

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/316011674>

Optimization of Feature Extraction Using Combined Image Centroid Zone and Zone Centroid Zone Method

Conference Paper · September 2016

CITATION

1

READS

166

3 authors, including:



[Tjut Awaliyah Zuraiyah](#)

Universitas Pakuan

6 PUBLICATIONS 1 CITATION

SEE PROFILE

Optimization of Feature Extraction Using Combined Image Centroid Zone and Zone Centroid Zone Method

T A Zuraiyah¹, A Qur'ania² and C R Pitoyo³

¹²³Computer Science Study Program, Pakuan University, BOGOR

¹tjut.awaliyah@gmail.com, ²arie.qurania@gmail.com, ³christyawan.ridanto@gmail.com

ABSTRACT

Computer vision is one of computer science part that imitates how the human visual works so that the computer can recognize the object. The process of object recognition by computer requires a certain process, one of the example is the handwriting recognition using zoning as a method for feature extraction and support vector machine as a method for classification. The method used was Image Centroid Zone (ICZ), Zone Centroid Zone (ZCZ), and the combined ICZ & ZCZ with 8 zone. Support Vector Machine used was Radial Basis Function (RBF) kernel. The Stages of research conducted including data collection, analysis, preprocessing, feature extraction, classification and testing. The highest accuracy value for the uppercase obtained by using combined zoning ICZ & ZCZ of 88.46%, while 76.92% for the lowercase. The highest accuracy value with slope of 10° and 20° for uppercase obtained using zoning ICZ of 71.15%, while the highest accuracy for lowercase with slope of 10° obtained using zoning ICZ of 76.92% and 59.61% with slope of 20°.

Keywords : Model Development, In-House Training, Learning Organization.

INTRODUCTION

Human handwriting has a style and a unique shape for each individual. Therefore, it was very often to occur a miss interpretation which is caused by the variation of writing from each individual. The ability of computer to recognize the handwritings requires data that represents the object of the handwriting in the form of digital data. The digital data then processed, so it can recognize the handwriting of the human being. One way of the computer to recognize the handwriting was using the zoning method for feature extraction and support vector machine for the classification. The training data used was 624 data (312 data for uppercase and lowercase respectively) and the testing data was 104 data (52 data for uppercase and lowercase respectively). The zoning method used was Image Centroid Zone (ICZ), Zone Centroid Zone (ZCZ), and combined ICZ & ZCZ [1] [2][3].

1. Stages of Research

Research stages to perform handwriting recognition using zoning methods and support vector machine can be seen in Figure 1.

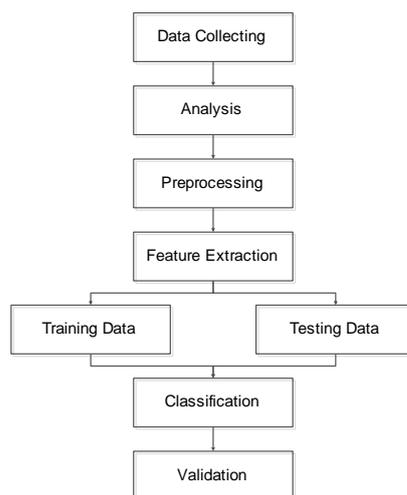


Figure 1. Stages of Research

IMPLEMENTATION

1.1 Preprocessing

In the preprocessing stage, the binary process was done to convert the image into a binary image, so that the image only has a value of 1 and 0. The binary image was complemented to obtain the pixel value of 1 for character pixel and value of 0 for the background pixel. After the binary image was done, the image was cropped to get the characters pixel. The next step was to rescaling the image into dimension of 32 x 32 pixels. Thinning process was also done to obtain the core framework of the character image.

1.2 Feature Extraction Zoning

The image would be over the stage of feature extraction zoning to get the value of a character trait. Feature extraction zoning will generate vector of 1 x 8 for each character, in order to obtain a vector of 156 x 8 for the training data and vector of 1 x 8 for the testing data and then stored in the extension of *.mat.

1.3 Support Vector Machine Classification

The classification process was done by comparing the data to be tested with previous training data [4] [5].

RESULT AND DISCUSSION

The image was obtained from the handwriting of 8 respondents. 728 data used was divided into training data and testing data. The training data was 624 data (312 data for uppercase and lowercase respectively) and the testing data was 104 data (52 data for uppercase and lowercase respectively). Zoning method was used as a method for feature extraction and Support Vector Machine method was used for the classification. The validation results using the zoning feature extraction ICZ for the uppercase obtained accuracy of 84.62%, using the zoning feature extraction ZCZ of 69.23%, and 88.46% by using the combination of feature extraction zoning ICZ and ZCZ. While for the lowercase using the zoning feature extraction ICZ obtained accuracy of 75%, using the zoning feature extraction ZCZ of 65.38%, and 76.92% by using the combination of feature extraction zoning ICZ and ZCZ. The validation was also done by changing the image slope of 10° and 20° from the normal. The display of the uppercase and lowercase handwriting recognition as shown in figure 2 and figure 3.

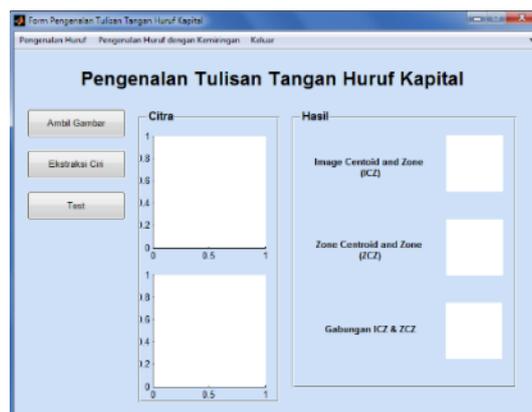


Figure 2 Display of handwriting recognition for uppercase

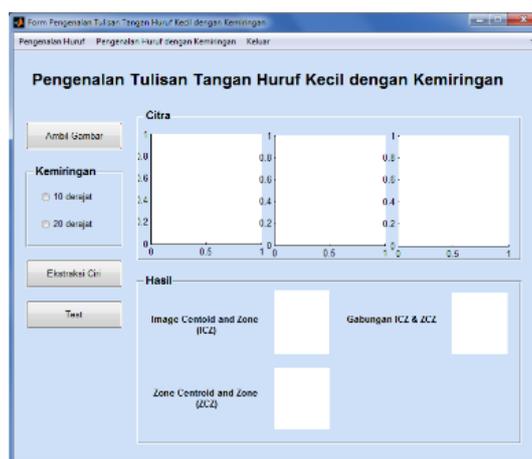


Figure 3 Display of handwriting recognition for lowercase with slope

Table 1. Uppercase Results Accuracy

Slope	Normal (0°)	10°	20°
Zoning			
ICZ	84,62%	71,15%	71,15%
ZCZ	69,23%	57,69%	46,15%
Combine ICZ & ZCZ	88,46%	69,23%	59,61%

Table 2. Lowercase Results Accuracy

Slope	Normal (0°)	10°	20°
Zoning			
ICZ	71,15%	76,92%	59,61%
ZCZ	65,38%	51,92%	34,62%
Combine ICZ & ZCZ	76,92%	67,31%	42,31%

Factors that led to the low accuracy was:

1. The image of handwriting letters was more varied than the image with printed letters because each person has different styles of writing.
2. Based on the confusion matrix table, some characters tested was classified into characters that do not correspond to the original class, it was because these characters had a shape similar to the other characters, such as:
 - a. O similar to D
 - b. G similar to C
 - c. Q similar to O
 - d. i similar to l
 - e. q similar to a
 - f. p similar to f

CONCLUSION

The testing for uppercase character recognition using zoning ICZ method reach into 84.62%, ZCZ method of 69.23%, and the combined method of 88.46%. As for the recognition of lowercase characters using zoning ICZ method reach into 71.15%, ZCZ method of 65.38%, and the combined method of 76.92%. The ICZ and ZCZ combined zoning method for uppercase and lowercase letters handwriting recognition was resulted as the highest accuracy than using only ZCZ and ICZ method.

REFERENCES

- [1] Arridho *et al.* 2013. Analisis Pen Pressure Tulisan Tangan Untuk Mengidentifikasi Kepribadian Seseorang Menggunakan Support Vector Machine (SVM). Diponegoro University. Semarang.
- [2] Rajashekararadhya SV & Ranjan Dr PV. 2008. Eficient Zone Based Feature Extraction Algorithm For Handwritten Numeral Recognition Of Four Popular South Indian Scripts. Journal of Theoretical and Applied Information Technology.
- [3] Syam RM. 2013. Pengenalan Aksara Jawa Tulisan Tangan dengan Menggunakan Ekstraksi Fitur Zoning dan Klasifikasi K-Nearest Neighbour. Skripsi. Bogor Agricultural University. Bogor.
- [4] Mulia I. 2012. Pengenalan Aksara Sunda Menggunakan Ekstraksi Ciri Zoning dan Klasifikasi Support Vector Machine. Skripsi. Bogor Agricultural University. Bogor.
- [5] Nugroho *et al.* 2003. Suport Vector Machine Teori dan Aplikasinya dalam Bioinformatika. IlmuKomputer.Com.