An open source reading system for print disabilities

Azadeh Nazemi
Electrical and Computer Engineering Department
Curtin University
Perth, WA, Australia
Azadeh.nazemi @postgrad.curtin.edu.au
Iain Murray
Electrical and Computer Engineering Department
Curtin University
Perth, WA, Australia
i.murray @curtin.edu.au

Abstract *:*

According to World Intellectual Property Organization (WIPO) estimation only 5% of the world's one million print titles that are published every year are accessible to the some 340 million around the world who are blind, vision impaired or who live with other print disabilities. Access to information and education is an established human right. Many with print disabilities struggle to achieve equality in this area due to the lack of accessible books and sources of information. This research describes an approach to design a comprehensive reading system for vision impaired people.

Keywords-component; *Reading System, Navigation, Vision impaired, Daisy player, OCR, TTS*

I. Introduction

Nowadays electronic documents become pervasive and people with print disabilities need to access essential information which is presented in these formats. Otherwise they are greatly disadvantaged when compared to the general population [1].Making electronic documents accessible to vision impaired people requires the convers these materials to a tactile or audio format. Using tactile reading methods for vision impaired is dependent on knowing Braille and having sufficient tactile sensitivity which not all vision impaired have. Learning Braille and remembering all its character sets, in some cases, is problematic particularly when using 8 dots Braille to communicate mathematics formulae [2]. Additionally achieved Braille reading rates for vision impaired users are around 100 words per minute which can be compared to an average visual reading rate of around 250 words per minute and preferred rates of around 200 words per minute of synthetic speech or even higher rates for experienced synthetic speech users [3]. Considering the cost, accuracy and static nature of tactile methods to deliver information the print disabled, audio methods appear to be far more promising if particular presentation issues (such as mathematical representation) are highlighted.

II. Research Modules

This research includes the basic following modules:

- Constructing the Block Diagram based on selecting appropriate components considering optimum performance and minimum cost
- Selecting the Boot Mode
- Confirming Electrical and Timing Compatibility
- Designing the Power Subsystem
- Designing the Clocking Sub system
- Testing /debugging[4]
- Development and testing software for prototype
- Confirming Electrical and Timing Compatibility

III. Hadware platform for reading System Prototye

The Initial prototype for Reading System has implemented using Beagle Board as hardware platform. The Beagle Board is the open source board and utilises the OMAP3530 (ARM-CORTEX-A8), it has been allowed to simplify and layout redevelopment. There are several accessories on the Beagle Board development environment which are unnecessary for Reading System, and will therefore be removed to cut cost to the user. Most of the Beagle Board's hardware that is required for this Reading System contains unused capabilities that increase cost which must be replaced with simpler hardware to decrease power consumption and keep cost as low as possible Figure 1 shows block diagram of customized board.

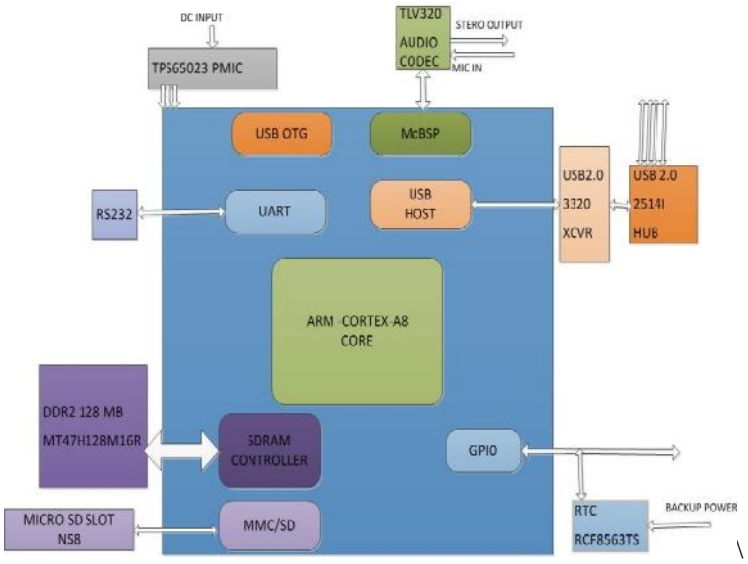

Figure1. Block diagram of customized board for reading system

As it is shown in figure 1 main components are:

- Microprocessor ARM-CORTEX-A8 core. AM3517 is specifically suitable for designing Reading System

- Multi-port USB hosts to connect USB storage for file transferring.
- Multichannel transmit/receive buffer to support TLV320 input/output audio codec
- Removable media interface to boot from mini SDRAM.

IV. Supported Formats

Developed software application written by C run under Arch Linux ARM supports to deal with many types of electronic formats. Reading System application designs a bookshelf and classifies different formats in terms of dealing methods. Reading system plays audio formats such as, WAV, OGG, MP3 and PCM. Besides it makes DOC, DOCX, ODT ,Structured PDF and HTML accessible utilizing open source packages doc2txt and odt2txt to convert them to text. Espeak an open source Text to Speech(TTS) converts text to Wav format and sends it to audio output port.

According a survey results which is conducted among 27 vision impaired users in July 2013 for more than 70% of respondents access to implicit information of non-textual components are very important. Additionally in the view of 86% of participants, full navigation ability is very significant to access visual components. Thus Reading System must be able to deal with JPG, GIF, PNG images of non-textual components such as line charts, bar charts, and pie charts.

Results of a survey among 124 students of Cisco Academy for the Vision Impaired in July 2011 indicate that they believe PDF is most common and the most inaccessible format. Reading System process to provide scanned PDF accessibility, involves image processing and optical character recognition (OCR) to extract text. After text extraction based on image processing text will be send to TTS.

Digital Accessible Information System (DAISY) is a format for Digital Talking Book and multimedia representation of a print publication. It contains navigation ability within sequential and hierarchical structure consisting of (marked-up).User can navigate via different levels, search, and bookmark.The DAISY books can be read by synthesized voice or pre-recorded human narration as mp3 files[5] .

In the opinion of respondent of 2011 survey DAISY navigation ability has very effective role, but complicated and expensive existing players hinder them to approach navigation ability.

Results of two survey among vision impaired users leads this research to focus on:

- PDF accessibility
- Making visual components and mathematics formulae accessible
- Utilizing DAISY Format navigation and bookmark ability

V. Text to Speech control play back functions

Text to Speech(TTS) can convert text file to audio file, it may take a few minutes to convert whole text file to wave audio format. After completing conversion control play back functions has been simply done by using wave player. Besides TTS has option to convert text to audio stream. In this case Reading System controls stream play back

functions by simulation using speech rate. Speech rate calculation of Espeak has been done by sampling and averaging. Table 1 indicates speech rate for six various samples and average value of them.

TABLE I CALCULATION SPEECH RATE

No of letters	No of words	Speech time	Word /second	Letter/second
3232	493	198	2.48	16.32
1515	217	103	2.1	14.7
1159	189	72	2.6	16.09
294	48	21	2.28	14
526	100	36	2.7	14.61
2697	401	164	2.44	16.44
			Average–2.43	Average–15.36

Reading System simulates:

- Pause: by killing TTS task when pause key pressed. Elapsed letters when pause key pressed is calculated by speech rate times elapsed time
- Resume: by starting TTS task from elapsed letters when pause key pressed again
- Forward: by killing TTS task when forward key pressed and starting TTS from :Elapsed letters when forward key pressed+10*speech rate (it means 10 seconds later)
- Rewind: by killing TTS task when rewind key pressed and starting TTS from: Elapsed letters when rewind key pressed-10*speech rate (it means 10 seconds before)

VI. Multi language talking OCR

OCR converts scanned images consisting of typewritten or printed text into machine-encoded text and makes it possible to apply TTS technique. This makes inaccessible print materials accessible to the vision impaired users in audio format. Based on this process one of the major part of Reading System application is Talking OCR which supports English, Hindi and Turkish completely (converts scanned image to audio format), and Bengali and Persian partially(converts scanned image to text).Tesseact 3.02 is an OCR engine which supports many languages even languages with writing from right to left such as Persian optionally by adding TRAINNED DATA language file to software .Also Tesseract has a feature to deal with scanned image contains more than one language. Another used OCR engine in this system is Cuneiform which by default supports most European languages.

VII. MathsReader

Reading and writing mathematics is inherently different from reading and writing text. Mathematics can even be considered a language of its own [6]. The lack of easy access to mathematical resources is a barrier to higher education for many blind students and puts them at an unfair disadvantage in school, academia, and industry [7]. The task of presentation of mathematical formula in an accessible form is very complex. Previous approach to make mathematical

expression accessible to vision impaired users was MATHSPEAK renders MathMl mathematics formulae and converts them to audio format [8]. MathsReader provides applicable input data for MATHSPEAK from image of formulae. MathsReader as a part of reading system is a developed application which represents image of mathematical expression as LaTex format in a logical and comprehensive manner. Used method in this application is based on several technique.OCR engines generally is used to convert image of text to plain text but dealing with the large range of mathematical symbols and two dimensional natures of mathematical expressions are problematic for OCR system. Several techniques running by MathsReader to convert image of math's formulae to LaTex format to be navigable and convertible to audio format, due to text nature of LaTex. Image scaling, gray scale conversion , binary transformation and creation good contrast are doing for preparation. After preparation, Segmentation or Primitive Extraction is running to solve two dimensional issue and transform to linear mode, Recursive Components Extraction (RCE)[9] is a basic technique for segmentation. RCE contains two sub modules: Vertical Components Extraction (VCE) and Horizontal Components Extraction (HCE).VCE first determines number of vertical segments in image by counting gaps (white pixels) and extracts these segments. HCE is applied for all VCE results.All single unit components which are extracted by RCE are applied to English and Greek OCR in order from left to right one by one. Tesseract [10] an open source multi language OCR is able to convert image to text considering more than one language by using this command:

tesseract foo.png text -l eng + ell

Since many Greek alphabets are used in math's expression, both English and Greek OCR run at the same time for every single unit components. Checking Tesseract results shows several symbol are not recognized by the default OCR process. Math's OCR (MOCR) is responsible to transform unrecognized components to text.

All non-alphanumeric math' symbol binary skeleton images have been generated using open source image processing package ImageMagick [11]. Each symbol in this database is represented by Peripheral Direction Contributively (PDC) parameters [12] which indicate image complexion and connectivity. PDC must be calculated for each single unit component extracted by RCE which is not recognized by normal English and Greek OCR. These values are searched in database and be replaced with description field of most appropriate one. Finally the result of the component recognition process is, the best match, without considering the role of the symbol in expression. Accurate semantic structural analysis module helps to transform single units consist two or more connected components and solve conflict problems, based on non-linearity and non-ordering nature of math's expression considering pixel coordinate position of each single unit component. Figure 2 illustrates all processing steps of MathsReader and MATHSPEAK.

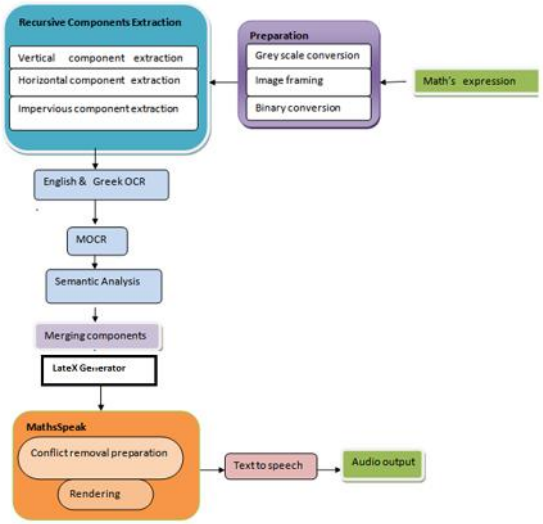

Figure 2. MathsReader block diagram

VIII. Non textual component accessibility

Non-textual graphical information (line graphs, bar charts, pie charts, etc.) are increasingly pervasive in digital scientific literatures and business reports which enabling readers to easily acquire the nature of the underlying information [13]. These graphical components are commonly used to present data in an easy-to interpret way. Graphs are frequently used in economics, mathematics and other scientific subjects. GRAPHREADER[14] as a part of Reading System extracts implicit information of Bar chart, Pie cart and line charts components of an electronic document and presents them to vision impaired users in audio format. Figure 3 shows flowchart of GRAPHREADER.

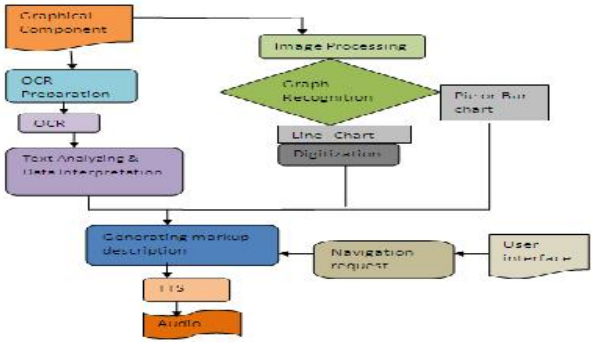

Figure 3. GRAPHREADER Flowchart

IX. Daisy player Features

Reading System contains software daisy player which supports ANSI/NISO Z39.86-2005, ANSI/NISO Z39.86-2002 , DAISY 2.02 and DAISY 3.0 with the following achieved features:

- Insert Bookmark
- Go to specific Bookmark

- Clock, Calendar
- Sleep mode by four different intervals
- Change of playback speed/volume
- Pause /Resume playback function
- Navigation by navigation levels such as: chapter ,section
- Go to next or previous page
- Jump forward or backward by specific time interval
- Use many navigation options
- Info button /title /remain time/total time /elapsed time
- Annotation ability (audio or text):insert ,access and delete
- Using joystick as an interface which is useful for the users with disability

X. Tactile method

A tactual computer monitor or Braille display includes rows and columns of rectangular cells. Each cell includes four rows and two columns of movable pins which are felt and read by a vision impaired user. The pins are driven by electromechanical impact drivers and are held in position by resilient elastomeric cords[15]. In order to represent text file to Braille an option has been considered for Reading System to communicate with Braille display as an auxiliary output. User can send the request to Reading System to use Braille for reading an specific part of text by pressing a key, then the system will interrupt normal reading session, send requested part of text to Braille display and wait for receiving acknowledge from user to send another part to terminal or back to normal reading. Aaccuracy and compatibility translation with multiple Braille codes and compatibility with vast majority of Braille displays are ccritical issues regarding transformation text to Braille . BRLTTY is a daemon provides access to the Linux console (text mode). It drives the Braille terminal and provides complete screen review functionality. Figure 4 shows Beagle Board as Reading System initial prototype connected to USB cartridge as source of data, keypad as user interface, Braille terminal as output for tactile method and speaker as output for audio method .

Figure 4.Reading System Prototype

XI. Proposed Operating System (OS) and software environment

The Reading System uses ARCH ARM Linux. As the system supports navigation or bookmarking ability and controlling playback function, the device must be controllable in real-time , allowing immediate interrupts to set and change the data being played (navigation) or how the data is played (speed, volume control) as well as insert ,delete and access to bookmarks .It is important that the OS kernel be open source (modifiable), free of license constraints, and lightweight which , requiring less memory, less power and cost .

XII. Further Development

Further development must be done in order to customize Linux Kernel to meet the requirements of Reading System. In addition portable electronics Daisy players are currently expensive and technically demanding. The overall aim of this research is to develop a portable, stand alone, low cost and fully functional simple to use reader device for people with print disabilities

References

[1] R.Archee ,M.Whitty, Enhancing student access to the University: the integration of online and course-based material for the visually impaired.,2001.Available at: stc.uws.edu.au/etext/blindpaper2.doc

[2] J. Dixon,(2007), "Eight-dot Braille", Available at http://www.brailleauthority.org/eightdot/eightdot.pdf, Retrieved 17 May2012

[3] I.Murray, Instructional eLearning technologies for the vision impaired .published doctoral thesis, 2008. Curtin University, Perth

[4] AM335x Hardware Design Guide, Available at: http://processors.wiki.ti.com/index.php/AM335x_Hardware_Design_Guide,retrievd 15 November 2012

[5] J.Russo,.DAISY is the Ticket.,2010

[6] A.I .Karshmer, G.Gupta, S.Geiger, and C.Weaver, Reading and Writing Mathematics: The MAVIS Project, BIT (Behavior &Information Technology),1999

[7] C. Jayant , (2006), "A Survey of Math Accessibility For Blind Persons and An Investigation on Text/Math Separation, Available at http://metrc.uoregon.edu/index.php/related-questions/122-how-canmathematical-text-be-made-accessible-to-students-with-printdisabilities.html, retrieved 17 May 2012

[8] A. Nazemi, I, Murray,"Mathspeak: An Audio Method for Presenting Mathematical Formulae to Blind Students", The proceeding of 5th International Conference on Human System Interaction (HSI 2012)

[9] A. Raja, M. Rayner, A.Sexton, V. Sorge ,"Towards a parser for Mathematical Formula Recognition", MKM'06 Proceedings of the 5th international conference on Mathematical Knowledge Management(pp 139-151*)*

[10] Tesseract(2006). Retrieved March 1, 2013, from https://code.google.com/p/tesseract-ocr/wiki/ReadMe

[11] ImageMagick.(2002). Retrieved Feb 15 ,2013, from http://www.imagemagick.org/script/index.php

[12] Ishikawa,C. Ashida,N. Enomoto,Y .Takata,M. Kimesaw,T. Kazuki (2009),"Recognition of Multi-Fonts Character in Early-Modern Printed Books" Nara Women's University, Japan, Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications, PDPTA 2009, Las Vegas,Vol2

[13] T Jinglun Gao , Yin Zhou, Kenneth E. Barner, Making Information Graphics Accessible for Visually Impaired Individuals ,University of Delaware Newark ,DE19716 USA, March 2012

[14] A. Nazemi, I. Murray, A method to provide accessibility for visual components to vision impaired , International Journal of Human-Computer Interaction,2013

Becker Deceased. V, Becker Legal Representative John David A. (late of New York;Brooklyn, NY), Hinton, Daniel E. (Arnold, MD), Anderson Jr., Hugh G. (Westminster, MD) 2004 Braille computer monitor United States 6700553,http://www.freepatentsonline.com/6700553.html