VISION BASED ASSITIVE SYSTEM FOR LABEL DETECTION

D.Naga Madhavi[1] and Mr.G.Subhramanya Sarma[2]

Student (M.Tech), Electronics and Communication Engineering, Lingayas institute of management and technology, Vijayawada[1]

Asst professor, Electronics and Communication Engineering, Lingayas institute of management and Technology, Vijayawada[2]

E mail: damarla.madhavi@gmail.com[1], sarma_g_s@yahoo.com[2]

ABSTRACT : This work is related to the principles of an Optical character Recognizer which Is also called as (O.C.R.).We propose a camera-based assistive text reading framework to help blind persons read text labels and product packaging from hand-held objects in their daily life. To isolate the object from untidy backgrounds or other surrounding objects in the camera vision.The final program runs in the linux environment, the mechanics is similar to a conventional (O.C.R.). To implement the O.C.R. we used several of image processing and identification of the images and conversion of image into text and Is automatically converted to speech. There are about 45 million blind people and 135 million visually impaired people worldwide. Disability of visual text reading has a huge impact on the quality of life for visually disabled people. Although there have been several devices designed for helping visually disabled to see objects using an alternating sense such as sound and touch, the development of text reading device is still at an early stage. Existing systems for text recognition are typically limited either by explicitly relying on specific shapes or colour masks or by requiring user assistance or may be of high cost. Therefore we need a low cost system that will be able to automatically locate and read the text aloud to visually impaired persons. The main idea of this project is to recognize the text character and convert it into speech signal.

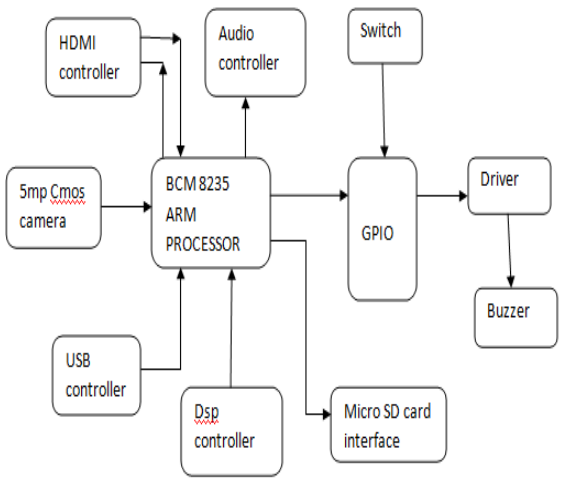
Keywords: Character recognition, Image capturing, Text, Speech signal, camera, Embedded systems.

I.INTRODUCTION

In worldwide the visually impaired peoples are 314 million, in that 45 million are visual impairment which was released by “World Health Organization” survey [1] for last few years, to help them or to make their life easy this project is very useful to them, according to this project is to develop an optical character recognizer (O.C.R.).We used several of image processing and identification of the images and converting the image into text and This text is automatically convert to audio tape by This Project we Can help the Blind people to recognise any objects and converting the image captured text into audio tape.Optical Character Recognition has been an active subject of research since a decade. The rapid growth of digital

libraries worldwide poses new challenges for document image analysis research and development. Speech is probably the most efficient medium for communication between humans. Optical character recognition has become one of the most successful applications of technology in the field of pattern recognition and artificial intelligence [3]. Character recognition or optical character recognition (OCR), is the process of converting scanned images of machine printed or handwritten text (numerals, letters, and symbols), into a computer format text. . Speech synthesis is the artificial synthesis of human speech. A Text-To Speech (TTS) synthesizer is a computer-based system[2] that should be able to read any text aloud, whether it was directly introduced in the computer by an operator or scanned and submitted to an Optical Character Recognition (OCR) system. Operational stages of the system consist of image capture, image preprocessing, image filtering, character recognition and text to speech conversion.

II.SYSTEM BLOCK DIAGRAM AND DESCRIPTION

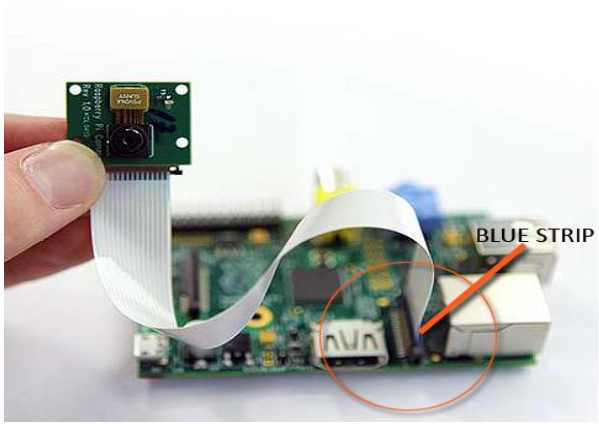


Block diagram

RASPBERRY PI: Raspberry pi comprises of two models Model A and Model B**.** The Raspberry Pi Model A is here, this is the ideal model for anyone on a budget but still wants a piece of the Pi[10]. The credit-card sized computer is capable of many of the things that your desktop PC does, like spread sheets, word-processing and games. It also plays high-definition video. The secret sauce that makes this computer so small and powerful is the Broadcom BCM2835, a System-on-Chip that contains an ARM 1176JZFS with floating point, running[11] at 700 MHz, and a Video core 4 GPU. The GPU provides Open GL ES 2.0, hardware-accelerated Open VG, and 1080p30 H.264 high-profile decode and is capable of 1Gpixel/s, 1.5Gtexel/s or 24 GFLOPs of general purpose compute. What's that all mean? It means that if you plug the Raspberry Pi into your HDTV, you could watch Blu-ray quality video, using H.264 at 40MBits/s.

MODEL B: This is the Raspberry Pi Model B 512MB RAM model with two USB ports and a 10/100 Ethernet controller.As typical of modern computers, generic USB keyboards and mice are compatible with the Raspberry Pi. The Raspberry Pi use Linux-kernel based operating systems. The Raspberry Pi[12] does not come with a real-time clock, so an OS must use a network time server, or ask the user for time information at boot time to get access to time and date info for file time and date stamping[14]. However a real time clock (such as the DS1307) with battery backup can be easily added via the I2C interface.

Raspberry Pi



Raspberry Pi camera

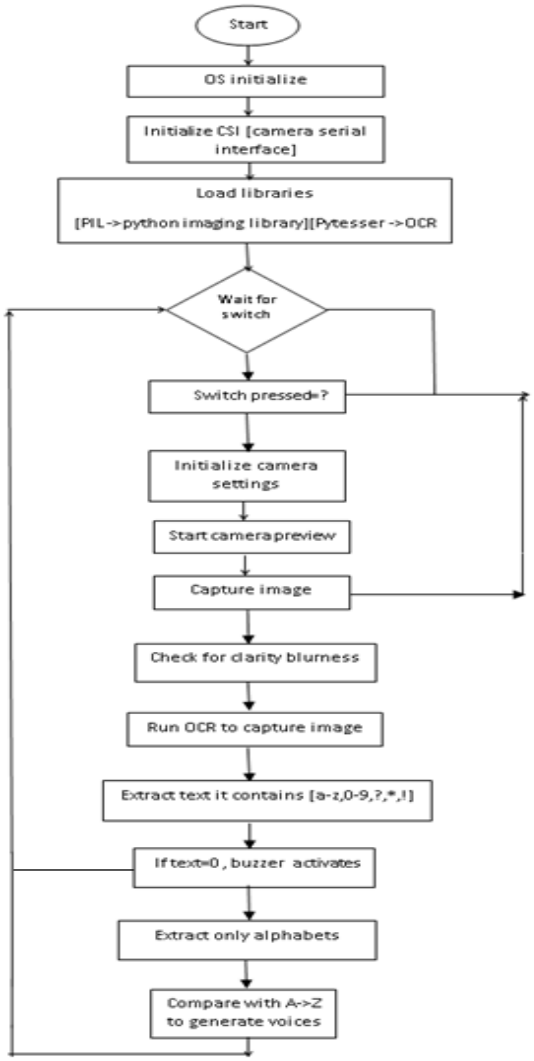
CMOS CAMERA: CMOS sensors have circuitry at the pixel level. This means that every pixel on the sensor is read and transmitted simultaneously, preparing voltage for the chip. The chip then uses additional technology, such as amplifiers, noise correction, and digitization, to convert the voltage to digital data. This means that CMOS sensors do not require a separate image processor. Because CMOS sensors are able to convert visual information to digital data more quickly than CCDs, they require less power, which preserves battery life. However, the extra technology on the sensor crowds the pixels, limiting their ability to capture light and resulting in generally poorer visual clarity in the final image. CMOS sensors are commonly designed with rolling shutters, especially on commercial applications. This means that the image frame is exposed from one side to the other, instead of all at once as on CCD sensors. For example, a video camera using a CMOS sensor may record data in a "rolling" sweep from left to right, or top to bottom. This results in the potential for a few types of distortion not found on CCD sensors.

III.WORKING

Optical character recognition (OCR) technology offers blind and visually impaired persons the capacity to scan printed text and then speak it back in synthetic speech. This technology can minimize the problems of blind people while shopping. There are three essential elements to OCR technology—scanning, recognition, and reading text. Initially, a printed text is scanned by a camera. OCR software then converts the images into recognized characters and words. The synthesizer in the OCR system then speaks the recognized text. Pytesser is an Optical Character Recognition module for Python. It takes as input an image or image file and outputs a string. Pytesser uses the Tesseract OCR engine, converting images to an accepted format and calling the Tesseract Executable as an external script. A Windows executable is provided along with the Python scripts. PCM/I2S Audio we use to convert the text into speech form and it is inbuilt. The PCM audio interface is an APB peripheral providing input and output of telephony or high quality serial audio streams. It supports many classic PCM formats including I2S. PCM is a serial format with a single bit data in and

single bit data out. Data is always serialized MS-bit first. The frame sync signal (PCM_FS) is used to delimit the serial data into individual frames. The length of the frame and the size and position of the frame sync are fully programmable. Frames can contain 1 or 2 audio/data channels in each direction. Each channel can be between 8 and 32 bits wide and can be positioned anywhere within the frame as long as the two channels don't overlap. The channel format is separately programmable for transmit and Receive directions. The PCM_CLK can be asynchronous to the bus APB clock and can be logically inverted if required. The direction of the PCM_CLK and PCM_FS signals can be individually selected, allowing the interface to act as a master or slave device. The input interface is also capable of supporting up to 2 PDM microphones, as an alternative to the classic PCM input format, in conjunction with a PCM output. The PCM audio interface contains separate transmit and receive FIFOs. Note that if the frame contains two data channels, they must share the same FIFO and so the channel data will be interleaved. The block can be driven using simple polling, an interrupt based method or direct DMA control.Let's suppose life was really simple and there was only one letter in the alphabet A, Even then you can probably see that OCR would be quite a tricky problem because every single person writes the letter A in a slightly different way. Even with printed text, there's an issue, because books and other documents are printed in many different typefaces (fonts) and the letter A can be printed in many subtly different forms. Broadly speaking, there are two different ways to solve this problem, either by recognizing characters or by detecting the individual lines. OCR can minimize all these problems in a simple manner and makes blind people to fulfil their requirements like shopping etc..,

IV.FLOW CHART



Initialize camera settings and then load libraries like python imaging library and pytesser OCR library and initialize camera settings and start camera preview then capture image and check clarity blurness of image, Run OCR to capture the image and extract the text it contains a-z,0-9,?,*,! Here extract only alphabets compare string 1 then after compare string1 with string 2, string 2 with string 3 likewise Compare with A->Z to generate voices.

V. CONCLUSION

We have built a complete text recognition system which is very useful for blind people in a friendly environment. This innovative application is able to automatically identify and recognize text zones in images taken from a camera. It performs well for a wide range of document images and no prior knowledge concerning document layout, character size, type, color, and orientation has been used. A new thresholding algorithm has been proposed and discrimination between kinds of documents enables to apply this new method on corresponding documents, such as strongly degraded ones. Segmentation and recognition steps aim at considering degraded characters among with touching and broken ones. A large study on low-level OCR error correction was presented. Results are already promising but this step will have a larger impact once character recognition rate will be higher thanks to some improvements, such as in character segmentation. Future work in text detection consists in modifying our approach to a real multi-resolution system by applying the same algorithm to different instances of the image at different resolutions. An expansion to text detection embedded into natural scenes is currently under investigation. A model could be created depending on some kinds of documents or types of degradations to improve the recognition rate drastically. For the time being, it is unrealistic to create a generic recognition system that reaches significant results for all kinds of text images.

VI. FUTURE SCOPE

We are using fixed focus camera, due to this the camera captures only for fixed distances with low clarity images. Instead of this if we use auto focus cameras then it will capture the images for long distances and with high clarity.

REFERENCES

[1] World Health Organization. (2009). 10 facts about blindness and visual impairment [Online]. Available: www.who.int/features/factfiles/blindness/blindness_facts/en/index.html

[2]D.Doermann, J. Liang, and H. Li, "Progress in camera-based document image analysis," in Proc. 7th IEEE International Conference on Document Analysis and Recognition (ICDAR'03), vol. 1, pp. 606–617, August 2003.

[3]R. Lienhart and A.Wernicke, "Localizing and segmenting text in images, videos and web pages," IEEE Trans. Circuits Syst.Video Technol., vol. 12, no. 4, pp. 256–268, 2002.

[4]X. Fernandez Hermida, F. Martin Rodriguez, J. L. Fernandez Lijo, F. Pita Sande, and M. Perez Iglesias, "An OCR for vehicle License plates," in Proc. International Conference on Signal Processing Applications & Technology (ICSPAT '97), SanDiego, Calif, USA, September 1997.

[5] J. Gao and J. Yang, "An adaptive algorithm for text detection from natural scenes," in Proc. IEEE Computer Society ConferenceOn Computer Vision and Pattern Recognition (CVPR '01), vol. 2, pp. 84–89, Kauai, Hawaii, USA, 2001.

[6] J. Ohya, A. Shio,s and S. Akamatsu, "Recognizing characters in scene images," IEEE Trans. Pattern Anal. Machine Intell.,vol. 16, no. 2, pp. 214–220, 1994.

[7] K. Wang and J. A. Kangas, "Character location in scene images from digital camera," Pattern Recognition, vol. 36, no. 10,pp. 2287–2299, 2003.

[8]P. Clark and M. Mirmehdi, "Recognizing text in real scenes," International Journal on Document Analysis and Recognition,vol. 4, no. 4, pp. 243–257, 2002.

[9] A. K. Jain and B. Yu, "Automatic text location in images and video frames," Pattern Recognition, vol. 31, no. 12, pp. 2055–2076, 1995.

[10] https://en.wikipedia.org/wiki/Raspberry_Pi

[11]http://lifehacker.com/5976912/a-beginners-guide-to-diying-with-the-raspberry-pi

[12]http://www.cs.unca.edu/~bruce/Fall14/360/RPiUsersGuide.pdf

[13]http://pi.cs.man.ac.uk/download/Raspberry_Pi_Education_Manual.pdf

[14]https://www.raspberrypi.org/products/model-b-plus/

[15]http://arstechnica.com/information-technology/2012/12/10-raspberry-pi-creations-that-show-how-amazing-the-tiny-pc-can-be/

BIOGRAPHIES

D.Naga Madhavi pursuing M.Tech in Lingayas Institute of Management &Technology, Vijayawada in the stream of Embedded Systems I received graduation with a Bachelor Degree in Electronics &

Communication Engineering from Priyadarshini Institute of Management and Technology, Near Tenali in 2013.

G.S.Sarma currently pursuning phd form K.L.Univeristy He completed M.Tech form K.L.CE in year 2008 he had 4 International journals and 15 National conference journal