

Multipurpose Guide for Visually Impaired People using Raspberry pi

Submitted in partial fulfillment of the requirements for the
award of Bachelor of Engineering degree in Electronics and
Communication Engineering

By

38130038 – Chandru K

38130077-Gunaseelan P V



DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

SCHOOL OF ELECTRICAL AND ELECTRONICS

SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY (DEEMED TO
BE UNIVERSITY) Accredited with Grade “A” by NAAC | 12B Status by UGC |
Approved by AICTE JEPPIAAR NAGAR, RAJIV GANDHI SALAI, CHENNAI - 600

SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY

(DEEMED TO BE UNIVERSITY) Accredited with “A” grade by NAAC I 12B Status by UGC I

Approved by AICTE Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai – 600 119

www.sathyabama.ac.in



DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

BONAFIDE CERTIFICATE

This is to certify that this Project Report is the Bonafide work of CHANDRU K (Reg.No: 38130038) and GUNASEELAN P V (Reg.No: 378130077) who carried out the project entitled Multipurpose Guide for Visually Impaired People Using Raspberry pi under my supervision from November 2021 to April 2022.

Internal Guide: DR. ARANGANATHAN, M.E., Ph.D.

Head of the Department: DR. RAVI, M.E., Ph.D.

Submitted for Viva voce Examination held on _____

Internal Examiner

External

Examiner

DECLARATION

We, **GUNASEELAN P V** (Reg. No.: 38130077) and **CHANDRU K**
(Reg. No.: 38130038) hereby declare that the Project Report entitled

Multipurpose Guide for Visually Impaired People Using

Raspberry pi done by us under the guidance of DR.

ARANGANATHAN, M.E., Ph.D. is submitted in partial fulfillment of the
requirements for the award of Bachelor of Engineering degree in
ELECTRONICS AND COMMUNICATION ENGINEERING.

CHANDRU K –38130038

GUNASEELAN P V –38130077

DATE

PLACE: CHENNAI

SIGNATURE OF CANDIDATES

ACKNOWLEDGEMENT

I am very pleased to acknowledge my sincere thanks to the **Board of Management of SATHYABAMA** for their kind encouragement in doing this project and for completing it successfully.

We are grateful to them. I convey our thanks to **Dr. N. M. NANDHITHA, M.E., Ph.D.**, Dean, School of Electrical and Electronics and **Dr. RAVI, M.E., Ph.D.**, Head of the Department, Department of Electronics and Communication Engineering for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my Project Guide DR. **ARANGANATHAN, M.E., Ph.D.** for her valuable guidance, suggestions, and constant encouragement paved way for the successful completion of my project work.

I wish to express our thanks to all Teaching and Non-teaching staff members of the Department of Electronics and Instrumentation Engineering who were helpful in many ways for the completion of the project.

ABSTRACT:

An OCR (Optical Character Recognition) system which may be a branch of computer vision and successively a sub-class of AI. Optical character recognition is the translation of optically scanned bitmaps of printed or hand transcription into audio output by using Raspberry Pi. OCRs are developed for many world languages and are already under efficient use. This method extracts moving object region by a mixture-of-Gaussian based-background subtraction method. A text localization and recognition are conducted to accumulate text information. To automatically localize the text regions from the image, a text localization and Tesseract algorithm by learning gradient features of stroke orientations and distributions of edge pixels in an Ada boost model. Text characters in the localized text regions are then converted into binary codes and recognized by optical character recognition software. The recognized text codes are given as output to blind users. Performance of the proposed text localization algorithm. As the recognition process is completed, the character codes within the document are processed using Raspberry pi device on which recognize character using Tesseract algorithm and python programming, the audio output is listened. This model is mounted on a shopping trolley to help visually challenged people to shop individually.

A Smart system concept is devised to provide a smart electronic aid for blind people. Blind and visually impaired people find difficulties to find the their navigation. The system is intended to provide artificial vision and object detection, real time assistance via GPS by making use of Raspberry Pi.

The system consists of a GPS module, and the feedback is receive through audio, voice output works through TTS (text to speech). The proposed system detects an object around them and sends feedback in the form of speech, warning messages via earphone and also provides navigation to specific location through GPS.

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE NO
	Abstract	5
1	INTRODUCTION	9
	1.1 Purpose	10
	1.2 Project Scope	11
	1.3 Existing System	12
	1.4 Benefit of Proposed System	13
	1.5 Architecture of the Proposed System	13
2	Literature Review	14
	2.1 A Navigation Assistance for the Blind	15
	2.2 Introduction	16
	2.3 Proposed System	17
	2.4 Results and Discussions	19
	2.5 Conclusion	19
3	OPTICAL CHARACTER RECOGNITION	20
	3.1 Introduction	20
	3.2 How Does OCR Work	21
	3.3 How does OCR Work	21
	3.4 Do I really need to proofread and Correct an OCR Output?	22
	3.5 What if My OCR Output is Really Bad?	22
	3.6 Blind and Visually Impaired Users	22
	3.7 Application	23
	3.8 Preprocessing	24
	3.9 Text Recognition	25

	3.10 Conclusion	26
	3.11 Proposed System	27
	3.11.1 Block Diagram	27
4	HARDWARE DESCRIPTION	29
	4.1 Raspberry Pi	29
	4.1.2 Overview	29
	4.2 Exploring the Raspberry Pi Board	29
	4.2.1 SD Card Slot	30
	4.2.2 Micro USB power	30
	4.2.3 HDMI Out	30
	4.2.4 Ethernet and USB Port	30
	4.2.5 RCA Video Out and Audio Out	30
	4.2.6 GPIO Headers(pins)	31
	4.2.7 Chips(Broadcom)	31
	4.3 Raspberry Pi 2 Model B	31
	4.4 Hardware Required for Raspberry Pi	31
	4.4.1 Storage	32
	4.4.2 Input	32
	4.4.3 Monitor	32
	4.4.4 Network	33
	4.5 Power Supply Unit	33

	4.5.1 Simple 5V power Supply for	
	Digital Circuits	35
	4.6 Webcam	35
	4.6.1 Lens	36
	4.6.2 Portability	36
	4.6.3 Features	37
	4.7 Image Processing on Raspberry Pi	37
	4.8 Basic Concept of Image Processing	39
	4.9 System Hardware Design	40
5	Software Description	41
	5.1 Start up with Raspberry Pi	41
	5.1.1 Installation of Operating System	
	On Raspberry Pi	41
	5.1.2 Use of Laptop Screen,Keyboard	
	and mouse for Raspberry Pi	41
	5.2 puTTY	42
	5.3 Installation of Required Applications	
	On Raspberry Pi	43
6	Results and Conclusion	51
	6.1 Result	51

CHAPTER 1

1. INTRODUCTION

Visually challenged people face many problems with accessing printed text, which includes problems with alignment, focus, accuracy, mobility, and efficiency. We present a smart device that assists the visually impaired which effectively and efficiently reads the paper-printed text. The proposed project uses the methodology of a camera based assistive device which will be employed by people to read Text document. The main idea of this project is to implement image capturing technique in an embedded system based on Raspberry Pi board. In this project, we proposed a text read out a system for the visually challenged. The proposed system uses a camera image as input and the image is processed using OCR software module and the output is given in audio format. A methodology is implemented to recognition sequence of characters, and the line of reading. As a part of the software development, the Open CV (Open-source Computer Vision) libraries are utilized to do image capture of text, to do the character recognition. Most of the access technology tools built for people with blindness and limited vision are built on the two basic building blocks of OCR software and Text-to-Speech (TTS) engines. Optical character recognition (OCR) is the translation of captured images of printed text into binary codes. OCR may be a process which associates a symbolic meaning with objects (letters, symbols a number) with the image of a character. It is defined because the process of converting scanned images of machine printed into a computer process able format. An optical Character recognition is additionally useful for visually impaired people that cannot read Text documents but have access to the content of the Text documents. An optical Character recognition is used to digitize

and reproduce texts that have been produced with the non-computerized system. Digitizing texts also helps reduce storage space. Editing and Reprinting of a Text document that was printed on paper are time-consuming and labor-intensive. It is widely wanted to convert books and documents into electronic files to be used in storage and document analysis. OCR makes it possible to use techniques like MT, text-to-speech, and text mining to capture/ scanned page. The final recognized text document is fed to the output devices, depending on the choice of the user. The output devices are often a headset connected to the raspberry pi or a speaker which may spell out the text document loud.

1.1 PURPOSE

- In this system we have described a prototype system to read printed or typed texts on hand-held objects for guiding blind people. To unravel the common aiming problem for blind users, we have proposed a motion-based method to detect the thing of interest. The automatic ROI detection and text localization algorithms were independently evaluated to make sure effectiveness and robustness of the entire system. We subsequently evaluated this prototype system of assistive text reading using images of hand-held objects captured by ten blind users face to face. Two calibrations were applied to organize for the system test. First, we instructed blind users to put hand-held objects within the camera view. Since it is difficult for blind users to aim their held objects, we employed a camera with a wide angle. Secondly, in an applicable blind-assistive system, a text localization algorithm might prefer higher recall by sacrificing some precision. By using ultrasonic sensor, we can measure the distance between the blind people and obstacle then the distance will be played through earphones. OCR is used for converting images of typed,

handwritten, or printed text into machine-encoded text, whether from a scanned document, a photo of a document. It is widely used to produce information entry from printed paper. It is a method of analyzing printed texts so that they can be easily edited, searched, stored, displayed online, and utilized in machine processes like cognitive computing, machine translation and text-to-text. This module can also be mounted on a shopping trolley which can help the visually challenged people shop individually at supermarkets. Four buttons are added along with this model. These four buttons are assigned with four basic food products and a DC motor, and a driver is fixed at the wheel of the shopping trolley. The buttons, DC motor and driver are connected to Raspberry Pi's General Purpose Input Output (GPIO) pins, and it is coded accordingly. If the buttons are pressed. It moves according to the program coded. The UV sensor calculates the distance between the object of interest and the camera. The camera captures the image of the product, and the output can be heard through the headphones.

1.2 PROJECT SCOPE

- The main objective of our project is getting the images and location of the blind persons when they are in the trouble
- to provide voice-based assistance to blind people. Here we have developed an intelligent system that helps blind people to travel independently and work efficiently. Current navigation device for the visually impair focus on travelling from one location to another.
- Our project focuses on designing a device for blind people that helps them to travel independently and it must be comfortable to use.

1.3 EXISTING SYSTEM

- Blind people use either the typical white cane or the guide dog to travel. The white cane is a widely used mobility aid that helps blind people to navigate in their surroundings.
- Although the white stick gives a warning about a few meters before the obstacle, for a normal walking speed, the time to react is very short. The idea of designing and manufacturing ultrasonic sensor combines the properties of sound monition and that benefit the blind and vibrating alert feature, which benefit from the experience of deafness.
- Sensor can detect obstacles within the designed range to avoid the blind person through the issuance of distinctive sound or vibration can be issued by the sense of the deaf by putting his finger on the button at the top of the device vibrate when there is a risk.
- This system involves more manual work, and it does not provide better results. The existing system does not provide proper navigation and is not highly effective.

1.4 BENEFIT OF PROPOSED SYSTEM

- It is Effective.
- It is Low cost.
- We can monitor the blind people with location by receiving the messages.
- Automatic character recognition.
- A prototype system to read printed text on hand-held objects for assisting blind persons.
- We have proposed a motion-based method to detect the object of interest, while the blind user simply shakes the object for a couple of seconds.

1.5 ARCHITECTURE OF THE PROPOSED SYSTEM

The Architecture of the optical character recognition system on a grid infrastructure consists of the three main components. They are: -

- Camera
- OCR Hardware or Software
- Output Interface

2. LITERATURE REVIEW

1. A Survey of Voice Aided Electronic Stick for Visually Impaired People. Author Name- Young Ho, Sung Jae Kang. Description- From this paper we got idea about latest technology like Graphics Positioning System (GPS) & Graphics System Messaging (GSM). Which will help for tracking the location & used for making module of smart stick for visually impaired people and it gives us idea about Voice message get from Android Phone to that blind person. [14].
2. Blind Navigation System Using Image Processing and Embedded System Author Name- Sarinah Jamaludin, Zul Azizi Haliana Description- We got idea from this paper for Navigation system has been developed which help to enhance mobility of blindness. This paper gives us idea for capturing live video of that person & grab video feed in front of blind person & this live video can be observed by admin itself. [4]
3. Smart Cane: Assistive Cane for Visually impaired People. Author Name-Amirul Talib, Mohd Helmy Wahab Description- We got idea from this paper for Voice message & Vibration when person detect obstacle with help of smart stick then blind person get aware to it by understanding Vibration alert & Voice message which comes from smart phone. [3]
4. Electronic Path Guidance for Visually Impaired People Author Name-Iwan Ulrich and Johann Borenstein Description- From this paper we got idea about the how range which is required for detecting obstacle or object from location of smart stick specific range is there we need to set threshold value if obstacle comes into that threshold range, then it can detect successfully otherwise cannot. [4]

5. Use of Ultrasonic Sensors in the Development of an Electronic Travel Aid Author Name- Alex Harold and Chris Gearhart Description- From this paper we got idea about for capturing video image processing is required by making use of some algorithm & method we make some processes on image for capturing it and live monitoring of that person is seen at admin side. All the processing data is stored on server side in serialized format.[5]

6. Automated Mobility & Orientation System for Blind or Partially Sighted people Author Name- Abdel Ilah Nour Alshbatat Description- From this paper we got idea about GSM, GPS of sensor like Integrated Ultrasonic sensor, Accelerometer.[14]

2.1 A Navigation Assistance for the Blind

In today's fast-moving world taking each step without obstacles is a challenge for the visually impaired. A smart assistant like 3S (Sensory Supervision Spectacles) can assist the visually impaired to read images, avoid obstacles and track routes. So, this paper presents a thought of building up a shrewd framework which can help the visually impaired individuals in their day-to-day life. Some of the main challenges include trouble in moving without the help of someone, difficulty in understanding text or pictures, detecting hindrances, etc. The proposed 3S (Sensory Supervision Spectacles) is capable of text recognition using Optical Character Recognition (OCR), speech synthesis using Text-To-Speech (TTS), and obstacle detection using ultrasonic sensors (HC-SR04) and a GPS tracker. The segments are associated with Raspberry Pi for being processed. The gadget is a voice empowered framework that would coordinate the visually impaired individual in their everyday works.

2.2 Introduction

The estimated number of people visually impaired in the world is 285 million. This number is huge to the point that it drastically affects the economy of the nation. If we look at this busy world, the average person does not have the opportunity to try and spend some time assisting these distinctively capable ones. In this manner, the visually impaired individuals continually need the help of somebody in their daily life,

particularly while on the streets, and this dependency sometimes leaves them to think about themselves in a helpless or loneliness state. Some other challenges include difficulty in reading signs, symbols, texts, or images, detecting obstacles, tracking geographical position, etc. This paper proposes a compact model which can be used everywhere without much difficulty. It incorporates an automated virtual text framework. This is completely processed on a Raspberry. The whole framework is integrated with a camera module, OCR (Optical Character Recognition) mechanical conversion, a TTS (Text-To Speech) engine, ultrasonic sensors (HC-SR04) and a GPS (Global Positioning System) satellite-based radio navigation system. The camera module snaps a photo when a user presses the button which is linked with the Raspberry Pi, the saved image is then taken care of by the OCR for electronic conversion. OCR accomplishes character recognition via performing some picture handling techniques, (for example, eliminating negative and positive spots). The extracted character will then be saved in a document format (e.g. .txt). This document will then be processed by the TTS software which reads out the text/characters to the user via an output speaker. This model also integrates an object detection module which utilizes ultrasonic sensors which can alert the user with the exact distance information of the object and the speaker is utilized to produce the alarm sound. Totally four ultrasonic sensors are utilized for covering four directions and alarm sound is produced via output speaker giving four different signals (example: 'left,' 'right,' 'front,' 'back') along with the distance of the object (like automobile sensors). Furthermore, the device also gives a vibration signal which aids the user in detecting obstacles in a noisy environment. Hence, the proposed framework is designed to caution the user through both voice and vibration. This framework utilizes an Arduino microcontroller, which gathers all the data from sensor unit and appropriately sends signals to response unit to alert the user. The model is additionally integrated with a GPS system, which utilizes internet service for navigation assistance and enables live position tracking by the user's family members.

2.3 PROPOSED SYSTEM

Visually impaired individuals experience various challenges in converting into text utilizing existing technologies, which face issues with arrangement, centering,

precision, portability, and productivity. We present a smart gadget that helps the visually challenged to read printed or written text successfully and effectively. The proposed project utilizes the technique of a camera based assistive device that can be utilized by individuals to read printed or written text. The structure is for executing image scanning method in an integrated framework on Raspberry Pi board. In this project, we have proposed a text reading framework for the visually impaired. The proposed completely coordinated framework has a camera as an input device to take care of the printed or written text for digitization, then the document is scanned and processed by a software module called the OCR (Optical Character Recognition engine). The majority of the accessibility tools developed for individuals with visual deficiency and impaired vision are programmed using OCR programming and Text-to-Speech (TTS) engine. OCR makes it conceivable to apply methods, for example, machine interpretation, text-to-speech, and text mining to the captured image. And finally, the recognized text document is fed to the output device as per user's preference, either a headset or a speaker. This output device reads out the text document loud. Below Fig. 1 shows the basic blocks of the image to Speech Processing Technique

The first step of the process is to read the image. This is done with the help of an inbuilt camera.

- The next step is preprocessing. Here the image is converted into a gray scale image. Then it is converted into a black and white image with some threshold value.
- Next the image is cropped into its edges by finding its boundary.
- Once it is finished, the image is converted into word matrix. So that we can match the letters with templates when template is loaded.
- Each time the extracted letter is compared with the templates. The output is written on a text file.
- Next step is to concatenate. Then it is given to espeak. Thus, text is converted into speech.

1) Optical Character Recognition: OCR is an electronic conversion tool which

recognizes the text characters in the image which can be in printed text as well as in written form. The OCR layouts of each character are used to recognize the character. Next, the character image is converted into ASCII code, to be used in data processing. ASCII code limits image recognition, where each character from the image is being picked and sent for recognition. There are some pre-measures required to make the picture noise free which includes processes like binarization, normalization and skew correction. Here we can see the image going through some enhancements, for example, contrast correction and filtering out noise. After these pre-processing, the real role of OCR begins with the segmentation process, in this process each and every character is separated. This level is assisted by some associated segment examination and projection analysis to again assist with the text segmentation process. Further the sequence is followed by feature extraction. All the real time processing in the OCR lies in the feature extraction.

- 2) Text to Speech: We realize that the artificial creation of human sound is generally 'speech'. A system is utilized to generate speech known as a speech synthesizer. In that, we can hear humans and robots. The text-to-speech framework is divided into two: namely a back-end and front-end. Here, the process begins with the conversion of the raw text like symbols, numbers, abbreviations, etc. This process includes text preprocessing & normalization. Next by allotting the phonetic transcription of the front ends of every word and divides the text into clauses into clauses into conditions, prosodic units, expressions, phrases, and sentences. The phonetic transcription is assigned into the word. The output of the front end is nothing but the prosody information and phonetic transcriptions of the symbolic semantic portrayal. Though back-end is transformed into symbolic linguistic portrayal into Audio structure, which is also referred to as the synthesizer

2.4 RESULTS AND DISCUSSIONS

The proposed system is tested with different input sets. In this work, the OCR system is implemented for the recognition of English capital alphabets (A to Z) and numbers 0 to 9. The first step is to capture the image. This is the input image which needs to undergo further process. Each character is recognized

at one time. The recognized character is saved as text with notepad file, as shown in Fig.9. Next, it converts that text into speech. The proposed system integrates the working of a HD camera, OCR tool, TTS engine, sensors, speakers, and a GPS. All these are integrated using Raspberry pi. The device has been designed in such a way that it is portable and easy for a blind person to use. The real time GPS tracker will assist the blind person to navigate his way to the predefined destination [for e.g., destination set from the home] with the help of the voice assistance feature in the GPS the person can walk with turn-by-turn voice guided navigation feature, on the way if there are any obstacles in the way, the sensors will give signal through voice both direction and the distance of obstacle from the person. This will help the blind person to avoid obstacles in his ways. The blind person can easily capture images with a push of a single button and the text in the images will be read aloud using the proposed system in a short time. As a result, the goal was to build and design a compact, convenient, and low-cost system that will assist visually impaired people in moving around in unfamiliar environments. The proposed system is built with all ages in mind, is user-friendly, and does not require any prior experience or knowledge of advanced technologies. Additionally, in future development, we can train the system to detect the person present in front of the user. Along with that the system's functionality can be upgraded by training the system with more text font. So that the user can easily detect any text like a normal person.

2.5 CONCLUSION

Machine learning and artificial intelligence are two of the most rapidly developing technologies. These technologies are critical to the growth of the IT industry. Blind people who have been denied the gift of sight have the right to live normal lives. In our research, we looked at developing a smart device for blind people to help them live more independently by using digital technology and artificial intelligence.

Chapter 3

OCR (Optical Character Recognition)

Optical character recognition or optical character reader (OCR) is the electronic or mechanical conversion of images of typed, handwritten or printed text into machine-encoded text, whether from a scanned document, a photo of a document, a scene-photo (for example the text on signs and billboards in a landscape photo) or from subtitle text superimposed on an image (for example: from a television broadcast).

Widely used as a form of data entry from printed paper data records – whether passport documents, invoices, bank statements, computerized receipts, business cards, mail, printouts of static-data, or any suitable documentation – it is a common method of digitizing printed texts so that they can be electronically edited, searched, stored more compactly, displayed on-line, and used in machine processes such as cognitive computing, machine translation, (extracted) text-to-speech, key data and text mining. OCR is a field of research in pattern recognition, artificial intelligence, and computer vision.

Early versions needed to be trained with images of each character and worked on one font at a time. Advanced systems capable of producing a high degree of recognition accuracy for most fonts are now common, and with support for a variety of digital image file format inputs.[2] Some systems can reproduce formatted output that closely approximates the original page including images, columns, and other non-textual components.

3.1 Introduction

OCR software can be used to convert a physical paper document, or an image into an accessible electronic version with text. For example, if you scan a paper document or photograph with a scanner, the scanner will most likely create a file with a digital image in it. The file could be a JPG/TIFF or PDF, but the new electronic file may still be only an image of the original document. You can then load this scanned

electronic document it created, which contains the image, into an OCR program. The OCR program will recognize the text and convert the document to an editable text file.

3.2 How Does OCR Work

OCR software processes a digital image by locating and recognizing characters, such as letters, numbers, and symbols. Some OCR software will simply export the text, while other programs can convert the characters to editable text directly in the image. Advanced OCR software can export the size and formatting of the text as well as the layout of the text found on a page.

3.3 How Does OCR Work

The short answer is no, not really. Some OCR programs allow you to scan a document and convert it to a word processing document in a single step, but this still may not be an accessible document. Once you use the OCR to process your document, you must select the text and read it to verify the process was successful, and the text makes sense. You may have to spell check it, add headings, add tags, reorder it and more. You can do this with your word processor such as Word, or Adobe Acrobat Pro

3.4 Do I really need to Proofread and Correct an OCR output?

Yes! Think of it this way: If your original had good contrast and readability a 99% success rate is possible with some OCR software, but what if the 1 % wrong was the tuition rate for the college? If the original image had poor contrast and readability the success rate could go down to 50% or even be unreadable. You will not know until you check it!

3.5 What if My OCR Output is Really Bad?

If your original document or image has poor contrast, fuzzy characters, overlaps, etc. the OCR software may recognize the text, but the text may not be accurate and be hard to read to make your OCR output is more successful with easier to verify and correct output, make sure your original is not a fuzzy reprint. The original should have good contrast and sharp letters. If you cannot get your hands on a better version of the original, the printer, or printer software, may have settings on it that will produce a better scan.

3.6 Blind and visually impaired users

In 1974, Ray Kurzweil started the company Kurzweil Computer Products, Inc. and continued development of omni-font OCR, which could recognize text printed in virtually any font (Kurzweil is often credited with inventing omni-font OCR, but it was in use by companies, including CompuScan, in the late 1960s and 1970s[3][6]). Kurzweil decided that the best application of this technology would be to create a reading machine for the blind, which would allow blind people to have a computer read text to them out loud. This device required the invention of two enabling technologies – the CCD flatbed scanner and the text-to-speech synthesizer. On January 13, 1976, the successful finished product was unveiled during a widely reported news conference headed by Kurzweil and the leaders of the National Federation of the Blind.[citation needed] In 1978, Kurzweil Computer Products began selling a commercial version of the optical character recognition computer program. LexisNexis was one of the first customers and bought the program to upload legal paper and news documents onto its nascent online databases. Two years later, Kurzweil sold his company to Xerox, which had an interest in further commercializing paper-to-computer text conversion. Xerox eventually spun it off as Scansoft, which merged with Nuance Communications.

In the 2000s, OCR was made available online as a service (WebOCR), in a cloud computing environment, and in mobile applications like real-time translation of foreign-language signs on a smartphone. With the advent of smart-phones and smart glasses, OCR can be used in internet connected mobile device applications that extract text captured using the device's camera. These devices that do not have OCR functionality built into the operating system will typically use an OCR API to extract the text from the image file captured and provided by the

device. The OCR API returns the extracted text, along with information about the location of the detected text in the original image back to the device app for further processing (such as text-to-speech) or display.

Various commercial and open source OCR systems are available for most common writing systems, including Latin, Cyrillic, Arabic, Hebrew, Indic, Bengali (Bangla), Devanagari, Tamil, Chinese, Japanese, and Korean characters.

3.7 Application

- Data entry for business documents, e.g. Cheque, passport, invoice, bank statement and receipt
- Automatic number plate recognition
- In airports, for passport recognition and information extraction
- Automatic insurance documents key information extraction [citation needed]
- Traffic sign recognition
- Extracting business card information into a contact list
- More quickly make textual versions of printed documents, e.g. book scanning for Project Gutenberg
- Make electronic images of printed documents searchable, e.g. Google Books
- Converting handwriting in real-time to control a computer (pen computing)
- Defeating CAPTCHA anti-bot systems, though these are specifically designed to prevent OCR. The purpose can also be to test the robustness of CAPTCHA anti-bot systems.
- Assistive technology for blind and visually impaired users
- Writing the instructions for vehicles by identifying CAD images in a database that are appropriate to the vehicle design as it changes in real time.
- Making scanned documents searchable by converting them to searchable PDFs

3.8 Pre Processing

- De-skew – If the document was not aligned properly when scanned, it may need to be tilted a few degrees clockwise or counterclockwise in order to make lines of text perfectly horizontal or vertical.
- Despeckle – remove positive and negative spots, smoothing edges
- Binarisation – Convert an image from color or greyscale to black-and-white (called a "binary image" because there are two colors). The task of binarisation is performed as a simple way of separating the text (or any other desired image component) from the background. The task of binarisation itself is necessary since most commercial recognition algorithms work only on binary images since it proves to be simpler to do so. In addition, the effectiveness of the binarisation step influences to a significant extent the quality of the character recognition stage and the careful decisions are made in the choice of the binarisation employed for a given input image type; since the quality of the binarisation method employed to obtain the binary result depends on the type of the input image (scanned document, scene text image, historical degraded document etc.).
- Line removal – Cleans up non-glyph boxes and lines
- Layout analysis or "zoning" – Identifies columns, paragraphs, captions, etc. as distinct blocks. Especially important in multi-column layouts and tables.
- Line and word detection – Establishes baseline for word and character shapes, separates words if necessary.
- Script recognition – In multilingual documents, the script may change at the level of the words and hence, identification of the script is necessary, before the right OCR can be invoked to handle the specific script.
- Character isolation or "segmentation" – For per-character OCR, multiple characters that are connected due to image artifacts must be separated; single characters that are broken into multiple pieces due to artifacts must be connected.
- Normalize aspect ratio and scale

3.9 Text recognition

- Matrix matching involves comparing an image to a stored glyph on a pixel-by-pixel basis; it is also known as "pattern matching", "pattern recognition", or "image correlation". This relies on the input glyph being correctly isolated from the rest of the image, and on the stored glyph being in a similar font and at the same scale. This technique works best with typewritten text and does not work well when new fonts are encountered. This is the technique the early physical photocell-based OCR implemented, rather directly.
- Feature extraction decomposes glyphs into "features" like lines, closed loops, line direction, and line intersections. The extraction features reduces the dimensionality of the representation and makes the recognition process computationally efficient. These features are compared with an abstract vector-like representation of a character, which might reduce to one or more glyph prototypes. General techniques of feature detection in computer vision are applicable to this type of OCR, which is commonly seen in "intelligent" handwriting recognition and indeed most modern OCR software. Nearest neighbour classifiers such as the k-nearest neighbors algorithm are used to compare image features with stored glyph features and choose the nearest match.

Software such as Cuneiform and Tesseract use a two-pass approach to character recognition. The second pass is known as "adaptive recognition" and uses the letter shapes recognized with high confidence on the first pass to recognize better the remaining letters on the second pass. This is advantageous for unusual fonts or low-quality scans where the font is distorted (e.g. blurred or faded).

Modern OCR software include Google Docs OCR, ABBYY FineReader and Transym. Others like OCRopus and Tesseract uses neural networks which are trained to recognize whole lines of text instead of focusing on single characters.

A new technique known as iterative OCR automatically crops a document into sections based on page layout. OCR is performed on the sections individually using variable character confidence level thresholds to maximize page-level OCR accuracy. A patent from the United States Patent Office has been issued for this method.

The OCR result can be stored in the standardized ALTO format, a dedicated XML schema maintained by the United States Library of Congress. Other common formats include OCR and PAGE XML.

3.10 Conclusion

Special fonts like OCR-A, OCR-B, or MICR fonts, with precisely specified sizing, spacing, and distinctive character shapes, allow a higher accuracy rate during transcription in bank check processing. Ironically, however, several prominent OCR engines were designed to capture text in popular fonts such as Arial or Times New Roman, and are incapable of capturing text in these fonts that are specialized and much different from popularly used fonts. As Google Tesseract can be trained to recognize new fonts, it can recognize OCR-A, OCR-B and MICR fonts.

"Comb fields" are pre-printed boxes that encourage humans to write more legibly – one glyph per box.[28] These are often printed in a "dropout color" which can be easily removed by the OCR system.

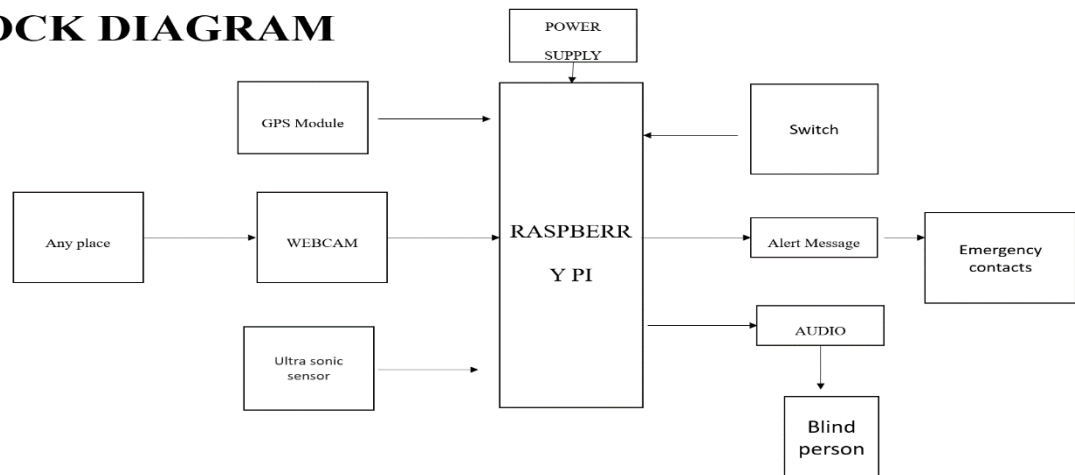
Palm OS used a special set of glyphs, known as "Graffiti" which are similar to printed English characters but simplified or modified for easier recognition on the platform's computationally limited hardware. Users would need to learn how to write these special glyphs.

Zone-based OCR restricts the image to a specific part of a document. This is often referred to as "Template OCR".

3.11 PROPOSED SYSTEM

3.11.1 Block Diagram

BLOCK DIAGRAM



We have described 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, we have proposed a motion-based method to detect the object of interest, while the blind user simply shakes the object for a couple of seconds. The automatic ROI detection and text localization algorithms were independently evaluated as unit tests to ensure effectiveness and robustness of the whole system. We subsequently evaluated this prototype system of assistive text reading using images of hand-held objects captured by ten blind users in person. Two calibrations were applied to prepare for the system test. First, we instructed blind users to place hand-held object within the camera view. Since it is difficult for blind users to aim their held objects, we employed a camera with a reasonably wide angle.

In future systems, we will add finger point detection and tracking to adaptively instruct blind users to aim the object. Second, in an applicable blind-assistive system, a text localization algorithm might prefer higher recall by sacrificing some precision

When our application starts running it first check all the devices and resources which it needs are available or not. After that it checks the connection with the devices and gives control to the user. The GUI for the user has the following options. An optional label is used for displaying the image taken from the camera. A status box is for representing the detected data from the image. The capture button is to detect the data from the image. The detect button is to detect the human from the video streaming

in front of the camera. The audio jack port is the output port here. The Raspberry board comes with integrated peripherals like USB, ADC and Serial etc. On this board we are installing Linux operating system with necessary drivers for all peripheral devices.

Hardware

- Raspberry Pi
- Web Camera
- Speaker

Software

- Raspian OS
- Python
- Open CV

CHAPTER 4

HARDWARE DESCRIPTION

4.1 RASPBERRY PI

4.1.2 Overview

Raspberry Pi seems to be new in the world and many people really don't know what the Raspberry Pi is. Raspberry Pi can be defined as a low cost, credit-card sized computer that plugs into a computer monitor or TV, and uses a standard keyboard and mouse. It is a capable little device that enables people of all ages to explore computing, and to learn how to program in languages like Scratch and Python. It's capable of doing everything you would expect a desktop computer to do, from browsing the internet and playing high definition video, to making spreadsheets, word-processing, and playing games. It is great bonding with Arduino and can do a lot with Arduino.

4.2 Exploring the Raspberry Pi Board

There are two models of Raspberry Pi, model A and model B. These two are bit similar with few advance features on model B compared to model A. Model B has 512 MB RAM, two USB port where as Model A has 256 MB RAM and just a USB port. Besides, Model B has Ethernet port while Model A does not.

The overview of the Raspberry Pi Model B is shown in Figure 5-1.

Figure 5-1: Raspberry Pi Model B(Sanders, 2013)

Different components of the Pi are named in the Figure 5-1 and brief description on each component is given in following sections.

4.2.1 SD Card Slot

Raspberry Pi doesn't have the real hard drive as in laptop and computer, SD card is taken as solid state drive (SSD) which is used to install operating system and all others software and store everything. This card is needed to insert into the slot for using the Raspberry Pi.

SD card may be 2GB, 4GB or 16GB.

4.2.2 Micro USB Power

The power port is a 5V micro-USB input and supply should be exactly 5v as it doesn't have onboard power regulator. So, power supply shouldn't exceed than 5V.

4.2.3 HDMI Out

This output port is used to connect the Raspberry Pi with a monitor via HDMI (High Definition Multimedia Interface). Hence, any screen or TV can be connected to it which consists of HDMI port.

4.2.4 Ethernet and USB port

Both the Ethernet port and USB port on Model B are supplied via the onboard LAN9512 chip. It is a high-speed USB 2.0 hub with a 10/100 Ethernet controller (Donat, 2014). USB ports are used to connect the inputs (keyboard, mouse). Almost everything that can connect to computer via USB also can connect with Raspberry Pi.

4.2.5 RCA Video Out and Audio Out

Audio and RCA video jacks are present on the board for audio and video out.

The Raspberry Pi does support sound over its HDMI output, but there is a standard 3.5-mm audio jack to plug in headphones but USM mikes may work or not.

For video, the RCA jack sends video to any connected RCA video device.

4.2.6 GPIO Headers(Pins)

GPIO pins stands for general purpose of input output pins. These pins are used to connect any number of physical extensions with the Raspberry Pi. Raspberry Pi has pre-installed libraries that allow us to access the pins using programming languages like C, C++ or python.

4.2.7 Chips (Broadcom)

The most important component in a Raspberry Pi is chip that is Broadcom which is placed at the middle of the board. The chip consists of ARM11 processor running at 700 MHz and a Videocore4 GPU and can be over clocked to at least 800 MHz without a problem.

4.3 Raspberry Pi 2 Model B

Recently, Raspberry Pi 2 Model B has been launched recently which Broadcom BCM2836 ARM Cortex-A7 Quad Core Processor has powered Single Board Computer running at

900MHz, 1GB RAM and 4 Quad USB ports. It is the advanced version of Model B and is 6 times faster than Model B Raspberry Pi. In addition, it has combined 4-pole jack for connecting your stereo audio out and composite video out and advanced power management. Figure 5-2 shows the top view of the board with labels of some important components (raspberrypi.org, 2015a).

Figure 5-2: Raspberry Pi 2 Model B

4.4 Hardware Required for Raspberry Pi

Raspberry Pi can't start alone, it needs many others peripherals (hardware). There is brief description of the hardware requirements in the following section (Bates, 2014).

4.4.1 Power Supply

As mentioned already in above theory portion, Raspberry Pi needs 5V power supply. If supply exceeds 5V then it can't guaranteed to work properly. And the power supply also need to supply at least 500 milliamps (mA), and preferably more like 1 amp (A). If the supply is 500 mA or less, it is likely to have the mal-function of keyboard and mouse. It is not good idea to power the Raspberry Pi from USB port of computer and hub as they mostly provide current less than required. Hence, the Raspberry Pi requires a Micro-USB connection which is capable of supplying at least 700 mA (or 0.7 A) at 5V.

4.4.2 Storage

A separate hardware is required for the storage purpose in Raspberry Pi. For this, SD card is used, mostly 4 GB and 8 GB if needed. The operating system and all files are stored in the card. We can buy blank SD card and install operating system or buy a pre-installed one.

4.4.3 Input

External keyboard and mouse are required to provide input to the Raspberry Pi. No any additional software is needed to use the keyboard and mouse.

4.4.4 Monitor

We can use monitor or TV with HDMI port or DVI inputs as the screen for the Raspberry Pi. For DVI inputs, HDMI-to-DVI converters are required which can be finding easily in a market. Monitor is most important for the Raspberry Pi as it is the only way to see what we have done on it.

4.4.5 Network

As in laptop or computer, we can access to internet and network in Raspberry Pi as well. For that, we can use wired Ethernet connection which is easier option or Wi-Fi module to access Wi-Fi in the Raspberry Pi.

4.5 Power supply unit

A power supply (sometimes known as a power supply unit or PSU) is a device or system that supplies electrical or other types of energy to an output load or group of loads. The term is most commonly applied to electrical energy supplies, less often to mechanical ones, and rarely to others.

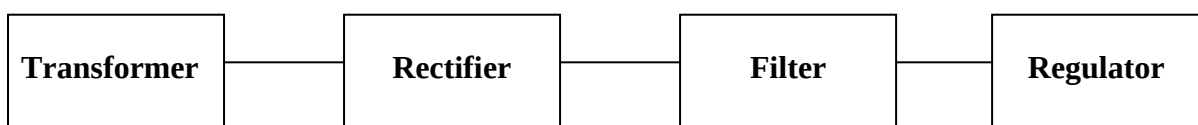
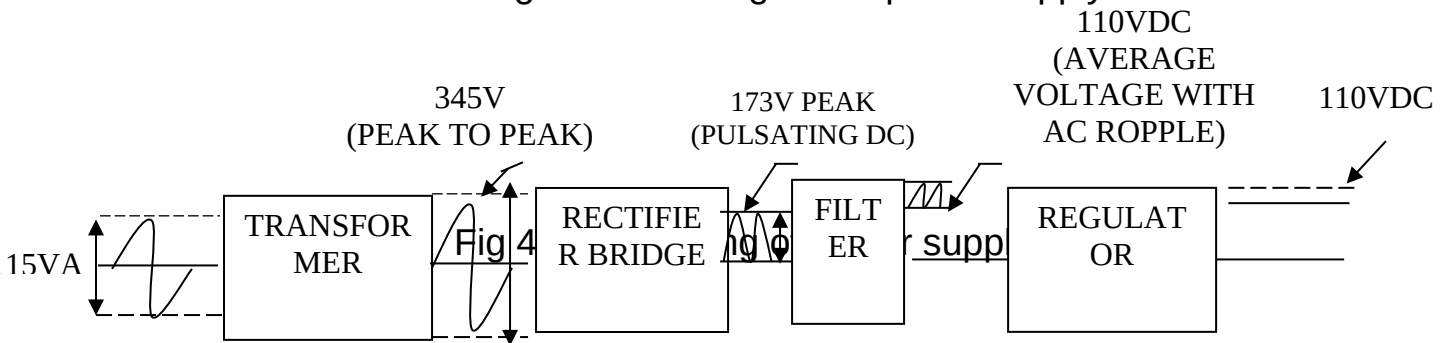


Fig4.1 Block diagram of power supply



The transformer steps up or steps down the input line voltage and isolates the power supply from the power line. The RECTIFIER section converts the alternating current input signal to a pulsating direct current. However, as you proceed in this chapter you will learn that pulsating dc is not desirable. For this reason a FILTER section is used to convert pulsating dc to a purer, more desirable form of dc voltage.

The final section, the REGULATOR, does just what the name implies. It maintains the output of the power supply at a constant level in spite of large changes in load current or input line voltages. Now that you know what each section does, let's trace an ac signal through the power supply. At this point you need to see how this signal is altered within each section of the power supply. Later on in the chapter you will see how these changes take place. An input signal of 115 volts ac is applied to the primary of the transformer.

The transformer is a step-up transformer with a turns ratio of 1:3. You can calculate the output for this transformer by multiplying the input voltage by the ratio of turns in the primary to the ratio of turns in the secondary; therefore, $115 \text{ volts ac} \times 3 = 345 \text{ volts ac}$ (peak-to-peak) at the output. Because each diode in the rectifier section conducts for 180 degrees of the 360-degree input, the output of the rectifier will be one-half, or approximately 173 volts of pulsating dc. The filter section, a network of

resistors, capacitors, or inductors, controls the rise and fall time of the varying signal; consequently, the signal remains at a more constant dc level. You will see the filter process more clearly in the discussion of the actual filter circuits. The output of the filter is a signal of 110 volts dc, with ac ripple riding on the dc. The reason for the lower voltage (average voltage) will be explained later in this chapter. The regulator maintains its output at a constant 110-volt dc level, which is used by the electronic equipment (more commonly called the load).

4.5.1 Simple 5V power supply for digital circuits

- Brief description of operation: Gives out well-regulated +5V output, output current capability of 100 mA
- Circuit protection: Built-in overheating protection shuts down output when regulator IC gets too hot
- Circuit complexity: Very simple and easy to build
- Circuit performance: Very stable +5V output voltage, reliable operation
- Availability of components: Easy to get, uses only very common basic components
- Design testing: Based on datasheet example circuit, I have used this circuit successfully as part of many electronics projects
- Applications: Part of electronics devices, small laboratory power supply
- Power supply voltage: Unregulated DC 8-18V power supply
- Power supply current: Needed output current + 5 mA
- Component costs: Few dollars for the electronics components + the input transformer cost

4.6 WEBCAM:-

A **webcam** is a video camera which feeds its images in real time to a computer or computer network, often via USB, Ethernet or Wi-Fi. Their most popular use is the establishment of video links, permitting computers to act as videophones or videoconference stations. This common use as a video camera for the World Wide Web gave the webcam its name. Other popular uses include security surveillance and computer vision. Webcams are known for their low manufacturing cost and flexibility, making them the lowest cost form of video telephony. They have also become a source of security and privacy issues, as some built-in webcams can be remotely activated via spyware.

4.6.1 LENS

Webcams typically include a lens, an image sensor, support electronics, and may also include a microphone for sound. Various lenses are available, the most common in consumer-grade webcams being a plastic lens that can be screwed in and out to focus the camera. Fixed focus lenses, which have no provision for adjustment, are also available. As a camera system's depth of field is greater for small image formats and is greater for lenses with a large f-number (small aperture), the systems used in webcams have a sufficiently large depth of field that the use of a fixed focus lens does not impact image sharpness to a great extent.

Image sensors can be CMOS or CCD, the former being dominant for low-cost cameras, but CCD cameras do not necessarily outperform CMOS-based cameras in the low cost price range. Most consumer webcams are capable of providing VGA resolution video at a frame rate of 30 frames per second. Many newer devices can produce video in multi-megapixel resolutions, and a few can run at high frame rates such as the PlayStation Eye, which can produce 320×240 video at 120 frames per second.

Support electronics read the image from the sensor and transmit it to the host computer. The camera pictured to the right, for example, uses a Sonix SN9C101 to transmit its image over USB. Typically, each frame is transmitted uncompressed in RGB or YUV or compressed as JPEG. Some cameras, such as mobile phone

cameras, use a CMOS sensor with supporting electronics "on die", i.e. the sensor and the support electronics are built on a single silicochip to save space and manufacturing costs.

4.6.2 PORTABILITY:

Most webcams feature built-in microphones to make video calling and videoconferencing more convenient. The USB video device class (UVC) specification allows for interconnectivity of webcams to computers without the need for proprietary device drivers. Microsoft Windows XP SP2, Linux and Mac OS X (since October 2005) have UVC support built in and do not require extra device drivers, although they are often installed to add additional features.

4.6.3 FEATURES (LOGITECH WEBCAM C100):-

- Plug-and-play setup (UVC)
- Video capture: Up to 640 x 480 pixels
- Photos: Up to 1.3 megapixels (software enhanced)
- Frame rate: Up to 30 frames per second (with recommended system)
- Hi-Speed USB 2.0 certified
- Fixed focus
- Universal clip fits notebooks, LCD or CRT monitor



Webcam

4.7 Image Processing on Raspberry Pi

Today image processing are used in various techniques, this paper presents the implementation of image processing operations on Raspberry Pi. The Raspberry Pi is a basic embedded system and being a low cost a single board computer used to reduce the complexity of systems in real time applications. This platform is mainly based on python. Raspberry pi consist of Camera slot Interface (CSI) to interface the raspberry pi camera. Here, the Dark and Low contrast images captured by using the Raspberry Pi camera module are enhanced in order to identify the particular region of image. This concept is used in the real time application of MAV, The MAVs are basically used to capture images and videos through the Raspberry pi camera module. Because of its credit card sized (small) and less weight in the design. However, the image captured by MAVs will consist of unwanted things due to atmospheric conditions; hence it is necessary to remove noise present in the MAVs images.

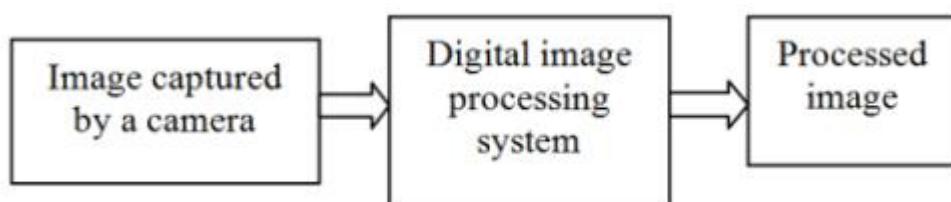
INTRODUCTION

The image processing is a form of signal processing where the input is an image, photograph or video frame, the output of an image processing may be either an image or a video frame or a set of characteristics or parameters related to the image. The acquisition of digital image usually suffers from undesirable camera shakes and due to unstable random camera motions. Hence image enhancement algorithms are required to remove these unwanted camera shakes. This image processing concepts are implemented in Raspberry pi in the application of MAV. The Raspberry Pi is a basic embedded system having a credit card-sized single board computers developed in the UK by the Raspberry Pi Foundation. The Raspberry Pi is based on the Broadcom BCM2835 system on a chip (SOC) which includes an ARM1176JZF-S Core (ARM V6K)700 MHz CPU processor, Broadcom Video Core IV GPU having 17 pins, 3.5W of power, and 512 MB of RAM memory. The Raspberry Pi system has Secure SD card reader (models A and B) or Micro SD card reader (models A+ and

B+) sockets for boot media and persistent storage. The system provides Debian Linux operating system Raspbian image for download. Python is used as main programming language for raspberry pi. A micro air vehicle (MAV) is a remote-controlled, Unmanned Aircraft Vehicle (UAV) significantly smaller than typical UAVs that have a size restriction. UAV is an aircraft without a human pilot. Its flight is controlled either autonomously on board computers or by the remote control of a pilot on the ground or in another vehicle. By having a Raspberry Pi camera module available on a MAV the efficiency of this air vehicle increases and new fields of applications are available. It is needed in military Operations, in which targets have to be identified. Such identification is often done by a human on ground, to reduce the probability of mistakes. But a Raspberry Pi camera module is also helpful if a MAV shall autonomously fly through an arch.

4.8 BASIC CONCEPT OF IMAGE PROCESSING

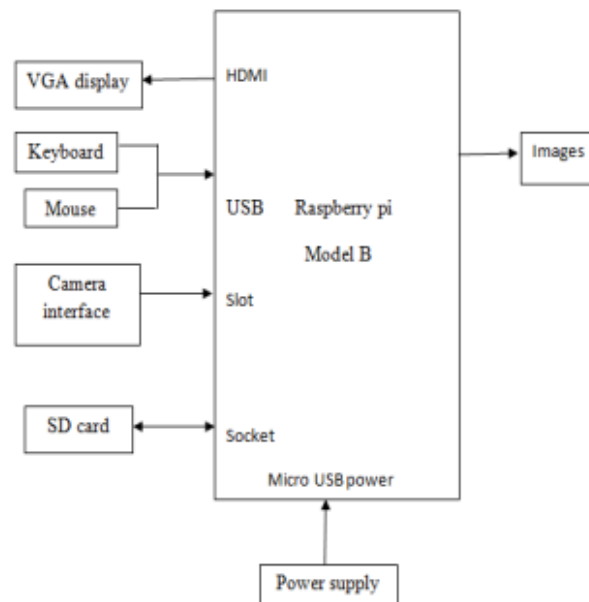
In general, any digital image processing algorithm consists of three stages: input, processor and output. In the input stage image is captured by a camera. It sent to a particular system to focus on a pixel of image that's gives, its output as a processed image.



General Block diagram of image processing

4.9 SYSTEM HARDWARE DESIGN

The Raspberry Pi board is the central module of the whole embedded image capturing and processing system as given in fig. 2. Its main parts include: main processing chip unit, memory, power supply HDMI Out i.e VGA display, Ethernet port, and USB ports.



System Block Diagram

CHAPTER 5

5. SOFTWARE DESCRIPTION

5.1 Start up with Raspberry Pi

5.1.1 Installation of operating system on Raspberry Pi

Raspberry Pi is a small computer; hence operating system (OS) should be installed. As the Raspberry doesn't have hard drive, OS is installed in the external memory. For that, memory card (SD card) is used for the installation of operating system and all the required software and supporting files are stored in the same SD card.

There are different types of operating system but we preferred to talk about NOOBS (New Out of the Box Software) as it is suitable for the beginners. We can either buy a preinstalled SD card or an empty SD card. In pre-installed SD card, NOOBS is already copied and ready to boot. The detailed description of installation of operating system in a blank SD card in step wise is given in Appendix B.

5.1.2 Use of laptop screen, keyboard, and mouse for Raspberry Pi

If you do not have a suitable screen for the Raspberry Pi, we can use the monitor of a laptop. Besides, we can use keyboard and mouse of laptop as well for Raspberry Pi. For that, two software programs are needed to install in a laptop which screen is going to use for Raspberry Pi. Installation and configuration of the software is discussed briefly in the following sections.

Xming

This is the first software need to be installed which can be downloaded from the link, [Download_Xming](#) and install it in the laptop. After completion of installation, run the application called 'XLaunch' and verify that the settings are as shown in three following figures; Figure 12-1, Figure 12-2 and Figure 12-3.

Finally, click "Next" button shown in Figure 12-3, it goes to last configuration dialogue box where "Finish" button is needed to click for the completion of the setting. After Completion of configuration, double click the application named as "Xming".

1 5.2 PuTTY

This is the primary software need to be installed. It can be downloaded in the provided link as follow, [Download_Putty](#). As, it is downloaded, it needs to be installed following some few normal steps of installation. For Configuration, double click the icon of Putty after the completion of installation and enter the IP address of Raspberry Pi as shown in Figure 12-4.

Finally, click on "open" button and as a result; window will be displayed as shown in Figure 12-6.

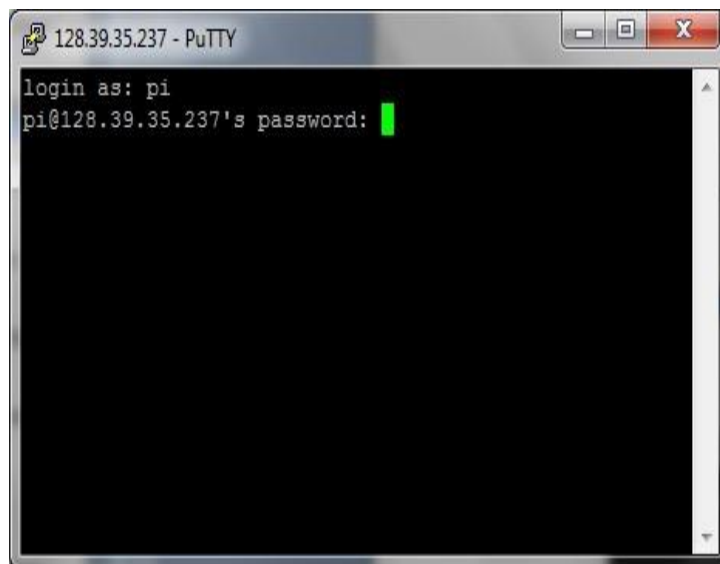


Figure 12-6: PuTTY login windows

In this window, username and password of the Raspberry Pi should be entered. The default username for a Raspberry Pi is pi. Press enter after entering correct username and password and you will see the windows where you should enter the text "lesson" as shown in Figure 12-7.

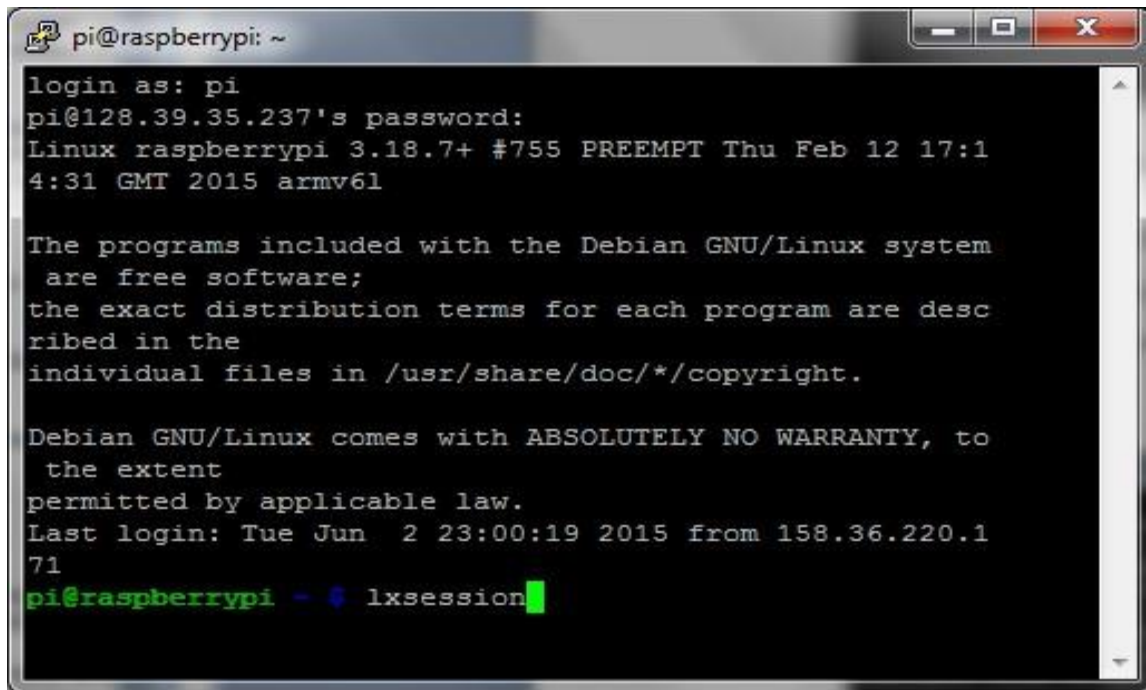


Figure 12-7: PuTTY success windows

Then, you will be able to see the desktop of Raspberry Pi in the laptop as shown in Figure 12-8 and mouse and keyboard can be used for the Raspberry Pi.

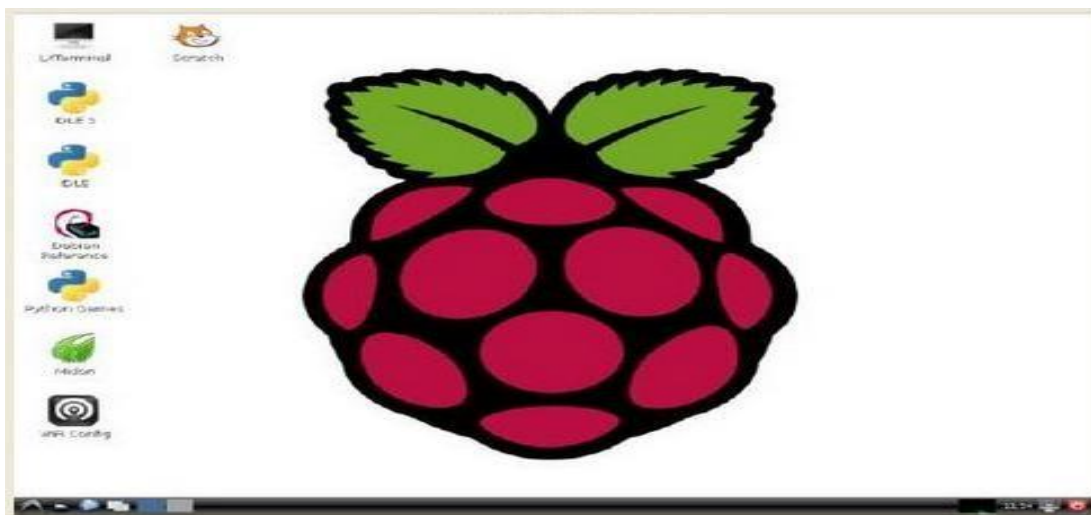


Figure 12-8: Raspberry Pi Desktop on your Laptop Display

5.3 Installation of required applications on Raspberry Pi

There are many applications that are needed to install in the Raspberry Pi for the completion of the thesis. For data logging, MySQL apache5 and php my admin are

needed to install whereas for the web-page development, PHP is needed to install. Web page is used for the monitoring and managing purpose

Raspbian OS

Raspberry Pi OS (formerly Raspbian) is a Debian-based operating system for Raspberry Pi. Since 2015, it has been officially provided by the Raspberry Pi Foundation as the primary operating system for the Raspberry Pi family of compact single-board computers. The first version of Raspbian was created by Mike Thompson and Peter Green as an independent project. The initial released build was completed on July 15, 2012.

Raspberry Pi OS is highly optimized for the Raspberry Pi line of compact single-board computers with ARM CPUs. It runs on every Raspberry Pi except the Pico microcontroller. Raspberry Pi OS uses a modified LXDE as its desktop environment with the Openbox stacking window manager, along with a unique theme. The distribution is shipped with a copy of the algebra program Wolfram Mathematica and a version of Minecraft called Minecraft: Pi Edition (note that Minecraft: Pi Edition is no longer installed as of the Debian bullseye update) as well as a lightweight version of the Chromium web browser.

User Interface

Raspberry Pi OS's desktop environment, PIXEL, looks similar to many common desktops, such as macOS, Microsoft Windows, and is based on LXDE. The menu bar is positioned at the top and contains an application menu and shortcuts to Terminal, Chromium, and File Manager. On the right is a Bluetooth menu, a Wi-Fi menu, volume control, and a digital clock.

Components

PCManFM is a file browser allowing quick access to all areas of the computer, and was redesigned in the first Raspberry Pi OS Buster release (2019-06-20).

Raspberry Pi OS originally used Epiphany as the web browser, but switched to Chromium with the launch of its redesigned desktop.

Raspberry Pi OS comes with many beginner IDEs, such as Thonny Python IDE, Mu Editor, and Greenfoot. It also ships with educational software like Scratch and Bookshelf.

Version Size

The Raspberry Pi documentation recommends at least a 4GB microSD card for Raspberry Pi OS Lite, an 8GB microSD card for Raspberry Pi OS, and a 16GB microSD card for Raspberry Pi OS Full.[19] The image files themselves are 1,875MB, 3,980MB, and 8,603MB respectively.

Python

What is Python?

Python: Dynamic programming language which supports several different programming paradigms:

- Procedural programming
- Object oriented programming
- Functional programming

Standard: Python byte code is executed in the Python interpreter (similar to Java) → platform independent code

Why Python?

Extremely versatile language Website development, data analysis, server maintenance, numerical analysis,

- Syntax is clear, easy to read and learn (almost pseudo code)
- Common language
- Intuitive object oriented programming
- Full modularity, hierarchical packages
- Comprehensive standard library for many tasks
- Big community Simply extendable via C/C++, wrapping of C/C++ libraries
- Focus: Programming speed

History

Start implementation in December 1989 by Guido van Rossum (CWI)

- 16.10.2000: Python 2.0
- Unicode support
- Garbage collector
- Development process more community oriented
- 3.12.2008: Python 3.0
- Not 100% backwards compatible 2007 & 2010 most popular programming language (TIOBE Index)
- Recommendation for scientific programming (Nature News, NPG, 2015)
- Current version: Python 3.9.2
- Python2 is out of support

Is Python fast enough?

- For user programs: Python is fast enough!
- Most parts of Python are written in C
- For compute intensive algorithms: Fortran, C, C++ might be better
- Performance-critical parts can be re-implemented in C/C++ if necessary
- First analyses, then optimize!

Conclusion

Python is a fully-fledged programming language. Any code you download from github, the Python Package Index, or anywhere else, can be malicious and would almost certainly go unnoticed by most people's firewalls.

OpenCV

OpenCV is a cross-platform library using which we can develop real-time computer vision applications. It mainly focuses on image processing, video capture and analysis including features like face detection and object detection. Let's start the chapter by defining the term "Computer Vision"

Computer Vision

Computer Vision can be defined as a discipline that explains how to reconstruct, interrupt, and understand a 3D scene from its 2D images, in terms of the properties of the structure present in the scene. It deals with modeling and replicating human vision using computer software and hardware.

Computer Vision overlaps significantly with the following fields:

- ☐ Image Processing: It focuses on image manipulation.
- ☐ Pattern Recognition: It explains various techniques to classify patterns.
- ☐ Photogrammetry: It is concerned with obtaining accurate measurements from images.

Computer Vision Vs Image Processing

Image processing deals with image-to-image transformation. The input and output of image processing are both images.

Computer vision is the construction of explicit, meaningful descriptions of physical objects from their image. The output of computer vision is a description or an interpretation of structures in 3D scene.

Applications of Computer Vision

Here we have listed down some of major domains where Computer Vision is heavily used. **Robotics Application**

- ☐ Localization — Determine robot location automatically
- ☐ Navigation

- ☐ Obstacles avoidance
- ☐ Assembly (peg-in-hole, welding, painting)
- ☐ Manipulation (e.g. PUMA robot manipulator)
- ☐ Human Robot Interaction (HRI): Intelligent robotics to interact with and serve people

Medicine Application

Classification and detection (e.g. lesion or cells classification and tumor detection)

- ☐ 2D/3D segmentation
- ☐ 3D human organ reconstruction (MRI or ultrasound)
- ☐ Vision-guided robotics surgery

Industrial Automation Application

- ☐ Industrial inspection (defect detection)
- ☐ Assembly
- ☐ Barcode and package label reading
- ☐ Object sorting
- ☐ Document understanding (e.g. OCR)

Security Application

- ☐ Biometrics (iris, finger print, face recognition)
- ☐ Surveillance — Detecting certain suspicious activities or behaviors Transportation

Application

- ☐ Autonomous vehicle
- ☐ Safety, e.g., driver vigilance monitoring

Features of OpenCV Library

Using OpenCV library, you can –

- Read and write images
- Capture and save videos
- Process images (filter, transform)
- Perform feature detection
- Detect specific objects such as faces, eyes, cars, in the videos or images.
- Analyze the video, i.e., estimate the motion in it, subtract the background, and track objects in it. OpenCV was originally developed in C++. In addition to it, Python and Java bindings were provided. OpenCV runs on various Operating Systems such as windows, Linux, OSx, FreeBSD, Net BSD, Open BSD, etc.

This tutorial explains the concepts of OpenCV with examples using Java bindings.

OpenCV Library Modules

Following are the main library modules of the OpenCV library.

Core Functionality

This module covers the basic data structures such as Scalar, Point, Range, etc., that are used to build OpenCV applications. In addition to these, it also includes the multidimensional array Mat, which is used to store the images. In the Java library of OpenCV, this module is included as a package with the name org.opencv.core.

Image Processing

This module covers various image processing operations such as image filtering, geometrical image transformations, color space conversion, histograms, etc. In the Java library of OpenCV, this module is included as a package with the name org.opencv.imgproc.

Video

This module covers the video analysis concepts such as motion estimation, background subtraction, and object tracking. In the Java library of OpenCV, this module is included as a package with the name org.opencv.video.

Video I/O

This module explains the video capturing and video codecs using OpenCV library. In the Java library of OpenCV, this module is included as a package with the name `org.opencv.videoio`.

calib3d

This module includes algorithms regarding basic multiple-view geometry algorithms, single and stereo camera calibration, object pose estimation, stereo correspondence and elements of 3D reconstruction. In the Java library of OpenCV, this module is included as a package with the name `org.opencv.calib3d`.

features2d

This module includes the concepts of feature detection and description. In the Java library of OpenCV, this module is included as a package with the name `org.opencv.features2d`.

Objdetect

This module includes the detection of objects and instances of the predefined classes such as faces, eyes, mugs, people, cars, etc. In the Java library of OpenCV, this module is included as a package with the name `org.opencv.objdetect`.

Highgui

This is an easy-to-use interface with simple UI capabilities. In the Java library of OpenCV, the features of this module is included in two different packages namely, `org.opencv.imgcodecs` and `org.opencv.videoio`.

OpenCV — Storing Images

To capture an image, we use devices like cameras and scanners. These devices record numerical values of the image (Ex: pixel values). OpenCV is a library which processes the digital images, therefore we need to store these images for processing. The `Mat` class of OpenCV library is used to store the values of an image. It represents an n-dimensional array and is used to store image data of grayscale or color images, voxel volumes, vector fields, point clouds, tensors, histograms, etc.

This class comprises of two data parts: the header and a pointer

- Header: Contains information like size, method used for storing, and the address of the matrix (constant in size).

□ Pointer: Stores the pixel values of the image (Keeps on varying).

The Mat Class

The OpenCV Java library provides this class with the same name (Mat) within the package `org.opencv.core`.

Constructors

The Mat class of OpenCV Java library has various constructors, using which you can construct the Mat object.

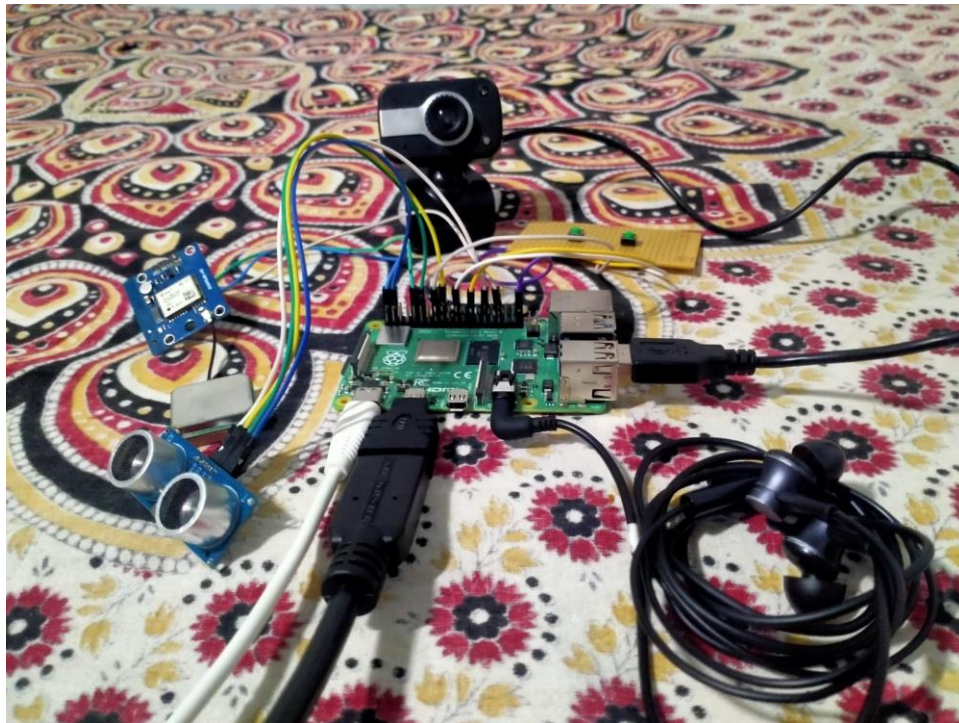
1. This is the default constructor with no parameters in most cases. We use this to constructor to create an empty matrix and pass this to other OpenCV methods.
 2. `Mat(int rows, int cols, int type)` This constructor accepts three parameters of integer type representing the number of rows and columns in a 2D array and the type of the array (that is to be used to store data).
 3. `Mat(int rows, int cols, int type, Scalar s)` Including the parameters of the previous one, this constructor additionally accepts an object of the class `Scalar` as parameter.
 4. `Mat(Size size, int type)`
3. OpenCV —
Storing Images

CHAPTER 6

RESULTS AND CONCLUSION

6.1 RESULT:

Past the pic



6.2 CONCLUSIONS:

In this analysis, we got described an epitome system to scan written communication and hand-held objects for serving to the blind people. To extract text regions from advanced backgrounds, we have got projected a totally distinctive text localization formula supported models of stroke orientation, and edge distributions. The corresponding feature maps estimate the worldwide structural feature of text at every element. Block patterns project the projected feature maps of an image patch into a feature vector. A adjacent character grouping is performed to calculate candidates of text patches prepared for text classification. Associate Ada boost learning model is used to localize text in camera-based footage. OCR is utilized to perform word recognition on the localized text regions and work into audio output for blind users. Throughout this analysis, the camera acts as input for the paper. As a result of the Raspberry Pi board is high-powered the camera starts streaming. The streaming data square measure attending to be displayed on the screen victimization interface application. Once the item for text reading is placed earlier than the camera then the capture button is clicked to supply image to the board. Mistreatment Tesseract library the image square measure attending to be converted into data, and conjointly the data detected from the image square measure attending to be shown on the standing bar. The obtained data square measure attending to be pronounced through.

REFERENCES:

- [1] X. Chen and A. L. Yuille, "Detecting and reading text in natural scenes," in Proc. Comput. Vision Pattern Rec-ognit., 2004, vol. 2, pp. II-366–II-373.
- [2] S. Kumar, R. Gupta, N. Khanna, S. Chaudhury, and S. D. Joshi, "Text extraction and document image seg-mentation using matched wavelets and MRF model," IEEE Trans Image Process., vol. 16, no. 8, pp. 2117– 2128, Aug. 2007.
- [3] K. Kim, K. Jung, and J. Kim, "Texture-based ap-proach for text detection in images using support vec-tor machines and continuously adaptive mean shift al-gorithm," IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 12, pp. 1631–1639, Dec. 2003.
- [4] N. Giudice and G. Legge, "Blind navigation and the role of technology," in The Engineering Handbook of Smart Technology for Aging, Disability, and Indepen-dence, A. A. Helal, M. Mokhtari, and B. Abdulrazak, Eds. Hoboken, NJ, USA: Wiley, 2008.
- [5] World Health Organization. (2009). 10 facts about blindness and visual impairment.

- [6] Advance Data Reports from the National Health Interview Survey (2008).
- [7] International Workshop on Camera-Based Document Analysis and Recognition (CBDAR 2005, 2007, 2009, 2011).
- [8] X. Chen, J. Yang, J. Zhang, and A. Waibel, "Auto-matic detection and recognition of signs from natural scenes," IEEE Trans. Image Process., vol. 13, no. 1, pp. 87–99, Jan. 2004.
- [9] D. Dakopoulos and N. G. Bourbakis, "Wearable obstacle avoidance electronic travel aids for blind: A survey," IEEE Trans. Syst., Man, Cybern., vol. 40, no. 1, pp. 25–35, Jan. 2010.
- [10] B. Epshtein, E. Ofek, and Y. Wexler, "Detecting text in natural scenes with stroke width transform," in Proc. Comput. Vision Pattern Recognit., 2010, pp. 2963–2970.
- [11] Y. Freund and R. Schapire, "Experiments with a new boosting algorithm," in Proc. Int. Conf. Machine Learning, 1996, pp. 148–156.
- [13] An overview of the Tesseract OCR (optical character recognition) engine, and its possible enhancement for use in Wales in a pre-competitive research stage Prepared by the Language Technologies Unit (Canol-fan Bedwyr), Bangor University April 2008.
- [14] A. Shahab, F. Shafait, and A. Dengel, "ICDAR 2011 robust reading competition: ICDAR Robust Reading Competition Challenge 2: Reading text in scene images," in Proc. Int. Conf. Document Anal. Recognit., 2011, pp. 1491–1496.
- [15] KReader Mobile User Guide, knfb Reading Technology Inc. (2008).
- [16] S. M. Lucas, "ICDAR 2005 text locating competition results," in Proc. Int. Conf. Document Anal. Recognit., 2005, vol. 1, pp. 80–84.

