Development of a Text to Braille Interpreter for Printed Documents through Optical Image Processing

Joshua L. Dela Cruz, Jonaida Angela D. Ebreo, Reniel Allan John P. Inovejas, Angelica Romaine C. Medrano, and Argel A. Bandala
Department of Electronics and Communications Engineering
Gokongwei College of Engineering
De La Salle University
argel.bandala@dlsu.edu.ph

***Abstract* – This paper presents the development of an optical text to braille converter device for aiding visually impaired individuals to read printed materials. This is a solution for the lag or even failure of translating or printing the braille version of everyday reading materials. The system utilized optical character recognition engine in which an image of the text to be translated into braille is captured. The digitized texts are then transferred electronically in a braille haptic device. This device are piezoelectric based haptic system which is composed of several haptic pins arranged in a way to resemble the braille writing system. Several experiments were conducted to determine the performance of the system. The overall system reliability obtained was 95.68%. The system is also capable of processing speed of 1 word in 2 seconds. The system performs at its best with a letter sized page reading material within the range of 15 to 20 cm from the camera, with the camera positioned at 0 degrees.**

Keywords – Haptic System, OCR, Image Processing, Text Conversion

I. INTRODUCTION

As the beginning of civilization, man wrote records and documents of transpired events and knowledge learned. These records are the foundations of our society today. Written information are studied and taught from primary school to advance studies. The footprint of technology can be seen clearly in the advancement of education. There are studies who analyzed the effects of modernization to the visually impaired, had foreseen the problems they could face in the future [1].

978-1-5386-0912-5/17/$31.00 ©2017 IEEE

Some have already started to start a cause by creating assistive technologies for communication with the blind. Likewise, there are studies which use refreshable braille display as a medium of communication. The ongoing improvements led to studies utilizing optical image processing, such as processing of image to text as well as conversion of ASCII file and displayed through a refreshable braille device [2]. Furthermore, an algorithm was proposed wherein segmented characters could be analyzed from motion pictures through an optical character recognition (OCR) software [3]. Another study about image processing and conversion was conducted about a kind of Braille paper automatic marking system. [4]. A method was proposed on recognizing low-resolution characters using a super-resolution technique before the recognition to enhance the resolution of images taken by the portable digital camera [5]. This method is important especially to those who will be using low resolution cameras. A step closer to braille reading devices was the two-way SMS-based communication technology, specifically designed for the blind [6]. There was also a study about a microelectromechanical Braille that helps the blind to communicate with the blind or deaf-blind through an SMS subsystem [7]. A more advanced technology presented a design and implementation of a portable keyboard and speaker with a braille refreshable display that is a comparatively low cost for the communication between the blind, deaf, deaf-blind and unimpaired [8].

As compared to these previous studies, there has been no exact work that has been conducted that is able to give assistance to the visually-impaired people regarding their reading education. What makes the research standout from other existing technology is that it adapts to the reading capability of the users introduced by the system's modes of reading. Furthermore, it uses Tesseract technology, one of the best OCR engine to have been developed for its high accuracy. Thus, the development of this braille interpreter with an image to text to braille conversion opens up possibilities to curious and interested visually impaired people to read any printed books and reading paraphernalia.

II. PROCESS FLOW OF THE SYSTEM

Fig. 1 shows the process flow of the system. From the figure, the camera will capture an image, in tiff format, which will be analyzed through an optical image processing integrated in a GUI. After analysis, a universal asynchronous receiver/transmitter (UART) device will interface the communication between the computer and the microcontroller. The microcontroller then sends the data to the refreshable braille cells.

A. *Camera*

See3Cam was the camera used mainly because of its high resolution and auto focus feature. The auto focus' functionality plays a vital role when the reading paraphernalia varies in thickness. It is the main input of the system and is connected to the computer via universal serial bus (USB) 3.0. The device produces thirteen megapixel images that are in *.tiff* format which can be directly fed to the OCR.

B. *Optical Image Processing*

OCR is a technology that enables one to convert different types of documents, such as scanned paper documents, PDF files or images captured by digital camera into editable and searchable data [9]. An open-source OCR engine was used, which is Tesseract. It was developed by HP and was originally a PhD research project as a possible software and/or hardware add-on for HP's line of flatbed scanners. It was released as an open-source OCR during late 2005.

Tesseract in terms of accuracy, has definitely set the standards. Its processing involves traditional step-by-step pipeline. The first step is a connected component analysis. It involves line and word finding through baseline fitting. After which is fixed pitch detection and chopping. Recognition then proceeds as a two-pass process. In the first pass, an attempt is made to recognize each word in turn. Each word that is satisfactory is passed to an adaptive classifier as training data. The adaptive classifier then gets a chance to more accurately recognize text lower down the page. A final phase resolves fuzzy spaces, and tries to find alternative for words that was not recognized. A graphical user interface (GUI) was created to integrate the camera and the OCR engine.

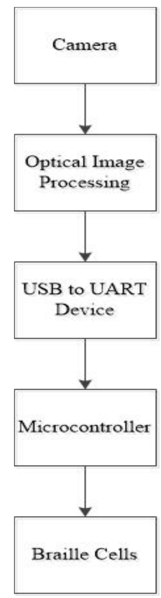

Fig. 1. Process Flow of the System

Furthermore, the serial communication between the computer and the microcontroller is programmed in the interface. After the analysis of the OCR, an output in ASCII code will be obtained, which will be sent to the microcontroller. The GUI will now wait for further commands from the microcontroller. Upon receiving a command from the microcontroller through the UART device, the GUI will send a word through the UART device to the microcontroller for the word to be displayed through the piezoelectric braille cells.

C. *USB to UART Device*

The device responsible for enabling the computer and microcontroller to communicate with each other is the USB to UART device. It is connected to the computer via Type B USB connector, and connected to the microcontroller via the transmitter and receiver pins. The converter has a rated transmission speed of 940 kbps.

D. Microcontroller

PIC16F877A was the microcontroller used to drive the braille cells. The actuation of the data, clock, and strobe signals was done to command the braille cells.

The input to the microcontroller will be coming from the PC, which is done through serial communication. The microcontroller is capable of serial communication through UART, which is used to transmit and receive data. The RS232 standard is used for the communication between the PC and the microcontroller. Since serial communication is the process of sending data one bit at a time, the braille device receives the most significant bit (MSB) first and least significant bit (LSB) last. The microcontroller is coded in such a way that it will process the LSB first so that the output to the braille device will be from the MSB to the LSB.

The microcontroller will convert the data from the PC into braille binary that will actuate the braille dots. The data from the PC that is received by the microcontroller is in ASCII hex format. The microcontroller is coded in such a way that it is capable of matching the ASCII hex data into each character's corresponding braille binary code. The braille binary code was obtained by tracing the grade 1 braille into the appropriate dot number of the braille cells. In addition, there are 3 buttons and 1 switch that will support the functionality of the microcontroller. These are the next and previous button for word-per-word mode, rescan button, and the switch for mode selection: word-per-word or continuous. Once these buttons are pushed, the microcontroller will send signals to the PC for it to correspond accordingly.

E. Braille Cells

The braille cells are the main display for the whole text to braille device. The proponents used 16 cells to display 16 separate characters at an instance. Each cell has 8 braille dots which move up or down, with the use of piezoelectric technology. The piezoelectric driven cells can provide 12 gram-force while activated. Furthermore, the device requires both a 5V and 200V supply, which is provided by the DC-DC converter which has an input of 5V and outputs 200V. Lastly, the braille device receives data from a single pin from the microcontroller, which is passed through an unshielded twisted pair (UTP) cable along with the power supplies, strobe and clock signals.

III. EXPERIMENT RESULTS

There were 180 tests gathered in this study; 30 tests each for the fonts Times New Roman, Arial, and Futura for word-per-word mode and continuous mode. Checking was done manually by letting the respondents spell out what they read. The accuracy of the system was measured by using equations (1) and (2).

$$\%\,Accuracy = \frac{Number\ of\ correct\ words}{Total\ number\ of\ words} x100\% \quad (1)$$

$$Average\ Accuracy = \frac{\sum Accuracy\ of\ each\ test}{Total\ number\ of\ tests} \quad (2)$$

From Fig. 2, the minimum number of correct words for font Times New Roman is 242, which corresponds to an accuracy of 93.43%, while the maximum number of correct words is 255, giving an accuracy of 98.46%. The average accuracy for the font Times New Roman for the Word-per-Word mode is 96.31%. As for the font Arial, as shown in Fig. 3, the minimum number of correct words is 247, thus having an accuracy of 95.37%. The maximum number of correct words on the other hand is 255, which corresponds to an accuracy of 98.46%. The average accuracy of 30 tests from the font Arial in the Word-per-word mode is 97.17%. Finally, for the font Futura in Word-per-word

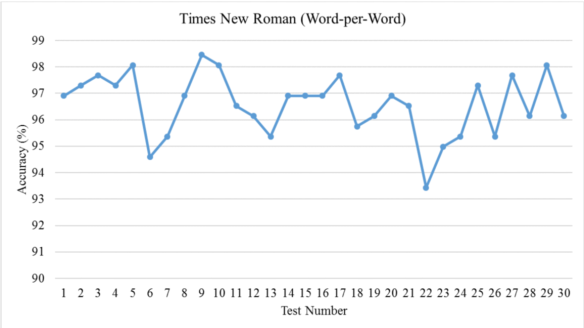

Fig. 2. Accuracy per Test for Times New Roman in Word-per-word Mode

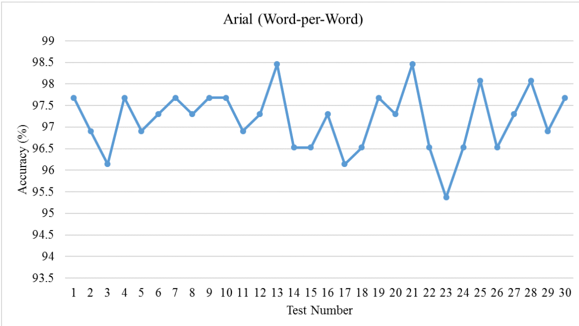

Fig. 3. Accuracy per Test for Arial in Word-per-word Mode

mode, as shown in Fig. 4, the minimum number of correct words is 238, having an accuracy of 91.89%. As for the maximum number of correct words, the result gathered is 253 correct words, with an accuracy of 97.68%. The average accuracy of 30 tests from the Futura font style in Word-per-word mode is 96.18%. The average accuracy for the Word-per-word mode is 96.55%. As for the continuous mode, the minimum number of correct words for the font Times New Roman, as shown in Fig. 5, is 238, thus having an accuracy of 91.89%. The most accurate test consisted of 258 correct words, thus having an accuracy of 99.61%. For the font Arial on the other hand, as shown in Fig. 6, the least accurate test result consists of 222 correct number of words, with an accuracy of 85.71%. The maximum number of correct words is 252 with an accuracy of 97.30%. The average accuracy for the font Arial in continuous mode is 93.72%. Lastly, for the font Futura, as shown in Fig. 7, the least accurate result consisted of 235 correct words, having an accuracy of 90.73%. The most accurate result on the other hand consists of 253 correct words, with an accuracy of 97.68%. The average accuracy for the font Futura in continuous mode is 94.74%. For the 90 tests conducted for the continuous mode, the average accuracy is 94.81%.

From the series of tests gathered in this study, it was observed that the sources of errors mainly came from the processing of image in the OCR. The differences in the output of this processing is caused by the recognition level of the OCR. External factors such as lighting, angle of the camera as it captures images, and orientation of the paper being scanned definitely has an effect to the image processing of the OCR. In addition, the font used in the test materials have different accuracy.Taking the average of the accuracy of each test, it was proven that the font Times New Roman has the most accurate output, while font Futura has the least accuracy.

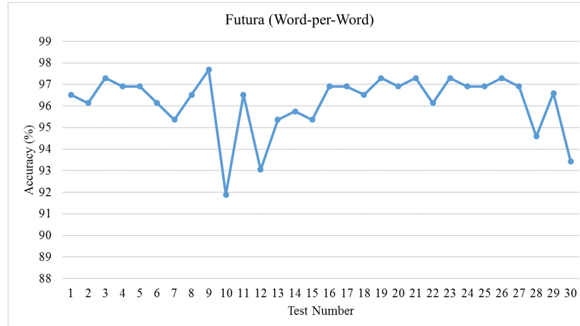

Fig. 4. Accuracy per Test for Futura in Word-per-word Mode

On the other hand, the microcontroller, as discussed in the previous sections, simply drives the braille cells based from what is given by the computer. The microcontroller compares the ASCII characters and then matches it with the corresponding braille binary equivalent output. In addition, the functionality of the word-per-word mode of reading is the same with the continuous mode of reading. The only difference is that the continuous mode automatically signals the computer to give the next word. Hence, the correctness of the output of the braille cells for any mode of reading will only depend on what is given by the computer. Therefore, it can be said that there are no errors caused by the microcontroller. When the ASCII character recognized by the OCR is incorrect and is not within the data matrix of the braille binary equivalent, the microcontroller simply outputs 0 to all 8 dots of a braille cell.

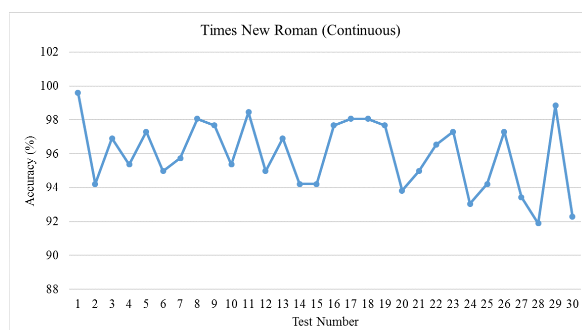

Fig. 5. Accuracy per Test for Times New Roman in Continuous Mode

The limitations of the system were obtained by capturing images of the test materials used from the earlier series of tests using different camera angles and heights. The results are summarized in a graph as seen in Fig. 8, Fig. 9, and Fig. 10, for the font styles Times New Roman, Arial, and Futura, respectively. As the camera is placed farther away from the reading paraphernalia, the result becomes less accurate. The cause of the

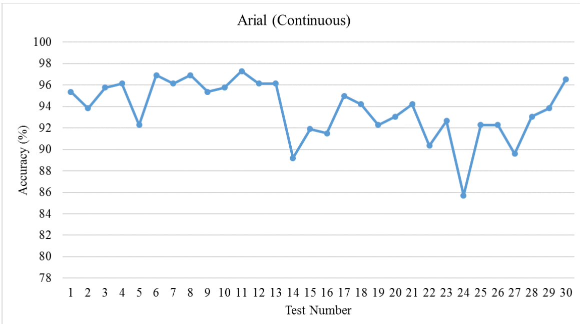

Fig. 6. Accuracy per Test for Arial in Continuous Mode

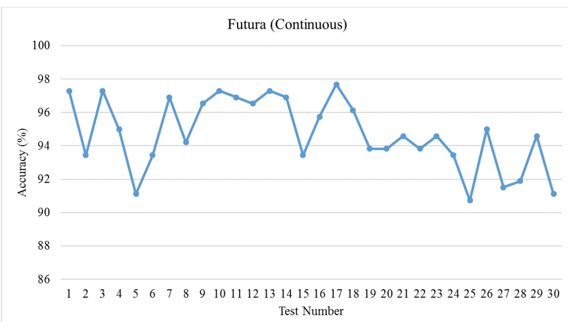

Fig. 7. Accuracy per Test for Futura in Continuous Mode

degrading accuracy is that the OCR is being fed by images with font sizes relatively smaller than font size 12, that it could not recognize the characters. Another factor that adds to the decline of accuracy, is the camera angle or tilt. Ranges below 20 cm would mark lower accuracies as the camera angle is increased. Particularly at a camera angle of ±20˚, the accuracies obtained are exceedingly lower compared to accuracies captured from a camera angle of 0˚. However, accuracies obtained from the height of 25cm and 30cm had an increase as the camera angle increased. This was because, at this height, as the camera is tilted, it tends to capture the half of the page with a closer distance. Hence, being able to recognize some of the characters that were captured more closely. Lastly, the optimal height for capturing images with a letter sized page would be in ranges within 15 to 20cm as at this height, the whole page of an 8" x 11" sized paper is captured.

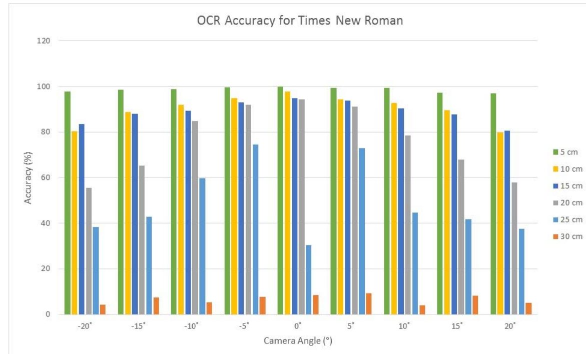

Fig. 8. OCR Accuracy with respect to Varying Height and Camera Angle for Times New Roman

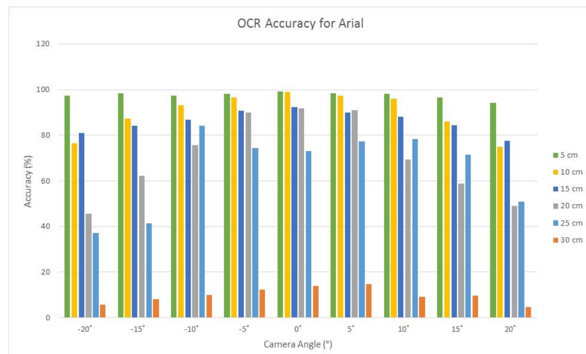

Fig. 9. OCR Accuracy with respect to Varying Height and Camera Angle for Arial

IV. CONCLUSION

Despite the existence of different laws, organizations, and movements to help the visually-impaired people to cope up with this time and age, there is still a hindrance for visually-impaired people to have a full accessible education due to limitations of resources. Also, the solution presented by this study is better than the similar and existing studies because the development of this system is definitely more modernized or smart compared to the existing ones. This system will give the visually-impaired people the chance to read any printed reading paraphernalia.

The system is able to achieve an overall system speed of 1 word in 2 seconds, and an overall system reliability of 95.68%. It is easy to use but is not stand-alone. The system is efficient both in character recognition and in actuation of braille cells, making it a good alternative to braille printed books.

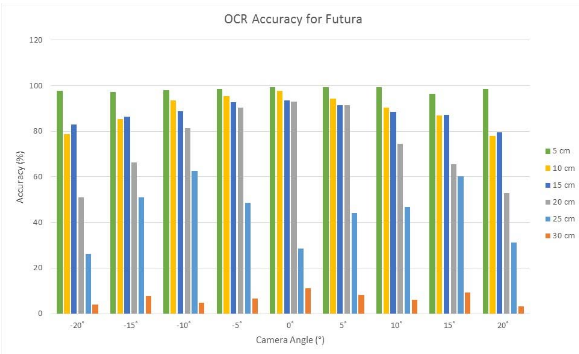

Fig. 10. OCR Accuracy with respect to Varying Height and Camera Angle for Futura

REFERENCES

[1] C. Barnes, "The Sociology of Disability and the Struggle for Inclusive Education," *British Journal of Sociology of Education,* pp. 603-619, September 2010.

[2] R. Sarkar, "A Low Cost Microelectromechanical Braille for Blind People to Communicate with Blind or Dead People Through SMS Subsystem," in *3rd IEEE International Advance Computing Conference*, Ghaziabad, India, 2013.

[3] J. Bernardo, D. Cawa, A. Chan and E. Galfo, "Text to Braille," De La Salle University, Manila City, 1996.

[4] R. Lienhart and F. Stuber, "Automatic Text Recognition in Digital Videos," in *Image and Video Processing IV*, San Jose, CA, 1996.

[5] L. Nian-feng and W. Li-rong, "A Kind of Braille Paper Automatic Marking System," in *International Conference on Mechatronic Science, Electric Engineering and Computer (MEC)*, Jilin, China, 2011.

[6] A. Ohkura, D. Deguchi, T. Takahashi, I. Ide and H. Murase, " Low-resolution Character Recognition by Video-based Super-Resolution," in *10th International Conference on Document Analysis and Recognition*, Beijing, China, 2009.

[7] A. Africa, S. Chua, L. D. Ansen, J. Gregorio and M. Quilnet, "Two-way SMS and Braille Communication for the Visually Impaired," De La Salle University, Manila, 2003.

[8] L. Balanag, J. Go, A. Gonzales, M. Romero and C. Sandoval, "PC-Based Videoke for the Blind," De La Salle University, Manila, 2006.

[9] F. Ramirez-Garibay, C. Olivarria, A. Aguilera and J. Huegel, "MyVox-Device for the Communication Between People: Blind, Deaf, Deaf-Blind and Unimpaired," in *IEEE Global Humanitarian Technology Conference (GHTC)*, USA Washington, DC, 2014.

[10] R. Smith, "An Overview of the Tesseract OCR Engine," in *9th International Conference on Document Analysis and Recognition (ICDAR)*, Parana, Brazil, 2007.

[11] L. S. Bartolome, A. A. Bandala, C. Llorente and E. P. Dadios, "Vehicle Parking Inventory System Utilizing Image Recognition Through Artificial Neural Networks," in *IEEE Region 10 Conference 2012 (TENCON 2012)*, Cebu City, Philippines, 2012.

[12] J. P. N. Cruz, M. L. Dimaala, L. G. L. Francisco, E. J. S. Franco, A. A. Bandala and E. P. Dadios, "Object Recognition and Detection by Shape and Color Pattern Recognition Utilizing Artificial Neural Networks," in *International Conference of Information and Communication Technology 2013 (ICoICT 2013)*, Bandung Indonesia, 2013.