International Journal of Computer Applications Technology and Research
Volume 4– Issue 7, 517 - 521, 2015, ISSN:- 2319–8656

Pattern Recognition of Japanese Alphabet Katakana Using Airy Zeta Function

Fadlisyah
Department of Informatics
Universitas Malikussaleh
Aceh Utara, Indonesia

Rozzi Kesuma Dinata
Department of Informatics
Universitas Malikussaleh
Aceh Utara, Indonesia

Mursyidah
Department of Multimedia and Networking
Politeknik Negeri
Lhokseumawe, Indonesia

Abstract: Character recognition is one of common pattern recognition study. There are many object used in pattern recognition, such as Japanese alphabet character, which is a very complex character compared to common Roman character. This research focus on pattern recognition of Japanese character handwriting, Katakana. The pattern recognition process of a letter of the alphabet uses Airy Zeta Function, with its input file is a .bmp file. User can write directly on an input device of the system. The testing of the system examines 460 letter characters. The first testing that examines 230 characters result in an accuracy of 55,65%, whilst the second testing that examines 460 characters produces an accuracy of 64,56% in recognizing the letters. These accuracy are much determined by the quantity of training. The approach of pattern recognition is a statistical approach, where more pattern of letters are trained and saved as a reference, more intelligent the system . The implementation of Airy zeta function methods in recognizing Japanese letter pattern is able to produce high accuracy level.

Keywords: Pattern recognition, katakana, airy zeta function, bitmap

1. INTRODUCTION

Advancement of information technology facilitates the way of working in various field of life. An issue that is main topic in present days research of information technology is image processing and computer vision. Both fields are researches in computer field to find a way or device to replace human eyes[1,2,3,].

Pattern recognition is a field of knowledge to classify or describe an object based on feature quantitative measurement or main characteristic of the object. Pattern is an defined entity and can be identified and given name. Pattern recognition can be executed on objects such as handwriting, eye, face and skin[4].

Pattern recognition can be applied to identify a peculiar character such as Japanese characters that is Katakana. The goal of character recognition of Japanese letter is as a learning[5].

Tool of studying Japanese for newcomers , especially in studying character Katakana. The simple use of the high recognition level of character can boost user attention in learning Japanese. Japanese character is a complex character compared to the common roman character, especially if the character is handwriting, where is produced various form of characters from different people[6].

One of the technology that is used in recognizing Japanese character Katakana is Airy Zeta Function. The first step in the recognition process is characteristic extracting, that is to find characteristic or special feature of an object.

In common, the pattern recognition using airy zeta function comprise of several step, that are image acquisition, grayscale process, segmentation using edge detection utilizing operator, identification using Airy Zeta Function method, and produces the result of Japanese character identification, Katakana.

The features in an image could be a pixel in a matrix that is from a digital image. This characteristic extraction process is implemented in pre-processing process on a digital image[7].

This is important for boosting the presence of successful matching of an object, such as the changing of size image in order to equalize the pixel of compared images, and thresholding process to make similar the pixel value of images along with abolishing existence of the noise[8].

After characteristic extraction process is done, the process of Katakana letter recognition starts using pattern recognition method. Structure of pattern recognition system is showed in figure 1. The system consists of censor (such as digital camera, the algorithm of feature searching, and algorithm for classification or recognition (depend on the approach). In addition, it is common that some classified datas is assumed already available to use in testing.

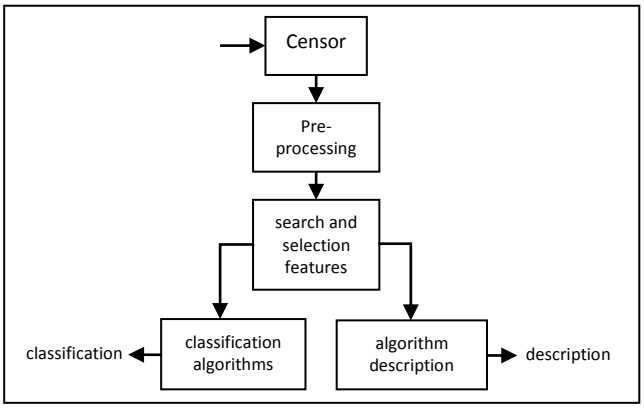

Figure 1. Structure of pattern recognition system

The steps in system training proses are :

1. Censor captures object from the real world and then change the object into digital signal, that is consist of a collection of number. This process is called digitalization.
2. Preprocessing is preparing images or signal in order to produces better characteristic at next level. In this stage,

the information signal is bumped and the interfering signal is minimized.

3. Feature finding and feature selection is useful for finding distinguishing characteristic that represents main characteristic of signal along with reducing signal dimension into a collection of less number, although it is still representative
4. Classification algorithm is functional for clustering features into suitable class
5. Description algorithm is useful to present signal description[4].

2. STUDY DESIGN

This study identifies patterns of handwriting. By applying the method Airy Zeta Function simple and complex as the achievement kearusian level pattern recognition with a more accurate pattern recognition.

This study identifies patterns of handwriting. By applying the method Airy Zeta Function simple and complex as the achievement kearusian level pattern recognition with a more accurate pattern recognition.

2.1.Airy Zeta Function

By applying the method Airy Zeta Function to see the level of accuracy with the value of the Zeta Function Airy transformation method is to use the equation.

$$Ai(x) = \frac{1}{n}\int_0^{\infty} \cos\left(\frac{1}{3}t^3 + xt\right) dt$$

Specification :

Ai(x) : Airy Value
n : Index Citra Value
t : Index Citra Value on airy Value

For the Airy function zeta function is defined by a series of zero order.

$$\zeta \mathrm{Ai}(s) = \sum_{i=1}^{\infty} \frac{F(i)}{|ai|^s}$$

This series converges when the real part of s is greater than 3/2, and can be extended by a further analysis for other values of s.[9]

Specification :

ζAi : Nilai airy zeta value
s : Transformation Index airy zeta function
$F(i)$: Index value images on airy zeta function

2.2. Letter Japanese Katakana

Katakana is derived from the Chinese characters are shortened and were used by Buddhist monks to show the proper pronunciation of Chinese characters in the 9th century. Katakana syllabary writing, consisting of 46 syllables and formerly called "paper man"[10].

Figure 2. Letter Basic Katakana

Collected reference on Image Processing and data required in the making of the application. Data or samples used in this study is a Japanese katakana letters pattern data scanning results of handwriting with a variant of the different writing difference[11]. The details are as follows:

1. Diagram Workflow System

Workflow diagrams which will be conducted in this study is illustrated in the following:

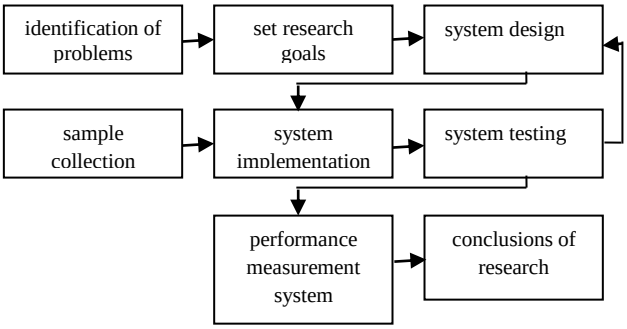

Figure 3. Workflow research in general.

2. System Scheme

The scheme of the overall system is as follows[7]:

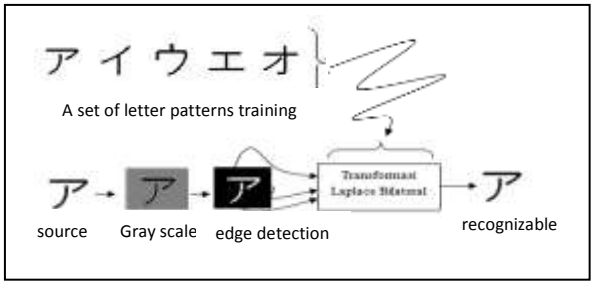

Figure 4. Schematic System Overall

The stages are performed after the system receives input is gray-scale stage, edge detection, and pattern recognition test letters through Airy Zeta Function. In the pre-processing stage, which becomes an input source image format file.bmp. In the main process, computing using Airy Zeta Function as follows[6]:

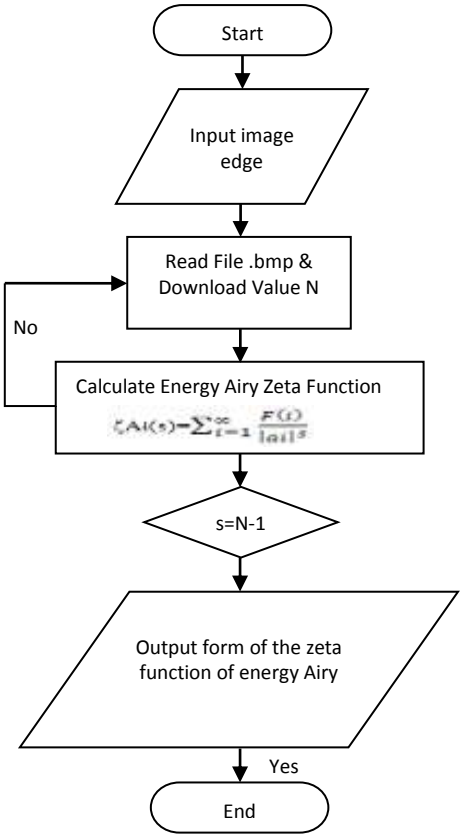

Figure 5. Process Flow Diagram Airy Zeta Function

3. ANALYSIS AND DISCUSSION

Samples of Japanese katakana letters training base used in this study gradually with the number of images from 230 training data with the data testing 460 then 460 training data with the data testing. The image of the pattern of Japanese katakana letters basis vectors that represent the characteristics of Japanese katakana letters pattern different basis. Figure 6 shows some sample patterns Japanese katakana letters are used as a training base. Training is done using the bilateral Laplace transform.

Figure 6. Some Japanese Katakana Sample Letter Writing Basics

3.1. Training Process

The process of training on this system will be described in repsentasi on the following pictures:

Figure 7. Process Painting Samples

Figure 7 illustrates the initial steps to be undertaken in this system that makes handwriting samples from the writings of different variants depending directly on the canvas that is available on the system. Generate output images of Japanese katakana character case basis with the rules of correct writing with bmp image formats.

After the painting process the sample, the following picture describes the process of training for extract characteristics of handwriting sample image of Japanese katakana letters that basis. Which further the values of the image will be saved into the database as a reference to an image pattern recognition.

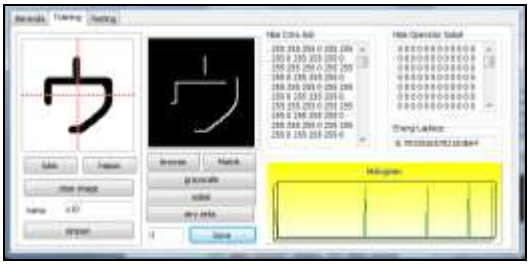

Figure 8. Direct Painting Process Training Samples

Direct Painting Process Training Samples describes sample letter patterns japanese katakana basic form of handwriting directly on the canvas that is available on this system.

The training process image input samples are as follows:

Figure 9. Sample Training Process Input Image

Figure 9 describes the process of training with the Japanese katakana letters pattern sample basis in the form of handwriting input image scan results.

3.2. Testing Process

Testing Process recognition system of Japanese katakana letters shown in the picture below base where in this process we will take a picture that has been painted and stored previously. Data testing is not the same image data with image data in the training process. And the form of handwritten images of different people, then the value of the image of the character pattern letters in this testing process will be compared with the value of the letters in the image of the character pattern prior training process. If energy is equal or close similarity of the pattern of the letters will be recognized and vice versa. The image data were tested as many as 460 images of Japanese katakana letters basic pattern.

Figure 10. Results of Pattern Recognition Letters Properly

Results Pattern Recognition Letters true of the testing process is case-sensitive pattern recognition. Where the Japanese katakana letters input in testing this basic form of handwritten images directly from the canvas are available in the system.

Figure 11. Results of Pattern Recognition Letters One

Figure 11 describes the results of the testing process pattern recognition incorrect letters. Where the Japanese katakana letters input in testing this basic form of handwritten images directly from the canvas are available in the system.

3.3. Work Systems

Measurement of the performance of the entire system is based measurement test data based on specifications or certain classification the correlated the number of training data is used.

Some of the results of the performance measurement system to test on letter recognition is presented as follows.

Table 1. Results of Performance Systems Pattern Recognition Letters

Jumlah Citra Pelatihan	Jumlah Citra Pengujian	Jumlah Pengenalan Pola yang Benar		Jumlah Pengenalan Pola yang Salah	
230	460	256	55,65 %	204	44,34 %
460	460	297	64,56 %	163	35,43 %

Test results for 46 Japanese katakana character letter basis, shows that the greater number of correct training data stored in the database as the image of a pattern recognition energy letter, the higher the level of accuracy of the letter pattern recognition. The following figure shows a graph of the results of the performance of the pattern recognition system of Japanese katakana letters basis. The graph Percentage Accuracy

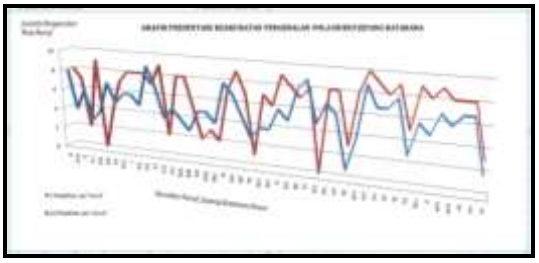

Figure 12. Graph Percentage Accuracy Japanese Katakana Basic Introduction Letter

Illustrating the accuracy of pattern recognition Japanese katakana letters training data base of 5 and 10 training data letter. It can be seen that the process of training data 5 average grade level each letter pattern recognition accuracy is lower than in the 10 training data. However, seeing a percentage a character letters on the 10 training data there are some letters that lower the level of accuracy of the process with 5 training data. This is due to the level of similarity approach or the energy generated from the same case characters are almost the same even there, the more the comparison value in the training system the harder it will take a decision to classify her character recognition letter patterns so that there was an error that letter pattern recognition. The highest level of accuracy contained in the letter patterns ***SO*** with an average accuracy rate of 95%. The graph Percentage inaccuracies are.

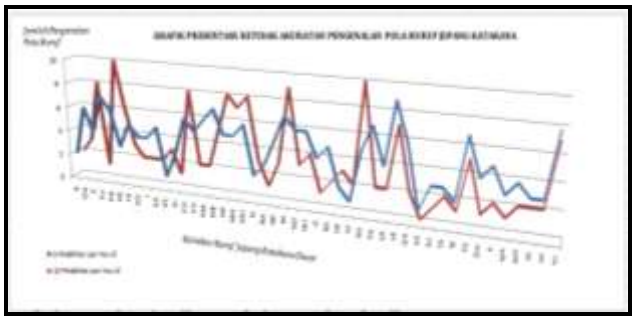

Figure 13. Percentage Graph inaccuracies Japanese Katakana Basic Introduction Letter

While in figure 13 above the level of illustrating inaccuracies Japanese katakana letters pattern recognition basis of training data 5 and 10 training data. It can be seen that the process of training data 5 average value inaccuracies rate each letter pattern recognition is higher than in the 10 training data. However, seeing a percentage character letters on the 10 training data there are some letters that lack accurated higher level than the process with 5 training data. This is due to the level of similarity approach or the energy generated from the same case characters are almost the same even there, the more the comparison value in the training system the harder it will take a decision to classify her character recognition letter patterns so that there was an error that letter pattern recognition. Accurate accuracy lack highest level found in the pattern of letters ***HA, NI*** and ***SE*** with an average error rate of 80%.

4. CONCLUSION

From the results of research and discussion that has been done, can be summed up as follows:

1. The pattern recognition system of Japanese katakana handwritten letters using Zeta Function Airy pattern recognition accuracy levels ranging from 55.65% to 64.56%. It is clear percentage handwriting pattern recognition truth Japanese katakana letters are very influential on the basis of training data.
2. The pattern recognition approach is a statistical approach, where a growing number of letters in the training pattern and stored as a reference, then the system will be more intelligent and percentage accuracy shows that Airy Zeta Function can be used as one method of pattern recognition on handwritten image.

5. REFERENCES

[1] Castleman, Kenneth R., 2004, Digital Image Processing, Vol. 1, Ed.2, Prentice Hall, New Jersey.

[2] Gonzalez. R. C, Woods. R. E., Digital Image Processing third Edition, Pearson Prentice Hall, New Jersey, 2008.

[3] Pitas, I., Digital Image Processing Algorithms, Prentice Hall, Singapore, 1993.

[4] Putra, Darma. 2010. Pengolahan Citra Digital. Andi. Yogyakarta.

[5] Puput Alit Resmika.2007. Conversion Application construction of Posts Japan shape Alphabet Using Wavelet Backpropagation With the transformation. Informatics engineering study program. University of Atma Jaya Yogyakarta.

[6] Masril, Mardhiah. 2013. Implementation of Neural Networks In Pattern Regonation (Studi Kasus : Huruf Jepang Katakana).

Journal of Information Technology & Education ISSN : 2086-4981, Vol. 6 No. 1 Maret 2013

[7] T.Sutoyo, dkk. 2009. Theory of Digital Image Processing.Andi. Yogyakarta.

[8] Manda Sari, Desita. 2014. Introduction Letter With Perceptron method. Informatics Engineering Program Malikussaleh University

[9] http://en.wikipedia.org/wiki/Airy Zeta Function. Accessed 5 May 2014.

[10] Afrillia, Yesy. 2014. Pattern Recognition Letters japanese use Bilateral Laplace transform. Thesis. Informatics Engineering Program Malikussaleh University

[11] Fadlisyah, Taufiq, Zulfikar, Fauzan. 2008. Image Processing Using Delphi. Graha Ilmu. Yogyakarta.

[12] Turban & Efraim, J.E. "Decision support systems and intelligent system-sevent edition", New Delhi, Prentice Hall of India, 2007.