Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 2, No. 2, August 2016.

ENHANCEMENT OF TRIANGLE COORDINATES FOR TRIANGLE FEATURES FOR BETTER CLASSIFICATION

Nur Atikah Arbain[1], Mohd Sanusi Azmi[2], Laith Bany Melhem[3], Azah Kamilah Muda[4] and Hasan Rashaideh[5]

[1,2,3,4] Faculty of Information Technology and Communication, Universiti Teknikal Malaysia Melaka, Malaysia.

[5] Al-Balqa Applied University, Al-Salt, Jordan.

nuratikah9.arbain@gmail.com[1], sanusi@utem.edu.my[2], lnm1989@gmail.com[3], azah@utem.edu.my[4], rashaideh@bau.edu.jo[5]

(Received: 25-Nov.-2015, Revised: 14-Jan.-2016, Accepted: 27-Jan.-2016)

ABSTRACT

Recently, the triangle features have been applied in digit recognition by adopting the angle as a part of the features. Most of the studies in digit recognition area which applied these features have given impressive results. However, the issue of big gap values that occurred between angles, ratios and gradients has shown a strong impact on the accuracy of the results. Therefore, we introduce our proposed method which is data normalization that has adopted the nature of triangle geometry in order to resolve this issue. Besides, we have applied other techniques, such as Z-score, Minimax and LibSVM function in the experiment. There are four digit datasets used which are HODA, MNIST, IFHCDB and BANGLA. The results of classification have shown that our proposed method has given better results compared to other techniques.

KEYWORDS

Triangle features, Triangle geometry, Feature extraction, Feature normalization, Feature scaling.

1. INTRODUCTION

1.1 Feature Geometry

Feature geometry is a compound of the basic geometric building blocks (e.g., points, lines, curves, surfaces and polygons) [1]. It was developed by [2], where natural classes are represented by hierarchical structures as well as by features themselves which have represented a major revision of the theory proposed by [3].

Feature geometry is an assortment of points that may produce a kind of multipoint, polyline or polygon geometry. Points are defined and created using the x, y coordinates, M is the measured value and Z is the vertical location. Every geometry kind has a definition that helps verify its validity. The tool should be designed to consider the geometry types. Whether or not it changes the geometry doesn't matter, however it has to handle the choice.

1.2 Properties of Triangle

A triangle is a closed figure that consists of three points which the three line segments have linked end-to-end. It is also known as a three sided polygon. There are eight properties or features for a triangle, which are: vertex, base, altitude, median, area, perimeter, interior angles and exterior angles. Generally, there are three types of triangle, which are: equilateral triangle, isosceles triangle and scalene triangle [4]. Nevertheless, there are three other types of triangle which can be formed from a scalene triangle. These are: acute scalene triangle, obtuse scalene triangle and isosceles right triangle. Table 1 shows the description of types of scalene triangle, while Figure 1 shows examples of types of triangle.

Table 1. Description of types of scalene triangle.

Types of Scalene Triangle	Description
Acute Scalene Triangle	This triangle has three angles each of which is less than 90 degrees.
Obtuse Scalene Triangle	This triangle has an angle which is more than 90 degrees.
Isosceles Right Triangle	This triangle has one right angle (90°).

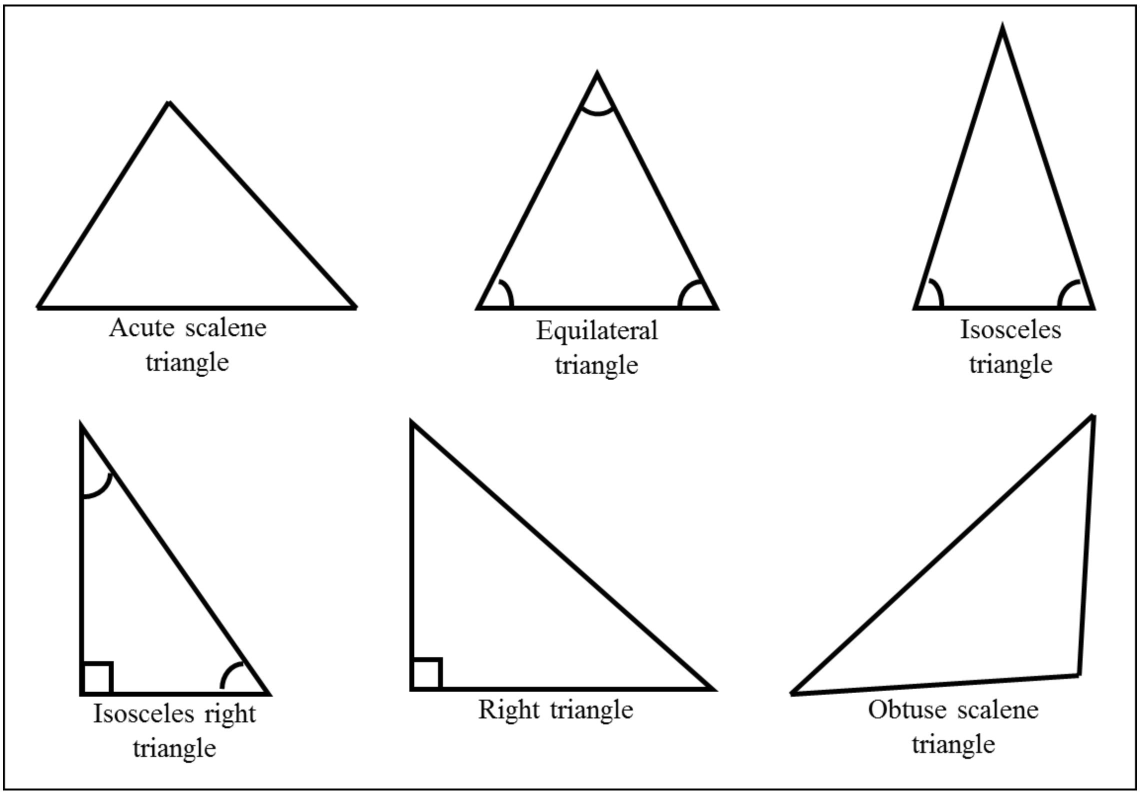


Figure 1. Examples of types of triangle.

Briefly, the vertex is known as a corner. It also refers to a point where lines meet. Usually, the vertex is used to mark the corners of a polygon. In addition, the included angle at each vertex is known as an interior angle of the polygon. Figure 2 shows an example of included angle [5].

However, in this paper, we are adopting scalene triangle, because the digit form is not symmetrical on the left and right sides like our eyes and nose, as stated in [6]. In with the [7], the author has reported that the Arabic digit and calligraphy do not have any pattern with the potential to form three edges for a triangle. Moreover, the author [7] has used nine features which will be described in detail in the feature extraction topic.

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 2, No. 2, August 2016.

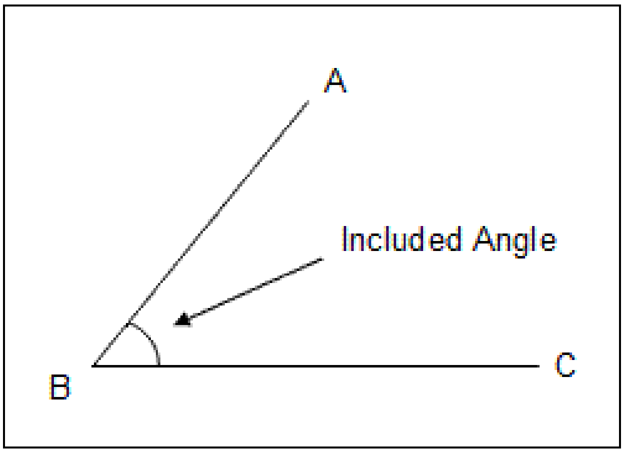


Figure 2. Included angle [5].

1.3 Feature Extraction

Feature extraction is an important task in image processing, because of the meaningful features extracted are vital in representing an object. An object is modelled and represented by geometrical properties. Geometry has properties that can be used in object recognition. The triangle properties have been adopted by many researchers to generate the proposed features for image classification. The triangle geometry is widely used in biometric research, such as face and fingerprint recognition [8]–[12]. Apart from that, other researchers have also adopted the similar geometry for intrusion [13], vehicle detection [14] and digit recognition [16]–[17].

In this paper, the triangle geometry features are used to extract the features from digit datasets. The triangle features have been proposed by [18]-[19]. These nine features of triangle are shown in Table 2.

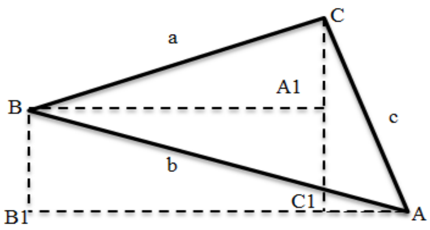
The features in Table 2 will be applied to the several zones in order to increase the accuracy of classification by using Support Vector Machine (SVM). In this case, the Cartesian Plane Zone has been chosen as the initial method in conducting the experiments. It also can be applied to other zoning methods. Therefore, this paper is focusing on how the triangle features are normalized or being scaled. In this paper, we also explicate the results of classification after using several techniques as aforementioned.

The triangle geometry used by [17]-[18] is based on the Scalene Triangle method. In [18], 21 features have been used based on the Scalene Triangle method. However, only three out of 21 features are directly used in the experiments. The features are: angles of corners which have been labeled as A, B and C, ratios of sides $(a \times b,\ b \times c,\ c \times a)$ and gradient of side for each angle added by authors in [17].

The feature values of angles A, B and C are big compared to the features in 1, 2, 3, 7, 8 and 9. This can be proven based on the example of results in Table 3. The huge gap between angles of corners, ratios of sides and gradients proposed by [18] has imposed a big impact on the accuracy of classification. Thus, the features need to be scaled in order to improve the accuracy of result. By using HODA train dataset which consists of tr_10009_0.bmp, tr_10013_0.bmp, tr_10014_0.bmp and tr_10022_0.bmp, the samples of extracted features that used the triangle features in [17] are illustrated in Table 3.

In Table 3, the values of $\angle$A, $\angle$B and $\angle$C have shown that both angle gaps between ratios and gradients are very obvious. Due to this issue, the angles must be scaled in order to lessen the gaps. Recently, there are many normalization algorithms which can be applied to extract the features; for example, Z-score, libSVM, etc. However, this paper has reported the normalization technique by using the original extracted values which has been obtained from triangle geometry features, Z-score [20], libSVM scale function [21] and our proposed method.

Table 2. Triangle features [20].

<table>
<tr><th>Triangle</th><th>No.</th><th>Feature</th><th>Formula</th></tr>
<tr><td rowspan="9">

$a = \sqrt{((A1(y) - C(y))^2) + ((A1(x) - B1(x))^2)}$
$b = \sqrt{((B1(y) - B(y))^2) + ((A(x) - B1(x))^2)}$
$c = \sqrt{((C1(y) - C(y))^2) + ((A(x) - C1(x))^2)}$</td><td>1</td><td>c:a</td><td>c:a = c/a</td></tr>
<tr><td>2</td><td>a:b</td><td>a:b = a/b</td></tr>
<tr><td>3</td><td>b:c</td><td>b:c = b/c</td></tr>
<tr><td>4</td><td>A</td><td>$A = \arccos\frac{b^2+c^2-a^2}{2bc}$</td></tr>
<tr><td>5</td><td>B</td><td>$B = \arccos\frac{a^2+c^2-b^2}{2ac}$</td></tr>
<tr><td>6</td><td>C</td><td>$C = \arccos\frac{a^2+b^2-c^2}{2ab}$</td></tr>
<tr><td>7</td><td>ΔBA</td><td>$\Delta BA = \frac{B(y)-C(y)}{B(x)-C(x)}$</td></tr>
<tr><td>8</td><td>ΔBC</td><td>$\Delta BC = \frac{B(y)-A(y)}{B(x)-A(x)}$</td></tr>
<tr><td>9</td><td>ΔCA</td><td>$\Delta CA = \frac{B(y)-C(y)}{B(x)-C(x)}$</td></tr>
</table>

Table 3. Example results of HODA train dataset.

c:a	a:b	b:c
0.63	0.62	2.55
1.05	0.49	1.92
2.06	0.89	0.54
0.97	0.51	2.02
0.63	0.62	2.55

i. Feature ratios

ΔBA	ΔBC	ΔCA
-0.67	-0.5	-1
-0.17	0	-0.33
0.25	0	0.5
-0.13	-0.25	0
-0.67	-0.5	-1

ii. Feature gradients

∠A	∠B	∠C
11.31	161.56	7.13
8.973	161.57	9.46
12.53	14.04	153.44
7.13	165.96	6.91
11.31	161.56	7.13

iii. Feature angles

The original features have been extracted by using triangle features shown in Table 3. These original features will be classified by using Support Vector Machine (SVM) without applying the normalization algorithm on the features. However, the triangle features will be normalized by using Z-score algorithm based on Equation (1). The normalized triangle features are worked out based on a linear scale and features will be scaled between ranges -1 to 1.

$$z = \frac{x - \mu}{\sigma} \quad (1)$$

2. Data Collection and Method

In this paper, we have used four different datasets that are taken from Arabic, Roman and Bangla handwritings. These datasets are known as HODA [22], Isolated Farsi Handwritten Character Database (IFHCDB) [23], MNIST [24] and BANGLA [25]. HODA and IFHCDB datasets are taken from Arabic handwriting. The MNIST dataset is taken from a Roman handwriting, while BANGLA dataset is taken from a Bangla handwriting. These datasets consist of alphabets and digits. However, in this paper, we are only focusing on digits instead of alphabets. The example of digits used in this paper is shown in Figure 3. Further details about these datasets can be referred to in [26].

2.1 Pre-processing

In the pre-processing, all datasets will be converted into a binary image by using Otsu's method [27]. After that, the datasets will get through the labelling process. The labelling process is a process to rename the images based on their type (test and train). In binary form, the foreground image is known as '1', while the background image is known as '0'. However, the binary images in HODA are different, whereas the foreground is white while the background is black. Thus, the images in HODA dataset will be inverted and the labelling process will take place thereafter. These datasets which are involved in pre-processing are shown in Figure 3.

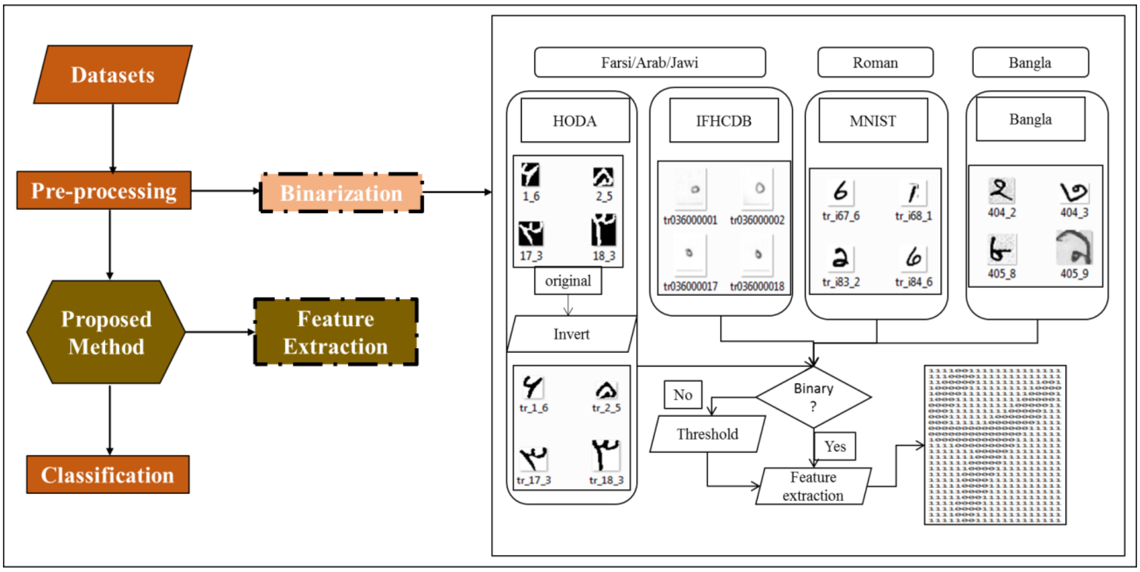


Figure 3. Pre-processing.

2.2 Proposed Method

In this paper, we have proposed a method based on the problem in [17]. The problem of huge gaps between the values of gradients, ratios of sides and angles of corners has imposed a strong impact on the accuracy of the results. However, we have only focused on the problem related to the angles of corners. As known, the total of a triangle's angles is 180 degrees. The total of a triangle's angles can be referred to in Equation (2).

$$\angle_{triangle} = \sum \angle A + \angle B + \angle C = 180 \tag{2}$$

In our proposed method, we have divided each angle by 180 degrees based on the total of a triangle's angles. This can be represented as shown in Equations (3), (4) and (5).

$$\angle A' = \angle A/180 \tag{3}$$

$$\angle B' = \angle B/180 \tag{4}$$

$$\angle C' = \angle C/180 \tag{5}$$

After diving each angle by 180 degrees, the triangle features represented in [18] have been changed. The triangle features in [17] are shown in Equation (6) while the new triangle features proposed are shown in Equation (7).

$$Ff = A_f, B_f, C_f, Rca_f, Rab_f, Rbc_f, \Delta BA_f, \Delta BC_f, \Delta CA_f \tag{6}$$

$$Ff' = A_f', B_f', C_f', Rca_f, Rab_f, Rbc_f, \Delta BA_f, \Delta BC_f, \Delta CA_f \tag{7}$$

In this paper, we have applied the Cartesian Plane Zone [28] as shown in Table 4. This zone has one main triangle and four zones represented as Zone A, Zone B, Zone C and Zone D.

Table 4. Cartesian plane zone.

<table>
<tr>
<td rowspan="2"><u>Main Triangle</u>

111111111100111111111
111111111000011111111
111111110000001111111
111111110000000111111
111111000000000111111
111110000000000011111
111100000111100001111
111100001111110001111
111000011111111000111
110000111111111000011
100000111111111000001
100001x111x1111100001
00001111111001x110000
000011111100001110000
000111111000001110000
000111111000001110000
000111110000000100001
101111000000000000001
111110000000001000011
111100000111111111111

Point B : (6 , 11)
Point C : (10 , 11)
Point A : (14 , 12)</td>
<td><u>Zone B</u>

Point B : (4 , 8)
Point C : (6 , 6)
Point A : (8 , 5)</td>
<td>11111111110
11111111100
11111111000
11111111000
11111100000
11111000x00
111100x0011
11110000111
1110x001111
11000011111
10000011111
10000111111</td>
<td>00111111111
00011111111
00001111111
00x00111111
00000111111
00000011111
1110x001111
11110001111
111110x0111
11111000011
11111000001
11111100001</td>
<td><u>Zone A</u>

Point B : (12 , 3)
Point C : (14 , 6)
Point A : (16 , 8)</td>
</tr>
<tr>
<td><u>Zone D</u>

Point B : (2 , 14)
Point C : (5 , 15)
Point A : (7 , 17)</td>
<td>10000111111
00001111111
00001111110
00x11111100
00011x11100
00011111000
1011110x000
11111000000
11110000011</td>
<td>11111100001
10011110000
00001110000
00001110000
00001x10x00
00x00100001
00000000001
00001000011
11111111111</td>
<td><u>Zone C</u>

Point B : (12 , 16)
Point C : (15 , 15)
Point A : (18 , 15)</td>
</tr>
</table>

Based on Table 4, the triangle form can be performed based on the coordinates of triangle that are represented by points A, B and C in each zone and the main triangle. Figure 4 shows the triangle form based on the coordinates, while Table 5 shows the summary of formulation for each zone including the main triangle.

Based on Table 5, C_x and C_y are the coordinates of point C for angles x and y of the main triangle, while N_x and N_y are the width and height of the image that has been converted into binary form. The coordinate of C_x is used as a border for the horizontal plane, while the coordinate of C_y is used as a border for the vertical plane. Based on the combination of the four zones, there are 36 features produced. This has caused that the total of features for Cartesian Plane Zone is 45 features, including the main triangle. As information, these zones are formed based on the coordinates of point C which is known as the centroid of the zone. We can distinguish between them by a shaded image. The mark 'x' on the red line is point C which is the centroid for Zone A as shown in Figure 5. The mark 'x' at the right side is point A, whereas at the left side it is point B. As information, the coordinates of point C need to be identified first before the coordinates of point A and point B. The coordinate of point C is used as a divider between right and left sections. The centroid at the right section represents point A, while the centroid at the left section represents point B. These features will be extracted by using the triangle concept. The results of feature extraction have produced 45 features after applying the

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 2, No. 2, August 2016.

Cartesian Plane Zone. These features have been applied to HODA, IFHCDB, MNIST and BANGLA datasets.

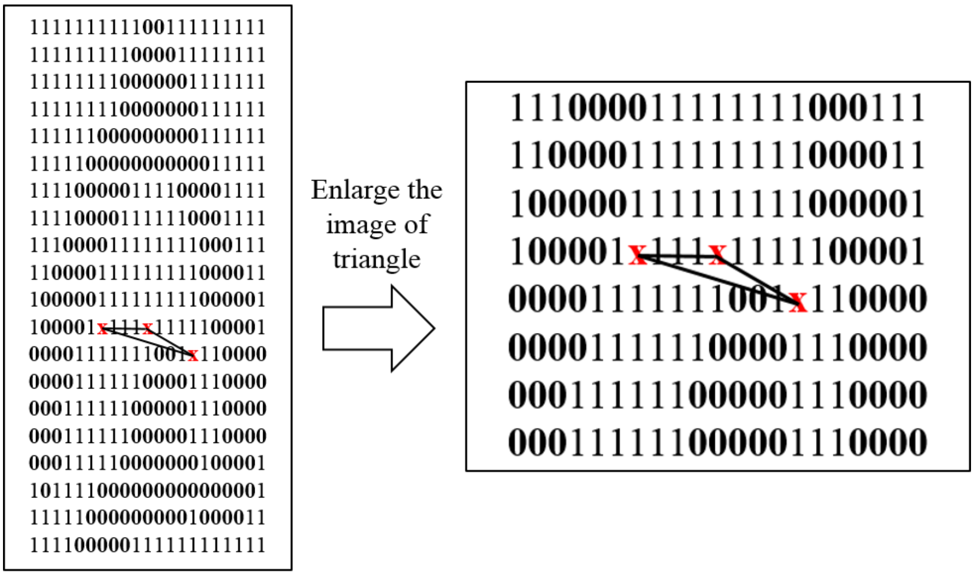


Figure 4. The form of triangle based on the coordinates.

Table 5. Summary of Cartesian plane zone.

Zone	Main Triangle	A	B	C	D
Binary Image	11111111110011111111111 11111111100001111111 1111111000000111111 11111110000000111111 11111000000000111111 11110000000000011111 1111000001111000011111 1111000011111110001111 1110000111111111000111 1100001111111111000011 1000001111111111000001 100001x111x1111100001 0000111111001x110000 000011111100001110000 0001111110000011100000 0001111110000011100000 000111110000000100001 101111000000000000001 111110000000001000011 1111000001111111111	0011111111 0001111111 0000111111 00x00111111 00000111111 00000011111 1110x001111 11110001111 111110x0111 11111000011 11111000001 11111100001	11111111110 11111111100 11111111000 11111111000 11111100000 11111000x00 111100x0011 11110000111 1110x001111 11000011111 10000011111 10000111111	11111100001 10011110000 00001110000 00001110000 00001x10x00 00x00100001 00000000001 00001000011 11111111111	10000111111 00001111111 00001111110 00x11111100 00011x11100 00011111000 1011110x000 11111000000 11110000011
Height (h)		$h = C_y$	$h = C_y$	$h = N_y - C_y + 1$	$h = N_y - C_y + 1$
Width (w)		$w = N_x - C_x + 1$	$w = C_x$	$w = C_x$	$w = N_x - C_x + 1$

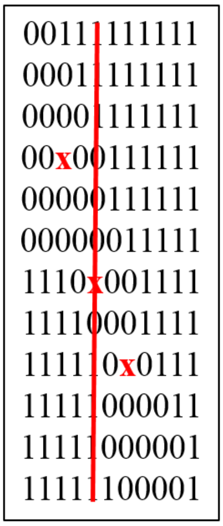


Figure 5. Zone A.

3. Experimental Setup and Results

In this paper, all the experiments were conducted by using Support Vector Machine (SVM). Table 6 shows the details of system requirements for SVM.

Table 6. System requirements for SVM.

Type	System Requirement
SVM Weka	Windows 7 and 8.1
	Ram size 4Gb and above
	Version 3.6.9
LibSVM tool (for grid search)	Windows 7 and 8.1
	Ram size 12Gb and above
	Python software (64bit, Version 3.5.0)
	Gnu plot: gp510-20150831 version
	Libsvm-3.17-GPU version, 64bit

The SVM classifier was chosen in order to compare the proposed method using the four datasets with the state-of-the-art works in this area. Cost and gamma are obtained from grid search using libSVM tool. There are three selected values of cost that were attained from grid search libSVM results, which are: c=32, c=8 and c=2, while gamma is g=0.00782. At this stage, the example of train and test images is provided by test datasets. The number of train and test images can be referred to in [26].

The first experiment has been conducted by using features without scaling, while the second experiment has been conducted by using Z-score algorithm. Table 7 shows the result without using normalization. It has been performed with an accuracy of not more than 70%. Among these datasets, HODA dataset has achieved the best result with an accuracy of 69.97%, while MNIST had the worst result with an accuracy of 52.06% in the first experiment. However, the result of classification shows some improvement that can be made by using either feature extraction or data normalization algorithm. Table 7 shows the results without using normalization, while Table 8 shows the results of normalization using Z-score.

Table 7 and Table 8 show that the triangle features without data normalization are slightly better compared to the results of data normalization by using Z-score algorithm. Thus, Z-score algorithm cannot be used as data normalization for triangle features. The third experiment has been conducted by using Minimax technique. We did not compare the results with those of

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 2, No. 2, August 2016.

SVM in [17]-[18], because the total of features used by [17]-[18] is not the same as used in this research. In [17]-[18], the researchers have used more than 200 features compared to only 45 features in this paper. Therefore, the results produced will affect the accuracy rate. Besides, this paper is introducing the data normalization technique that can be applied before the process of classification. Table 9 shows the results of normalization using Minimax.

Table 7. Results without normalization.

Dataset/Cost (C)	C=32	C=8	C=2
BANGLA	55.75	56.1	55.65
HODA	69.33	69.41	69.67
MNIST	52.06	52.14	52.34
IFHCDB	65.47	65.62	66.09

Table 8. Results of normalization using Z-Score.

Dataset/Cost (C)	C=32	C=8	C=2
BANGLA	55.7	56.03	55.85
HODA	69.34	69.41	69.65
MNIST	52.05	52.11	52.33
IFHCDB	65.4138	65.62	66.10

Table 9. Results of normalization using Minimax.

Dataset/Cost (C)	C=32	C=8	C=2
BANGLA	71.2	66.775	58
HODA	87.93	87.93	87.93
MNIST	74.02	71.86	69.23
IFHCDB	89.81	87.66	84.15

The fourth experiment has been conducted by using our proposed method. Table 10 shows that HODA and IFHCDB for Arabic digit datasets have achieved an accuracy of 90.35% and 91.72%, respectively compared to other datasets. All datasets which utilized our proposed normalization method outperformed other methods. Next, Table 11 shows the comparison results using a cost value of 32 and a gamma value of 0.078125, while Table 12 shows the training time taken to build each technique.

Table 10. Results of proposed method (normalization using LibSVM).

Dataset/Cost (C)	C=32	C=8	C=2
BANGLA	77.3	75.1	71.65
HODA	90.35	89.1	87.55
MNIST	77.91	76.42	73.86
IFHCDB	91.72	91.25	90.57

Table 11. Comparison results.

Dataset	Without normalization using LibSVM	Normalization using Z-Score	Normalization using Minimax	Normalization using LibSVM (Proposed Method)
BANGLA	55.75	55.7	71.2	77.3
HODA	69.33	69.34	87.925	90.35
MNIST	52.06	52.05	74.02	77.91
IFHCDB	65.47	65.41	89.81	91.72

Table 12. Training time taken for each technique (in seconds).

Dataset	Without normalization using LibSVM	Normalization using Z-Score	Normalization using Minimax	Normalization using LibSVM (Proposed Method)
BANGLA	33.16	277.36	508.69	343.1
HODA	103.68	6203.69	8154.28	6005.79
MNIST	492.07	8304.22	8747.97	7771.47
IFHCDB	7.24	233.94	333.79	138.7

Based on Table 12, the training time taken for our proposed method was remarkably longer compared to the case without normalization using LibSVM technique. However, the results without normalization were worse compared to other techniques. Besides, the time taken for our proposed method has shown to be shorter in general compared to normalization using Z-Score and Minimax techniques.

4. CONCLUSION

This paper is extended from our previous article [28]. Overall, we report that our normalization method using triangle features is suitable to our feature extraction method. Previously, we have published several papers on techniques that have been used in the extraction process. However, we have never mentioned any normalization techniques which are used to enhance the accuracy of classification for digit datasets. Thus, in this paper, we have stated that we have used normalization technique in our features before the process of classification. We used our proposed normalization algorithm in [18] without reporting it. Thus, any benchmarking should be referred to in paper [18]. In [18], we have introduced 25 zones in order to improve the accuracy of classification. Based on Table 11, the results have shown that our proposed method has better results compared to other techniques. Thus, it has been proven that our proposed method which is based on the nature of summation of angles A, B and C that are equal to 180 degrees has given better results using LibSVM technique. The next step will consider further improvements in different rotation invariants.

ACKNOWLEDGEMENTS

The authors would like to thank the Ministry of Education for funding this research (grant FRGS/2/2014/ICT/02/FTMK/02/F00246). Thanks are also extended to Zamalah Scheme Scholarship for providing the scholarship for this research. Besides, thanks are due to Universiti Teknikal Malaysia Melaka and the Faculty of Information Technology and Communication for providing excellent research facilities.

REFERENCES

[1] G. Gartner, L. Meng and M. P. Peterson, Digital Terrain Modelling, Lecture Notes in Geoinformation and Cartography, 2007.

[2] G. N. Clements, The Geometry of Phonological Features, Cambridge University Press, vol. 2, pp. 225–252, 1985.

[3] M. L. T. Cossio, L. F. Giesen, G. Araya, M. L. S. Pérez-Cotapos, R. L. VERGARA, M. Manca, R. A. Tohme, S. D. Holmberg, T. Bressmann, D. R. Lirio, J. S. Román, R. G. Solís, S. Thakur, S. N. Rao, E. L. Modelado, A. D. E. La, C. Durante, U. N. A. Tradición, M. En, E. L. Espejo, D. E. L. A. S. Fuentes, U. A. De Yucatán, C. M. Lenin, L. F. Cian, M. J. Douglas, L. Plata and F. Héritier, Uma ética para quantos?, vol. XXXIII, no. 2, pp. 81–87, 2012.

[4] E. W. Weisstein, "Triangle," Wolfram Research, Inc., 1999-2016.

[5] "Math Open Reference Citations," [Online], Available: http://www.mathopenref.com/site/cite.html, [Accessed: 19-Nov-2015].

[6] C. Lin and K. Fan, "Human Face Detection Using Geometric Triangle Relationship," Proc. 15th Int. Conf. Pattern Recognition (ICPR-2000), vol. 2, pp. 941–944, 2000.

[7] Moh'd Sanusi Azmi, Khairuddin Omar, M. F. Nasrudin and A. K. Muda, Fitur Baharu Dari Kombinasi Geometri Segitiga Dan Pengezonan Untuk Paleografi Jawi Digital, Ph.D. dissertation, Universiti Kebangsaan Malaysia, 2013.

[8] L. Gao and Y. Xu, "Face Orientation Recognition Based on Multiple Facial Feature Triangles," in Proc. of the 2012 Int. Conf. Control Eng. Commun. Technol. (ICCECT 2012), pp. 928–932, 2012.

[9] Z. Liu and J. An, "A New Algorithm of Global Feature Matching Based on Triangle Regions for Image Registration," pp. 1248–1251, 2010.

[10] X. Chen, J. Tian, X. Yang and Y. Zhang, "An Algorithm for Distorted Fingerprint Matching Based on Local Triangle Feature Set," IEEE Transactions on Information Forensics and Security, vol. 1, issue. 2, pp. 169–177, 2006.

[11] Z. Zhang, S. Wang and A. I. Morphing, "Multi-feature Facial Synthesis Based on Triangle Coordinate System," Proc. of the IEEE 2nd International Conference on Computer Science and Network Technology (ICCSNT 2012), no. 2, pp. 141–145, 2012.

[12] M. M. M. Tin and M. M. Sein, "Multi-Triangle Based Automatic Face Recognition System By Using 3D Geometric Face Feature," in IEEE International Instrumentation and Measurement Technology Conference (I2MTC 2009), Singapore, pp. 895–899, 5-7 May 2009.

[13] P. Tang, R. Jiang and M. Zhao, "Feature Selection and Design Olintrusion Detection System Based on k-means and Triangle Area Support Vector Machine," Proc. of the IEEE 2nd International Conference on Future Networks, 2010. (ICFN’10), pp. 144–148, 2010.

[14] A. Haselhoff and A. Kummert, "A Vehicle Detection System Based on Haar and Triangle Features," in Intelligent Vehicles Sysmposium, pp. 261–266, 2009.

[15] S. Mozaffari, K. Faez and V. Margner, Application of Fractal Theory for On-Line and Off-Line FarsiDigit Recognition, Chapter: Machine Learning and Data Mining in Pattern Recognition, Vol. 4571 of the series Lecture Notes in Computer Science, pp. 868–882, 2007.

[16] R. Ebrahimpour, A. Esmkhani and S. Faridi, "Farsi Handwritten Digit Recognition Based on Mixture of RBF Experts," IEICE Electron Express, vol. 7, no. 14, pp. 1014–1019, 2010.

[17] M. S. Azmi, K. Omar, M. F. Nasrudin, B. Idrus and K. Wan Mohd Ghazali, "Digit Recognition for Arabic/Jawi and Roman Using Features from Triangle Geometry," AIP Conf. Proc., vol. 1522, pp. 526–537, 2013.

[18] M. S. Azmi, K. Omar, C. Wan and S. Bahri, "Exploiting Features from Triangle Geometry for Digit Recognition," Proc. of the IEEE 2013 International Conference on Control, Decision and Information Technologies (CoDIT), Hammamet, Tunisia, pp. 876 - 880, 6-8 May 2013.

[19] Norm Matloff, From Algorithms to Z-Scores: Probabilistic and Statistical Modeling in Computer Science, University of California, p. 464, 2013.

[20] C.-C. Chang and C.-J. Lin, "LIBSVM: A Library for Support Vector Machines," ACM Transactions on Intelligent Systems and Technol., vol. 2, pp. 27:1–27:27, 2011.

[21] H. Khosravi and E. Kabir, "Introducing a Very Large Dataset of Handwritten Farsi Digits and a Study on Their Varieties," Pattern Recognition Letters, vol. 28, no. 10, pp. 1133–1141, 2007.

[22] S. Mozaffari and H. Soltanizadeh, "ICDAR 2009 Handwritten Farsi/Arabic Character Recognition Competition," Proc. of the IEEE 10th International Conference on Document Analaysis and Recognition (ICDAR), pp. 1413–1417, 2009.

[23] Y. LeCun and C. Cortes, MNIST Handwritten Digit Database, 1992.

[24] U. Bhattacharya and B. B. Chaudhuri, "Handwritten Numeral Databases of Indian Scripts and

Multistage Recognition of Mixed Numerals," IEEE Transactions on Pattern Analaysis and Machine Intelligence, vol. 31, issue 3, pp. 444–457, 2009.

[25] M. S. Azmi, K. Omar, M. F. Nasrudin and A. K. Muda, "Features Extraction of Arabic Calligraphy Using Extended Triangle Model for Digital Jawi Paleography Analysis," International Journal of Computer Information Systems and Industrial Management Applications, vol. 5, pp. 696–703, 2013.

[26] P. Smith, D. B. Reid, C. Environment, L. Palo, P. Alto and P. L. Smith, "A Threshold Selection Method from Gray-Level Histograms," IEEE Transactions on Systems, Man and Cybernetics, vol. 9, issue 1, pp. 62–66, Jan. 1979.

[27] M. S. Azmi, K. Omar, M. F. Nasrudin, A. Abdullah and K. Wan Mohd Ghazali, "Features Extraction of Arabic Calligraphy Using Extended Triangle Model for Digital Jawi Paleography Analysis," International Journal of Computer Information Systems and Industrial Management Applications, vol. 5, pp. 696-703, 2013.

[28] M. S. Azmi, N. A. Arbain, A. K. Muda, Z. Abal Abas and Z. Muslim, "Data Normalization for Triangle Features by Adapting Triangle Nature for Better Classification," 2015 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), pp. 1–6, 3-5 Nov. 2015.

ملخص البحث:

تم حديثاً استخدام خصائص المثلث في تمييز الأرقام عن طريق اعتبار الزاوية جزءاً من الخصائص. وقد أعطت غالبية الدراسات في مجال تمييز الأرقام التي استخدمت تلك الخصائص نتائج مشجّعة. ومع ذلك، فإنّ مسألة الفجوة الكبيرة التي حدثت بين الزوايا ونسب الأضلاع والميلانات كان لها تأثير قويّ على دقة النتائج. ولهذا، نقترح طريقةً في هذه الدراسة هي عبارة عن تسوية البيانات باستخدام طبيعة هندسة المثلث من أجل حلّ هذه المسألة. إلى جانب ذلك، طبّقنا تقنياتٍ أُخرى، مثل (Z-Score)، و(Minimax) و(LibSVM)، في التجارب المتعلقة بهذه الدراسة. وقد تم استخدام أربع مجموعاتٍ من البيانات الرقمية هي: "HODA" و"MNIST" و"IFHCDB" و"BANGLA". وبينت نتائج التصنيف أن الطريقة المقترحة أعطت نتائج أفضل مقارنة بالتقنيات الاخرى.