Camera-Captured Writing System Recognition of Logosyllabic Han Character

Maria Christina Aragon, Melissa Juanillo, and Rosmina Joy Cabauatan

Abstract-The study is focused on the recognition of Logosyllabic Han writing system such as Cantonese for Chinese, Hiragana and Katakana for Japanese and Hangul for Korean from a camera-captured image using image processing algorithms: Peripheral Direction Contributivity and 3-level matching, 2D Array and Recursion methods. In order to improve performance and to reduce memory usage of a mobile device, a remote web server has been used to communicate information from the user to the server and vice versa using Simple Object Access Protocol. Experiments were conducted to test the accuracy of the application such as capturing the same image several times, checking the percentage of the features of the text in the captured image, and capturing text from images in different circumstances such as low lighting and skewed images. The overall result showed the application has 80-100% accuracy in identifying Chinese, Japanese, and Korean writing systems.

Index Terms—Image processing, logosyllabic han character, optical character recognition, peripheral direction contributivity.

I. INTRODUCTION

The field of handwriting recognition has gained numerous benefits from the research in Optical Character Recognition (OCR) technology. Handwriting technology makes use of a recognition process which focuses on the dynamic motion created during the handwriting process. With the research in OCR technology as a basis, some applications have become successful in interpreting hand gestures. Implementation of this technology can be seen in most Tablet PC machines, as well as PDAs and some mobile phone devices. It is expected that the continued research in OCR technology in relation to this field will lead to advancements in Intelligent Character Recognition and real time character recognition implementations. At present, research in OCR technology has successfully implemented printed text recognition for some commercial devices in the market [1], [2].

Over the past decades, researches about Optical Character Recognition have been rampantly conducted on text, handwriting and font recognition systems. However, such research only focuses on the recognition of the font type, styles and are only limited to one writing system.

Han writing systems originated in China and are adopted by the Koreans and Japanese in the 13th century [3]. The Koreans borrowed a huge number of Chinese words, gave Korean readings and/or meanings to some of the Chinese characters and also invented about 150 new characters, most of which are rare or used mainly for personal or place names until 1444-1446 AD where the Korean alphabet was invented [4], [5]. On the other hand, the Japanese started to learn Classical Chinese during 7th and 8th Century. Over time, Japanese writing systems have evolved and developed two systems namely Hiragana to be used for words without kanji representation and, Katakana which are used to write foreign words. Kanji writing system is usually used by both the Chinese and Japanese.

Logosyllabic Han characters such as in Chinese, Japanese and Korean writing systems are not common to people who were only familiar with western writing system. They are closely identical in features and in strokes. People may have the tendency of interchanging the characters using their own perception. This type of dilemma occurs when a tourist or consumer see these writing systems in a product or in a certain place but does not have an idea where the characters or writing system originated from. By looking at characters written on an imported product, though unreadable, the consumer will have an idea where these products came from.

The objective of this study is to help users recognize Logosyllabic Han characters using sets of theories and algorithms from a camera-captured image. Combining the knowledge in these languages and current technologies, this study utilizes a mobile phone to capture the Han character and develop a mobile application that is capable of identifying the type of writing system used, whether it is written in Chinese, Japanese or Korean. This study may also serve as an input to multi-language translation researches such as Chinese to English language translation and the like.

II. ALGORITHMS OF THE STUDY

The set of algorithms used and implemented are programmed in a Java server to minimize memory usage of the mobile device and using a web server to give the user the result of the query. The captured image is sent to the server where all the algorithms and processes are used in the identification process.

A. Image Processing

In image processing, Otsu's thresholding method (1979) is used for automatic binarization level decision, based on the shape of the histogram [6]. The algorithm assumes that the image is composed of two basic classes: Foreground and Background. It then computes an optimal threshold value that minimizes the weight within class variances of these two classes. It is mathematically proven that minimizing the intra-class variance is same as maximizing inter-class variance.

Manuscript received October 29, 2013; revised February 25, 2014. This work was supported in part by Technological Institute of the Philippines, Quezon City.

The authors are with Technological Institute of the Philippines, Quezon City, Philippines (e-mail: christina.aragon@tip.edu.ph, melissa_tipqc@yahoo.com, rjmcabauatan@ yahoo.com).

In image processing using Otsu's method, the captured image will be converted first into grayscale image. From a grayscale image, thresholding will be used to convert the current image to binary image. This method identifies the foreground and background of the image then calculate the optimum threshold separating those two classes so that their combined spread is minimal. In this process, it will identify the prominent color in the image, this color is assigned a value of 255 (intensity value of white) and is identified as the background and the rest are assigned a value of 0 (black) and are identified as the foreground. Fig. 1 shows the binarization method that is implemented on the image. In (A) the portion is the original image and in (B), grayscale. In Otsu's method it does exhaustively search for the threshold that minimizes the intra-class variance (the variance within the class), defined as a weighted sum of variances of the two classes:

$$\sigma_{w}^{2}(t) = w_{1}(t)\sigma_{1}^{2}(t) + w_{2}(t)\sigma_{2}^{2}(t)$$

Weights w_i are the probabilities of the two classes separated by a threshold t and σ_i^2 variances of these classes.

Otsu [7] shows that minimizing the intra-class variance is the same as maximizing inter-class variance:

$$\sigma_b^2(t) = \sigma^2 - \sigma_w^2(t) = w_1(t)w_2(t)[\mu_1(t) - \mu_1(t)]^2$$

Which is expressed in terms of class probabilities w_i and class means μ_i .

The class probability $w_1(t)$ is computed from the histogram as t:

$$w_1(t) = \sum_{0}^{t} p(i)$$

While the class mean $\mu_1(t)$ is:

$$\mu_1(t) = \sum_{0}^{t} p(i) x(i)$$

where x(i) is the value at the center of the *i*th histogram bin.

Similarly, $w_2(t)$ and μ_t can be computed on the right-hand side of the histogram for bins greater than *t*.

The class probabilities and class means can be computed iteratively. This idea yields an effective algorithm.

Otsu's Algorithm:

- 1) Compute histogram and probabilities of each intensity level
- 2) Set up initial $w_i(0)$ and $\mu_i(0)$
- 3) Step through all possible thresholds t = 1... maximum intensity
 - Update W_i and μ_i
 - Compute $\sigma_b^2(t)$
- 4) Desired threshold corresponds to the maximum $\sigma_b^2(t)$
- 5) Compute two maxima (and two corresponding thresholds). $\sigma_{b_1}^2(t)$ is the greater max and $\sigma_{b_2}^2(t)$ is the greater or equal maximum

6) Desired threshold =
$$\frac{threshold_1 + threshold_2}{2}$$



Fig. 1. Image processing using Otsu's thresholding.

B. Peripheral Direction Contributivity (PDC)

Peripheral Direction Contributivity (PDC) is an efficient technique for Chinese character recognition [8]. PDC checks the strokes using eight direction arrows which is the process for determining the strokes used in the text and image binarization converts the image to a computer readable data with 0's for white and 1's for black. For different text images, the sizes of characters are usually not uniform. Peripheral Direction Contributivity (PDC) will play an important role in the development of the application. The spaces between characters and text lines in a given text image may occur which will seriously affect the character recognition result. PDC checks the 8 directions of the image as shown in Fig. 2.



Fig. 2. Peripheral direction contributivity (PDC).

C. 3-Level Matching

Next, three levels of matching are performed to recognize characters in bitmap. In this phase, the program will check the grid sets in 0's or 1's. Using this method, the application determines the strokes of the image of the writing system and checks the given parameters to properly and accurately give the results when parameters are met.



Fig. 3. PDC and 3-level matching interaction.

Fig. 3 illustrates how PDC and 3-level matching interacts with each other. After PDC checks all the 1's and 0's in the image, 3-level matching will now check the given parameters to evaluate the strokes of the captured image level of the identification.

The characteristics of the Chinese, Japanese, and Korean writing systems can be determined by their features [9]. For Korean characters, it is observed to have more lines and has loops. For Chinese characters, it is observed to have more

lines and intersection, referred to this paper as forks. For Japanese characters it is observed to have more lines and corner.

III. METHODS

A. Project Design

The design of the study is focused on the ability of the system to recognize Logosyllabic Han writing systems specifically Cantonese for Chinese characters, Katakana and Hiragana for Japanese and Hangul for Korean writing systems. Logosyllabic or complex script is a character which generally represents one syllable of spoken Chinese and may be a word on its own or a part of a polysyllabic word. Applying the fundamentals of image processing and optical character recognition, the application will be able to recognize the captured image whether it is Chinese, Japanese or Korean characters.

B. System Architecture

The diagram shows the System Architecture used in the application. (See Fig. 4).



Fig. 4. System architecture.

The application will start when a user opens the application from the mobile device and is ready for capturing images. Using an appropriate web service, the captured data will be sent to the server and will be processed and assessed. The following methods are applied to achieve the necessary results.

Otsu Method. The first part of Otsu's algorithm [10] is the conversion of the original image into grayscale image. In this process, the system converts the colors from 4 million colors or 32 bit to 32k or 16 bit then from 32k or 16bit to 256 colors or 8 and from 256 colors or 8 bit as shown in Figure 5. This method is used to optimize the process of the application thus eliminating all unnecessary colors before the recognition process. Otsu's algorithm uses a constant conversion of color to grayscale through linear combination, based on pure colorimetric considerations and the definition of luma coefficients HDTV according to Rec. 709 forms luma (Y') using R'G'B' coefficients 0.2126, 0.7152, and 0.0722.



Original Image Grayscale Image Fig. 5. Image conversion to grayscale.

In this part, the system gets the opacity and the color of the image multiplying the value of red to 0.21, green to 0.71 and blue to 0.07 using Computed below as:

 $RED = (red \times 0.21 + green \times 0.71 + blue \times 0.07)$

Setting all the pixel values of RGB to one value, in this portion, the system maintained Otsu's to red.

Otsu's Threshold gets the dominant color and setting it to 255 or white as background and the less dominant color to 0 or black as the foreground. Otsu's threshold assumes that the image has only two colors which is black or white. In Figure 6 the dominant color gray will be converted to white as the background and the darker color will be converted to black as the foreground.



The last and final stage of the image processing is the conversion of the grayscale image to binary image using Otsu's thresholding method above. Where gray is set to white as it assumes it is the background because it is the dominant color and black for the darker color for it assumes it is the foreground.

After image processing, the output image then go through the next process.

Peripheral Direction Contributivity (PDC). This phase of the application is where the program will check the image properties and features of the character. This phase will check the strokes, patterns and pixels that represent the distances from the eight directions to the point boundaries and point strokes defined by the application and is described as 8-dimentsion vector <d0,d1,...,d7> calculated according to

$$d_1 = \frac{l_i}{\sqrt{\sum_{k=0}^7 l_k^2}}$$

where i = 0, ..., 7, and l_i is the number of black pixels in the ith direction to the boundary of stroke. Peripheral Direction

Contributivity (PDC) corresponds to the distance from a point on the character frame to the first background to foreground transition. PDC will check the distance between the edges of the strokes to the closest second background to foreground transition. As a solution, the application will read the binary data from left to right and from top to bottom to search for the nearest black pixel. Using two dimensional array, the application will set a grid for the image to properly define its features and strokes.

In the application, the systems set 50×50 grid to lessen the process time and to get an accurate reading. Once the system located the nearest black pixel, the application will now perform the 8-direction arrows to look for the next black pixel. The 8-direction arrows are right, upright, up, upleft, left, downleft, down and downright. When the application detects the next black pixel, it will then move where it is and start another 8-direction arrow to search for the following black pixel. While searching for the black pixels of the image, PDC is also checking and counting the features of the character such as lines, curves, corners, loops and forks to be used in determining the writing system of the character.

3-level Matching. This phase is where the application checks all the points that have been referenced by PDC and then convert them into machine readable codes to properly distinguish the properties required to match the captured image. This phase has three levels.



Fig. 7. General features for Chinese, Japanese and Korean writing systems.

1st level Matching of the identification process is the checking of the general features such as lines, corners and curves shown in Fig. 7. The characters to be identified use these common features in their writing systems. In general Chinese, Korean and Japanese writing systems uses strokes with straight, diagonal, horizontal lines and curves.

Table I shows the given condition of the matching. If line is greater than the corner, the application will add 1 point to Japanese and Chinese and 2 points for Korean. However, if the condition, line is greater than corner is not met, 2 points will be added for Japanese and Chinese and 1 point for Korean. This is to say that a Korean character has more lines than corners compared to Japanese and Chinese characters. The second condition stated on Table I is if the corner is greater than the curve, the application will add 1 point to Japanese and if the condition is not met, the application will then add 1 point for Korean. This is to prove that Japanese characters have more corners and Korean characters have more curves. Third condition is if the curve is greater than the line, then the application will add 2 points for Japanese and 1 point for Chinese but if the curve is less than the line it will add 1 point for Korean. This condition proves that Japanese characters have more curves than lines.

 2^{nd} level Matching of the identification process is the checking of the loop as shown in Fig. 8. By this feature, the application can conclude that the character being identified is a Korean writing system.

Table II shows the given condition of the second stage of the matching. This condition is simple to check the loop and if Loop is greater than corner it will add 2 points for Korean and if the condition fails, statistics will add additional 1 point for both Japanese and Chinese.

| TABLE I: 1 ST LEVEL MATCHING | | | | | |
|---|-----------|--------|----------|---------|--------|
| | Condition | | JAPANESE | CHINESE | KOREAN |
| Line | > | Corner | 1 | 1 | 2 |
| Line | < | Corner | 2 | 2 | 1 |
| Corner | > | Curve | 1 | | |
| Corner | < | Curve | | | 1 |
| Curve | > | Line | 2 | 1 | |
| Curve | < | Line | | | 1 |



LOOP Fig. 8. Loop.

| TABLE II: | 2^{ND} | LEVEL | MATCHING |
|-----------|----------|-------|----------|
|-----------|----------|-------|----------|

| Condition | | | JAPANESE | CHINESE | KOREAN |
|-----------|---|--------|----------|---------|--------|
| Loop | > | Corner | | | 2 |
| Loop | < | Corner | 1 | 1 | |

 3^{rd} level Matching of the identification process is the checking of the Fork. Fork is the intersection of lines as shown in Fig. 9. The more forks that the system identifies, the higher possibility that the system will conclude that the character is Chinese Cantonese writing system. Any other result will be Japanese.



Fig. 9. Fork or intersection of lines.

| TABLE III: 3 RD LEVEL MATCHING | | | | | | |
|---|------|--------|----------|---------|--------|--|
| | Cond | ition | JAPANESE | CHINESE | KOREAN | |
| Fork | > | Curve | | 1 | | |
| Fork | < | Curve | | | 1 | |
| Fork | > | Corner | 1 | 1 | | |
| Fork | < | Corner | | | 1 | |
| Fork | > | 2000 | | 1 | | |
| Fork | < | 2000 | 1 | | | |

Table III shows the given condition on the 3rd and final matching level, where it mainly checks the forks or the intersection. First condition is, if the fork is greater than the curve, then it will add 1 point to Chinese. If the fork is less than the curve, it will only add 1 to Korean. Next condition is if the fork is greater than the corner, it will add 1 point for Chinese and Japanese and if the fork is less than the corner, it will only add 1 for Korean. Lastly is if the fork is greater than 2000 counts, it will add 1 point for Chinese otherwise, 1 point for Japanese if the value is less than 2000.

Results of recognition shall be presented to the user by sending the data back to the mobile device.

IV. TESTING, EXPERIMENTATION, AND RESULTS

A. Testing

The application was rigorously tested and executed several times in the emulator to check whether the modules are running accordingly. After integrating the modules and individual applications of the system, the overall requirements specification has been tested. The application was deployed on a mobile phone where an image was captured for testing the codes.

A network testing was also done to check the connectivity of the mobile device to the tethering device or router to the server. This test is very essential for the application because the server will perform all the necessary compilation and translation of the system.

B. Experimentation Procedures

In order to test the application, three experiments were conducted.

- Experiment Number 1–Capturing same image three times to check if the application will return the same result
- Experiment Number 2–Capturing an image to check the dominant feature of the text.
- Experiment Number 3–Capturing images to test the accuracy of the application in different conditions.

1) Experiment No. 1

In the first experiment, the user took a picture of a character three times. The user then checks if the result is the same on every test. Capturing three images of identical writing scripts and check whether the application will return the same results test after test. The characters were captured at least three times and the application gave the correct result in each trial. In this experiment, the results from the application were verified through human effort, the developers, checking the result's accuracy.

2) Experiment No. 2

Given the characteristics of each writing systems, the application was tested for the accuracy of the results. The second experiment shows the data gathered on the experimentation stage. Through PDC and 3-level matching, the researchers proved that Korean characters have more lines and curves than corners, Japanese characters have more corners than lines and Chinese characters have more lines than curves. The researchers ran the server in debug mode to get the values of the features of the character then computed for the percentage of each trait by using the formula.

> $P(a) = a/(a+b+c) \times 100$ $P(b) = b/(a+b+c) \times 100$ $P(c) = c/(a+b+c) \times 100$

where *a*=lines; *b*=corner; *c*=curve; *P*=percentage.

3) Experiment No. 3

In experiment 3, the researchers tested the application in different capture conditions such as normal lighting, low light and conclave test. The test started with normal lighting where returned answers were almost all correct; then the application is tested with low light where returned answers were mostly incorrect because the image produced a high dense color of black producing mostly black image, in this condition PDC was not able to return an accurate result. Finally, the researchers tested the application in concaved image and it returned at least half correct answers. Table IV-Table VI show the statistical percentage of the third experiment where (\checkmark) mark means the result is correct and (X) mark means the result is incorrect.

| | Image Capture Condition | | | |
|-------------------|-------------------------|-----------|---------|--|
| Captured Image | Normal | Low Light | Concave | |
| 1^{st} | ✓ | ✓ | Х | |
| 2 nd | ✓ | X | ✓ | |
| 3 rd | Х | Х | ✓ | |
| 4^{th} | ✓ | Х | ✓ | |
| 5 th | ✓ | ✓ | Х | |
| 6 th | ✓ | ✓ | ✓ | |
| 7 th | ✓ | ✓ | ✓ | |
| 8 th | ✓ | ✓ | Х | |
| 9 th | ✓ | ✓ | Х | |
| 10 th | ✓ | ✓ | ✓ | |
| Percentage | 90% | 70% | 60% | |

Table IV shows the result of experiment 3 on Chinese characters. The researchers tested 10 different characters in different conditions, a total of 30 characters per writing system. Results indicates the results for Chinese character on a normal lighting condition is 90% accuracy, on a low lighting condition 70% accuracy and for concave condition 60% accuracy.

Table V shows the result of experiment 3 on Japanese characters. The researchers tested 10 different characters in different conditions, a total of 30 characters per writing system. Results indicates the results for Japanese character on a normal lighting condition is 80% accuracy, on a low lighting condition 30% accuracy and for concave condition 40% accuracy.

TABLE V: EXPERIMENT 3 ON JAPANESE CHARACTERS

| | Image Capture Condition | | |
|-------------------|-------------------------|-----------|--------------|
| Captured Image | Normal | Low Light | Concave |
| 1 st | ✓ | ✓ | Х |
| 2 nd | \checkmark | X | Х |
| 3 rd | Х | Х | \checkmark |
| 4 th | ✓ | Х | Х |
| 5 th | \checkmark | Х | Х |
| 6 th | \checkmark | Х | \checkmark |
| 7 th | Х | Х | ✓ |
| 8^{th} | ✓ | Х | Х |
| 9 th | ✓ | ✓ | Х |
| 10 th | \checkmark | ✓ | ✓ |
| Percentage | 80% | 30% | 40% |

Table IV shows the result of experiment 3 on Korean characters. The researchers tested 10 different characters in different conditions, a total of 30 characters per writing system. Results indicate the results for Korean character on a normal lighting condition is 100% accuracy, on a low lighting condition 20% accuracy and for concave condition 60% accuracy.

| | Image Capture Condition | | | |
|-----------------------|-------------------------|-----------|---------|--|
| Captured | Normal | Low Light | Concave | |
| Image | | | | |
| 1 st Image | \checkmark | Х | Х | |
| 2^{nd} | ✓ | Х | Х | |
| $3^{\rm rd}$ | ✓ | ✓ | ✓ | |
| 4 th | ✓ | Х | ✓ | |
| 5 th | ✓ | Х | ✓ | |
| 6 th | ~ | Х | Х | |
| 7 th | \checkmark | ✓ | Х | |
| 8 th | \checkmark | Х | ✓ | |
| 9 th | ✓ | Х | ✓ | |
| 10^{th} | ✓ | Х | ✓ | |
| Percentage | 100% | 20% | 60% | |

V. FINDINGS AND CONCLUSION

The Chinese, the Koreans, and the Japanese have shared multiple characters in their writing systems due to their cultural and geographical influences. Their writing systems became identical. At first glance, Chinese, Japanese, and Korean characters may be difficult to tell apart, but there are differences between each of them that can help identify these characters. The system used the features of the Han character namely lines, curves, corner, loops and forks to distinguish the difference between Chinese, Japanese and Korean characters using Peripheral Direction Contributivity and 3-level matching algorithms. The researchers then conducted three experiments to test the accuracy of the application using a mobile phone to capture the Han character image. Based on these experiments, the researchers concluded that:

- 1) The application has 80-100% accuracy in identifying Chinese, Japanese and Korean writing systems.
- 2) The accuracy of results is affected by the lighting condition used to capture the image; the application will most likely return an incorrect result if the surrounding is dark or if the image is distorted. But has very accurate results on a normal lighting condition.
- 3) Users should avoid capturing an image with more than one color of background.
- To process an accurate binary image, users should avoid capturing an image with inverted colors such as white font and black background.

VI. FUTURE WORKS

Based on the findings and conclusions of the study, the following are recommended:

- 1) The study can be improved by looking for more means of illumination to satisfy the constraints stated above.
- 2) Improving the image processing phase of the application, the application should be able to recognize blurred and skewed images and users can select and crop an image they would like to test and be able to get the same results.
- 3) The output of the system can be an input to Natural Language Processing which includes reading and translating captured images.

REFERENCES

- P. Charles, V. Harish, M. Swathi, and C. Deepthi, "A review on the various techniques used for optical character recognition," *International Journal of Engineering Research and Applications*, vol. 2, issue 1, pp. 659-662, Jan-Feb 2012.
- [2] Research in OCR leads to higher accuracy. [Online]. Available: http://www.cvisiontech.com/ocr/file-ocr/research-in-ocr.html
- [3] China-language, culture, customs, and etiquette. [Online]. Available: http://www.kwintessential.co.uk/resources/global-etiquette/china-cou ntry-profile.html
- [4] K. Renaud and Y. Key. *The Korean Alphabet: Its History and Structure*, Honolulu, Hawaii: University of Hawai'i Press, 1997, ch. 12, p. 317.
- [5] P. Lee, *A history of Korean Literature*, Cambridge, UK: Cambridge University Press, 2003, pp. 26-30.
- [6] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62-66, January 1979.
- [7] H. Farid. (2010). Fundamentals of image processing. Dartmouth College. [Online]. Available: http://www.cs.dartmouth.edu/farid/downloads/tutorials/fip.pdf
- [8] A. Greensted. (June 2010). Otsu thresholding. The Lab Book Pages: An Online Collection of Electronics Information. [Online]. Available: http://www.labbookpages.co.uk/software/imgProc/otsuThreshold.htm
- [9] C. Perfetti, Y. Liu, J. Fiez, J. Nelson, D. Bolger, and L. Tan, "Reading in two writing systems: Accomodation and assimilation of the brain's reading network," in *Bilingualism: Language and Cognition*, UK: Cambridge University Press, 2007, vol. 10, pp. 131-146.
- [10] C. Poynton, "The magnitude of nonconstant luminance error," in A Technical Introduction to Digital Video, New York: John Wiley & Sons, 1996.



Maria Christina Aragon received her BS and MS in computer science degrees at AMA Computer University-Makati. She finished her doctorate degree at Technological University of the Philippines in 2011 and her research interests are image processing and natural language processing. She is currently the chair of the CS Program at Technological Institute of the Philippines, Quezon City. Dr. Aragon is a member of ACM, IEEE, Philippine Society of IT Educators and

Computing Society of the Philippines.



Melissa Juanillo is a graduate of BS in computer science from STI College in the year 2000. She obtained her master in information technology from the Technological University of the Philippines-Manila in 2006. In 2007, she finished her teaching certificate program and passed the licensure examination for teachers. Three years after, she took 15 units in doctor of technology program from TUP-Manila. Currently, she is taking. doctor of

information technology, where she already finished the academic units from Technological Institute of the Philippines-Quezon City. She became a full-time faculty member of the TIPQC-Computer Science Department in 2004



Rosmina Joy Cabauatan is an assistant professor of both graduate and undergraduate levels in Information Technology Programs of Technological Institute of the Philippines. She is also the research coordinator of the College of Information Technology Education, a Center of Excellence in Information Technology Education in the Philippines. As thesis advisor, she has co-authored published researches in data mining, game development, algorithm applications in

computing and education. Now, she is pursuing her doctoral studies in information technology, she is currently undertaking her dissertation on multidisciplinary applications of data mining