# An Alternative Reading Eye for The Visual Impaired

Shih-Lun Chen[1], Hsin-Ju Tsai[2], Chiao-Yi Chien[3], and Ting-Lan Lin[4]

*Department of Electronic Engineer, Chung Yuan Christian University, Taiwan*
*200 Chung Pei Rd., Chung Li City, Taiwan, R.O.C.*

[1]chrischen@cycu.edu.tw
[2]plaisirgirl@hotmail.com
[3]winnie19910929@gamil.com
[4]tinglan@cycu.edu.tw

*Abstract*—**In this generation, everyone has the right to receive real-time and great amount of information. However, it is difficult for the visually impaired people to have the benefit of using this great information. In this paper, a reading system is proposed as an alternative eye for the visual impaired people. By this reading system, the information of words is transformed into the information of voice. It helps the visually impaired people to get all kinds of information quickly and easily. This system was developed based on the image processing technique, by which the captured image is converted into the text information by a character recognition technique. In order to achieve the demands of fast computing, multiple texts identification, huge database accessing, and real-time application, the very large scale integration (VLSI) technology was used to improve the performance of this system. This design was implemented by an Altera DE2-115a field programmable gate array (FPGA) board with a 5 Mega Pixel lens to capture the image of the book. For the future development, this system will be connected with the cloud systems, by which the recognized words can be translated into voice information by the cloud computing and tools. We hope this system can help the visually impaired people to get information from the books and papers independently.**

***Keywords*—FPGA; image segmentation; image zooming; text detection; visual impaired people; VLSI; word identification**

## I. INTRODUCTION

Recently, people can get a great amount of information from Internet, books, new papers, etc. Nevertheless, for the visually impaired people, they have difficulty in reading and have fewer resources compare to others. Although the Braille books and audio books are available for them to get the information by voice, the information contains on Braille and audio books are limited and it spends a lot of the manpower and time to develop these tools [1]. In other words, the visually impaired people needs somebody or tools to help them reading the books or papers.

Taipei Public Library [2], for example, buys 100 audio books per month which is obvious less than other kinds of books. Although the Taipei branch has a library for the blind which can provides a convenient method but they cannot provide up to date information to the blind readers. The information provided to them will be limited to the contents of the books in the library. Audio books are more expensive because it needs more manpower to build it. Moreover, since the poor eyesight of visually impaired people, they are unable to operate complex machines. Hence, a simple and easy to operate reading system was developed for the visually impaired people. This system is designed to be able to translate the text information into voice information, which will help the visually impaired people read books and papers independently. It also means that the visually impaired people will have more choices to reading more kinds of books and papers. They can also read independently and save much time for reading.

The proposed innovative system was developed based on the vehicle license plate recognition techniques [3]-[12]. It uses the device of lens to capture the images and then identifies the text information from the images as shown in Fig. 1. There are four steps to identity the words of the captured images. First, the noise of the captured images is reduced by the gray level, binarization, and the edge detection. Second, the figure of each letter is divided from the whole image. Third, the divided letter image is scaled up/down to the size the same as the size of the letter image in the database. Finally, the letter can be obtained by comparing the divided letter image with the database.

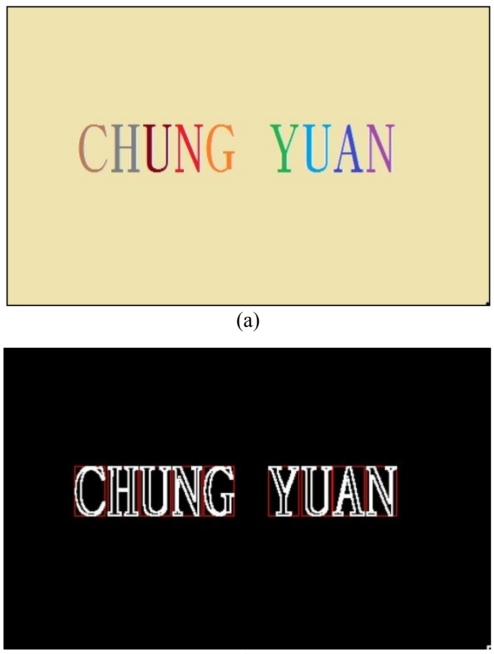



Fig. 1 (a) The original captured image. (b) The image processed by Example of an unacceptable low-resolution image

After identifying process, the captured image will be converted into text file. We are also planning to link the text file to Google to convert the text to audio. Audio books will be stored in the cloud and it provides a convenient tool for the visually impaired people to get the word information quickly. We hope to provide the visually impaired friends to have easier and faster way of reading the books and getting knowledge rather than just relying on the dictionary learning or volunteer services such as newspapers. Creating a combined technology and barrier-free living will provide the disabled groups more equal learning opportunities.

## II. SWOT Analysis and Future Outlook

The four indexes strength, weakness, opportunity, and threat (SWOT) were used to evaluate the feasibility of the proposed alternative eye system. The details of each index are described as follow.

### *A. Strength*

The concept of the system is to provide an alternative eye for the visually impaired people. It helps the visually impaired people have the ability to read independently. In comparison with the present products, this system has the characteristics of simple operating, easy obtaining, and high compatibility since all lenses can be compatible to this system.

### *B. Weakness*

The technology of recognition is widely used in the license plate recognition applications and it is similar to this system. Moreover, the accuracy of the recognition in this system is not yet better than the products in the market. To summarize, since the license plate recognition system have been developed for a long time, this system is missed the opportunity to be the first one.

### *C. Opportunity*

By using portable lens, the applications of this system are widespread. The applications are not limited to visually impaired people. It can be widely used for everybody. With the development of the cloud technology, the books will be translated into the voice by the cloud computing and stored into the virtual voice library in the cloud system. It will make the information more accessible and available for the human in the world.

### *D. Threat*

For quite some time, many countries in European had paid more attention to the disabled groups. The countries in North America had spent more resources in such kinds of the researches. Hence, to develop in this field, we need to face the threats from other countries.

This system is not only applicable in reading but also placed in the Google Glasses in the future. Through the lens, the road sign or signpost can be captured and then sent to cloud for identification. By cooperating with the Google map, it can also be used to recognize and then give the location information for the users. In addition, this system can also be a mobile guide dog to help the visually impaired people to find the right direction and distinguish of the busy stops by voice.

To develop this system, it is not only to help the visually impaired people but also contribute to the consumer products such as mobile phone, tablet PC and so on. It also offers an alternative way to get the information by ears when the eyes are overused or tired. The applications of this system will be widespread in the future.

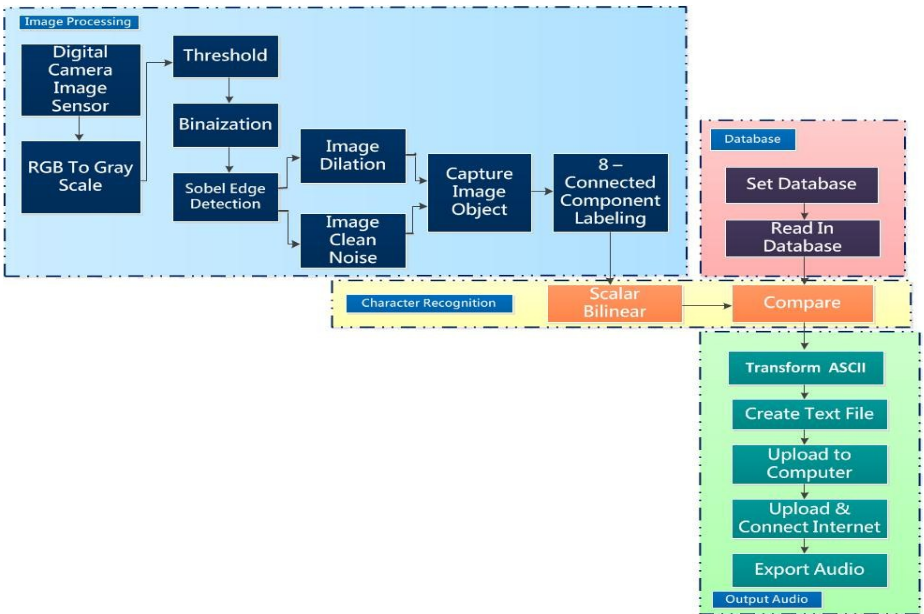



Fig. 2 Block diagram of the software structure

## III. SOFTWARE STRUCTURE

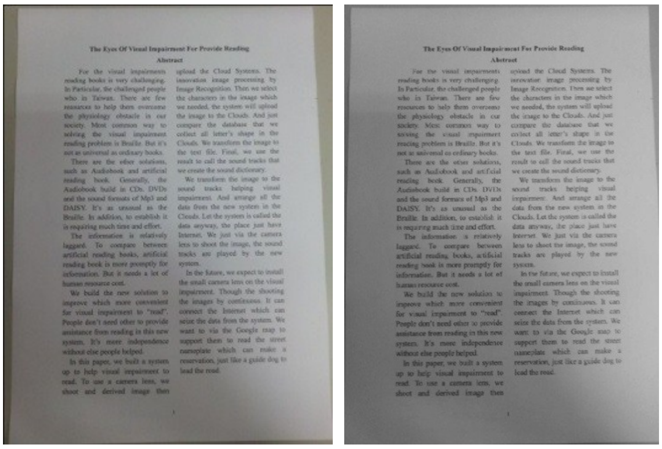
To implement this system, first we used a software environment to develop the image processing algorithms. Second, some partitions which have the characteristics of complex computing and timing limitation were replaced by hardware designs. Finally, this system was completed as the hardware and software co-designed system. The hardware structure will be introduced in the next section. Fig. 2 shows the block diagram of the software structure in this system. The details of each step in the software structure are described as follow.

(a) Grayscale Translation

The system used D5M 5 Mega Pixel Digital Camera to capture the image. The captured image was stored in the memory. In order to decrease the obstruction of the backdrop in the image, the captured image was translated from Red, Green, and Blue (RGB) to the gray scale format. The grayscale can be obtained by

$$Gray= 0.299*R+ 0.587*G + 0.114*B \qquad (1)$$

where R, G, and B are the values of red, green, and blue color in the captured image. Fig. 3 (a) shows the original captured image in RGB format. After grayscale translation, the RGB format image was translated into grayscale as shown in Fig. 3 (b). It is easy to obtain that the image in grayscale format is clearer than that in RGB format.




(a) (b)

Fig. 3 (a) The original image in RGB format. (b) The translated image in the grayscale format.

(b) Binarization

Before doing the binarization process, it is necessary to find the threshold of the image. Next, the binarization process can be done according to the relations between the values of the pixels with the threshold. The threshold is used to separate each pixel into 0 when the value of the graylevel is less than the threshold as well as 255 when the value of the graylevel is more than the threshold, as shown in Fig. 4. After this process, the image is translated into black and white image which is better to highlight the letters. Fig. 5 shows the distribution of binarization results for each pixel in whole image. Fig. 6 (a) shows the real image results after binarization. To compare with the images in Fig. 3, the image after binarization has less noise, which provides a well base for letter identification.

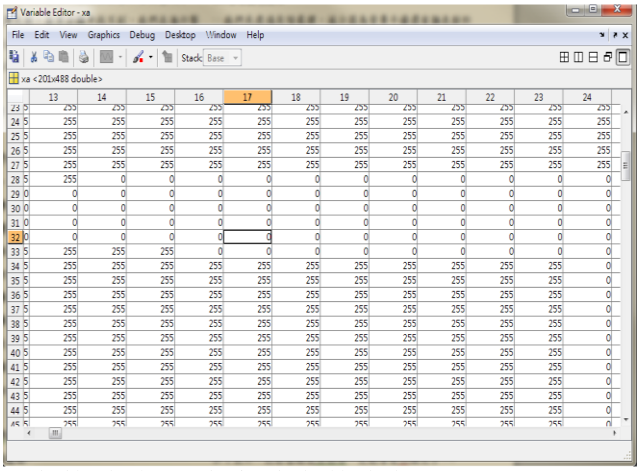



Fig. 4 The binarization result of each pixel in the image.

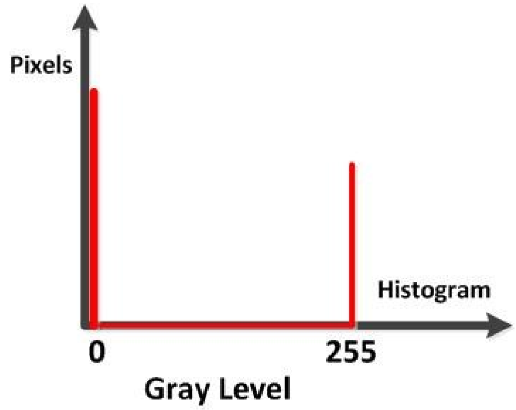
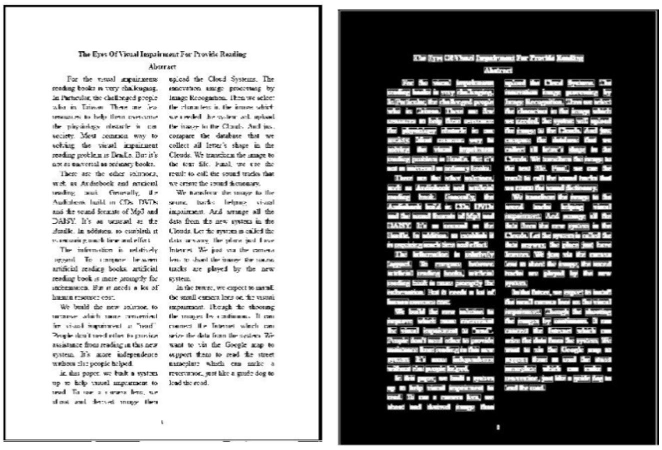



Fig. 5 The distribution of the results after binarization




(a) (b)

Fig. 6 (a) Real image results after binarization. (b) Real image results after Sobel edge detection

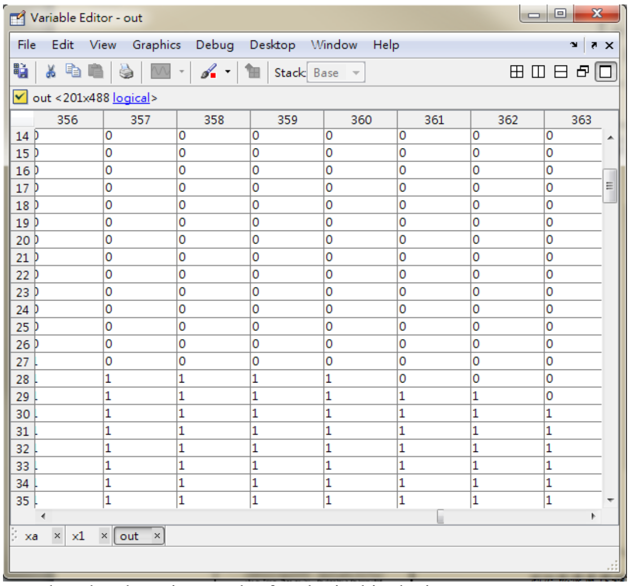
(c) Edge Detection

The Sobel edge detection algorithm was used to find the edge in the image. It detects the edges by calculating the vertical and horizontal gradients first. And then summaries the absolute values of the values of gradients in both vertical and horizontal directions. The horizontal gradient ($G_x$), vertical gradient ($G_y$), and total gradient (G) can be obtained as

$$Sx = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * F \qquad (2)$$

$$Sy = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * F \qquad (3)$$

$$S = \sqrt{Sx^2 + Sy^2} \qquad (4)$$




| | 356 | 357 | 358 | 359 | 360 | 361 | 362 | 363 |
|---|---|---|---|---|---|---|---|---|
| 14 | [illegible] | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 15 | [illegible] | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16 | [illegible] | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 | [illegible] | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 18 | [illegible] | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 19 | [illegible] | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 20 | [illegible] | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 21 | [illegible] | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 22 | [illegible] | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 23 | [illegible] | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 24 | [illegible] | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 25 | [illegible] | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 26 | [illegible] | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 27 | [illegible] | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 28 | [illegible] | 1 | 1 | 1 | 1 | 0 | 0 | 0 |
| 29 | [illegible] | 1 | 1 | 1 | 1 | 1 | 1 | 0 |
| 30 | [illegible] | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 31 | [illegible] | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 32 | [illegible] | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 33 | [illegible] | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 34 | [illegible] | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 35 | [illegible] | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

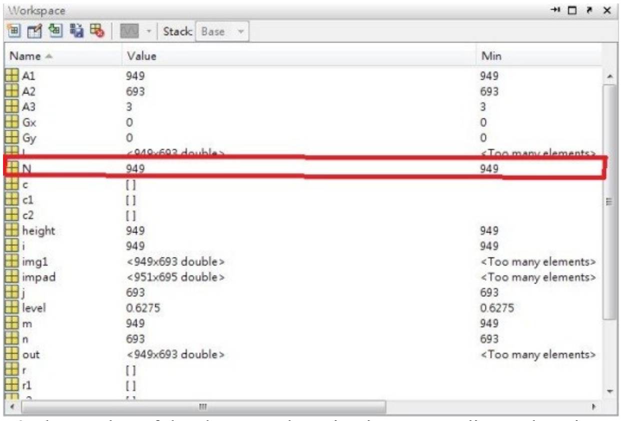
Fig. 7 The edge detection result of each pixel in the image.

Fig. 7 shows the results after the edge detecting. Each pixel can be separated into the value of 0 or 1 according to find the edge information in each pixel or not. Fig. 6 (b) shows the real image results after the edge detection.

(d) Image rounding and Cutting

After finding the edge in the image, we can use the edge information to detect the graphic and box of each letter. First, we used the 8-connected regions algorithm to detect the boundaries of the characters. Second, the characters were rounded by the circles. Finally, the graphic of each figure was cut from the image according the rounded circle of each character. Fig. 8 shows the real image results by using the 8-connected regions algorithm and image rounding technique. By using image rounding and cutting techniques, the number of the characters can be obtained as shown in Fig. 9.

Fig. 8 Real image results after image rounding




| Name | Value | Min |
|---|---|---|
| A1 | 949 | 949 |
| A2 | 693 | 693 |
| A3 | 3 | 3 |
| Gx | 0 | 0 |
| Gy | 0 | 0 |
| I | <949x693 double> | <Too many elements> |
| N | 949 | 949 |
| c | [] | |
| c1 | [] | |
| c2 | [] | |
| height | 949 | 949 |
| i | 949 | 949 |
| img1 | <949x693 double> | <Too many elements> |
| impad | <951x695 double> | <Too many elements> |
| j | 693 | 693 |
| level | 0.6275 | 0.6275 |
| m | 949 | 949 |
| n | 693 | 693 |
| out | <949x693 double> | <Too many elements> |
| r | [] | |
| r1 | [] | |

Fig. 9 The number of the characters by using image rounding and cutting techniques.

(e)Image Scaling

Since the size of the cut character image is different with the size of that in the database. The image size of cut character image should be scaled into the size the same as that in the database. The bilinear interpolation algorithm was selected as the scaling algorithm for this system since it has the benefits of low complexity and high performance [13]. Fig. 10 shows the principle of the bilinear interpolation. The interpolated result P'(x, y) can be calculated by linear interpolation in the horizontal directly first and then linear interpolation again in the vertical direction. The values of P'(x, y) can be calculated by equations (5), (6), (7), and (8).

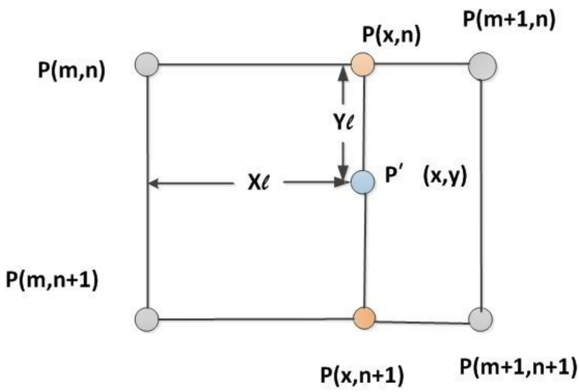



Fig. 10 Bilinear interpolation.

$$P(x, n) = (1-dl)\times P(m, n) + dl\times P(m+1, n) \quad (5)$$

$$P(x, n+1) = (1-dl)\times P(m, n+1) + dl\times P(m+1, n+1) \quad (6)$$

$$P'(x, y)=(1-yl)\times P(x, n) + yl\times P(x, n+1) \quad (7)$$

$$P'(x, y) = (1-dl)\times(1-yl)\times P(m, n) + (1-dl)\times yl\times P(m, n+1) + dl\times(1-yl)\times P(m+1, n) + xl\times yl\times P(m+1, n+1) \quad (8)$$

(f) Library construction and database comparison

There is a difference between the length and width of every figure of characters. In order to improve the accuracy of identification, a database of each letter was created and set up. To compare the figure of character with the database, first it will compare the length and width of the figure. If there is no recognition result by the length and width comparing, we will compare the figure of the character with each of character figure in the database. To compare length and width first can greatly reduce the comparison time and increase the efficiency. Fig. 11 shows the length and width of each character figure stored in the database. The figure of each character in the database was illustrated in Fig. 12.

| | 寬度 | 長度 | | | 寬度 | 長度 | |
|---|---|---|---|---|---|---|---|
| 0 | 280 | 450 | 1.607143 | | | | |
| 1 | 219 | 440 | 2.009132 | | | | |
| 2 | 270 | 440 | 1.62963 | | | | |
| 3 | 270 | 440 | 1.62963 | | | | |
| 4 | 270 | 440 | 1.62963 | | | | |
| 5 | 270 | 440 | 1.62963 | | | | |
| 6 | 280 | 450 | 1.607143 | | | | |
| 7 | 270 | 430 | 1.592593 | | | | |
| 8 | 282 | 450 | 1.595745 | | | | |
| 9 | 280 | 450 | 1.607143 | | | | |
| A | 440 | 440 | 1 | a | 270 | 310 | 1.1481481 |
| B | 400 | 430 | 1.075 | b | 310 | 440 | 1.4193548 |
| C | 392 | 450 | 1.147959 | c | 238 | 308 | 1.2941176 |
| D | 418 | 430 | 1.028708 | d | 318 | 440 | 1.3836478 |
| E | 388 | 440 | 1.134021 | e | 250 | 310 | 1.24 |
| F | 359 | 440 | 1.225627 | f | 250 | 440 | 1.76 |
| G | 460 | 450 | 0.978261 | g | 290 | 440 | 1.5172414 |
| H | 478 | 440 | 0.920502 | h | 318 | 430 | 1.3522013 |
| I | 220 | 440 | 2 | i | 160 | 290 | 1.8125 |
| J | 312 | 440 | | j | 170 | 430 | 2.5294118 |
| K | 488 | 428 | 0.877049 | k | 330 | 432 | 1.3090909 |
| L | 400 | 430 | 1.075 | l | 158 | 428 | 2.7088608 |
| M | 588 | 430 | 0.731293 | m | 510 | 298 | 0.5843137 |
| N | 434 | 430 | 0.990783 | n | 318 | 298 | 0.9371069 |
| O | 450 | 450 | 1 | o | 278 | 310 | 1.1151079 |
| P | 360 | 428 | 1.188889 | p | 318 | 436 | 1.3710692 |
| Q | 538 | 450 | 0.836431 | q | 320 | 440 | 1.375 |
| R | 460 | 430 | 0.934783 | r | 250 | 298 | 1.1481481 |
| S | 302 | 450 | 1.490066 | s | 200 | 314 | 1.57 |
| T | 390 | 440 | 1.128205 | t | 198 | 410 | 2.0707071 |
| U | 428 | 440 | 1.028037 | u | 320 | 302 | 0.94375 |
| V | 440 | 440 | 1 | v | 298 | 288 | 0.966443 |
| W | 618 | 438 | 0.708738 | w | 462 | 298 | 0.6450216 |
| X | 430 | 430 | 1 | x | 278 | 288 | 1.0359712 |
| Y | 440 | 428 | 0.972727 | y | 310 | 424 | 1.3677419 |
| Z | 400 | 430 | 1.075 | z | 270 | 290 | 1.0740741 |

Fig. 11 The length and width of the character figure in the database.

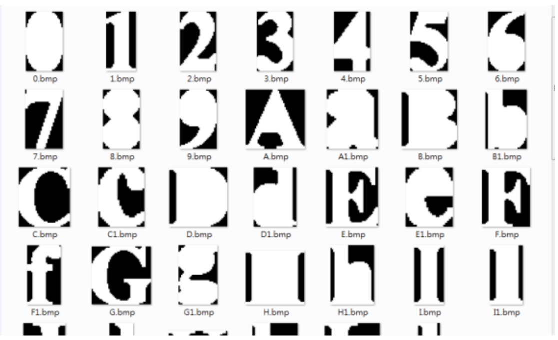
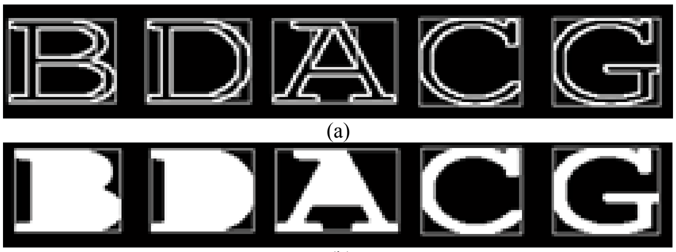


Fig. 12 The figure of the character in the database.

(g) Image Correction

After rounding the characters, some of the character figures include two or more blocks. As shown in Fig. 13 (a), there are two blocks rounded the third letter "A". It will cause the mistake when we want to compare the figure with the database. In order to avoid this mistake, an expansion technique was used for image correction. After expansion, the character was filled as shown in Fig. 13 (b). After image correction, the figure of the character can be stored in the memory and then compared with figures in the database for identification. Finally, the recognition results are exported into a text file stored in the computers. Fig. 14 shows the flow chart of the identification process in this system.



(a)

(b)

Fig. 13 (a) The character images before expansion (b) The character images after expansion

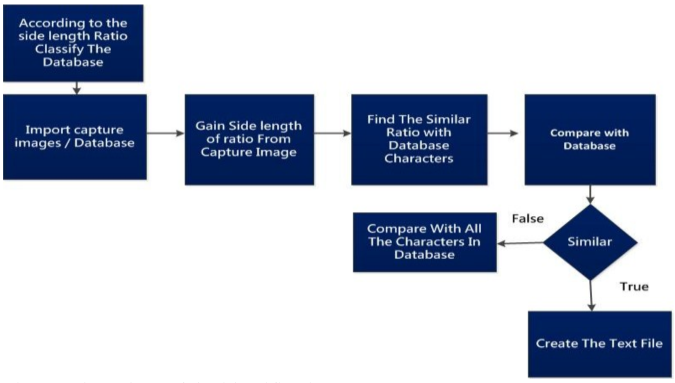



Fig. 14 Flow chart of the identification process

## IV. Hardware Architecture

The hardware partition of this system was implemented by the Altera DE2-115a field programmable gate array (FPGA) development board. Fig. 15 shows the block diagram of the hardware architecture in this system. The image processor includes a RGB to Y circuit, a binarization module, an edge detector, an object detector, a scalar, and a comparator. It also includes two SDRAM devices SDRAM1 and SDRAM2. The SDRAM1 stores the captured image in RGB format. The SDRAM2 stores the database of the character figures.

The input image was captured by the 5 Mega pixels digital camera built in the FPGA development board. The image can be captured by this camera and then stored in the SDRAM1. Two memory controllers were designed for communicating between the SDRAM and the image processor. The image processor can read the RGB data stored in the SDRAM1 or database stored in the SDRAM2 through the memory controllers.

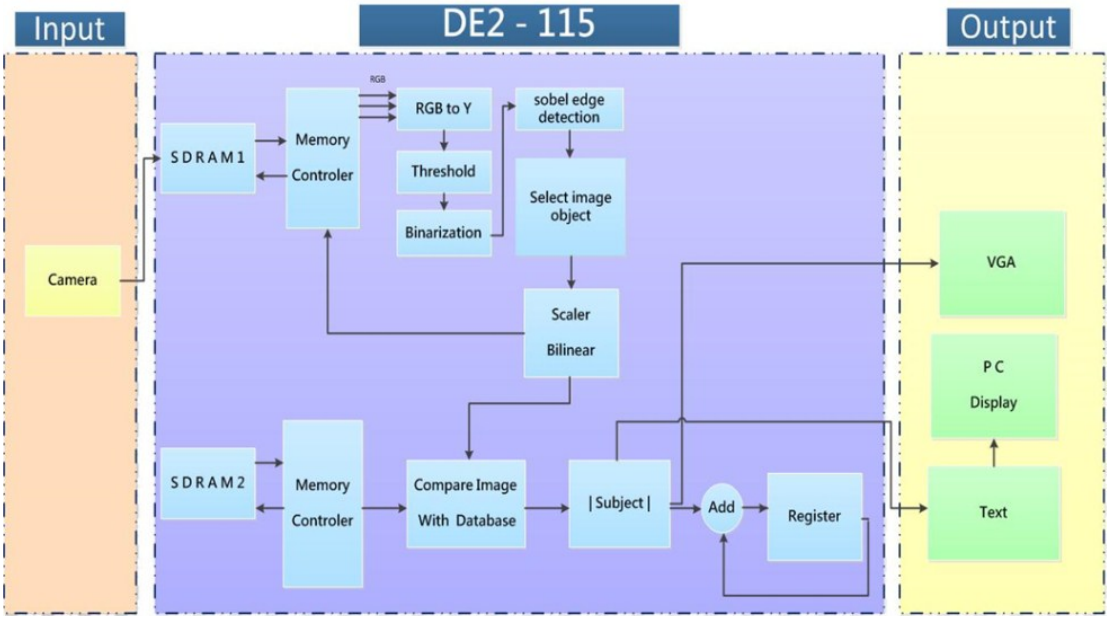



Fig. 15 block diagram of the hardware architecture of this system.

The image processor read the RGB image data from the SDRAM1 first and then the image data were transferred to the grayscale format by a RGB to Y circuit. Second, the grayscale data were processed by the binarization module, edge detector, object detector, and scalar. After processed by these four circuits, the figures of characters are produced. To compare with the database stored in SDRAM2 through the comparator, the text information is produced. The text information is sent to computer and stored as a text file. Finally, the sound files are broadcasted by the software in the computer.

## V. Design Features

Reading is a very important way to get the knowledge. Reading by braille books is not enough to satisfy the blind people because it spends more money and cost more manpower to publish the book. Although there are a wide variety of different products available for them, the information contents are still very limited and not up to the date. The feature of this design is easily operating for the visually impaired people. It provides an easy way to read the contents of any book or paper by listening immediately. By the pure software system, the processing time for a whole image takes about 7 seconds, which makes it impossible to achieve the immediate identification and fluency from image to audio information immediately. By the hardware design with the DE2 FPGA development board, it make possible to identify the rate within 0.1 second when processes one whole image. We hope this design is not only a simple recognition but also an accessible system for the human.

## VI. Conclusions

Several months ago, we heard about many issues concerning disabled people. Hence, we decided to start this research and hoped to do some contribution for the visual impaired people. Although this system is not complete enough for using immediately, we will keep going to develop this paper better and better. We hope the visually impaired people can enjoy their reading without any obstacles by this system in the future.

## Acknowledgment

This work was supported by the National Science Council, R.O.C, under Grant numbers of NSC-101-2221-E-033-036, NSC-101-2221-E-033-066, NSC-102-2221-E-033-018, NSC-102-2221-E-033-063 and NSC-102-2622-E-033-009-CC2 and the National Chip Implementation Center, Taiwan. The authors are also grateful to ALTERA and Terasic companies to hold such a good FPGA design competition for us.

## References

[1] (2013) Wiki website. [Online]. Available: http:// zh.wikipedia.org/wiki/%E8%A6%96%E9%9A%9C%E8%AE%80%E8%80%85%E6%9C%8D%E5%8B%99
[2] (2013) [Online]. Available:http://www.tpml.edu.tw/
[3] Sheng-Long Lee, "Using motion information and multiple classifiers in license plate recognition system for moving vehicles on the road," M. Eng. thesis, Department of Electronic Engineer, Chung Yuan Christian University, Taiwan, Jun. 2005.
[4] Y. P. Huang, S. Y. Lai, and W. P. Chuang, "A template-based model for license plate recognition," *Proceedings of IEEE International Conference on Networking, Sensing and Control*, Vol. 2, pp. 737-742, 2004.
[5] C. N. E. Anagnostpoulos, I. E. Anagnostpoulos, I. D. Psoroulas, V. Loumos, and E. Kayafas, "License plate recognition from still image and video sequences: a survey," *IEEE Transactions on Intelligent Transportation System*, Vol. 9, No. 3, pp. 377-391, 2008.
[6] S. L. Chang; L. S. Chen, Y. C. Chung, S. W. Chen, Automatic license pl ate recognition," IEEE Trans. on Intelligent Transportation Systems, vol. 5, pp. 42-53, 2004.
[7] A. Taleb-Ahmed, D. Hamad, and G. Tilmant, "Vehicle license plate recognition in marketing application," Proceedings of IEEE Intelligent Vehicles Symposium, pp. 90-94, 2003.
[8] V. Koval, V. Turchenko, V. Kochan, A. Sachenko and G. Markowsky, "Smart license plate recognition system based onimage processing using neural network," *in Proc. The Second IEEE International Workshop on Intelligent Data Acquisition andAdvanced Computing Systems: Technology and Applications*, pp.123-127, Sep. 2003.
[9] C. A. Rahman, W. Badawy and A. Radmanesh, "A real time vehicle's license plate recognition system," *in Proc. IEEE Conference on*

*Advanced Video and Signal Based Surveillance*, pp.163-166, July, 2003.

[10] H. L. Bai, J. M. Zhu and C. P. Liu, "A fast license plate extraction method on complex background," *in Proc. IEEE Intelligent Transportation Systems*, vol. 2, pp.985-987, Oct. 2003.

[11] P. V. Suryanarayana, K. M. Suman, B. Asim and K. R. Anil, "A morphology based approach for car license plate extraction," *INDICON, 2005 Annual IEEE,* pp. 24-27, 2005.

[12] Chien-Ben Wang, "Digital Image Segmentation Techniques on. Comparison and Applications," M. Computer Science and Information, thesis, Cheng Kung University, Taiwan, Jul. 2013.

[13] Shih-Lun Chen, Hong-Yi Huang, and Ching-Hsing Luo, "A low-cost high-quality adaptive scalar for real-time multimedia applications," *IEEE Transaction on Circuits and Systems for Video Technology,* Vol. 21, no. 11 , pp. 1600-1611, Nov. 2011.