Feature Extraction of Character Image using Shape Energy

Galih Hendra Wibowo, Riyanto Sigit, Aliridho Barakbah
Department of Information and Computer Engineering
Graduate Program of Engineering Technology
Politeknik Elektronika Negeri Surabaya
Surabaya, Indonesia
gabo@student.pasca.pens.ac.id, riyanto@pens.ac.id, ridho@pens.ac.id

***Abstract*— One effort to maintain documents or records is to make changes in the form of digital images. The drawings further processing needs to be done so that the text or sentences therein can be operated as do the search, analysis, or manipulation of the contents of the text. The treatment process is known as optical character recognition (OCR) and continues to develop. OCR is generally divided into three main stages, namely preprocessing, feature extraction and classification. Feature extraction is one of the essential or fundamental processes in character recognition. The purpose of feature extraction is to obtain the characteristics of each character. The results at this stage can affect the quality of character recognition. Generally, feature extraction on character is done by a complex calculation so as to cause the necessary time computing is not a little, especially in real time recognition case. In this paper, feature extraction can be done simply proposed as an alternative, called Shape Energy. This method uses the approach of how humans are able to distinguish between characters or numbers in a simple. It results in three elements which are elasticity, curvature, and texture. The elasticity is first derivative and the curvature is second derivative of each pixel in the frame of the character, which is obtained from thinning. While the texture value is 4-direction chain-codes. This method testing has been done on some type of character by using back propagation neural network as a method on classification stage. This testing resulted in average value accuracy rate of success in identifying these characters by 90.3%.**

Keywords*— *Character recognition, Feature Extraction, Shape Energy, Neural Network

I. INTRODUCTION

Science and technology is currently experiencing rapid development. One of them is pattern recognition. Pattern recognition (character recognition) is a field of science in order to define or classify an object by measuring features (characteristics) that a quantitative or main properties (characteristics). Character recognition or known as Optical Character Recognition (OCR) is one specific field of pattern recognition. Over the last half century there are a lot of research effort has been done in this field. The purpose of the research is a process that is able to convert or change the text on the digital image into text, in the form of data, which can be changed [1]. In general, there are several major mechanism in the OCR process, namely the characteristic or feature extraction mechanism and the mechanism of recognition. Before entering the feature extraction mechanism, there is an additional mechanism in the form of preprocessing such as noise filtering, image smoothing, and binarization. This mechanism also has an important rather than the other mechanisms, one example in binarization. Bad binarization can result in loss of information characteristic of a character and the lead character so difficult to recognize [2], so there is research [3] [4] on the issue. The next is a feature extraction mechanism. This mechanism is made to obtain the characteristics or the unique identity of the character as a differentiator. Results of the extraction feature known as data models. This data is used to match the input data in the introduction of a mechanism so that the input data can be recognized. The success rate of the character recognition is largely influenced by the characteristics extraction mechanism. Good feature extraction can facilitate the process of introducing a character so that the probability of recognition success rate would be good. Therefore, feature extraction or the establishment of model data on a character is an important point and interesting in the world of research in OCR.

Hossain et al (2012) [6] serves as an alternative rapid new feature for feature extraction in the field of OCR. The method presented author using statistical approaches and compares with some methods that use a similar approach such as histogram projection, crossing, zoning, and momments. The technique used in this method is dividing the character image into multiple (k) section vertically. Furthermore, the projection was made from parts of the picture so that it becomes the features of the image. By using a similar technique, cutting diagonally are also made in the implementation of this new method. Gaurav et al (2012) [8] an investigation of feature extraction based on local and global geometry of a frame or skeleton of characters as an alternative. Preprocessing carried out beforehand in the form of separation between letters, binarization, background noise removal, and skeletonizing to get new features with this method. Pressure is applied to the image size to obtain a length and width corresponding to the picture and then do zoning. Zoning is used in this method for divide the 9 area with the same size. Feature extraction performed on each of these areas. In each area, there are nine types of features are extracted, namely the number of horizontal and vertical lines, the number of diagonal lines to the right and to the left [9], and the normalization of all four values. In addition, more research is done related to the extraction of

978-1-5090-1640-2/16/$31.00 ©2016 IEEE

features in the OCR, as Huette et al [10] said that the main features of the structural approach consists of a number of vertical and horizontal lines, the position of cavities or holes, the end point (approach vector) , the number of intersections and junctions. Similarly, the Kahan et al [11] also perform feature extraction to calculate the position of cavities or holes. However, added with other features such as the number, the endpoint of the vertical viewing angle, and the box that surrounds the character.

This study tries to create a new alternative in the process of feature extraction through a structural approach by calculating the energy value of the character. Energy is composed of elasticity, curvature, and texture (described in more detail in Chapter II). Furthermore, recognition mechanism is performed by using neural network, which is mostly used in OCR [7] [1].

II. BASIC CONCEPT

A. Preprocessing

Preprocessing has an important role in OCR. The role lies in the ability to recover the problems that appear in the image like noise, image resolution, and contrast. In addition, the purpose of preprocessing is to prepare images for the extracted information. It has a sense that the values which is retained are the important values or represents the object to be recognized. Many of the techniques that can be used in the OCR preprocessing are histogram equalization, noise filtering, the value of gray, binerization, thinning, erosion and dilation. The technique is used in this research is thinning.

Thinning is a method, by morphological approach, to get the skeleton form of a character that has been processed binerization first. General process of this method is to scrape the inside or outside edges of characters little by little to obtain a thickness of one pixel which is located in the middle [15]. Generally, there are two approaches thinning are repetitive and non-repetitive. The non-repetitive approach is the technique used to obtain frame without performing checks on all pixels. There are several methods that have been raised in a non-repetitive approaches such as neural networks, Voronoi diagrams, and wavelet transforms [15]. While the repetitive approach is done by checking against all pixels and reduce or eliminate the pixels that are not needed. Zhang-Suen thinning method [16] is one example of implementation of this technique. In addition, this method is also known as an efficient and quick to get to the main frame.

B. Feature Extraction

Feature extraction can be divided into two main categories: the statistical features and structural features [5]. The main purpose of statistical features is to establish decision limits or feature both linear and non-linear to distinguish each character. Some examples of features with a statistical approach is a histogram projection, zoning, invariant moment, crossing, rapid [6] etc. In addition, the structural features approach are the features of characters that includes such as curvature of the line, end points, branching, intersection point, and connectivity between points. In hierarchical approach these features can be described as primitive features [5]. This feature can be obtained by taking the main skeleton or builder of a character, which can be obtained by morphological approach such as thinning or skeletonizing [7]. In addition, this feature can also be obtained by taking the edges of characters [7]. Among two of the feature extraction approach, a structural approach is more resilient to different writing style than the statistical approach.

In this paper proposed feature extraction through a structural approach, named by Shape Energy. In addition, the basic idea of this method is how to distinguish between different objects, such as triangle, square, and circle (Fig. 1), by using human eye perspective. It stands out that differences of these objects are out of shape. Circle, triangle, and square has a shape which is not similar at all. This form can be categorized as the texture of the object.

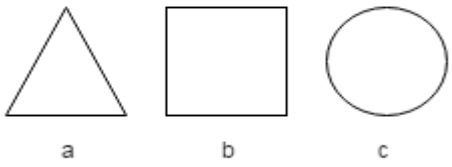


Fig. 1. Sample object image

In addition, curvature of each object can determine the differences. These objects have a different curvature, for example a triangle there are three angles which values is less than 90 degrees, the square is formed on the four corners with 90 degrees values, and a circle formed on the corners very tender in conformity with each other. Hence, the degree of curvature is a way to distinguish a simple objects. And the last is the length of the object itself. By using similar scale, the object has different length contours. Hence, this method includes three structural components, namely the length value of an object in the image (elasticity), curvature value, and texture value.

1) Elasticity

Elasticity (*El*) is a structural feature that can be obtained in the object of image, Fig. 1.

$$El = \sum_{i=1}^{n-1}(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2 \quad (1)$$

The value of this feature can be obtained by calculating the first derivative of each pixel, or by calculating the neighbors thereafter at each pixel except the pixels of most recent. In other words, *El* can be obtained with the sum of all pixel (*p*) that form or compose the object (1). In addition by using the formulations, length of pixels can be obtained by calculating the sum of the values of neighborhood with the condition, adjacency has a value of 1 for the neighboring pixels that directed horizontally or vertically and has a value of √2 for diagonal direction. To calculate the *El* on each object in Fig. 1, the first thing to do is to determine the starting point pixel and applies to all objects. One of the examples where the starting point is the lowest position x and the highest position y.

2) Curvature

The curvature of an object can be used as one of the features which is able to distinguish of each object. There are several ways to calculate the value of curvature such as calculating the

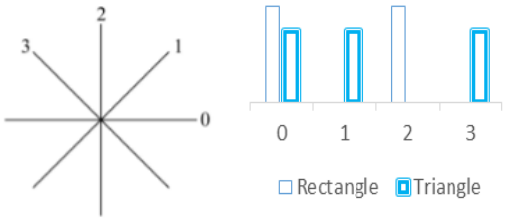
value of the slopes formed, calculating the angle value builder, calculate the value of a two-level derivatives, etc.

$$Ec = \sum_{i=2}^{n-2}(x_{i+1} - 2x_i + x_{i-1})^2 + (y_{i+1} - 2y_i + y_{i-1})^2 \quad (2)$$

The approach used in this method is based on structural, with sense of curvature value calculation (*Ec*) can be obtained by calculating the two neighbors of each pixel (2) or a derivative level two. This technique can be used in other fields such as filtering. Gaussian filtering is one example of a filtering method that uses this technique. Determination of the starting point of this feature is similar to elasticity.

3) Texture

The third structural component on Shape Energy is texture. Texture is the relationship between the points of identifying the characteristics of an object. This feature uses a structural approach such as chain-code. Chain-code is a relationship between two pixels that make up the direction of the wind as Fig. 2 (a) and then shaped into a histogram, Fig. 2 (b).



(a) (b)

Fig. 2. Texture Feature. (a) 4 wind-direction for calculating histogram, (b) histogram of Fig. 1 (b) and Fig. 1 (a)

In Fig. 2 (a) shows that there are eight directions of pixels that can be used. Various studies have also used this technique with a variety of different directions [17] [18] [19]. This study uses a 4-way Fig. 2 (a) those successfully used in face recognition [17]. Therefore this feature will produce 4 value.

III. System Design

The system design in this research is illustrated in Fig. 3. An important point in this study lies in the feature extraction (located in the large box) as new methods are proposed.

A. Data Acquisition

The first stage of this research is to collect data to be used as research objects. The object is in the form of numeric characters or digits from a computer font which obtained from Chars74K [20]. This dataset has 62 classes with details of 10 class numbers, 26 lowercase class and 26 uppercase class. Examples of classes held this dataset is presented in Fig. 4 (a). Besides, this dataset have number of variants of as many as 1016 pieces at each grade measing. Type variant on this dataset in the form of font types and combinations of, the normal format, italic, and bold. Fig. 4 (b) is one example of a variant of the class number zero.Characters in this dataset is 128 x 128 pixels and black and white. Besides Chars74K also have other character types such as natural character image and the image of handwritten characters which also used in [12] [13] [14] in the same field.

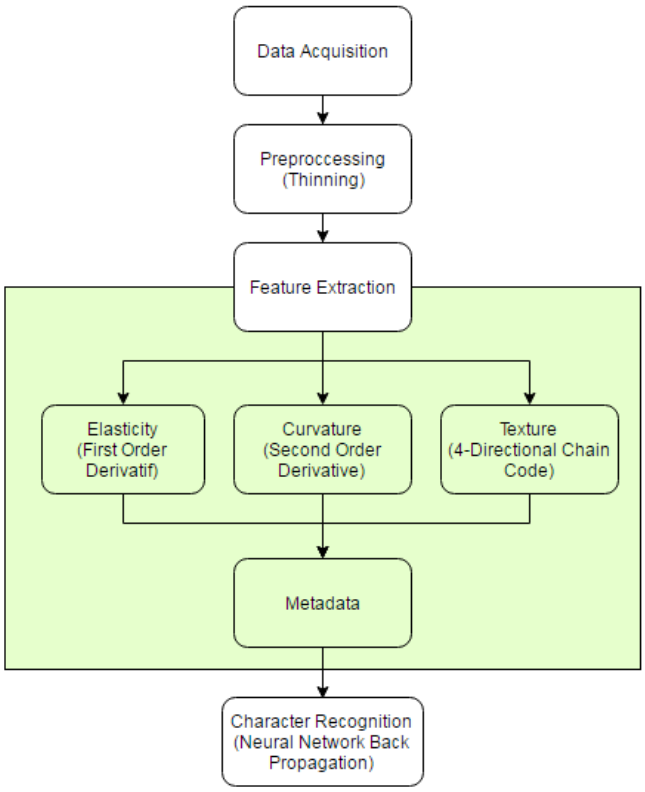
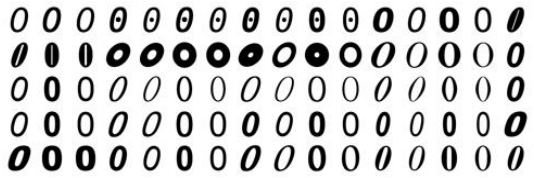


Fig. 3. System Design

(a)



(b)

Fig. 4. (a) Type of character font computer Chars74K dataset and (b) Type of variance of character zero

B. Preprocessing

In this research, a preprocessing stage performed prior to feature extraction. The technique used on preprocessing is thinning because the image dataset used to have good image quality, already formatted in black and white (binary), no noises, have the same size, and clear. As explained in Chapter II, the method which known as the Zhang-Suen is used in this research because this method is efficient and fast in getting the skeleton of character. Implementation examples of this stage is presented in Fig. 5 (a) and (b).

C. Feature Extraction

The main stages in this study is the feature extraction using Shape Energy. The first thing to do is to determine the starting point. In this study, the starting point used is the smallest position of the x-axis and greatest y-axis of x-axis. The starting point is valid on all the objects to be extracted so that the movement of feature extraction on the starting point to the last point is similar.

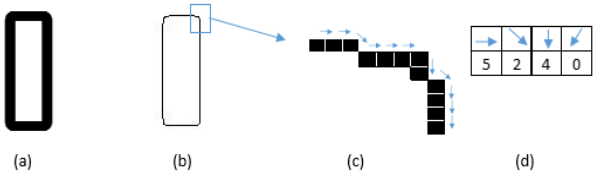


Fig. 5. (a) Original object image. (b) Thinning result from (a). (c) Magnification of the image at a certain point of (b). (d) Sample calculation texture features of (c).

Features elasticity and curvature to the object image can be extracted by using the formula (1) and (2). It is important in the extraction of pixels that have been passed or are calculated are not allowed to be recalculated. This is because the value generated can cause a less than ideal. Within the skeleton of the pixel, allowing a pixel has a neighbor which is coincide, so that the flow or the calculation must be in the form of advance.

The last feature is the texture. Objects that have been obtained skeleton figures, calculating the histogram of the 4-direction Fig. 2 (a). Extraction of texture features on zeros number is illustrated in Fig. 5 (c) and (d). Each image in the dataset to do the feature extraction in order to obtain the value of its features and is referred to as metadata.

In the metadata, the value of each object has a high variance. The point of the variants is the maximum value is different from each data and differences value among the features is high. This can lead to inaccuracies of data so that the character recognition process becomes not good. Hence needs to be normalized in order to align or balance the values that exist in the metadata. Normalization z-score is a solution used to overcome this problem.

D. Character Recognition

In this study, the character recognition is done by using a neural network back-propagation. The structure of the method used can be seen in Fig. 7. This method have standard 3 layer with 1 hidden layer therein. This layer consists of several hidden units (can be manipulated in number) and is used to recognize the characters. The input layer of network containing the value of the features that have been extracted by the proposed method as 6 value. And in the last layer is the output layer which have 10 value. The output value is represented by binary digits to represent the output (the numbers zero through 9). Example output number zero is 1,0,0,0,0,0,0,0,0,0 and output number nine is 0,0,0,0,0,0,0,0,0,1.

The purpose of this research is to present a new approach for feature extraction in the field of OCR. Necessary to test the metadata to see whether the results of features that are produced can be used in detecting these characters.

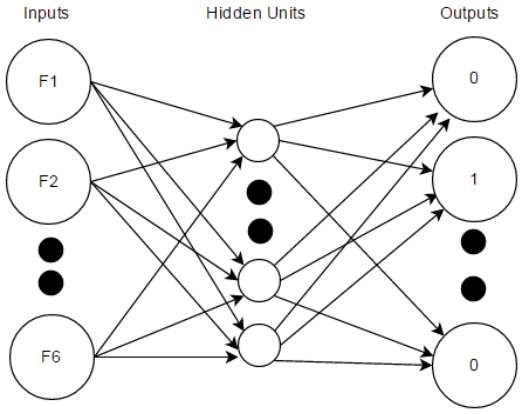


Fig. 6. Neural Network Structure

Testing was conducted using artificial neural networks (NN) and observed the SSE of the learning process.

IV. Experimantal

In this experiment, NN controlled some variable such as the iteration values is 5000, the number of hidden units is 30, and the learning rate is 0.1. Fig. 7 shows the value of 3 times SSE experiments on metadata. The graph shown in the results of these experiments indicate that changes are approaching with a value of zero or it is called convergent condition. From these results indicate that the metadata generated by the proposed method can be used as an alternative in the feature extraction character recognition (OCR). Additionally, based on Fig. 7 it can be observed that the value iteration iterations can be reduced from 5000 to 2500 to 3000 iterations.

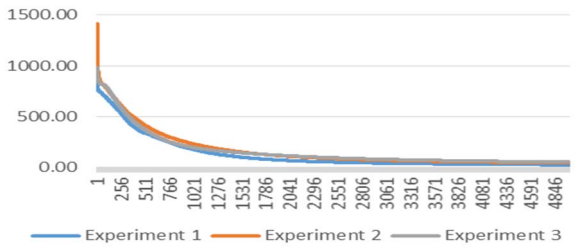


Fig. 7. Experiment Result of SSE

After the success obtained by the proposed method, the next experiment is to find out the results of the performance by using metadata. In this research, there are some variable that is controlled like number of iterations is 3000 and the number of hidden units remained is 30, and also there is a variable manipulation which is the learning rate with some value are 0.1, 0.2, and 0.3. The results of these experiment are presented in Table 1.

To find out the experiment results that represent each learning rate, then calculating the average is performed. This is because the weighting of back propagation neural network method is done randomly at the beginning of each trial that caused recognition results different. Observation on the above experiment results is conducted and it was found that the value of the highest success of each learning rate is over 90% with the details of 92.25%, 96.25% and 96.25%. However, after thorough

observation by calculating the average value of each learning rate, the highest value obtained success in the form of 92.25% on the value of learning rate 0.1. This shows that by using the values the learning rate is able to produce relatively stable weighting in each experiment.

TABLE I. EXPERIMENT RESULT

Experiment	Learning Rate		
	0.3	0.2	0.1
1	87.50%	*96.25%*	86.25%
2	88.75%	86.25%	*96.25%*
3	90.00%	92.50%	91.25%
4	88.75%	87.50%	93.75%
5	88.75%	87.50%	91.25%
6	*92.50%*	90.00%	95.00%
7	91.25%	85.00%	90.00%
8	90.00%	95.00%	92.50%
9	82.50%	93.75%	93.75%
10	90.00%	95.00%	92.50%
Average	89.00%	90.88%	**92.25%**

V. CONCLUSION

The new method which is proposed that Shape Energy, with its main components, namely elasticity, curvature, and texture, can be used as an alternative to the extraction of character features, especially on the figures. The level of performance of this method is able to achieve 92.25% by using a back propagation neural network classification. For further research, this method can be tested on the character of the handwriting, natural images, or ancient characters such as Javanese script.

REFERENCES

[1] S. Singh, "Optical Character Recognition Techniques : A Survey," J. Emerg. Trends Comput. Inf. Sci., vol. 4, no. 6, pp. 545–550, 2013.

[2] K. V Kale, Advances In Computer Vision And Information Technology. IK International Pvt Ltd, 2008.

[3] B. Wang, X. F. Li, F. Liu, and F. Q. Hu, "Color text image binarization based on binary texture analysis," Pattern Recognit. Lett., vol. 26, no. 11, pp. 1650–1657, 2005.

[4] O. D. Trier and A. K. Jain, "Goal-directed evaluation of binarization methods," IEEE Trans. Pattern Anal. Mach. Intell., vol. 17, pp. 1191–1201, 1995.

[5] S. Mozaffari, "A Hybrid Structural/Statistical Classifier for Handwritten Farsi/Arabic Numeral Recognition.," 2005.

[6] M. Hossain, M. Amin, and H. Yan, "Rapid Feature Extraction for Optical Character Recognition," arXiv Prepr. arXiv1206.0238, pp. 1–5, 2012.

[7] O. D. Trier, A. K. Jain, and T. Taxt, "Feature extraction methods for character recognition - A survey," Pattern Recognit., vol. 29, no. 4, pp. 641–662, 1996.

[8] D. D. Gaurav and R. Ramesh, "A feature extraction technique based on character geometry for character recognition," pp. 1–4, 2012.

[9] I. No and P. E. Ajmire, "Available Online at www.ijarcs.info Structural Features for Character Recognition System-A Review," vol. 3, no. 3, pp. 766–768, 2012.

[10] L. Heutte, T. Paquet, and J. Moreau, "A structural/statistical feature based vector for handwritten character recognition," Pattern Recognition, vol. 19, no. 7, pp. 629–641, 1998.

[11] S. Kahan, T. Pavlidis, and H. S. Baird, "On the recognition of printed characters of any font and size.," IEEE Trans. Pattern Anal. Mach. Intell., vol. 9, no. 2, pp. 274–288, 1987.

[12] A. J. Newell and L. D. Griffin, "Natural image character recognition using oriented basic image features," Proc. - 2011 Int. Conf. Digit. Image Comput. Tech. Appl. DICTA 2011, pp. 191–196, 2011.

[13] Kai Wang, B. Babenko, and S. Belongie, "End-to-end scene text recognition," 2011 Int. Conf. Comput. Vis., pp. 1457–1464, 2011.

[14] A. J. Newell and L. D. Griffin, "Multiscale histogram of oriented gradient descriptors for robust character recognition," Proc. Int. Conf. Doc. Anal. Recognition, ICDAR, pp. 1085–1089, 2011.

[15] W. Abu-Ain, S. N. H. S. Abdullah, B. Bataineh, T. Abu-Ain, and K. Omar, "Skeletonization Algorithm for Binary Images," Procedia Technol., vol. 11, no. Iceei, pp. 704–709, 2013.

[16] Y. S. Chen and W. H. Hsu, "A modified fast parallel algorithm for thinning digital patterns," Pattern Recognit. Lett., vol. 7, no. 2, pp. 99–106, 1988.

[17] Y. T. Lisa, B. Arun, H. Sharat, P. Andrew, and R. Bolle, "Real World Real-time Automatic Recognition of Facial Expressions."

[18] S. K. Pradhan, "A Character Recognition Approach using Freeman Chain Code and Approximate String Matching," vol. 84, no. 11, pp. 38–42, 2013.

[19] G. Gohil and M. Goyani, "Chain Code And Holisticf Eatures Based Ocr System For Printed D Evanagaris Cript," vol. 3, no. 1, 2012.

[20] T. de Campos, "The Chars74K dataset." 2009.