Detection on Straight Line Problem in Triangle Geometry Features for Digit Recognition

N. A. Arbain#, M. S. Azmi#, S. S. S. Ahmad#, R. Nordin*, M. Z. Mas'ud#, M. A. Lateh#

#Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, Durian Tunggal, 76100 Melaka, Malaysia
E-mail: nuratikah9.arbain@gmail.com, {sanusi, sakinah, zaki.masud}@utem.edu.my, masitah.a.lateh@gmail.com

*Centre for Languages and Human Development, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, Durian Tunggal, 76100 Melaka, Malaysia
Email: radzuan@utem.edu.my

Abstract—Geometric object especially triangle geometry has been widely used in digit recognition area. The triangle geometry properties have been implemented as the triangle features which are used to construct the triangle shape. Triangle is formed based on three points of triangle corner A, B, and C. However, a problem occurs when three points of triangle corner were in parallel line. Thus, an algorithm has been proposed in order to solve the straight line problem. The Support Vector Machine (SVM) and Multi-Layer Perceptron (MLP) were used to measure based on the classification accuracy. Four datasets were used: HODA, IFCHDB, MNIST, and BANGLA. The comparison results in classification demonstrated the effectiveness of our proposed method.

Keywords—classification; digit recognition; feature extraction; image processing; triangle geometry

I. INTRODUCTION

Research in digit recognition has been actively explored since four decades ago [1]. Many methods and techniques have been used to recognize digit recognition such as Hidden Markov Model, Neural Network, et cetera. Nowadays, triangle method has become a popular method that has been used in recognizing the digit images. The study in digit recognition has been expanded due to demand to identify the pattern, style, and signature of handwriting. Generally, the study in digit recognition has been explored in numerous languages, for example, Hebrew, Roman, Chinese, Indian, Jawi, Arabic, and Latin.

The Modified NIST dataset (MNIST) is a ROMAN handwriting that has known as the largest dataset compared to HODA dataset. The MNIST dataset was developed as a result of handwritten digit classification competition which has been held in summer of 1992 [2]. The studies of digit recognition had evolved rapidly along with improvement made on prior methods and techniques [3]–[10].

In 2009, characterization loci and mixture of experts were proposed by [11] in order to recognize the Farsi handwritten digits. The characterization loci features were used as a main feature and mixture of experts as the classification stage. The characteristic loci features were extracted while mixture experts were used for recognizing the digit handwriting. The HODA database was used to be extracted using proposed method by [11] in the experiment.

[12] Proposes new features from triangle properties by using the triangle geometry method to extract digit images. Geometrical properties are the medium to form geometric by modeling them as an object. In object recognition, the properties of geometry can be implemented. The triangle geometry method also has been widely applied in biometric research such as face recognition and fingerprint recognition [13]–[16]. Besides, the triangle geometry also had been implemented in intrusion, vehicle detection and digit recognition [12], [17]–[20]. Research in digit recognition has been conducted on several datasets such as MNIST [2], HODA [33] and IFCHDB [34].

This paper focuses on solving straight line problem that occurs when three triangle points are inline. Thus, the proposed method to solve the straight line problem is proposed in this paper.
This paper is organized as follows. In Section II, we discuss the material and method used in this paper. In Section III, the result and discussion are covered. Finally, in Section IV, we conclude the paper.

A. Triangle Geometry

Triangle shape forms based on three points where the three segments are linking to each other. Besides, it also is known as a three-sided polygon. The types of triangles are divided into four types, namely equilateral triangle, isosceles triangle, right triangle, and scalene triangle. For scalene triangles, there are three types of scalene triangles namely acute scalene triangle, obtuse scalene triangle, and isosceles right triangle. The triangle properties are divided into eight types which are vertex, base, altitude, median, area, perimeter, interior angles and exterior angles [21]. In triangle geometry, each of triangle type has different properties. For example, three angles of an equilateral triangle are equal which is congruent to each other. Unlike isosceles triangle, it only has two equal angles and two equal sides.

The triangle properties can be used to produce some features vector. Vertex or vertices are a corner of the triangle. Every type of triangle has three vertices. Base in context of the triangle is a drawn at the bottom. The base of a triangle can be any one of the three sides. Usually, the base is used as a reference side for calculating the area of a triangle. For example, the base of an isosceles triangle is at unequal side. Altitude of a triangle is the perpendicular from the base to the opposite of vertex. Since there are three possible bases, thus there are also three possible altitudes. These altitudes are intersecting at a single point, which called as orthocenter of the triangle. Orthocenter of a triangle also known as one of a triangle’s points of concurrency. Commonly, orthocenter is not always within the triangle. For example, the orthocenter for obtuse is at outside. Nevertheless, there are three other types of triangle centers namely incenter, circumcenter, and centroid.

The median of a triangle is a contour from a vertex to the midpoint of the contradictory side. Since there are three points of corner thus there are also three medians that meet at a single point which called as centroid of the triangle. Area of a triangle is the number of square units it takes to fill the interior of a triangle. The perimeter of a triangle is the distance around the triangle. The interior angles of a triangle are the angle between a side of a triangle and the extension of an adjacent side. Lastly, exterior angles of a triangle are the angle between a side of a triangle and the extension of an adjacent side.

The triangle points of the corners can be calculated using the basic formula, which is Pythagorean Theorem. The theorem was a fundamental importance in Euclidean Geometry where it assists as a root for the description of distance between two points [22]. Fig. 1 shows the triangle shape using Pythagorean Theorem. In algebraic terms, \( a^2 + b^2 = c^2 \) where \( c \) is the hypotenuse while \( a \) and \( b \) are the legs of a triangle.

B. Feature Extraction

Feature extraction is an important process because the extracted features are crucial in representing an object.
Before performs the feature extraction process, images need to be processed by transforming them into binary form through a binarization process. The binarization process was performed in the pre-processing stage. The outcome of the pre-processing stage will be used as the input to the feature extraction process.

In [12], the dataset images have been converted into binary form through a binarization process in a pre-processing stage. In binarization process, the Otsu threshold method [35] is used to convert the images into a binary form. Then, the outcome of the pre-processing stage is used as an input to the next process, which is feature extraction. In the feature extraction, the zoning method is used to extract the features from triangle properties. The zoning method is also known as 33 zones [12]. The 33 zones are represented as 33 triangles for an image.

The zoning method used are Cartesian Plane Zone (CPZ), Vertical Zone (VZ), Horizontal Zone (HZ) and Zoning Based 45 Degree (Z45d). Each of zoning method produces a different total number of features. The total number of features produced by CPZ is 45 features, VZ is 54 features, HZ is 126 features and Z45d is 72 features. Therefore, the total number of features for zoning method is 297 features.

On the other hand, the total number of triangles is produced as 33 triangles which have 297 features for an image. According to [12], one triangle can produce nine features. These features are angle, ratio, and gradient. The formula can be expressed for each of triangle points as shown in Table 1, Table 2, and Table 3. An example of a triangle object denoted A, B and C is shown in Fig. 3. Next, the formula to calculate the length of a triangle is shown in equation 4, 5 and 6.

$$a = \sqrt{(A1(y) - C(y))^2 + ((A1(x) - B1(x))^2)}$$ (4)

$$b = \sqrt{(B1(y) - B(y))^2 + ((A(x) - B1(x))^2)}$$ (5)

$$c = \sqrt{(C1(y) - C(y))^2 + ((A(x) - C1(x))^2)}$$ (6)

C. Straight Line Problem in Constructing Triangle Shape

A straight line problem triggered when the chosen points of geometry in a triangle shape do not comply with the geometrical properties. This also can be referred when the value of triangle’s gradient is zero. Hence, a formation of a triangle geometry cannot be formed when the coordinates of point A, B and C are in one line.

In the formation of a triangle, the corner points of triangle A, B and C need to be in different positions to construct a triangle line between the points. A problem in constructing a triangle shape is located at three triangle points that may cause the straight line occurs when constructing the triangle shape.

There are six shapes whereas the triangle geometry’s coordinates can be formed [12]. This also can be referred as a normal condition when forming a triangle shape. For an abnormal condition, there are four possible conditions that may trigger the straight line problem when forming the triangle. Fig. 4 shows the normal condition when forming a triangle shape while Fig. 5 shows the abnormal condition of straight line problem that occurs in forming a triangle shape. Table 4 shows the position of y coordinate for point C.

<table>
<thead>
<tr>
<th>Position</th>
<th>Position of y coordinate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>yA ≥ yC ≥ yB</td>
</tr>
<tr>
<td>2</td>
<td>yA ≥ yB ≥ yC</td>
</tr>
<tr>
<td>3</td>
<td>yA ≤ yB ≤ yC</td>
</tr>
<tr>
<td>4</td>
<td>yA ≤ yC ≤ yB</td>
</tr>
<tr>
<td>5</td>
<td>yA ≥ yB ≤ yC</td>
</tr>
<tr>
<td>6</td>
<td>yA ≥ yC ≤ yB</td>
</tr>
</tbody>
</table>
II. MATERIAL AND METHOD

A. Data Collection

In this paper, four types of digit dataset were used. There are as follows:
- HODA [33]
- IFCHDB [34]
- MNIST [2]
- BANGLA [36]

In 1992, MNIST dataset was developed and known as the National Institute of Standard Technology (NIST) in beginning. The dataset, then was naming as MNIST after some improvement were made [2]. The MNIST dataset is the handwritten Roman digits. The sample of MNIST dataset can be downloaded at http://yann.lecun.com/exdb/mnist/. An example of MNIST dataset has shown in Fig. 8.

HODA dataset was developed in 2005. The article about it was published in 2007 by [33]. HODA dataset contains Arabic handwriting. The sample of HODA dataset can be downloaded at http://FarsiOCR.ir. An example of HODA dataset has shown in Fig. 6.

In 2006, IFCHDB dataset was introduced by the Department of Electric Engineering, AmirKabir University in Tehran [34]. IFCHDB is known as Isolated Farsi Character Handwritten Database. IFCHDB dataset contains characters and digits in Farsi/Arabic which were gained from high school examination entry forms [34]. The IFCHDB can be acquired by filling the agreement form before the samples are sent through email. An example of IFCHDB dataset has shown in Fig. 7.

BANGLA dataset was developed by Pattern Recognition Unit and Vision Computer, Institute of Indian Statistic. The BANGLA dataset consists of digit Bangla handwriting. Similar as IFCHDB dataset, an agreement form is required to be filled in before the samples will be sent through email. An example of BANGLA dataset has shown in Fig. 9.

Each dataset consists of 10 test classes and 10 training classes. Besides, these datasets have general characteristic and contain the benchmark for this study. Table 5 shows the description for each dataset.
### TABLE V

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution Sample</td>
<td>200 dpi</td>
<td>300 dpi</td>
<td>200 dpi</td>
<td>300 dpi</td>
<td></td>
</tr>
<tr>
<td>Scale</td>
<td>Binary</td>
<td>Gray</td>
<td>Binary</td>
<td>Gray</td>
<td></td>
</tr>
<tr>
<td>Total of Sample</td>
<td>102,352</td>
<td>70,120</td>
<td>102,352</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Number of digits</td>
<td>-</td>
<td>17,740</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Number of training sample</td>
<td>56,790</td>
<td>12,292</td>
<td>60,000</td>
<td>19,392</td>
<td></td>
</tr>
<tr>
<td>Number of test sample</td>
<td>20,000</td>
<td>5,268</td>
<td>10,000</td>
<td>4,000</td>
<td></td>
</tr>
</tbody>
</table>

![Fig. 10 The flow of pre-processing phase](image)

#### B. Proposed Method

The proposed method was applied to solve the straight line problem when forming the triangle shape for images. The proposed method was developed using java code and stand-alone as a basic platform.

A straight line detection was performed by identifying which coordinates of point A, B and C would lead to the straight line problem. The straight line problem was identified based on the position of y coordinate which has shown in Table 4. After identifying which points and coordinates, the proposed method would be applied to solve the straight line problem. The proposed method was applied in a pre-processing stage as shown in Fig. 10.

Based on Fig. 10, the pre-processing process began when the datasets were converted into a binary image using the Otsu’s method [35]. In binary form, the foreground image is known as ‘0’ while the background image is known as ‘1’ [27]. Then, the zoning method was applied to extract the features. The further explanation for zoning method can be referred in [12].

In this paper, the triangle features from [12] were used. The 297 features were produced after applying zoning method. Based on the explanation in Section I (B), 33 zones were represented as 33 triangles for an image. Each zone (33 zones) was divided into four parts. In the pre-processing stage (refer to Fig. 10), the digit image from HODA dataset was converted into a binary form.

For example, the binary form (refer to Fig. 10) was split into four zones using Cartesian plane zone. These four zones were named as Zone A, Zone B, Zone C and Zone D as shown in Fig. 10. Then, each zone was divided into four parts namely part1, part2, part3, and part4. In this case, part1, part2, part3 and part4 will be used in solving straight line problem. The number of pixels (only foreground images) for each partition will be count. The comparison will be performed between partition (part1, part2, part3 and part4) in order to produce new coordinate for selected triangle point (that has straight line problem). The new coordinate for selected triangle point will be determined by a total number of pixel of partition. The highest total of pixels will move the coordinate of x and y to up (add by one). However, if the total number of pixels were less, then the coordinate of x and y will be going down (minus by one). The proposed method is only applied to the triangle point that has a straight line problem. The triangle point that has no straight line problem will be skipped.

Table 6 shows the pseudo code for the proposed method which is used to cater the straight line problem.
III. RESULT AND DISCUSSION

The experiments were conducted using Support Vector Machine (SVM) and Multi-Layer Perceptron (MLP). For SVM technique, the cost and gamma values were obtained from a grid search using libSVM function. For MLP technique, the learning rate (LR) was 0.2 which was obtained using heuristic search from [12]. All datasets (HODA, IFCHDB, MNIST and BANGLA) were used in this experiment. The results of classification accuracy were compared with previous research. The comparison was made between [12] and the present proposed method. This is because the proposed method used in [12] implements the triangle geometry and zoning methods to extract the features. Table 7 shows the comparison results of classification accuracy.

Based on the findings, the results for the present proposed method using SVM technique showed some improvement compared to the result for proposed method in [12]. The classification accuracy for SVM technique increased for HODA (98.3342%), IFCHDB (95.794%) and BANGLA (90.95%). However, the result of MNIST dataset showed decreased which was 94.595%. The results for our proposed method using MLP technique produced poor results except for BANGLA dataset which has shown an improvement. The results of the classification accuracy might possibly be biased because of the pattern of handwritten itself.

Our proposed method was successfully applied to solve the straight line problem that has been triggered in a triangle formation. The results in Table 7 has shown good results for SVM technique while poor results were obtained using MLP technique. The formula used in MLP technique was complex which is one of the factors that influenced the results of classification accuracy. Besides that, the nature of handwritten itself also has effected the results.

IV. CONCLUSION

This paper presents the proposed method to cater the straight line problem in a triangle formation in order to increase the result of classification accuracy. The straight line problem was detected through three triangle points. By solving the straight line problem, the triangle shape can be remodeled. By focusing in digit recognition area, there are four datasets (HODA, IFCHDB, MNIST and BANGLA) which contain digit form used in this experiments.

Two techniques namely SVM and MLP have been used in the experiments. The results of classification accuracy demonstrate the effectiveness of our proposed method. In general, while SVM technique is more suitable to be applied for HODA, IFCHDB and BANGLA to measure the accuracy, MLP technique is suitable to measure the accuracy for BANGLA dataset only.

Feature extraction is an important factor in the presentation of digit recognition. An influential classifier such as SVM and MLP may produce much different accuracy based on the different features of patterns. The high recognition accuracy can be shaped by selecting the suitable low-complexity classifier and suitable data extraction. The best final result is attained not only from the combination of good classifier but also from the features itself. The enhancement on features supports to achieve high recognition accuracy. The proposed method is a suitable method to solve the straight line problem when triangle formation in digit images. The proposed method of this study can be used in other recognition work that is based on the pixel of images (binary form). Further research is needed to increase performance when constructing a triangle shape.

ACKNOWLEDGMENT

The authors would like to thank the Ministry of Education for funding the study through FRGS/2/2014/ICT/02/FTMK/02/F00246 and PJP/2016/PBPI/HI4/S01476. Besides, thank you to the Universiti Teknikal Malaysia Melaka and Faculty of Information Technology and Communication for providing excellent research facilities.