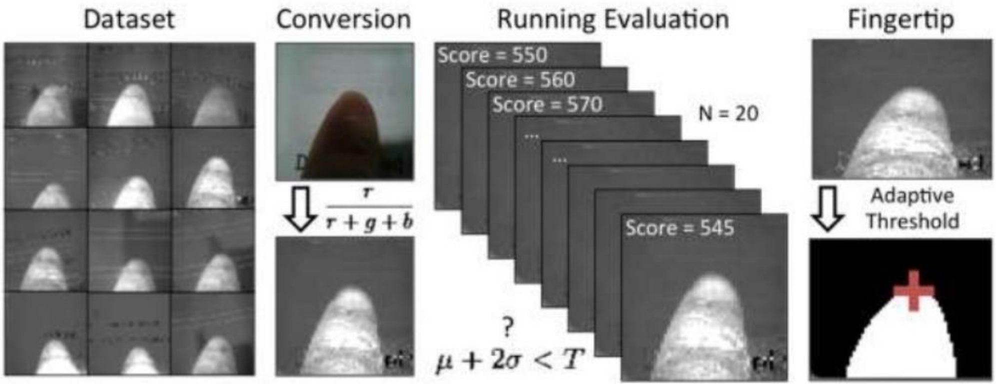

Figure 12: Scene and Finger Detection

The initial calibration step tries to ascertain whether the camera sees a finger on a constrating paper. The input camera image is converted to the normalized-RGB space:(R,G,(B)=rg),, b, however we keep only r+g+br+g+br+g+b thenormal-ized red channel(R) that corresponds well with skin colors nad ameliorates lighting effects. The monochromatic image is downscaled to 50x50 pixels and matched to a dataset of prerecorded typical images of fingers and proper perspective of the device camera. To score an incoming example image, we perform a nearest neighbor matching and use the distance to the closest database neighbor. Once a stable low score is achieved(by means of a running-window of 20 samples and testing if Uscore +20 < threshold) the system deems the scene to be a well-placed finger on a paper, issues an audio command and advanced the state machine.

In the finger detection state we binarize the R channel image using Otsu adaptive thresholding and line scan for the top white pixel, which is considered a candidate fingertip point. During this process the user is instructed not to move, and our system collects samples of the fingertip location from which we extract a normal distribution. In the next working states the fingertip is tracked in the same fashion from the R channel images, however, in this case, we assign each detection with a probability measure based on the learned distribution to eradicated outliers. The inlying fingertip detection guides a local horizontal focus region, located above the fingertip, within which the following states perform their operations. The focus region helps with efficiency in calculation and also reduces confusion for the line extraction algorithm with neighboring lines. The height of the focus region may be adjusted as a papameter, but the sysyem automatically determines it onces a text line is found.

ADVANTAGES:

- The Finger Reader wearable health device can help them gain access to a greater number of learning resources and contribute greatly to their quality of life.
- The index-finger worn device is more flexible.
- A lot more immediate than any solution that they have right now.
- The main advantage of this is to detect the text and gives the voice output for visually impaired.

- It helps visually impaired people (VI) to read the text in a sequence order if it misses any line the device vibrates and gives the voice output.

DISADVANTAGES:

- The camera doesn’t auto-focus, it takes time to focus on the text and also making it hard to adjust to different finger lengths.
- There is lot of effort that is required to make trainer data of different languages and implement that.
- It doesn’t work with text as small as, say, on a medicine bottle, but it can read 12-point printed text.
- Finger reader just reads the text and does not specify whether it is headline are sub head line.
- The voice is clipped but work is going on in order to improve the quality of sound.

SOLUTION TO THE EXISTING PROBLEMS:

In order to combat the existing problems, a novel hardware and software can be used that includes quick response, videoprocessing algorithms and different output mechanisms. The ring prototype adjusts the camera at a fixed distance and utilizes the sense of touch when scanning the surface. The device can be made to contain few buttons and a simple user interface thus making it compact and user-friendly.

FUTURE WORK

Mobile phones and laptops are very fragile and at times become very complex to use. There is a growing need of more and more user-friendly devices which are sturdy and innovative. We expect in the coming years we will see a lot more wearable devices such as glasses, bracelets and watches enabling us to glance at some relevant information without running to places.

Future Applications will included that adding more hardware such as a microphone, an infrared light source or a laser module, a second camera, a depth sensor or inertial sensors, will open up a multitude of new uses for this specific wearable design.

APPLICATIONS

- **Virtual Walking Cane:**

  Used for navigation is certainly fashionable and appealing.

- **Reading Non-Braille:**

  Visually impaired people are mostly bound to reading Braille or listening to audio books but are limited. This will allows the user to read regular printed material using this ring.

- **Tourist Helper :**

  Tourists visiting a new city often rely on maps and landmarks for navigation.

- **Shopping Assistant:** e.g. price tag recognition, currency recognition.
- Indoor navigation.
- Reading non- braille text.
- A tool to empower children in pre-reading stage (read text on their own).
- **Currency Detector :**

Figure 13: Currency Detector.

Intended to help the user to identify currency bills(1$,5$,10$,20$,50$) or(1,2,5,10,20,100,500,1000,2000) rupees. Simply point index finger to a currency note and click the button to deliver an image for processing. A detection algorithm based on a Bag of Visual Words (BoVW) scans the image and makes a decision. We use Opponent Space SURF features to retain color information for notes detection.

- **Color Detector:**

Figure 14: Colour Detector

Aids a visually impaired person to understand the color of an object. Users simply touch point to an object and click the button to delivery an image for processing. Analyses the image and returns the average color via audio feedback. We use a calibration step to help the system adjust to different lighting conditions.

CONCLUSION

We contributed Finger Reader, a novel concept for text reading for the blind, utilizing a local-sequential scan that enables continuous feedback and non-linear text skimming. Motivated by focus group sessions with blind participants, our method proposes a solution to a limitation of most existing technologies: reading blocks of text at a time. Our system includes a text tracking algorithm that extracts words from a close-up camera view, integrated with a finger-wearable device. A technical accuracy analysis showed that the local-sequential scan algorithm works reliably. Two qualitative studies with blind participants revealed important insights for the emerging field of finger-worn reading aids.

First, our observations suggest that a local-sequential approach is beneficial for document exploration–but not as much for longer reading sessions, due to troublesome navigation in complex layouts and fatigue. Access to small bits of text, as found on business cards, pamphlets and even newspaper articles, was considered viable. Second, we observed a rich set of interaction strategies that shed light onto potential real-world usage of finger-worn reading aids. A particularly important insight is the direct correlation between the finger movement and the output of the synthesized speech: navigating within the text is closely coupled to navigating in the produced audio stream.

Our findings suggest that a direct mapping could greatly improve interaction (e.g. easy “re-reading”), as well as scaffold the mental model of a text document effectively, avoiding “ghost text”. Last, although our focus sessions on the feedback modalities concluded with an agreement for cross-modality, the thorough observation in the follow-up study showed that user preferences were highly diverse. Thus, we hypothesize that a universal finger-worn reading device that works uniformly across all users may not exist and that personalized feedback mechanisms are key to address needs of different blind users.

REFERENCES:

- **https://www.slideshare.net/SarveshMeena/finger-reader-thesis-and-seminar-report**

- **Blindsight, Text detective, 2013 http://blindsight.com/textdetective/.**

- **https://www.seminarsonly.com/computer%20science/finger-reader-seminar-report-ppt.php**

- **https://www.afb.org/aw/16/7/15438**

- **G. Tejaswani, Afroz. B, Shruthi. Y, Sunitha .S, Agriculture Robotics , International Journal of Engineering and Computer Science: Vol. 7 No. 03 (2018).**

- **http://ijiit.logicinside.net/archive/volume%204/issue%202/V4I2P01.pdf**