Automatic Assessment Mark Entry System Using Local Binary Pattern (LBP) and Salient Structural Features

Lim Lam Ghai[1], Suhaila Badarol Hisham[2], Norashikin Yahya[3]
Department of Electrical and Electronics Engineering
Universiti Teknologi PETRONAS
Tronoh, Malaysia
limlamghai@gmail.com[1], suhaiba@petronas.com.my[2], norashikin_yahya@petronas.com.my[3]

***Abstract*—Offline handwritten digit 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 time consuming, not cost effective and prone to human mistakes. Thus, 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 training data contain three sets of LBP values 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 scaling the digits '0 to 9' from the highest to the lowest similarity score as compared with the sample using chi square distance. The recognition rate can be greatly improved to about 95% by verifying the ranking of chi square distance with the salient structural features of digits.**

***Index Terms*—handwritten recognition, local binary pattern, chi square distance, structural feature**

I. Introduction

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

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 [2]. Generally, OCR can be divided into three main stages which are preprocessing of the input, feature extraction and classification of the characters [3]. The principle of handwritten recognition is similar with OCR.

Feature extraction is one of the processes in handwritten recognition 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 [4]. 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 [4].

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% [5].

A 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 where the binary image of handwritten character is transformed into points in polar coordinates [6]. Besides, the concept of eigenvalues has been used as features extraction for individual offline handwritten digit recognition in which this method is named as 3 points feature extraction [7].

In addition, 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 [8]. An experimental result shows that conventional features with back propagation network yields recognition accuracy of 91.2% [8]. 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 by combining eight classifiers with ten feature vectors [9]. 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) [9].

II. Brief Review of LBP

LBP is an image operator used to describe and characterize the texture patterns of an image in a binary number [10] and it transforms the gray scale image into an array which commonly represented in histogram [11].

The basic LBP operates in a 3 x 3 block of a pixel which consists of 8 neighbors [12]. Its center gray value, denoted as g_c is the threshold value for the surrounding neighborhood [5]. 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 Fig. 1.

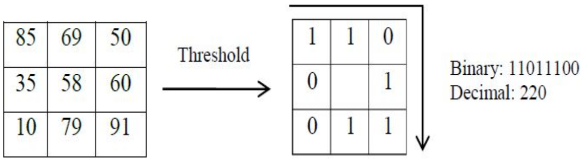

Fig. 1. The basic LBP operator.

The above LBP can be generalized to include more sampling points and radius as shown in Fig. 2. Let g_n denotes the gray value of the sampling point in an evenly spaced circular neighborhood of P sampling points and radius R around the center point (x, y) [12]. Its coordinate is $(x + R\cos(2\pi p/P),\ y - R\sin(2\pi p/P))$ where $n = 0, \ldots\ldots, P-1$. The coordinate of g_n can be estimated using bilinear interpolation if the sampling point is not located in the center of the block [12].

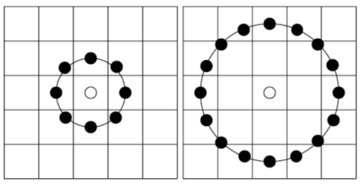

Fig. 2. The circular (8,1) and (16,2) neighborhoods [12].

Generally, the LBP value for pixel (x_c,y_c) can be calculated as follow where $s(z)$ is a signum function [12]:

$$\mathrm{LBP}_{\mathrm{P,R}}(x_c, y_c) = \sum_{n=0}^{P-1} s(g_n - g_c)2^n \tag{1}$$

$$s(z) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0 \end{cases} \tag{2}$$

Given an image of size N x M. The whole texture image can be represented in histogram for each pixel (i, j) using Equation (3) and (4) [13]:

$$\mathrm{H}(k) = \sum_{i=1}^{N}\sum_{j=1}^{M} f\left(\mathrm{LBP}_{\mathrm{P,R}}(i,j), k\right),\ k \in [0, 2^{\mathrm{P}} - 1] \tag{3}$$

$$f(x,y) = \begin{cases} 1, & x = y \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

There are several mapping types for LBP labels; namely uniform patterns, rotational invariance and rotation invariance uniform pattern [12], [14]. 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 [12]. 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 [5], [11]. The non-uniform output labels are omitted and this reduces the number of output labels for P bits to P(P – 1) + 3 [12].

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 [12]. It has been proved that chi square distance performs better than other methods due to its stability when dealing with small sample sizes [15]. For that reason, the dissimilarity between S and M can be calculated as follow using chi square distance [13]:

$$\chi^2(\mathrm{S,M}) = \sum_{b=1}^{B} \frac{(\mathrm{S}_b - \mathrm{M}_b)^2}{\mathrm{S}_b + \mathrm{M}_b} \tag{5}$$

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 [5].

III. Salient Structural Features of Digits

The salient structural features of digits are defined in TABLE I and it is not related with LBP. The reasons of adding these structural features are explained in section V. Circle is one of the salient structural features of digits 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 across the binary image as it is split into half vertically is measured by calculating the total times of pixel transition from value 0 to 1 [7]. 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. Digit ‘8’ is used as 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 a closed loop. This concept also applies to the left side pixels between node 1 and node 2. It can be safely assumed that 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 3 illustrates the application of this concept on digits ‘3’ and ‘8’ and it is also applied to digits ‘2’, ‘5’, ‘6’, ‘9’ and ‘0’.

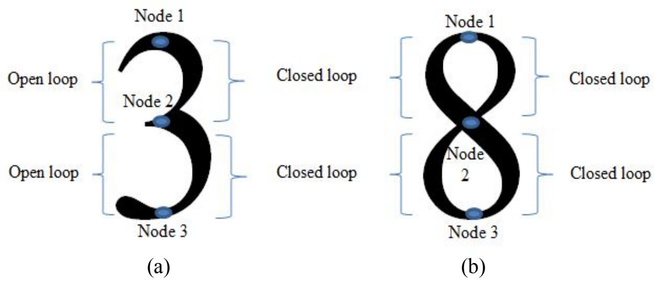

Fig. 3. Circle detection on (a) digit 3, (b) digit 8.

TABLE I. 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

IV. METHODOLOGY

Figure 4 shows the process flow of offline handwritten digit recognition using LBP and salient structural features of digits as feature extraction methods. The scanned image of assessment form is converted into black and white image using Otsu's method [16]. The pixel location of a student's identity number and the associated score are identified.

Next, the numbers are segmented and image noise is removed. Digits are then corrected for slant and the size of segmented digit is normalized before extracting its features. 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 score. These outputs are then verified using the salient structural features of the digits going through the ranking until a match is found.

A score sheet in A4 size is designed to collect the samples of handwritten digits. 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 and 50 score sheets are randomly distributed to lecturers and undergraduate students of Universiti Teknologi PETRONAS. The average handwritten digit recognition rate and execution time of a score sheet are calculated, recorded, tabulated and analyzed.

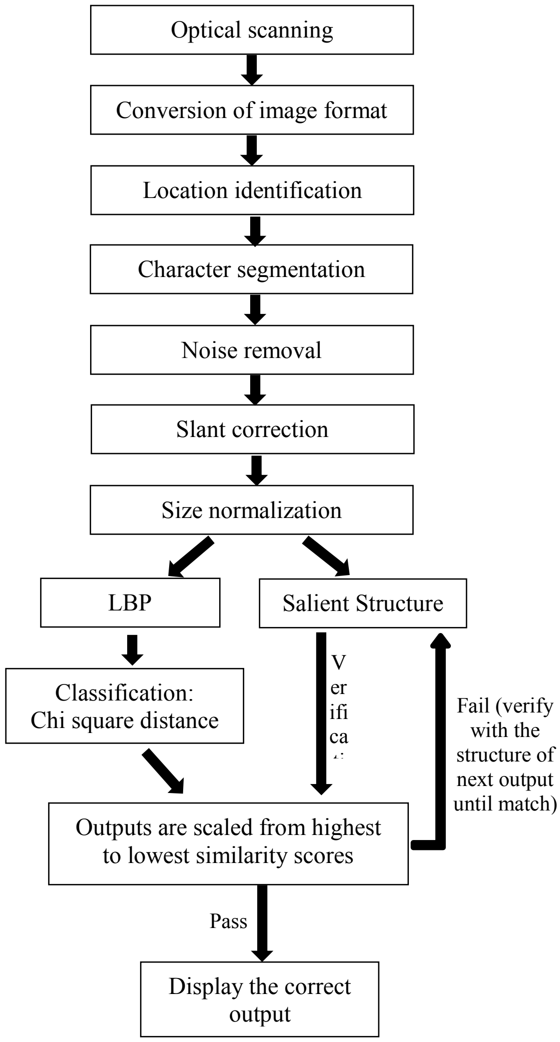

Fig. 4. Offline handwritten digit recognition process.

V. RESULT AND DISCUSSION

A. Implementation of LBP on Binary Image

The basic LBP using 8 pixels in a 3 x 3 pixel block is implemented on the binary image of handwritten digit where the number of sampling points P is 8 with radius R of 1 around center pixel as shown in Fig. 5. 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

Fig. 5. Description of 3 x 3 pixel blocks using LBP.

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

$$
\begin{aligned}
LBP_{P,R}(x_c, y_c) = {} & s(g_0 - g_c) + s(g_1 - g_c)2 + s(g_2 - g_c)4 \\
& + s(g_3 - g_c)8 + s(g_4 - g_c)16 \\
& + s(g_5 - g_c)32 + s(g_6 - g_c)64 \\
& + s(g_7 - g_c)128
\end{aligned} \tag{5}
$$

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 remaining bins are omitted due to the following 3 cases:

Case 1: when $g_c = 0$

- The signum function $s(g_n - g_c)$ is always equal to 1 when $g_c = 0$ because $g_n - g_c \geq 0$. Hence, LBP value is 255 and this feature is stored in bin 255. This information is not much of use because it only indicates the black spaces surrounding the handwritten digit.

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

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

Case 3: when the uniformity measure is more than 2

- 90% of the binary images of handwritten digits have uniform pattern where the uniformity measure is at most only 2. The non-uniform patterns are omitted because they do not have useful information to be used in identifying the digits. Moreover they may affect the overall result if taken into consideration.

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 II. The binary value of the center pixel is equal to 1 for all these 32 bins.

The LBP histogram is now simplified to only these 32 bins instead of 256 bins. The simplified histogram is then normalized to 100% by dividing the value of each bin with the summation of values in 32 bins. Figure 6 illustrates the binary image of printed digits from 0 to 9 and Figure 7 shows the simplified version of LBP histograms for printed digits ‘1’, ‘5’ and ‘8’. Henceforth in the discussion, the new bin numbers will be used.

Based on the LBP histogram, majority of LBP values for all the printed digits are located at bin 5, bin 26, bin 11 and bin 19. This is because bin 5 and bin 26 represent the vertical lines of the digit whereas bin 11 and bin 19 represent the horizontal lines of the digit. As a result, digits ‘0’ and ‘1’ have highest percentage on bin 5 and bin 26 while digits ‘2’, ‘3’, and ‘5’ have highest percentage on bin 11 and bin 19.

TABLE II. THE IMPORTANT BINS AND IT BINARY VALUES

New bin value	LBP bin	Binary value ($g_7g_6g_5g_4g_3g_2g_1g_0$)	New bin value	LBP bin	Binary value ($g_7g_6g_5g_4g_3g_2g_1g_0$)
1	7	00000111	17	193	11000001
2	15	00001111	18	195	11000011
3	28	00011100	19	199	11000111
4	30	00011110	20	207	11001111
5	31	00011111	21	223	11011111
6	60	00111100	22	225	11100001
7	62	00111110	23	227	11100011
8	63	00111111	24	231	11100111
9	112	01110000	25	240	11110000
10	120	01111000	26	241	11110001
11	124	01111100	27	243	11110011
12	126	01111110	28	247	11110111
13	127	01111111	29	248	11111000
14	135	10000111	30	249	11111001
15	143	10001111	31	252	11111100
16	159	10011111	32	253	11111101

Fig. 6. Binary image of printed digits.

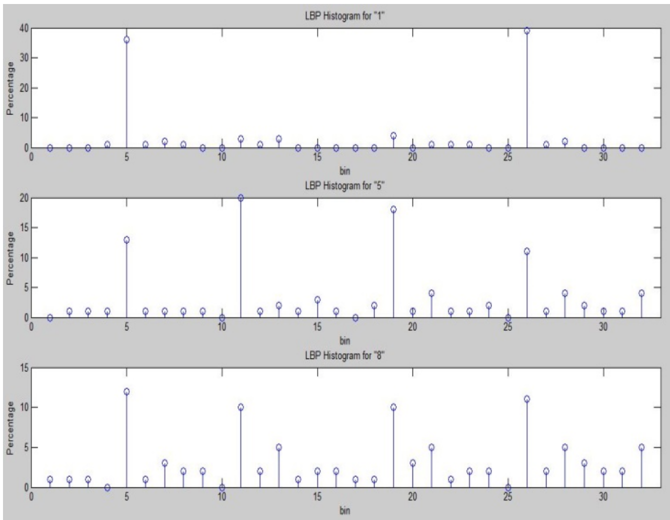

Fig. 7. LBP histograms of digits ‘1’, ‘5’ and ‘8’.

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 digits and these features are distributed in the identified 28 bins. A case example in the opposite, digit ‘1’ has the least curvy shape and hence only a small percentage falls in those 28 bins. Generally, the percentages distributed on these 32 bins contain the features of digits such as horizontal straight line, vertical straight line and curvy shape. However, the exact locations of these features on the digits are not known.

B. Training Data of Handwritten Digits

Figure 8 displays three different styles of handwritten digits '1', '4', and '7' - labeled 'a' to 'i' - and their respective LBP percentages are tabulated in TABLE III. The LBP percentage of vertical line for 'b' and 'c' are higher as compared to 'a' because 'a' has an extra horizontal line for digit '1'. Digit '4' as shown in 'd' has higher LBP percentage on vertical line due to two straight lines whereas 'e' and 'f' have higher LBP percentages on curvy line, especially in bins 13 and 28. On the other hand, digit '7' shown in 'h' has higher LBP percentage on horizontal line and 'g' has very small LBP percentage on curvy line.

C. Handwritten Digit Recognition Rate based on LBP

Chi square distance is used as a 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.

Based on the results, the average recognition rate is low, about 45% and about 75% 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

- The LBP histograms of all digits have almost similar pattern. For instance, the LBP histograms of digits '6' and '9' are quite closely matched.
- LBP cannot detect circle, crossing point, and is unable to determine the exact location of vertical and horizontal straight lines. It only shows the number of occurrence of that particular binary number within the sampling points.
- A slanted handwritten number would change the distribution of LBP values in those 32 bins.

D. 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 IV. 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 test handwritten digits have huge variations from the digits used in the training data. This causes them to fall in the ranking below three. However, the recognition rate is greatly improved to about 95% by verifying all the rankings until the structural features are matched. The drawback of this approach is the program execution is slightly increased.

The recognition rate using structural features only is considered high, about 90% 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 structures of digits selected in this project are only used to assist the verification of ranking in chi square distance. This explains why the combination of LBP and structure has the highest recognition rate compared to other approaches.

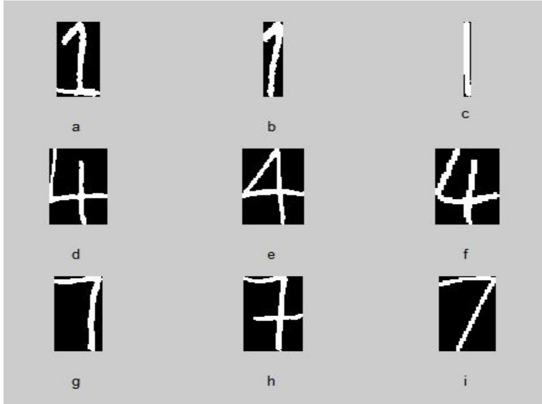

Fig. 8. Three different styles of handwritten digits '1', '4' and '7'.

TABLE III. LBP PERCENTAGE OF HANDWRITTEN DIGITS DISPLAYED IN FIGURE 8

Bin value	LBP percentage (%)								
	a	b	c	d	e	f	g	h	i
1	1	1	0	0	0	1	1	0	0
2	1	0	1	1	1	1	0	0	0
3	1	2	0	0	1	1	0	1	2
4	2	2	0	1	3	1	2	1	2
5	17	26	42	28	16	20	22	18	11
6	0	1	0	1	0	1	1	1	1
7	3	2	0	1	3	2	1	1	4
8	3	3	0	1	4	2	2	1	4
9	1	1	0	1	0	0	0	1	0
10	0	0	1	0	1	1	0	1	0
11	9	2	1	9	8	7	13	16	11
12	1	2	0	1	1	2	1	2	3
13	6	5	1	3	7	6	4	3	9
14	1	0	1	0	1	1	0	2	0
15	1	1	1	1	1	1	1	1	0
16	2	1	0	1	1	1	1	0	1
17	1	2	0	1	0	1	1	1	3
18	1	1	1	1	1	2	1	1	1
19	9	3	0	10	8	6	11	13	11
20	1	1	0	0	1	1	1	2	0
21	4	2	2	2	2	3	2	3	1
22	1	2	0	1	3	2	2	1	2
23	2	2	0	1	3	2	1	1	4
24	1	2	0	1	1	2	1	2	3
25	1	1	1	1	1	1	0	0	0
26	18	24	41	27	17	19	24	18	11
27	2	3	0	1	3	2	2	2	4
28	5	5	1	2	7	6	4	4	9
29	1	1	0	1	0	1	0	0	0
30	2	2	0	1	1	1	0	1	1
31	1	1	1	1	1	1	0	1	0
32	3	2	2	2	2	3	1	2	1

TABLE IV. RECOGNITION RATE OF HANDWRITTEN DIGIT 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

VI. CONCLUSION AND RECOMMENDATION

In this paper, handwritten digit recognition is used in mark entry system to solve the problems arise in conventional method of extracting the mark. The problems are time consuming, not cost effective and prone to human mistakes when assigning a person to key in the students' identity numbers and marks into a computer.

The LBP histograms for the digits in binary format are analyzed and only 32 bins contain useful information instead of 256 bins. LBP is found to perform 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 verification of salient structure of digits to increase the recognition rate. An average recognition rate of about 50% is achieved using LBP only as a feature extraction method and it is greatly improved to about 95% by combining LBP and salient structure of digits.

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 implement 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.

ACKNOWLEDGMENT

The authors would like to thank the Electrical & Electronics Department of Universiti Teknologi PETRONAS and its Final Year Project committee for their support and guidance throughout the research.

REFERENCES

[1] R. Plamondon and S. N. Srihari, "Online and off-line handwriting recognition: a comprehensive survey," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 22, pp. 63-84, 2000.

[2] L. Eikvil, "Optical Character Recognition," *citeseer. ist. psu. edu/142042. html,* 1993.

[3] G. S. Taylor, "Method of optical mark recognition," ed: Google Patents, 2004.

[4] M. Cheriet, N. Kharma, C.-L. Liu, and C. Suen, *Character recognition systems: a guide for students and practitioners*: John Wiley & Sons, 2007.

[5] B. Julsing, "Face Recognition with Local Binary Patterns," *Research No. SAS008-07, University of Twente, Department of Electrical Engineering, Mathematics & Computer Science (EEMCS),* 2007.

[6] W. Xianjing and A. Sajjanhar, "Using a circular grid for offline handwritten character recognition," in *Image and Signal Processing (CISP), 2011 4th International Congress on*, 2011, pp. 945-949.

[7] q. Wang, 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, pp. 6327-6330.

[8] S. B. Hallale and G. D. Salunke, "OFFLINE HANDWRITTEN DIGIT RECOGNITION USING NEURAL NETWORK," *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering,* vol. 2, 2013.

[9] C.-L. Liu, K. Nakashima, H. Sako, and H. Fujisawa, "Handwritten digit recognition: benchmarking of state-of-the-art techniques," *Pattern Recognition,* vol. 36, pp. 2271-2285, 2003.

[10] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 28, pp. 2037-2041, 2006.

[11] D. Huang, C. Shan, M. Ardabilian, Y. Wang, and L. Chen, "Local binary patterns and its application to facial image analysis: a survey," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on,* vol. 41, pp. 765-781, 2011.

[12] M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen, "Local binary patterns for still images," in *Computer Vision Using Local Binary Patterns*, ed: Springer, 2011, pp. 13-47.

[13] Z. Guo, D. Zhang, and X. Mou, "Hierarchical multiscale LBP for face and palmprint recognition," in *Image Processing (ICIP), 2010 17th IEEE International Conference on*, 2010, pp. 4521-4524.

[14] Z. Guo, L. Zhang, and D. Zhang, "Rotation invariant texture classification using LBP variance (LBPV) with global matching," *Pattern recognition,* vol. 43, pp. 706-719, 2010.

[15] S. Noh, "Metric learning for nearest neighbor classification and its analysis," in *Pattern Recognition (ICPR), 2012 21st International Conference on*, 2012, pp. 991-995.

[16] R. Farrahi Moghaddam and M. Cheriet, "AdOtsu: An adaptive and parameterless generalization of Otsu's method for document image binarization," *Pattern Recognition,* vol. 45, pp. 2419-2431, 2012.