ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Portable Camera Based Assistive Text and Label Reading For Blind Persons

Undurthi Tanuja
PG Scholar,
BVC Engineering College,
Odalarevu, Amalapuram.

P.Harika
Assistant Professor,
BVC Engineering College,
Odalarevu, Amalapuram.

J.Umamaheswari
Assistant Professor,
BVC Engineering College,
Odalarevu, Amalapuram.

ABSTRACT:

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 lives. To isolate the object from cluttered backgrounds or other surrounding objects in the camera view, we first propose an efficient and effective motion based method to define a region of interest (ROI) in the video by asking the user to show the object or text. This method extracts object or text region by a mixture-of-Gaussians-based background subtraction method. In the extracted ROI, text localization and recognition are conducted to acquire text information. To automatically localize the text regions from the object ROI, we propose a novel text localization algorithm by learning gradient features of stroke orientations and distributions of edge pixels in a text reading algorithm. Text characters in the localized text regions are then binarized and recognized by off-the-shelf optical character recognition software. We explore user interface issues and assess robustness of the algorithm in extracting and reading text from different objects with complex backgrounds

Keywords:

RASPBERRY PI ,OCR, Qt.

INTRODUCTION:

Recent developments in computer vision, digital cameras, and portable computers make it feasible to assist these individuals by developing camera-based products that combine computer vision technology with other existing commercial products such optical character recognition (OCR) systems. Reading is obviously essential in today's society.

Printed text is everywhere in the form of reports, receipts, bank statements, restaurant menus, classroom handouts, product packages**,** instructions on medicine bottles, etc. And while optical aids, video magnifiers, and screen readers can help blind users and those with low vision to access documents, there are few devices that can provide good access to common hand-held objects such as product packages, and objects printed with text such as prescription medication bottles. The ability of people who are blind or have significant visual impairments to read printed labels and product packages will enhance independent living and foster economic and social self-sufficiency.

Today, there are already a few systems that have some promise for portable use, but they cannot handle product labeling. For example, portable bar code readers designed to help blind people identify different products in an extensive product database can enable users who are blind to access information about these products through speech and braille. But a big limitation is that it is very hard for blind users to find the position of the bar code and to correctly point the bar code reader at the bar code. Some reading-assistive systems such as pen scanners might be employed in these and similar situations.

Such systems integrate OCR software to offer the function of scanning and recognition of text and some have integrated voice output. However, these systems are generally designed for and perform best with document images with simple backgrounds, standard fonts, a small range of font sizes, and well-organized characters rather than commercial product boxes with multiple decorative patterns.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Most state-of-the-art OCR software cannot directly handle scene images with complex backgrounds. The aim of the project is to help the blind persons to read labels from hand-held objects. The main contribution of this paper is to develop a low cost system to help the blind persons. The camera which is connected with hardware is used to capture the image. The image is processed internally in the raspberry pi hardware to separate the label from the captured image by using OPENCV (open source computer vision) library. The desired letters in the label is identified by using Tesseract-OCR (optical character recognition). The identified product name is announced through voice. The FLITE software is used to convert the obtained text into speech. The ARM processor available in the Raspberry pi is reduced the cost and size of the hardware

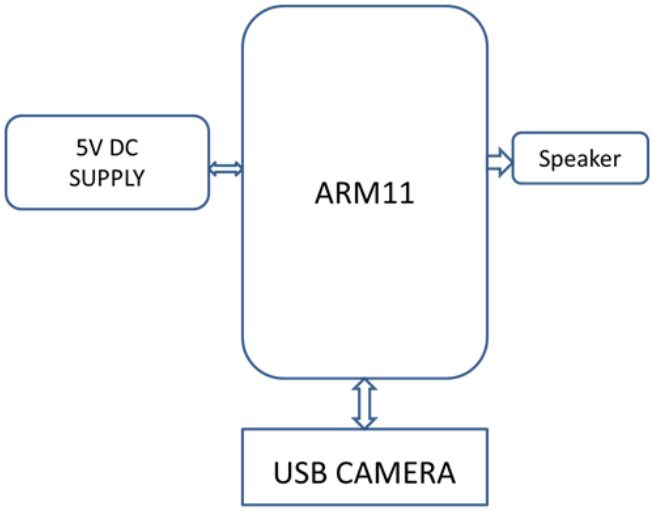

Figure 1. Block Diagram

The ARM processor in the hardware is to be programmed. The camera connected through USB port(Universal Serial bus) is used to get the image. The image is then processed internally in the hardware. The identified words of the label is announced through earphone which is connected in the audio port of the hardware.

ARM Processor:

ARM is a 32-bit RISC processor architecture developed by the ARM Corporation. ARM processors possess a unique combination of features that makes ARM the most popular embedded architecture today. First, ARM cores are very simple compared to most other general-purpose processors, which means that they can be manufactured using a comparatively small number of transistors, leaving plenty of space on the chip for application specific macro cells. A typical ARM chip can contain several peripheral controllers, a digital signal processor, and some amount of on-chip memory, along with an ARM core. Second, both ARM ISA and pipeline design are aimed at minimizing energy consumption — a critical requirement in mobile embedded systems.

Third, the ARM architecture is highly modular: the only mandatory component of an ARM processor is the integer pipeline; all other components, including caches, MMU, floating point and other co-processors are optional, which gives a lot of flexibility in building application-specific ARM-based processors. Finally, while being small and low-power, ARM processors provide high performance for embedded applications. For example, the ARM11 processor running at 700MHz provides performance comparable to Pentium 2 at 300MHz, while using fifty times less energy.

Operating System:

LINUX is a free and open source software operating system for computers. The operating system is a collection of the basic instructions that tell the electronic parts of the computer what to do and how to work. Some components of an installed Linux system are: A boot loader: for example GNU GRUB this is a program which is executed by the computer when it is first turned on, and loads the Linux kernel into memory. An init program: This is the first process launched by the Linux kernel, and is at the root of the process tree: in other terms, all processes are launched through in it. It starts processes such as system services and login prompts (whether graphical or in terminal mode). Software libraries which contain code which can be used by running processes. User interface programs such as command shells or windowing environments.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

QT Embedded Frame Work:

Qt is a cross-platform application framework that is widely used for developing application software with a graphical user interface (GUI) (in which cases Qt is classified as a widget toolkit), and also used for developing non-GUI programs such as command-line tools and consoles for servers. It uses standard C++ but makes extensive use of a special code generator (called the Meta Object Compiler, or moc) together with several macros to enrich the language.

Open CV Open CV (Open Source Computer Vision) is a library of programming functions for real time computer vision. Advantages of Open CV over MAT LAB (Collected from various blogs/forums. See references below)

Resources needed:

Due to the high level nature of Mat lab, it uses a lot of your systems resources. Mat lab code requires over a gig of RAM to run through video. In comparison, typical Open CV programs only require ~70mb of RAM to run in real-time.

TESSERACT-OCR Tesseract is probably the most accurate open source OCR engine available. Combined with the Leptonica Image Processing Library it can read a wide variety of image formats and convert them to text in over 60 languages.

FLITEFlite offers text to speech synthesis in a small and efficient binary. Flite is written in ANSI C, and is designed to be portable to almost any platform, including very small hardware.

GCC Compiler:

GCC, formerly for "GNU C Compiler", has grown over times to support many languages such as C++, Objective-C, Java, Fortran and Ada.GCC is a key component of "GNU Tool chain", for developing applications, as well as operating systems. The GNU Tool chain includes: GNU Compiler Collection (GCC): a compiler suit that supports many languages, such as C/C++, Objective-C and Java.GNU Make: an automation tool for compiling and building applications.GNU Binutils: a suit of binary utility tools, including linker and assembler.GNU Debugger (GDB).GNU Auto tools: A build system including Auto conf, Auto header, Auto make and Libtool.GNU Bison: a parser generator GCC is portable and run in many operating platforms. GCC (and GNU Tool chain) is currently available on all Unixes. They are also ported to Windows by MinGW and Cygwin. GCC is also a cross-compiler, for producing executable on different platform.

Hardware RASPBERRY PI

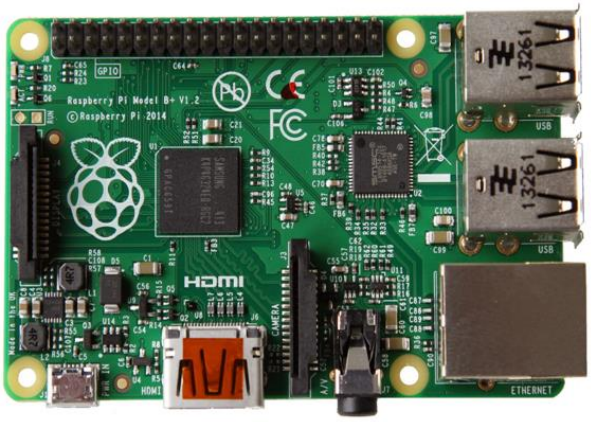

Figure 2.Raspberry pi

The Raspberry Pi board is a credit card sized computer that can be widely used for real-time applications. It contains a processor and graphics chip, program memory (RAM) and various interfaces and connectors for external devices. Some of these devices are essential, others are optional. It operates in the same way as a standard PC, requiring a keyboard for command entry, a display unit and a power supply.

Since raspberry Pi board operates like PC it requires ‘mass-storage’, but a hard disk drive of the type found in a typical PC is not really in keeping with the miniature size of RPi. Instead we will use an SD Flash memory card normally used in digital cameras, configured in such a way to ‘look like’ a hard drive to RPi’s processor. RPi will ‘boot’ (load the Operating System into RAM) from this card in the same way as a PC ‘boots up’ into Windows from its hard disk .

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Camera

A UVC (or Universal Video Class) driver is a USB-category driver. A driver enables a device, such as your webcam, to communicate with your computer's operating system. And USB (or Universal Serial Bus) is a common type of connection that allows for high-speed data transfer. Devices that are equipped with a UVC driver, such as the Logitech® Quick Cam® Pro 9000 for Business, are capable of streaming video. In other words, with a UVC driver, you can simply plug webcam into computer and it'll be ready to use.

Results

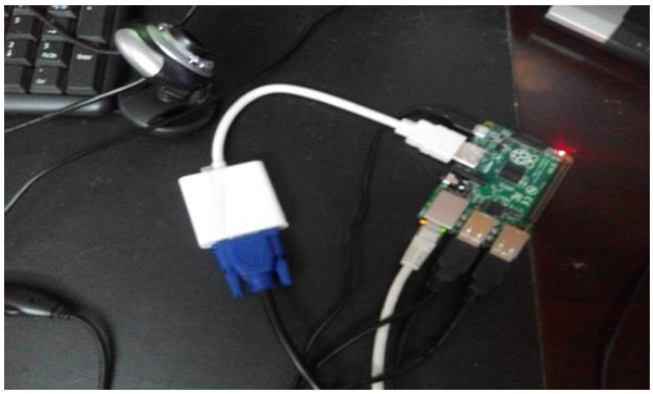

Figure 4. Hardware connection

The simulation output of the camera based assistive text and label reading has been designed. The hardware which is connected with monitor having Raspbian OS is used to get the simulated output results. The Raspbian OS is installed in memory of the hardware. The software which is installed on it is used to recognize the character from the object. The recognized character is then converted and announced through speech over the earphone which is connected into the audio port of the raspberry pi.

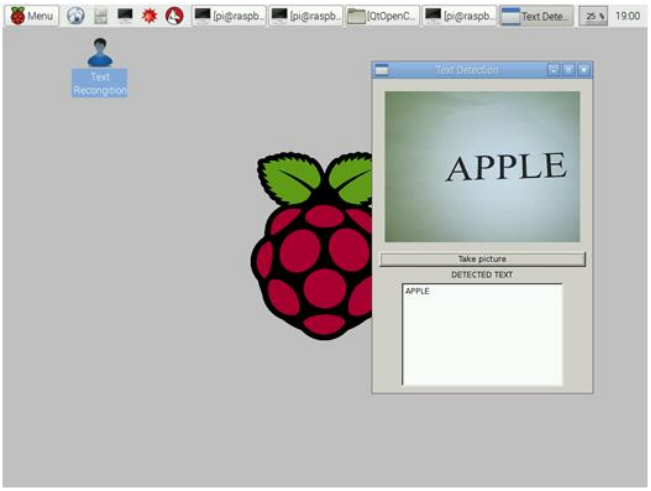

Figure 5.output

Scope of the Study

The system is designed for detecting text through camera from certain distance and it is pronounced through voice. This project can be useful for the blind persons to know the name of the product.

Conclusion:

We have reviewed different types of algorithm for detection of target object in image using different types of image processing techniques accompanied with designed image processing technique in different application areas. Then we focused on the importance of hardware and other environmental factors.

References:

[1]X. Chen, J. Yang, J. Zhang, and A. Waibel, "Automatic detection and recognition of signs from natural scenes," IEEE Trans. Image Process., vol. 13, no. 1, pp. 87–99, Jan. 2004.

[2] 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.

[3] 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.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

[4] Y. Freund and R. Schapire, "Experiments with a new boosting algorithm,"in Proc. Int. Conf. Machine Learning, 1996, pp. 148–156.

[5] N. Giudice and G. Legge, "Blind navigation and the role of technology," in The Engineering Handbook of Smart Technology for Aging, Disability, and Independence, A. A. Helal, M. Mokhtari, and B. Abdulrazak, Eds. Hoboken, NJ, USA: Wiley, 2008.