AUTOMATIC ASSESSMENT MARK ENTRY SYSTEM USING LOCAL BINARY PATTERN (LBP)

By

LIM LAM GHAI

FINAL PROJECT REPORT

Submitted to the Electrical & Electronics Engineering Programme
in Partial Fulfillment of the Requirements
for the Degree
Bachelor of Engineering (Hons)
(Electrical & Electronics Engineering)

Universiti Teknologi Petronas
Bandar Seri Iskandar
31750 Tronoh
Perak Darul Ridzuan

CERTIFICATION OF APPROVAL

AUTOMATIC ASSESSMENT MARK ENTRY SYSTEM USING LOCAL BINARY PATTERN (LBP)

by

Lim Lam Ghai

A project dissertation submitted to the
Electrical & Electronics Engineering Programme
Universiti Teknologi PETRONAS
in partial fulfilment of the requirement for the
Bachelor of Engineering (Hons)
(Electrical & Electronics Engineering)

Approved:

____

Suhaila Binti Badarol Hisham
Project Supervisor

UNIVERSITI TEKNOLOGI PETRONAS
TRONOH, PERAK

May 2014

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.

Lim Lam Ghai

ABSTRACT

Offline handwritten recognition continues to be a fundamental research problem in document analysis and retrieval. The common method used in extracting handwritten mark from assessment forms is to assign a person to manually type in the marks into a spreadsheet. This method is found to be very time consuming, not cost effective and prone to human mistakes. In this project, a number recognition system is developed using local binary pattern (LBP) technique to extract and convert students' identity numbers and handwritten marks on assessment forms into a spreadsheet. The template of the score sheet is designed as in Appendix 1 to collect sample of handwritten numbers. The training data contain three sets of LBP histograms for each digit. The recognition rate of handwritten digits using LBP is about 50% because LBP could not fully describe the structure of the digits. Instead, LBP is useful in term of arranging the digits '0 to 9' from highest similarity score to the lowest similarity score as compared to sample using chi square distance. The recognition rate is greatly improved to about 95% by verifying the output of chi square distance with the salient structural features of digits.

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my supervisor Ms Suhaila binti Badarol Hisham, lecturer of Universiti Teknologi PETRONAS for the continuous support of my final year project , for her patience, motivation, enthusiasm, and immense knowledge. Her guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better supervisor and mentor for my final year project.

My sincere thanks also go to my fellow classmates who assist me in this project and to those who give continuous support to me when I faced obstacles in completing the project. This includes Mr. Yeo Lip Wee for his kindness and generous in lending me his scanner to test my prototype. Besides that, I would like to thank the students and lecturers who willing to spend their times to fill up the score sheet form for my experiment data.

Furthermore, I would like to thank my parents Lim Chuan Sen and Kho Siew Chu, for giving birth to me at the first place and supporting me spiritually throughout my life. I end my acknowledgement to everyone that contributes directly and indirectly in my successfulness to complete this final year project. Without them, I would not be able to accomplish my final year project entitled “Automatic Assessment Entry Mark System Using Local Binary Pattern (LBP)”.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

LBP	Local Binary Pattern
PDA	Personal Digital Assistant
OCR	Optical Character Recognition
MLP	Multi-Layer Perception
OMR	Optical Mark Recognition
MCQ	Multiple Choice Question
SVC	Support Vector Classifier
JPEG	Joint Photographic Experts Group
PDF	Portable Document Format
DPI	Dots Per Inch
MATLAB	Matric Laboratory

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

Handwritten recognition continues to be a fundamental research problem in document analysis and retrieval with application in document indexing, recording, translation and search. It can be off-line handwritten recognition or on-line handwritten recognition. Off-line recognition converts the handwritten texts on the paper using image as an input to letter codes. On the other hand, on-line recognition automatically converts the handwritten text as it is written, usually on personal digital assistant (PDA) by sensing the movement of the pen-tip [1].

Off-line recognition is a type of static representation of handwriting whereas on-line recognition is considered as digital representation of handwriting. The principle behind the handwritten character recognition is similar to those of optical character recognition (OCR). Nowadays, OCR technology mainly focuses on machine printed text as it is very difficult to recognize handwritten text due to infinite writing styles from various people. Generally, OCR can be divided into three main steps which are preprocessing of the input, feature extraction and classification of the characters [2].

Local binary pattern (LBP) is a technique used in texture analysis and it has been applied in many applications such as face recognition, biometric and remote sensing application. The features of a particular face can be extracted using LBP and the researchers have shown that the recognition rate can increase up to 95% [3]. In this project, a digit recognition system is developed using LBP technique to extract and convert students’ identity numbers and handwritten marks on assessment forms into a spreadsheet.

1.2 PROBLEM STATEMENTS

The common method used in extracting examination or assessment mark from evaluation forms written by hand is to have dedicated persons manually type in the marks into a spreadsheet. This method is very time consuming, not cost effective and prone to mistakes or errors. For example, the dedicated persons might mistakenly key in the marks of student ‘A’ under student ‘B’ and such mistake tends to be repeated if the number of students increase or if the process has to be rushed.

In this project, LBP technique is used to extract features of different handwriting styles of digits 0 to 9. The features extracted are compared with the training data to convert the handwritten digit into values in a spreadsheet. However, the current limitations in image preprocessing of handwritten digits are segmenting connected characters and correcting the skew of scan image. Thus, this project is focused more on the feature extraction of handwritten digits using LBP.

1.3 OBJECTIVES AND SCOPE OF STUDY

The fundamental theory of LBP is studied and evaluated on handwritten marks. In addition, the basic steps of OCR such as preprocessing of image, feature extraction and classification of characters are combined and implemented in this project together with LBP technique.

The objectives of the project are:

- To develop an automatic assessment mark entry system using LBP as a feature extraction method
- To evaluate the performance and recognition rate of LBP on handwritten marks
- To improve the recognition rate of LBP by combining with the salient structural features of digits.

CHAPTER 2

LITERATURE REVIEW

2.1 THEORY

The local binary pattern (LBP) is an image operator used to describe and characterize the texture patterns of an image in a binary number [4]. LBP transforms the gray scale image into an array which commonly represented in histogram and it has been used in many applications such as face expression recognition, biomedical image analysis, video retrieval and motion analysis [5].

The basic LBP operates in a 3 x 3 block of a pixel which consists of 8 neighbors [6]. Its center gray value, denoted as g_c is the threshold value for the surrounding neighborhood [3]. If the gray value of the neighbor exceeds g_c, a binary value of 1 is represented or else 0 is given as shown in Figure 1.

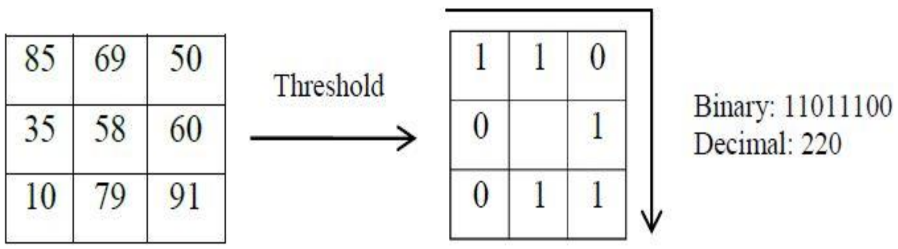

Figure 1: The basic LBP operator

The above LBP can be generalized to include more sampling points and radius as shown in Figure 2. The local neighborhoods are defined as the evenly circular spaced of the sampling points, P with radius R from the center pixel of an image, $I(x,y)$ [6]. The notation (P, R) is used to represent the neighborhoods. Let g_n

denotes the gray value of the sampling point and its coordinate is $(x + R\cos(2\pi p/P)\,,\ y - R\sin(2\pi p/P)\,)$ where $n = 0, \dots\dots, P-1$. The coordinate of g_n can be estimated using bilinear interpolation if the sampling point is not located in the center of the block [6].

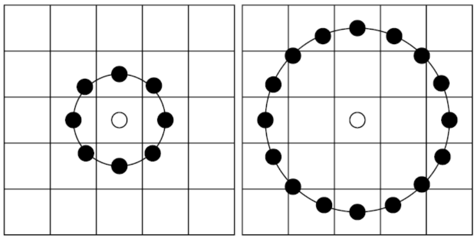

Figure 2: The circular (8,1) and (16,2) neighborhoods

Generally, the LBP value for pixel (x_c,y_c) can be calculated as follow where $s(z)$ is a signum function [6]:

$$\mathrm{LBP_{P,R}}(x_c, y_c) = \sum_{n=0}^{P-1} s(g_n - g_c)2^n, \qquad s(z) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0 \end{cases} \tag{1}$$

Given an image of size N x M. The whole texture image can be represented in histogram for each pixel *(i, j)* as shown in Figure 3 using equation (2) and (3) [7]:

$$\mathrm{H}(k) = \sum_{i=1}^{N}\sum_{j=1}^{M} f\big(\mathrm{LBP_{P,R}}(i,j)\,, k\big), \qquad k \in [0, 2^{\mathrm{P}} - 1] \tag{2}$$

$$f(x, y) = \begin{cases} 1, & x = y \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

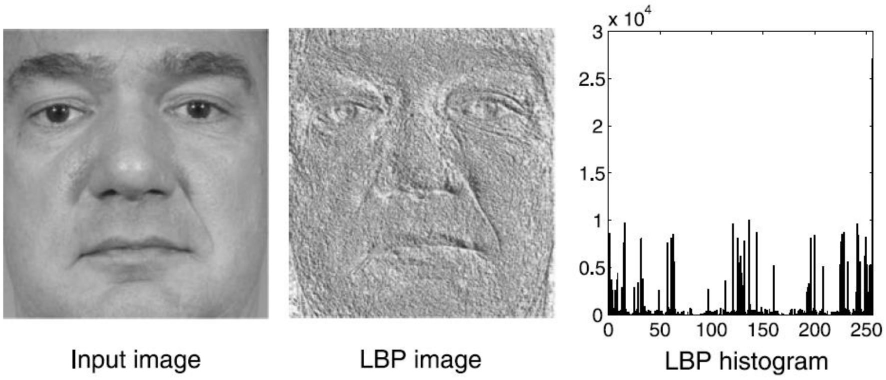

Figure 3: LBP image and histogram for an input image [6]

Thus, 256 different labels of features will be extracted for P=8, sampling points.

There are several mapping types for LBP labels; namely uniform patterns, rotational invariance and rotation invariance uniform pattern [6], [8]. The local neighborhoods are moved into other block and the orientations of sampling points are changed due to the rotation of input image. Basically, uniform patterns are the most commonly used in LBP mapping because natural images are mostly uniform and statistical robustness of uniform pattern gives more stable result [6]. A pattern is called uniform if the bitwise transition from 0 to 1 or 1 to 0 is at most 2. For instance, 00000000 (0 transitions) and 00011000 (2 transitions) are uniform whereas 10011001 (4 transitions) and 10101010 (7 transitions) are non-uniform [3], [5]. The non-uniform output labels are omitted and this reduces the number of output labels for P bits to P(P – 1) + 3 [6].

The difference between feature vectors of the training data, S and the input image, M can be measured using non-parametric statistic test such as G statistic, log-likelihood statistic and chi square distance [6]. It has been proved that chi square distance performs better than other methods due to its stability when dealing with small sample sizes [9]. For that reason, the dissimilarity between S and M can be calculated as follow using chi square distance [7]:

$$\chi^2(\mathrm{S,M}) = \sum_{b=1}^{B} \frac{(\mathrm{S}_b - \mathrm{M}_b)^2}{\mathrm{S}_b + \mathrm{M}_b} \tag{4}$$

where B is the number of bins, S_b and M_b are the LBP values displayed in the histogram at b^{th} bin. The similarity between a sample and a model is high when the value of χ^2 is reduced [3].

2.2 CRITICAL ANALYSIS

Optical character recognition (OCR) is a technology used to convert and transform scanned images of printed text into readable and editable text using computer program [10]. The research and development of OCR technology has begun since 1940. Feature extraction is one of the processes in OCR where the input data is transformed into set of features. Analysis with a large number of variables generally requires a huge amount of memory and computation power which exceeds the size of the training sample [11]. Therefore, the purpose of feature selection is to describe a large set of data accurately by simplifying the amount of resources required. The feature extraction methods can be grouped into structural features, geometric features, and feature space transformation methods [11].

Nowadays, many OCR software are available in the internet either for free use or commercial application. Some of the most commonly used OCR freeware are FreeOCR, WeOCR and OCRopus [10]. People can also perform online OCR by just uploading the document through website such as onlineOCR.net. Tesseract is an OCR engine integrated into FreeOCR and WeOCR software. Initially, blob filtering and line construction are used to find text lines and then the baseline is fit using quadratic spline. Next, the pitch is used to separate the characters of the word [12]. Besides that, Tesseract uses two types of classifiers which are static character classifier and adaptive classifier. In static character classifier, the input features are matched with the training data using segment of polygonal approximation and then the characters are classified by computing the similarity. On the other hand, isotropic baseline or x-height normalization techniques are used in adaptive classifier to obtain better recognition rate when the information of fonts are not available in training data [12]. Tesseract gives high characters recognition rate for clean printed input but at the moment, it has limited handwritten characters recognition.

OCRopus is an open-source OCR program that provides better layout analysis compared to FreeOCR and WeOCR [13]. It is a feed-forward system which consists of layout analysis, text line recognition and statistical language modeling. Tesseract is not only the OCR engine of OCRopus. Multi-layer perceptions (MLPs) and HMM-based recognizer are integrated in OCRopus to extract the feature and recognize the text line. Dynamic programming algorithm is used in MLPs to perform over segmentation of the input and then recognizes the characters [13].

On the other hand, optical mark recognition (OMR) is a method used to detect discrete information indicated with pencil marks on specific OMR sheets [14]. OMR sheets are widely used in education for multiple choice questions (MCQ) in examination. However, it has some disadvantages where OMR sheets usually thicker and expensive than normal A4 paper [14]. Moreover, high precision layout of OMR sheet is needed to achieve zero error.

In addition, a research has been conducted to identify number display from a seven-segment light emitting diode (LED) of measuring instrument on computer through captured images from a digital camera. In this work, the number is divided into 12 equal parts and features such as binary edges and corners are extracted and correlated with the number templates [15]. 92% of number recognition rate is achieved when testing 100 images from a digital multi-meter with average simulation time of 0.4 seconds. Nevertheless, this research faces problems in the skew image and it has certain limitation in the distance between camera and display [15].

Another research conducted by Xianjing Wang and Atul Sajjanhar using circular grid zoning as feature extraction for offline handwritten characters recognition has shown high recognition rate [16]. In this research, the binary image of handwritten character is transformed into points in polar coordinates. The standard zoning method in polar coordination and circular grid zoning method are then used to extract the feature from the image as shown in Figure 4 and Figure 5. The image is divided into 16 equal zones with equal size, distance and angle in circular grid zoning in which the features of the image are extracted in individual zones rather than the whole image [16]. It is found that circular grid zoning achieves higher recognition rate (92.31%) compared to standard zoning (86.63%) in polar coordination. However, some transformed images might appear empty in certain zone which can lead to low recognition rate and rotated images might not be suitable for this method due to different direction and angle [16].

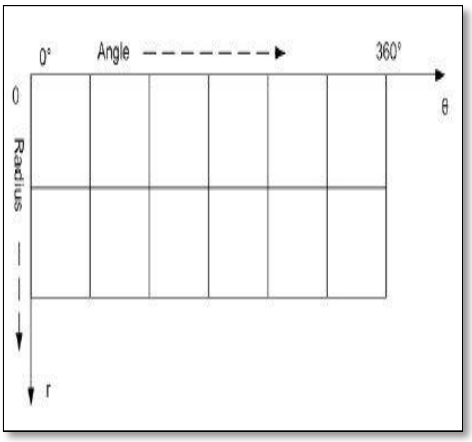

Figure 4: Standard zoning [16]

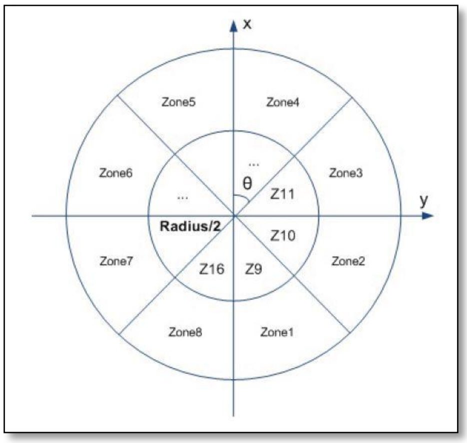

Figure 5: Circular grid zoning [16]

Furthermore, the concept of eigenvalues is used as features extraction for individual offline handwritten digit recognition where this method is named as 3 points feature extraction [17]. 3 points stands for eigenvalues A, B and C in which A is the number of nodes intersects through the curve as the image is split into half vertically, B is the state value of the upper part of the handwritten digit and C is the state value of the lower part of the handwritten digit [17]. In this study, the researchers conclude that a digital character can be decomposed into six basic states as shown in Figure 6. This method is proved to be simple and no complex processing is involved as only 3 eigenvalues are needed to be extracted from an image.

state	C-L	C-R	C-U	CC	S-C	W-C	NT
state value	1	2	3	4	5	6	0

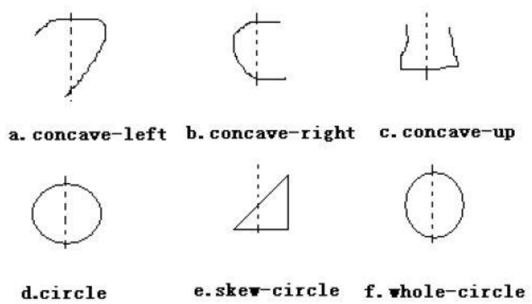

Figure 6: The figures of states and its state value [17]

Neural network is a commonly used method in offline handwritten character recognition where it is an interconnected group of natural or artificial neurons which uses mathematical or computational model for information processing [18]. The adaptive system of neural network can change its structure based on external or internal information that flows through the network which makes it useful in statistical data modeling or decision making tool [19]. They can be used to model complex relationships between inputs and outputs or to find patterns in the data. A multilayer feed forward neural network has been implemented with one hidden layer and back propagation algorithm to train the network for offline recognition of handwritten isolated digits [19]. An experimental result shows that conventional features with back propagation network yields good classification accuracy of 100% and recognition accuracy of 91.2% [19]. However, the work can be extended to increase the results by adding some more relevant features.

A well-known image databases using state-of-the-art feature extraction and classification techniques also have been applied on handwritten digit recognition [20]. The tested databases are CENPARMI, CEDAR, and MNIST. On the test data set of each database, 80 recognition accuracies are given by combining eight classifiers with ten feature vectors [20]. The features include gradient feature, profile structure feature, and chaincode feature. The classifiers include the k-nearest neighbor classifier, three neural classifiers, a learning vector quantization classifier, a discriminative learning quadratic discriminant function (DLQDF) classifier, and two support vector classifiers (SVCs) [20]. All the classifiers and feature vectors give high recognition accuracies.

CHAPTER 3

METHODOLOGY

3.1 SCORE SHEET TEMPLATE

The score sheet template as in Appendix I is designed for this project. There are 8 spaces provided for students' identity number and 8 spaces for their scores. The student's identity number is specified to 5 digits only and the score is within 0 to 100. Overall, each score sheet can have a total of about 56 handwritten numbers.

Several guidelines need to be followed where they are stated in the score sheet template. The guidelines are only pen or 2B pencil can be used and numbers are allowed to be written in the spaces provided. 50 score sheets are distributed to lecturers and undergraduate students of Universiti Teknologi PETRONAS to collect the samples of handwritten numbers. The average handwritten digit recognition rate and execution time of a score sheet are calculated, recorded, tabulated and analyzed.

3.2 HANDWRITTEN DIGIT RECOGNITION PROCESS

The process of offline handwritten digit recognition is shown in Figure 7. The training data consists of three sets of LBP histogram for each digit. Initially, the score sheet that contains handwritten numbers is scanned using an optical scanner and it is saved as Joint Photographic Experts Group (JPEG) format. The Portable Document Format (PDF) is not used because the existing MATLAB software could not automatically convert it to JPEG format. The default resolution of the optical scanner is 300 dots per inch (DPI) which is used in this project. The quality of scanned image is poor below 300 DPI and it should be avoided to achieve high recognition rate. The scanned image is then manually inserted as an input to the programming code and it is converted to logical format using suitable threshold calculated by Otsu's method [21].

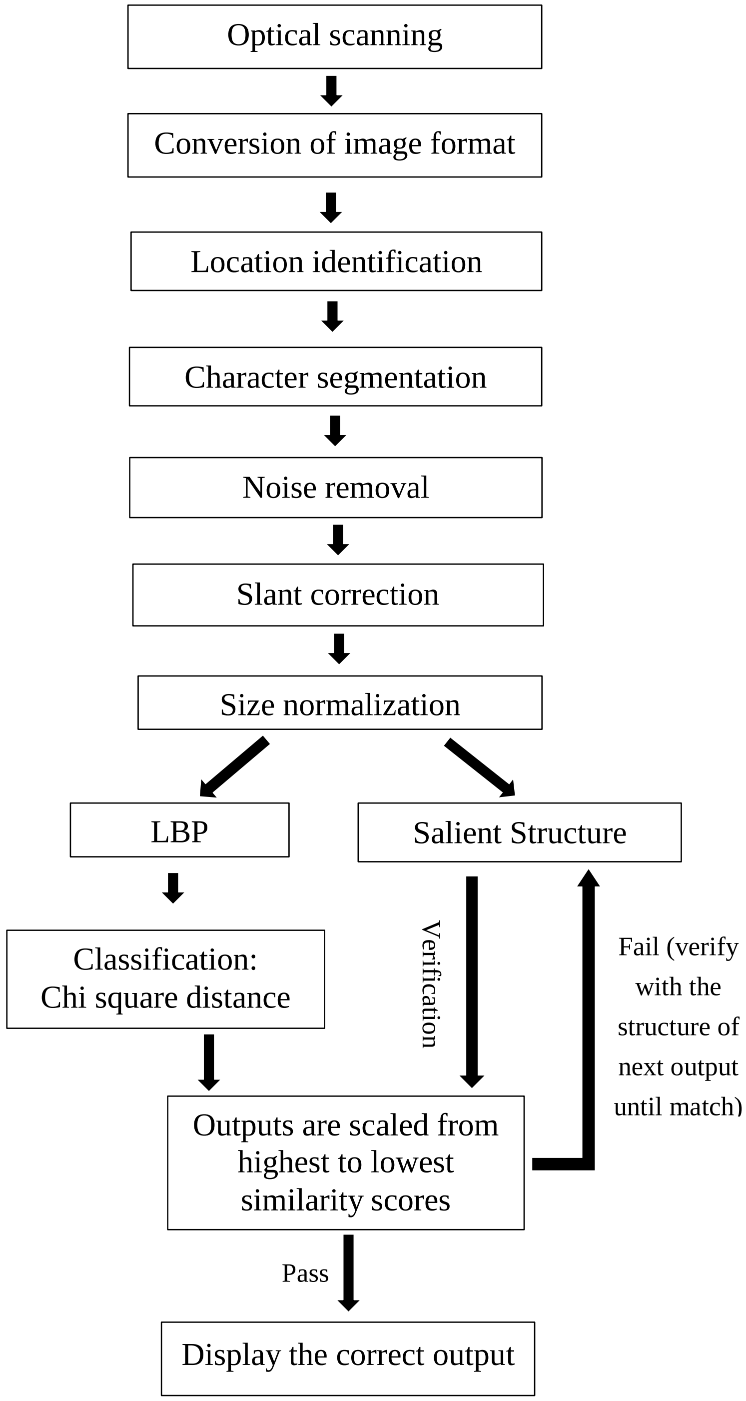

Figure 7: Offline handwritten digit recognition process

Next, the pixel locations of the student's identity number and the associated score are identified and cropped as shown in Figure 8. The segmentation process is then applied on the cropped image to separate the handwritten digits individually using the Image Processing Toolbox in MATLAB called "regionprops" [22]. However, the drawback of this method is it cannot segment connected numbers where the connected numbers are identified as an image. The segmented image is inverted and converted into logical format which consist of values '0' and '1' only where '0' stands for black and '1' stands for white. Figure 9 shows the output of an image after it is inverted and converted into logical format.

Figure 8: Cropped image of student's identity number

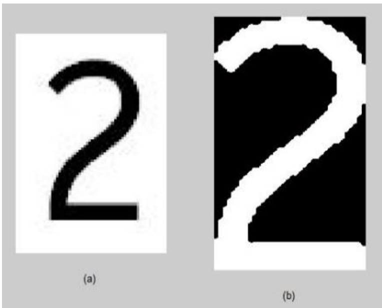

Figure 9: (a) Original image, (b) Binary image (inverted)

Afterward, the image is preprocessed to remove irrelevant information or noise and slanted image is corrected. Prior to the process of LBP, the area of the image is normalized to 10000 pixel^2 by increasing the width and height of the image in same ratio. The purpose of area normalization is to facilitate feature extraction process without distorting the original structure of the image [11].

The features of the handwritten marks and students' identity numbers are extracted using LBP technique and these feature vectors are compared with the training data to identify the correct number. Chi square distance is used to find the closest match of the sample with the training data. The outputs (0 to 9) are scaled using their similarity scores from the highest to the lowest similarity. These outputs are then verified with the salient structural features of the digit till it pass by following the ranking. The salient structural features of digit are defined in the following section 3.3 and it is not related with LBP. The reasons of adding these structural features are later explained in Chapter 4. The identity numbers of the student with his/her assessment mark are displayed as an output in the spreadsheet.

3.3 SALIENT STRUCTURAL FEATURES OF DIGIT

Circle is one of the salient structural features of digit where digit ‘8’ should have two circles and digits ‘0’, ‘6’ and ‘9’ should have one circle only, either upper or lower circle. The number of nodes that intersects through the binary image as it is halved vertically is measured by calculating the total times of pixel transition from value 0 to 1. For instance, digits ‘2’, ‘3’, ‘5’ and ‘8’ have three nodes whereas digits ‘0’ and ‘7’ have two nodes.

The circle is calculated using the location of nodes as reference points. Let digit ‘8’ be an example, where it has 3 nodes. If the right side pixels between node 1 and node 2 contain binary value of ‘1’ to a certain threshold, then it is considered as closed loop. This concept also applies to the left side pixels between node 1 and node 2. It can be safely assumed the digit contains a circle when the pixels of both (left and right) sides are closed loops. Again, this concept is applied to node 2 and node 3 and the circles can be categorized as top circle (node1 & node 2) and bottom circle (node 2 & node 3). Figure 10 illustrates the application of this concept on digits ‘3’ and ‘8’ and it is also applied to digits ‘2’, ‘5’, ‘6’, ‘9’ and ‘0’.

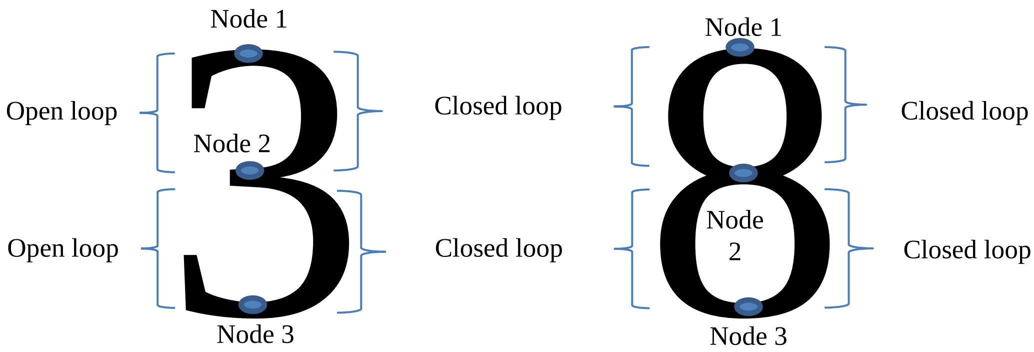

Figure 10: The concept of closed loop and open loop for circle detection

In addition, the ratio of the height over the width of the sample is measured to identify digit ‘1’ where it should have a minimum ratio of 3. The horizontal straight line of digits ‘7’ and ‘4’ are also included as salient features. The salient structural features are selected in a way that can match different styles of handwritten digits ‘1’, ‘2’, ‘4’, ‘7’ and ‘9’ as shown in Figure 11. Table 1 summarizes the salient structural features of the digit used in this project.

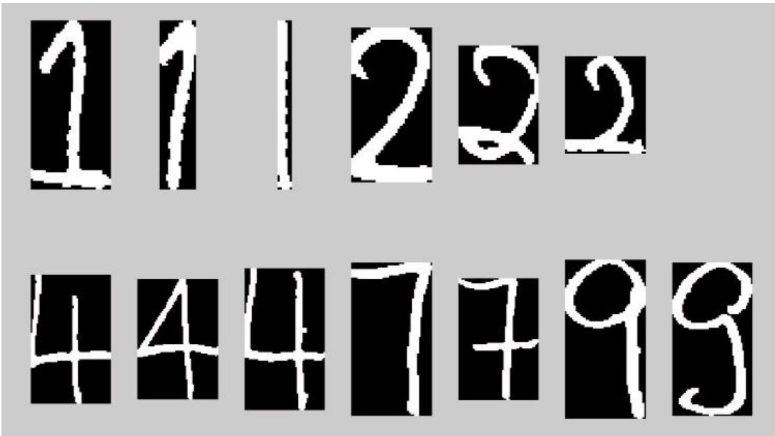

Figure 11: Different styles of handwritten numbers ‘1’, ‘2’, ‘4’, ‘7’, and ‘9’

Table 1: Salient structural features of digits

Digit	Salient Structural Features
0	➢ 2 nodes intersect vertically in the middle ➢ One big circle
1	➢ The ratio of height over width must be greater than 3
2	➢ Interval of node 1 and node 2 • Closed loop for right side and open loop for left side ➢ Interval of node 2 and node 3 • Open loop for right side and closed loop for left side
3	➢ Interval of node 1 and node 2 • Closed loop for right side and open loop for left side ➢ Interval of node 2 and node 3 • Closed loop for right side and open loop for left side
4	➢ The position of horizontal straight line
5	➢ Interval of node 1 and node 2 • Open loop for right side and closed loop for left side ➢ Interval of node 2 and node 3 • Closed loop for right side and open loop for left side
6	➢ Interval of node 1 and node 2 • Open loop for right side and closed loop for left side ➢ Bottom circle
7	➢ The position of horizontal straight line
8	➢ 3 nodes intersect vertically in the middle ➢ Top circle and bottom circle
9	➢ Top circle ➢ Interval of node 2 and node 3 • Closed loop for right side and open loop for left side

3.4 RESEARCH INSTRUMENTS

The software used in conducting the experiment are MATLAB R2013a, Microsoft WORD and EXCEL 2010. The hardware are personal computer and Canon MP250 scanner.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 IMPLEMENTATION OF LBP ON BINARY IMAGE

The basic local binary pattern (LBP) using 8 pixels in a 3 x 3 pixel block are implemented on the binary image of handwritten number where the number of sampling points P is 8 with radius R of 1 around center pixel (x, y) as shown in Figure 12. g_0 to g_7 are the binary value (0 or 1) of the sampling points and g_c is the binary value of the center pixel.

g_7	g_0	g_1
g_6	g_c	g_2
g_5	g_4	g_3

Figure 12: Description of 3 x 3 pixel block using LBP

The $LBP_{8,1}$ operator from Equation (1) can be expended as

$$\mathrm{LBP}_{\mathrm{P,R}}(x_c, y_c) = s(\mathrm{g}_0 - \mathrm{g}_c) + s(\mathrm{g}_1 - \mathrm{g}_c)2 + s(\mathrm{g}_2 - \mathrm{g}_c)4 + s(\mathrm{g}_3 - \mathrm{g}_c)8 + s(\mathrm{g}_4 - \mathrm{g}_c)16 + s(\mathrm{g}_5 - \mathrm{g}_c)32 + s(\mathrm{g}_6 - \mathrm{g}_c)64 + s(\mathrm{g}_7 - \mathrm{g}_c)128 \quad (5)$$

Overall, 256 different labels of features or bins ranking from 0 to 255 are extracted and displayed in LBP histogram. However, only 32 bins contain the essential information of handwritten number while the remaining bins are omitted. The reasons why the remaining bins are omitted can be explained by the following 3 cases:

Case 1: when g_c = 0

- The signum function $s(\mathrm{g_n} - \mathrm{g_c})$ is always equal to 1 when g_c = 0 because $g_n - g_c$ is always greater than or equal to zero. Hence, LBP value is 255 and this feature is stored in bin 255th. For example, if 1000 pixels of the binary image have value of 0, then the value of bin 255th will equal to 1000. This information is not much of use because binary value 0 is the black spaces surround the handwritten digit.

Case 2: when g_c = 1 and all sampling points, g_n = 1

- The signum function $s(\mathrm{g_n} - \mathrm{g_c})$ is always equal to 1 because $g_n - g_c$ is always equal to zero and this information is stored in bin 255th. Since it only represents the thickness of handwritten digit, therefore it does not help in identifying the digit.

Case 3: when the uniformity measure is more than 2

- It is found that 90% of the binary images of handwritten digits have uniform pattern where the uniformity measure is at most 2 only. The non-uniform patterns are omitted because they do not have useful information to be used in identifying the digits, where they may affect the overall result if taken into consideration. In fact, the non-uniform bins are empty and the calculation can be simplified by removing them.

The uniform mapping of LBP labels produces 59 output labels for neighborhoods of 8 sampling points deriving from formula P(P – 1) + 3. The 32 bins that contain essential information with its binary values are tabulated in Table 2. The binary value of the center pixel is equal to 1 for all these 32 bins.

Table 2: 32 bins are selected from LBP histogram of 8 sampling points

Bin	Binary value ($g_7g_6g_5g_4g_3g_2g_1g_0$)	Bin	Binary value ($g_7g_6g_5g_4g_3g_2g_1g_0$)
7	00000111	193	11000001
15	00001111	195	11000011
28	00011100	199	11000111
30	00011110	207	11001111
31	00011111	223	11011111
60	00111100	225	11100001
62	00111110	227	11100011
63	00111111	231	11100111
112	01110000	240	11110000
120	01111000	241	11110001
124	01111100	243	11110011
126	01111110	247	11110111
127	01111111	248	11111000
135	10000111	249	11111001
143	10001111	252	11111100
159	10011111	253	11111101

LBP histogram is simplified to these 32 bins only instead of 256 bins where first bin is referred to decimal value of 7, second bin is referred to decimal value of 15, third bin is referred to decimal value of 28 and so forth. The simplified histogram is then normalized to 100% by dividing the value of each bin with the summation of values in 32 bins. Figure 13 shows the binary image of printed digits from 0 to 9 which are selected as training data for handwritten digit and their LBP histograms are displayed in the following figures.

Figure 13: Binary image of printed digits

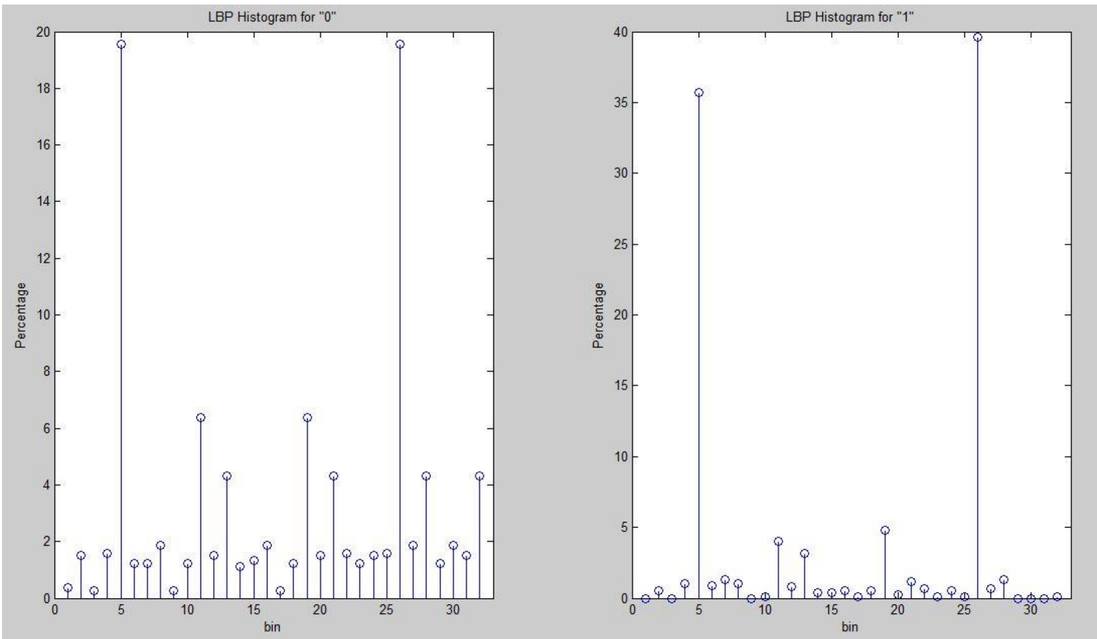

Figure 14: LBP histogram for images ‘0’ and ‘1’

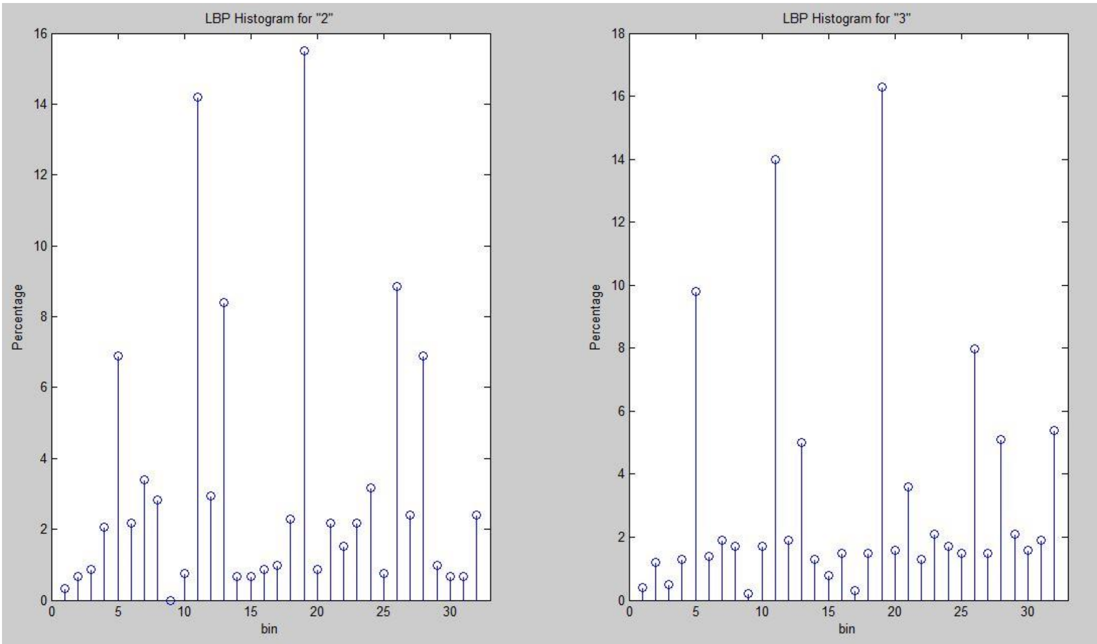

Figure 15: LBP histogram for images ‘2’ and ‘3’

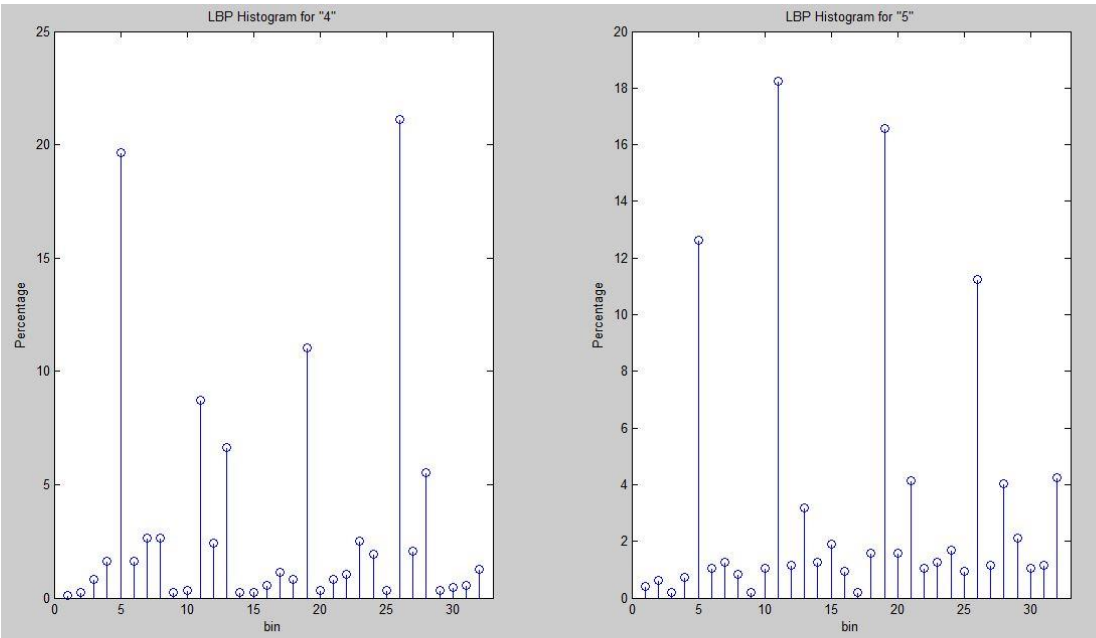

Figure 16: LBP histogram for images ‘4’ and ‘5’

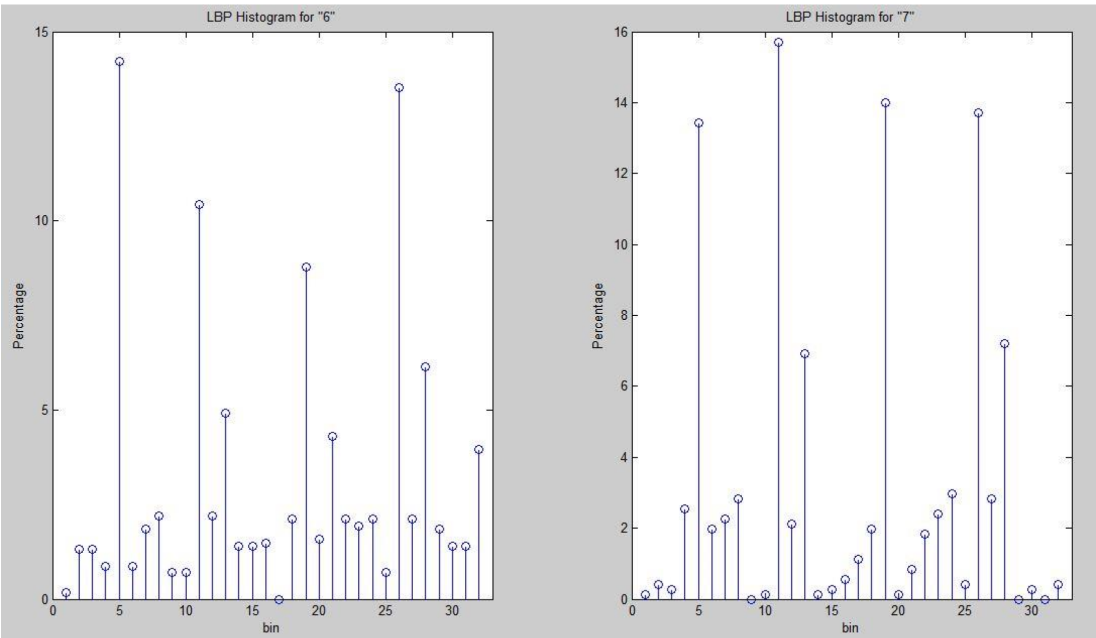

Figure 17: LBP histogram for images ‘6’ and ‘7’

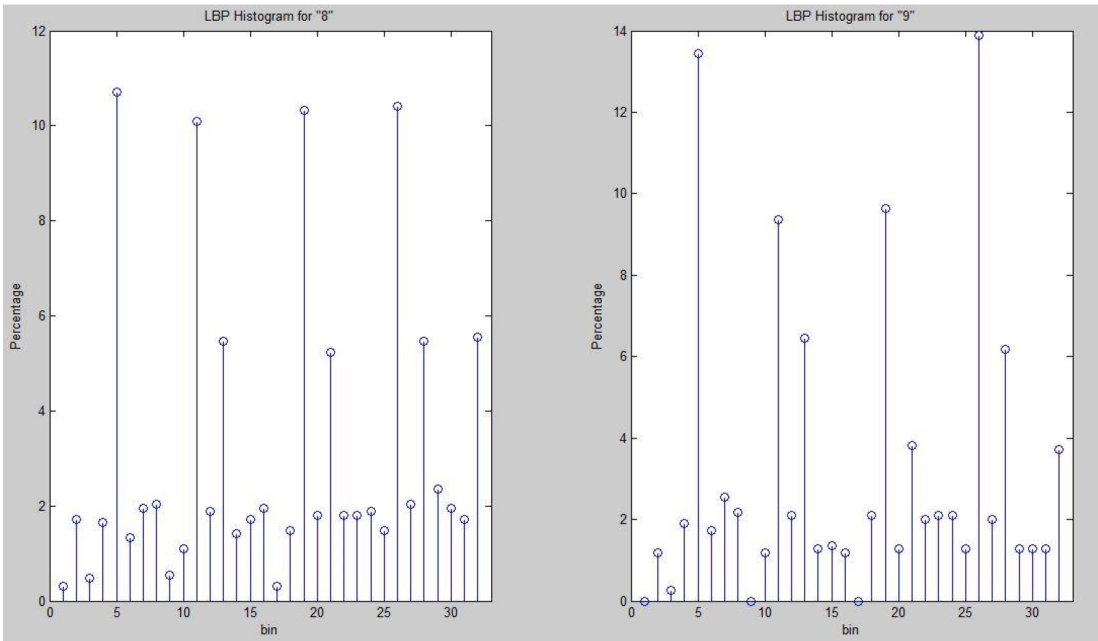

Figure 18: LBP histogram for images ‘8’ and ‘9’

Based on the LBP histogram, majority of LBP values for all the digits are located at bin 5^{th} (actual bin is 31), bin 26^{th} (actual bin is 241), bin 11^{th} (actual bin is 124) and bin 19^{th} (actual bin is 241). By looking at the binary values of the sampling points at those bins as shown in Figure 19, it is found that bin 5^{th} and bin 26^{th} represent the vertical lines of the printed number whereas bin 11^{th} and bin 19^{th} represent the horizontal lines. This explains the images ‘0’ and ‘1’ have highest percentage on bin 5^{th} and bin 26^{th} while images ‘2’, ‘3’, and ‘5’ have highest percentage on bin 11^{th} and bin 19^{th}.

0	1	1		1	1	0		0	0	0		1	1	1
0	1	1		1	1	0		1	1	1		1	1	1
0	1	1		1	1	0		1	1	1		0	0	0
Bin 5^{th}				Bin 26^{th}				Bin 11^{th}				Bin 19^{th}		

Figure 19: The binary values of g_n at bin 5^{th}, 26^{th}, 11^{th}, and 19^{th}

The remaining 28 bins basically represent the curvature of the printed number. For instance, digits ‘6’, ‘8’ and ‘9’ have more curvy lines compared to other numbers, and these features are distributed in the identified 28 bins as shown in LBP histogram. A case example in the opposite, digit ‘1’ has the least curvy shape and hence only very small percentage falls in those 28 bins as displayed in Figure 14. Generally, the percentages distributed on these 32 bins contain the features of numbers such as horizontal straight line, vertical straight line and curvy shape. However, the exact locations of these features on the digits are unknown or could not be identified. Only the LBP values calculated on the outer shell of the handwritten numbers are taken into consideration as highlighted in the binary image shown in Figure 20.

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	0	1	1	1	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0
0	0	0	0	1	1	1	1	1	1	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 20: LBP values are calculated on highlighted binary values of 1 only

4.2 HANDWRITTEN DIGIT RECOGNITION RATE BASED ON LBP

Chi square distance is used as the classification method to find the closest match of the sample with the training data. The training data consist of LBP histograms of the printed digits and general different styles of nicely written digits. The minimum chi square value means that the sample has the highest similarity score with one of the digit of training data and this digit is displayed as an output.

The average recognition rate is low, about 30% to 55% based on the result and about 65% to 85% of the samples are listed in the top three ranking of minimum chi square distance where one of these three closest matched digits is the correct output. However, chi square distance could not differentiate the correct output in the top three rankings due to limitations of LBP.

These limitations are

i. The LBP histograms of all digits have almost similar pattern. As a result, the values of chi square distance of different digits are found to be quite close to each other for difference categories of training data.

ii. LBP cannot detect circle, crossing point, and is unable to determine the location of vertical and horizontal straight lines. It only shows the number of occurrence of that particular binary number within the sampling points. For example, the handwritten numbers ‘1’ and ‘7’ show in Figure 21 are matched to printed number ‘2’ instead of ‘1’ and ‘7’.

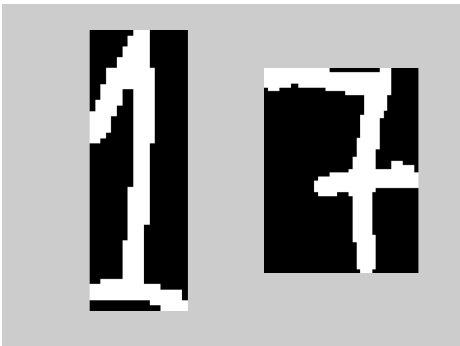

Figure 21: Handwritten numbers ‘1’ and ‘7’ are matched to digit ‘2’

iii. A slanted handwritten number would change the distribution of LBP values in those 32 bins.

In summary, LBP alone does not have enough information to identify the handwritten digit and it can only be used as a ranking by positioning the chi square distance of each training data from minimum to maximum. Thus, the total ranking is ten because the training data have ten sets of different digits. The salient structural features of the handwritten number must be included as a secondary feature extraction to verify the ranking of the chi square distance. For instance, if the sample is ‘6’ and the first ranking is ‘8’, then it will check whether the sample has two circles? Only when the structural features are matched with the sample, then the output is displayed as ‘8’, otherwise it moves to second ranking and continue verifying the structural features. Since about 75% of the handwritten numbers fall in the top three ranking, there is no significant increase in program execution.

4.3 HANDWRITTEN DIGIT RECOGNITION RATE OF COMBINED FEATURES EXTRACTION

The average recognition rates of handwritten digits of a score sheet using different approaches are summarized in Table 3. The recognition rate is at most 85% only when three lowest chi square distances are verified with the structural features. This is because some of the handwritten digits vary so much from the digits used in the training data and they fall in the ranking below than three. However, the recognition rate is greatly improved to about 95% by verifying all the rankings till the structural features are matched. The drawback of this approach is the program execution of a score sheet is slightly increased.

Table 3: Recognition rates of handwritten digits using different approaches

Features Extraction Methods	Classification	Average Recognition Rate	Execution Time (seconds)
LBP only	The lowest chi square distance is the output	30% - 55%	5
LBP + structure	Only three lowest chi square distances are verified with the structural features	65% - 85%	5.8
LBP + structure	Consider all rankings till the structural features are matched	85% - 100%	6
Structure only	Salient features such as circles, crossing points, ratio and straight line	80% - 90%	3

The recognition rate of using only structural feature is considered high but it is still lower than the recognition rate of LBP with the structure. The difference in recognition rate between these two approaches is just about 5 to 10% and the execution time is definitely shorter when LBP is removed. The salient structural features of digits selected in this project are used to assist the verification of ranking in chi square distance only, which explains the combination of LBP and structure has the highest recognition rate compared to other approaches.

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.1 CONCLUSION

The training data is very important as it will determine the recognition rate. The LBP histograms for the printed numbers in logical format are analyzed and only 32 bins contain useful information instead of 256 bins. LBP is found to be performed poorly as it does not provide all the salient structures of the digits such as circle detection, crossing junction and location of horizontal and vertical lines. Instead, LBP can be used to find the closest match of the sample with the training data with the aid of chi square distance and verification of salient structure of digits to increase the recognition rate.

5.2 RECOMMENDATION

Future work includes studying other methods such as k-nearest neighbor, standard deviation, mean and Bayes theorem to improve the classification of the features extracted by LBP. Furthermore, future improvement of such technique may aid the retrieval of value and data from handwritten forms such as cheques, final exam script, admission forms and many more.

REFERENCES

1. Plamondon, R. and S.N. Srihari, *Online and off-line handwriting recognition: a comprehensive survey.* Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2000. **22**(1): p. 63-84.
2. Taylor, G.S., *Method of optical mark recognition*. 2004, Google Patents.
3. Julsing, B., *Face Recognition with Local Binary Patterns.* Research No. SAS008-07, University of Twente, Department of Electrical Engineering, Mathematics & Computer Science (EEMCS), 2007.
4. Ahonen, T., A. Hadid, and M. Pietikainen, *Face description with local binary patterns: Application to face recognition.* Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2006. **28**(12): p. 2037-2041.
5. Huang, D., et al., *Local binary patterns and its application to facial image analysis: a survey.* Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 2011. **41**(6): p. 765-781.
6. Pietikäinen, M., et al., *Local binary patterns for still images*, in *Computer Vision Using Local Binary Patterns*. 2011, Springer. p. 13-47.
7. Guo, Z., D. Zhang, and X. Mou. *Hierarchical multiscale LBP for face and palmprint recognition*. in *Image Processing (ICIP), 2010 17th IEEE International Conference on*. 2010. IEEE.
8. Guo, Z., L. Zhang, and D. Zhang, *Rotation invariant texture classification using LBP variance (LBPV) with global matching.* Pattern recognition, 2010. **43**(3): p. 706-719.
9. Noh, S. *Metric learning for nearest neighbor classification and its analysis*. in *Pattern Recognition (ICPR), 2012 21st International Conference on*. 2012.
10. Eikvil, L., *Optical Character Recognition.* citeseer. ist. psu. edu/142042. html, 1993.
11. Cheriet, M., et al., *Character recognition systems: a guide for students and practitioners*. 2007: John Wiley & Sons.
12. Smith, R. *An Overview of the Tesseract OCR Engine*. in *ICDAR*. 2007.
13. Breuel, T. *Recent progress on the OCRopus OCR system*. in *Proceedings of the International Workshop on Multilingual OCR*. 2009. ACM.
14. Hui, D., W. Feng, and L. Bo. *A Low-Cost OMR Solution for Educational Applications*. in *Parallel and Distributed Processing with Applications, 2008. ISPA '08. International Symposium on*. 2008.
15. Ghugardare, R.P., et al. *Optical character recognition system for seven segment display images of measuring instruments*. in *TENCON 2009 - 2009 IEEE Region 10 Conference*. 2009.
16. Xianjing, W. and A. Sajjanhar. *Using a circular grid for offline handwritten character recognition*. in *Image and Signal Processing (CISP), 2011 4th International Congress on*. 2011.
17. Wang, q., A. Yang, and W. Dai. *An improved feature extraction method for individual offline handwritten digit recognition*. in *Intelligent Control and Automation (WCICA), 2010 8th World Congress on*. 2010.
18. Gupta, A., M. Srivastava, and C. Mahanta. *Offline handwritten character recognition using neural network*. in *Computer Applications and Industrial Electronics (ICCAIE), 2011 IEEE International Conference on*. 2011. IEEE.
19. Hallale, S.B. and G.D. Salunke, *OFFLINE HANDWRITTEN DIGIT RECOGNITION USING NEURAL NETWORK.* International Journal of

Advanced Research in Electrical, Electronics and Instrumentation Engineering, 2013. **2**(9).

20. Liu, C.-L., et al., *Handwritten digit recognition: benchmarking of state-of-the-art techniques.* Pattern Recognition, 2003. **36**(10): p. 2271-2285.
21. Farrahi Moghaddam, R. and M. Cheriet, *AdOtsu: An adaptive and parameterless generalization of Otsu's method for document image binarization.* Pattern Recognition, 2012. **45**(6): p. 2419-2431.
22. Zheng, W., et al. *Feature Extraction of X-ray Fracture Image and Fracture Classification*. in *Artificial Intelligence and Computational Intelligence, 2009. AICI '09. International Conference on*. 2009.

APPENDICES

APPENDIX I – SCORE SHEET TEMPLATE

FINAL YEAR PROJECT: HANDWRITTEN NUMBER RECOGNITION

SCORE SHEET TEMPLATE

Notes:

- Please use **PEN** or **2B PENCIL** only.
- Please write **NUMBERS ONLY** in the space provided.

Student's ID (5 digits)	Score (0 to 100)

APPENDIX II – PROJECT TIMELINE (GANTT-CHART)

Detail / Week	FYP I (January 2014)														FYP II (May 2014)													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Selection of project topic																												
Literature review (LBP & OCR)																												
Collection of project samples																												
Preparation of training data																												
Design of basic LBP coding																												
Identification of mark’s location																												
Segmentation of characters																												
Testing using various LBP techniques																												
Implementation of OCR system																												
Addition of structural features																												
Analysis and evaluation on recognition rate																												

APPENDIX III – KEY PROJECT MILESTONES

Detail / Week	FYP I (January 2014)														FYP II (May 2014)														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Submission of extended proposal						■																							
Completion of training data								■																					
Proposal defense (presentation)									■																				
Completion of basic LBP coding										■																			
location identification program												■																	
Submission of interim draft report													■																
Submission of interim report														■															
Characters segmentation program														■															
Various LBP techniques testing result																				■									
Submission of progress report																					■								
Structural feature program																							■						
Completion of OCR system																							■						
Pre-SEDEX																								■					
Submission of draft final report																											■		
Submission of dissertation (soft bound)																												■	
Submission of technical paper																												■	
Viva																													■
Submission of project dissertation (hard bound)																													■