Identification of Speed and Unique Letter of Handwriting Using Wavelet and Neural Networks

Esmeralda C. Djamal, Febriyanti
Department of Informatics
Universitas Jenderal Achmad Yani
Cimahi, Indonesia
Email : esmeralda.contessa@lecture.unjani.ac.id

Abstract—Graphology is scientific method to evaluation personality and emotion condition through handwriting and signature. There are many features to identify personality so that previous researches made handwriting analysis automatically. There are page margins, spacing, baseline, vertical zone, font size, slant, pen pressure, and the type t letter. While other studies used features of signature. As image, the analysis of graphology is divided into two approaches that graphics features and segmentation digit each character. This research integrated both of approach to identify personality of handwriting. It used speed and the type of a, d, i m, t letters as features using structure analysis and artificial neural networks. Type of letter recognition was done after character segmentation. Wavelet transform was used to improve recognition. The proposed methods could be used to identify personality of handwriting. Identification of speed feature using structure analysis toward page margin, spacing between lines, and spacing between words that gave 81% accuracy. While identification of unique letters using neural network with multilayer perceptron architecture, which gave 74% accuracy. Variations training data greatly affect recognition.

Keywords— speed and unique letter features, handwriting identification; personality; wavelet extraction; multilayer perceptron.

I. Introduction

Handwriting are image that have certain pattern reflecting anything such personality prediction. Handwriting analysis or Graphology is a scientific method of identifying, evaluating and understanding personality through the strokes and patterns revealed by handwriting. All systems of handwriting analysis are based on the assumption that as a brain controlled, expressive body movement, and handwriting reflects in some way, the dynamics of the human mind or called the personality. Graphology can identify the qualities, traits, attitudes, sentiments or postures that seem indicated in the handwriting; they further seek insight into how these aspects of selfhood may integrate together to constitute the dynamic organization that we recognize as the personality of that writer [1] . Some of guidelines for handwriting analysis are seven basic elements: speed, pressure, shape, dimensions, continuity, direction, and order. Personality overview was obtained from the research on handwriting psychiatric patients [1]. Handwriting can be classified toward various aspects of personality. There are two approaches in graphology i.e. graphical analysis of the structure type of writing and analysis of the type of symbol or letter.

If the graphology test is still done manually it takes a long time considering the aspects reviewed in graphology very much. Besides, accuracy of handwriting analysis depends on how skilled the analyst is. Development in image processing and pattern recognition lead to analyzing of handwriting can be done automatically.

Several research handwriting analysis automatically with the aid of a computer without the human intervention to predict personality traits have been conducted. Some of them, using baseline, the pen pressure and the height of the T-bar on the stem of the letter 't' are considered for predicting the personality of the writer[2]. Other research about personality analysis considered six main different types of features i.e. size of letters, slant of letters and words, baseline, pen pressure, spacing between letters and spacing between words in a document to identify the personality of the writer. It used SVM that generated various parameters are then calculated by simple use of trigonometry and threshold technique [3]. While other research using the polygon method for pattern recognition based handwriting baseline, letter slant and pen pressure, and the identification of letters i and f using template matching as the input of neural networks [4], and toward the size of 'a' letters using backpropagation algorithm [5]. There is also research that used Farsi handwriting with the shape of the page margins, line spacing, line skew, word slant, corner sharpness, size of letters, text density, writing speed and regularity of writing. It used support vector machine and identified certain pixel [6]. This research is a continuation of previous research with different features, complementary features for handwriting analysis. Last research used baseline, page margin, spacing, dominance of vertical zone and signature [7], [8]. Others, used form application to identification type of A-Z letters using learning vector quantization [9].

Handwriting is image so that could use some methods toward other image like finger print as biometric analysis [10], [11]. The research study using a neural network with multilayer perceptron.

The method is widely used considering correction weights done for all training data in training [11].

This research using speed feature and type of letter a, d, i, m, and t, it integrated two approach i.e. structure and type of character. Analysis was done of handwriting image of A4 paper.

II. HANDWRITING FEATURES DESCRIPTION

In this research handwriting structure analysis carried out on speed, as in Table I [1].

TABLE IV.   DESCRIPTION SPEED FEATURE

No	Type	Personality	Examples
1	Fast	Respond and think fast, impatient	
2	Normal	Able to arrange something well	
3	Sloe	Calm, happy, but slow in execution	

Meanwhile type of letter identification using five letters that are a, d, i, m, t. It is illustrated as Table II.

TABLE V.   DESCRIPTION TYPE OF FIVE LETTERS

No	Letter	Type	Examples of Features each class
1	a	5	
2	d	1	
3	i	2	
4	m	1	
5	t	5	

In this study, type of font that identified tend to not good personalities. So that, it may not appear in certain handwriting.

III. PROPOSED METHODS

Pattern recognition performed on a sheet of A4 handwritten scanned using a scanner in jpg or bmp format as Fig 1. Speed feature used last analysis of margin, spacing between baseline and word with decision function. Meanwhile, type of letter analysis after character segmentation. Each features recognized in parallel that indicates the distinct personality, in order to obtain a review until six personalities. The unique letters identified tend to reflect not good personality, so that in certain handwriting may will not be appear. Of the five letters that were identified have various types, all of which have 14 types, Identification system has 15 classes, with the addition of one class to type or other letters.

A. Data Aquisition

Handwriting of different individuals are used in this research which is digitally collected by scanning the handwriting of 25 different writers (training data) and 100 different writers (testing data). Each of them was asked to write a text document of simple in running hand. Most of the handwritings are printed but few of them are cursive handwriting. The samples were written on A4 size paper without any lines. In pre-processing stage, image processing was done with gray scale and threshold so the handwriting image in a bit then segmented.
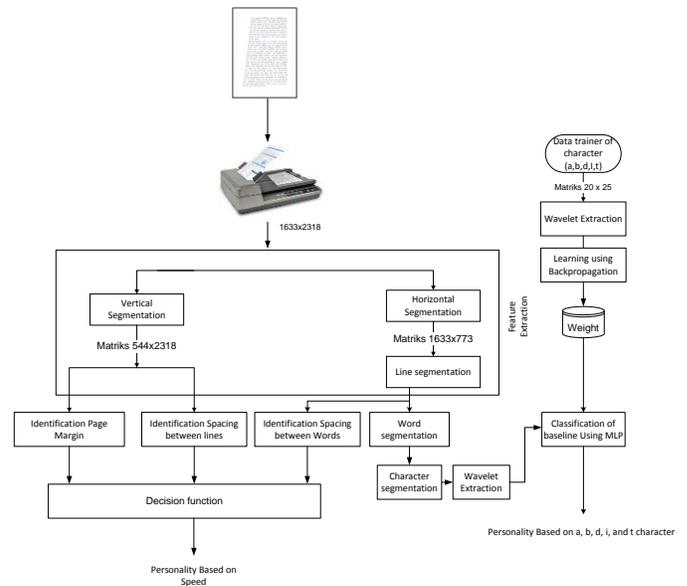

Fig. 3.   Identification handwriting system using speed and a, d, i, m, t

B. Segmentation

The process of feature extraction is done before type of font identification. Segmentation performed on three stages, ie. vertical segmentation, horizontal segementation, lines segmentation as Fig 3 and character segmentation as Fig. 4
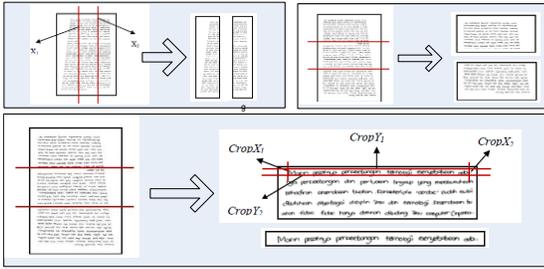
Fig. 4.   Three step segmentation

Vertical segmentation divide the image into three sections that right and left side of the image used to analyze page margin. Whereas spacing between lines analysis using left side. Horizontal segmentation to divide into three parts which middle segment are processed into line segments. The process begins by taking the coordinates of x in the upper left corner and y coordinates of the lower left corner. Having found the black pixel value of x will be stored as the value $CropX_1$. Then look for the value of the value of x $CropX_2$ started last form to the right y-axis until no black pixels are found. Then look $CropY_1$ and $CropY_2$ the same process starting from the bottom point. It used to classify dominance zone, baseline patterns and spaces between words features.

Segmentation of the character will be processed from the word segmentation. Rules for segmentation letters will start from the top left of the word. Checking used pixel analysis from the top left vertically. Segmentation of the character for some distance between letters.
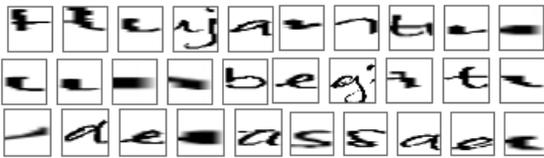

Fig. 5.   Character Segmentation

C. Wavelet Extraction

Wavelet transformation is used to feature extraction of an image. Discrete wavelet transform is a good form of transformation is used to image and easy implementation process. Family wavelet used is Daubechies4 (DB4), low pass filter and high pass filter DB4 wavelet in Eq. 1.

Using wavelet transform, image of letter was extracted using wavelet transform acted as multiple band pass filters. Discrete wavelet transform is described as follows:

$$C(j,k) = \frac{1}{\sqrt{\left|2^j\right|}} \sum_n x(n)\psi * \left(2^{-j}n - k\right) \tag{1}$$

Where $C(j, k)$ is wavelet coefficient and $\psi\left(2^{-j}n-k\right)$ is wavelet base function. While wavelet synthesis can be written as follows

$$\hat{x}(n) = \sum_j \sum_k c(j,k)\psi_{j,k}(n) \tag{2}$$

Or

$$= \sum_k c(j,k)\varphi_{j,k}(n) + \sum_j \sum_k d(j,k)\psi_{j,k}(n) \tag{3}$$

Where $\varphi_{j,k}(n)$ and $\psi_{j,k}(n)$ are wavelet functions for low frequency component and high frequency component respectively. Using (3) signal can be decomposized into j scale level with narrower frequency interval, either for high frequency or low frequency groups. Daubechies 4 has four low-pass filter coefficients (denoted by h_n) and four high-pass filter coefficients (denoted by g_n).

a) Coefficient Function Scale (low-pass filter).

$$h_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}, h_1 = \frac{3+\sqrt{3}}{4\sqrt{2}} \ h_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}, h_3 = \frac{1-\sqrt{3}}{4\sqrt{2}}$$

b) Coefficient Function of Wavelet (high pass filter)

$$g_0 = \frac{1-\sqrt{3}}{4\sqrt{2}}, g_1 = -\frac{3-\sqrt{3}}{4\sqrt{2}} \ g_2 = \frac{3+\sqrt{3}}{4\sqrt{2}}, g_3 = -\frac{1+\sqrt{3}}{4\sqrt{2}}$$

D. Recognition of Handwriting System

Identification of speed used three features identification before. They are page margin, space between line and word. Then used decision function, as Table III.

TABLE VI.   DECISION FUNCTION OF SPEED CLASSIFICATION

No.	Type	Decision Function			
		Margin		Space of word	Space of line
		Left	Right		
1	Fast	Width	Narrow	>35 pixel	>50 pixel
2	Normal	Width	Width	> 18, < 35 pixel	≥ 20, ≤ 50 pixel
3	Slow	the narrowed	Widened	≤ 18 pixel	< 20 pixel

The image of the input on multilayer perceptron is a, d, i, m, t unique image which has been normalized by the size (MLP) of 20 x 25 pixels. Through this normalization process, then each matrix being input vector on MLP networks have the same number, as many as 500 nodes.
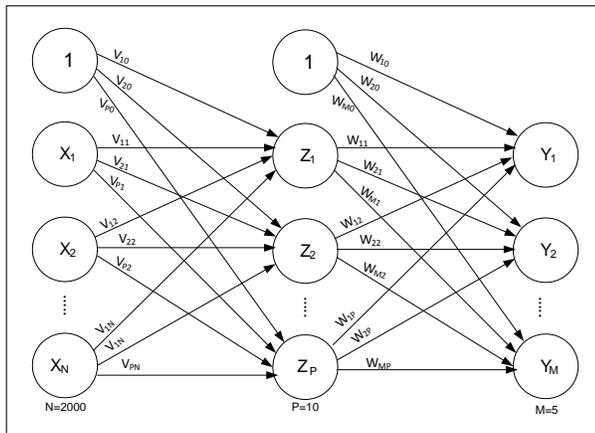
Fig. 6.   Architecture of Classification System Using MLP Artificial Neural Networks

Each character classified in 14 type of letter that five type of "a" letter, one type of "d", two type of "i", a type of "m", and five type of "t" letter. Total of all is 14 type of letter. In addition a class, the system become 15 classes. Last class is for other type or letter. It consists of input layer (500 neurons), hidden layer (10 neurons) and output layer (15 neurons). The error between the output of the feed-forward network and the target became the basis of the correct weight, which is done by back propagation algorithm. It is illustrated in Fig 4. Unique identification letters conducted all character that has segmented, thus allowing more than one type of letters from around the handwriting paper. Five dominant types of letters written as output personality identification system.

IV. Result and Discussion

Testing of identification system performed comparing visual observation. Tests carried out on speed feature and unique letter. So testing is done separately. Training of ANN identification is also conducted each unique letter that take half hour to finish training.

Speed identification result 81% accuracy. It may less decision function and variation of data training. Meanwhile, identification testing of unique letter gave 62-87% accuracy and 74% average. It is showed as Table IV, using 20 testing data.

TABLE VII.   IDENTIFICATION OF UNIQUE LEETTER

Letter	Class	Number of Training Data	Accuracy (%)
a	5	150	73
d	2	60	62
i	3	90	87
m	2	60	82
t	6	180	67
Average			74.2

The system has been implemented to software so that easy to use it. Pattern recognition handwriting and signature will result until personalities. But, it will be without unique letter identification. It gives more of an overview of the author's personality.

Weakness of the identification in word segmentation and character segmentation. It was when the handwritten form of writing connecting

V.   CONCLUSION

This research proposed methods to develop system that can identify the type of handwriting from a variety of features personalities using multi structure and ANN algorithm. It is useful to obtain an overall picture of the personality. The system using combination of structure approach and symbol approach. It provides convenience to our analysis on a piece of paper like a hand writing curriculum vitae, without the need for specially created application form.

Using multilayer perceptron gave identification accuracy 74% average. Type of "d" letter less than other letter. While identification of unique letter gave 81% accuracy.

This system has been implemented in software to provide convenience to the public in identifying personality easily and quickly. It can be used in the selection of the employee or job application letter. Besides the system can apply in determination of an appropriate field of study.

For the future, handwriting pattern recognition system can increase the number of features in order to provide a more complete picture of personality. It should also consider variation of training data and more optimal parameters of training in order to provide better accuracy. In addition, it is necessary to find the relation between personality pictures of the handwritten signatures for more accurate prediction of personality

REFERENCES

[1]   B. Ludvianto, "Handwriting Analysis", Gramedia Pustaka Utama, 2011.

[2]   H.N Champa, K.R AnandaKumar, "Artificial Neural Network for Human Behavior Prediction through Handwriting Analysis", International Journal of Computer Applications (0975 - 8887) Volume 2 - No.2, May 2010.

[3]   S. Prasad, V.K. Singh, A. Sapre, "Handwriting Analysis based on Segmentation Method for Prediction of Human Personality using Support Vector Machine", International Journal of Computer Applications (0975 - 8887) Volume 8- No.12, October 2010.

[4]   P.K.Grewal, D Prashar, "Behavior Prediction Through Handwriting Analysis", IJCST Vol. 3, Issue 2, April - June 2012

[5]   Z. Mohd Zam, "Handwriting Analysis for Employee Selection Using Neural Network", Thesis of Intelligent System Faculty of information Technology And Quantitative Science of University Technology MARA, 2006.

[6]   S. Hashemi, B. Vaseghi, F. Torgheh, "Graphology for Farsi Handwriting Using Image Processing", Journal of Electronics and Communication Engineering, Volume 10, Issue 3, Ver. I (May - June), pp: 1-7, 2015.

[7]   E.C. Djamal, R Darmawati, "Recognition of Human Personality Trait Based on Features of Handwriting Analysis Using Multi Structural Algorithm and Artificial Neural Networks", 2013 IEEE Conference on Control, Systems & Industrial Informatics (ICCSII), 22-24 June, 2013.

[8]   E.C. Djamal, R Darmawati, "Application image processing to predict personality based on structure of handwriting and signature", International

Conference on Computer, Control, Informatics and its Applications, Jakarta, Indonesia Nov 19-20, 2013.

[9] E.C. Djamal, S.N. Ramlan, J. Saputra, "Recognition of Handwriting Based on Signature and Digit of Character Using Multiple of Artificial Neural Networks in Personality Identification", Information Systems International Conference (ISICO), 2 - 4 December 2013.

[10] M. Fakir, "Fingerprint classification using fuzzyneural network and other methods", IAES International Journal of Artificial Intelligence, Vol 3 No 3, 2014.

[11] A. A. Abbood, G. Sulong, "Fingerprint Classification Techniques: A Review", International Journal of Computer Science Issues, Vol. 11, Issue 1, No 1, January, pp. 111-122, 2014.