ICIC Express Letters
Part B: Applications
Volume **11**, Number **6**, June **2020**

ICIC International ©2020 ISSN 2185-2766
pp. **587–593**

USABILITY OF TABLET MOBILE DEVICES FOR OFFLINE HANDWRITTEN CHARACTER RECOGNITION

Thi Thi Zin and Takato Otsuzuki

Graduate School of Engineering
University of Miyazaki
1-1 Gakuen Kibanadai-nishi, Miyazaki 889-2192, Japan
thithi@cc.miyazaki-u.ac.jp

Received December 2019; accepted March 2020

Abstract. *In recent years, offline handwritten character recognition has been one of the major topics and challenging research areas in the field of pattern recognition and image processing. Moreover, it is a very problematic research area due to the nature of hand writing styles which can vary from one user to another. On top of that, writings by early learners in formal and non-formal education make the problem more complex and challenging. Generally speaking, the early school age children have diversity in handwriting style, variation in angle, size and shape of characters, making the problems of character recognition more difficult. The number of students has been increasing year by year worldwide. However, due to lack of teachers, many children cannot access high quality education. Therefore, this paper proposes offline character recognition for handwritten characters written on tablet mobile devices. This paper can be considered as an additional supporting system for education of preschool children. This proposed method firstly performs character segmentation process on words acquired from the tablet. In the feature extraction process Histogram of Oriented Gradients (HOG) and Bag of Visual Words (BOVW) are used. Support Vector Machine (SVM) is applied in classification process. Some experimental results are shown to confirm the proposed method.*
Keywords: Handwritten character recognition, Early school years, Styles of handwriting, Tablet mobile devices, Morphological operations, Histogram of Oriented Gradients (HOG), Support Vector Machine (SVM)

1. **Introduction.** Today global initiatives around the world highlight the development of Education for All (EFA). As consequences, the number of students who enroll to schools becomes significantly larger. The main idea behind is to create an environment in which everyone in the world can receive at least basic education. To be precise the number of students enrolled in 2012 has been increased by two thirds compared to 1999 before EFA began in worldwide [1]. However, still many are left behind due to various reasons such as economic situation, unstable environments and lack of communications. According to UNESCO EFA Global Monitoring Port 2015 [2], there are approximately 58 million children around the world who cannot attend school due to various reasons. Among them, many are located in the developing countries, and developing countries do not have sufficient IT infrastructure. Based on this current situation and the Sustainable Development Goals (SDGs) [3], the high-quality education system that fosters literacy and literacy skills, which is the foundation of education, for preschool and younger children, is needed.

Thanks to today's advancement of mobile technologies, in the first quarter of 2010 alone, sales of mobile devices grew by 56.7% according to figures from the International Data Corporation (IDC). These numbers are boosted with the 21.7% growth of the overall mobile market [4]. Therefore, it is worthwhile to explore the utilization of mobile tablets

DOI: 10.24507/icicelb.11.06.587

in an environment where the Internet cannot be used so that some barriers in basic education especially in developing countries can be removed. In order to do so, some research activities are on their way to explore and examine the usability of tablet mobile devices for character recognition systems.

Handwritten character recognition can be classified based upon two major criteria: the data acquisition process (on-line or off-line) and the text type (machine-printed or handwritten) [5]. In this paper, we will emphasize only offline character recognition system. Offline character recognition system is that in which the handwritten character images are scanned from a surface (such as a sheet of paper) and stored into computer file in the form of two dimensional image representations which is called digital format and after that it is further processed to allow superior recognition [6].

In online recognition, handwritten characters are recognized by using information obtained from an image and information obtained from a time-series brush point coordinate sequence. In this research, we propose a method to divide and recognize images using offline handwritten character recognition. In the experiment, English words and 2-digit numbers written on the tablet were treated, and division processing was performed for each letter. The division process was performed on the assumption of two types of connected characters and non-connected characters. Then, the feature amount was extracted from the shape of the obtained divided image to perform character recognition. The target characters were extracted and classified for 26 letters in total of alphabet letters a to z and 10 numbers from 0 to 9, and character recognition was tried.

The rest of the paper is organized as follows. Section 2 describes the proposed method of this research. Section 3 explains experimental environment and evaluation method. Section 4 shows experimental result and discussions. Finally, Section 5 concludes the paper.

2. Proposed Character Recognition Method.

The proposed handwritten character recognition system includes various phases namely, Data acquisition, Preprocessing, Character segmentation & Normalization, Feature extraction, and Classification and Recognition. The overview of proposed system is described in Figure 1.

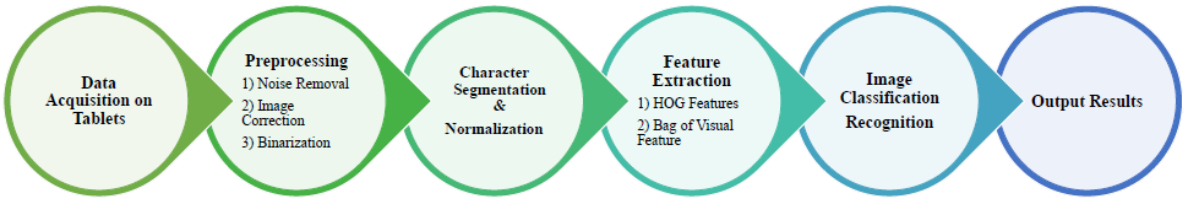

FIGURE 1. Overview of character recognition system

2.1. Data acquisition.

In this stage, we have developed handwritten database for recognition of English handwritten characters collected from the people of different age groups having different backgrounds. The collected characters are needed to be cleaning before moving to next stage. This will be done in the preprocessing phase.

2.2. Preprocessing phase.

In preprocessing phase, noise and other undesirable effects are reduced to improve the data for the recognition process. Noises occur during the image procurement or hand writing process and transmission process. Mostly, noise can be different types such as Gaussian noise, Gamma noise, Exponential noise, Uniform noise, Salt and pepper noise. Noise reduction and normalization techniques are applied for the noise removal process. Noise reduction: During the input, undesired data can also be registered. For example, if the user accidently touches the screen, such "wild points" have to be corrected. Size normalization: During the input process, the size of a character

can vary. To achieve a better recognition, the characters have to be normalized to a general size. Various techniques like gray scale, binarization, normalization and thinning are applied so that proper features are extracted.

In this system, the characters written on the tablet are also represented by RGB images. Next, a gray image is an image that represents between black (0) and white (1) in 256 levels by deleting hue information and saturation information while holding luminance. In binarization process, color or grey-scale image is converted into binary image with the aid of thresholding. Binary image can be achieved using adaptive thresholding, global thresholding, variable thresholding, Otsu's method, etc.

2.3. **Image or character segmentation phase.** Silhouette images after binarization may have missing characters and noise. The lack of letters and noise makes accurate segmentation difficult. To that end, morphological processing is performed to correct missing characters and remove noise. Morphological processing is a generic term for shape-based image processing operations in the process of processing an image, and various special processing can be performed on an image such as expansion or contraction [7].

In this aspect, we perform majority processing. Majority processing is a method of setting a pixel to 1 when five or more pixels in the 3-by-3 neighborhood are one. If the condition is not met, set the pixel to zero. Morphological processing removes noise and makes it possible to create a smooth silhouette. Thereafter, segmentation processing is performed using the resultant image. Figure 2 shows an example of morphological processing.

FIGURE 2. Example of morphological processing

In segmentation process, normalization is performed in two stages before and after segmentation. The input image before the segmentation processing is resized depending on the tablet because the obtained image size is different, thereby enhancing versatility.

In addition, since the image size is large, resizing is performed to increase the processing speed. The size was 28×28. As a procedure for normalization after segmentation processing, maximum and minimum values of the horizontal axis and maximum and minimum values of the vertical axis are detected with respect to white pixels of the divided binary image. Resize each detected value as four vertices. In this way, extra parts can be removed, and more accurate feature extraction can be performed. The state of resizing is shown in Figure 3.

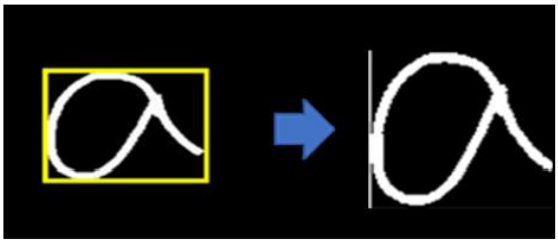

FIGURE 3. State of resizing

Before we move to next step, the labeling process is performed. It refers to a process in which white numbers (or black areas) assign the same number to consecutive pixels in a binarized image. By performing the labeling process, it becomes possible to calculate the width, height, aspect ratio, area and the like of the silhouette of the same number, and it is used for feature quantity extraction and classification process. There are two types of processing in labeling: 4-connection in which longitudinal and horizontal continuous parts are made the same label, and 8-linkage in which longitudinal, horizontal and diagonal continuous parts are made the same label. In this research, since pixels connected diagonally are highly likely to have different characters, we used a 4-connected method in which pixels connected vertically and horizontally are regarded as the same label.

2.4. **Feature extraction phase.**

2.4.1. *HOG feature.* The feature extraction is one of the most important parts of character recognition system. In this paper, we use two type features: HOG (Histogram of Oriented Gradients) and BOVW (Bag of Visual Words). For HOG feature, the gradient measures the direction and magnitude of the maximum variation in intensity in a small zone of each pixel; gradients are calculated using gradient filter. Specifically, the brightness gradient is determined by Equation (1) and Equation (2), and the brightness gradient direction is determined by Equation (3). Further, the luminance intensity is shown in Equation (4), and the number of dimensions of the feature amount is shown in Equation (5). An example of the relationship between Cell size and Block size is shown in Figure 4. In this case, processing is performed with Cell size = 4 (pixels) and Block size = 2 for an image of 200 verticals by 100 horizontals.

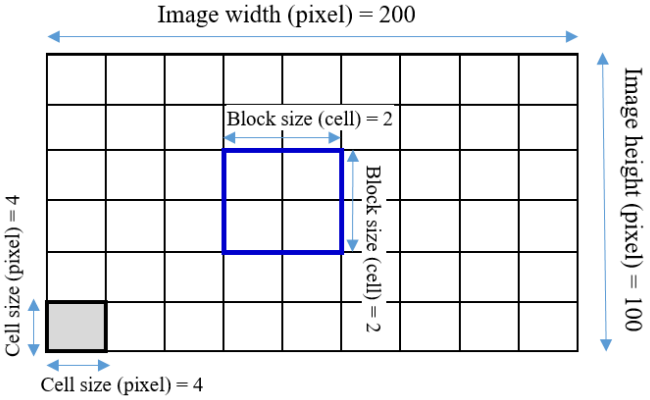

FIGURE 4. Relationship between Cell size and Block size

The gradients of image $I(x, y)$ along x-direction and y-direction are given as:

$$g_x = I(x+1, y) - I(x-1, y) \tag{1}$$

$$g_y = I(x, y+1) - I(x, y-1) \tag{2}$$

We then have the orientation angle and magnitude as described in Equations (3) and (4).

$$\tan\theta = \frac{g_y}{g_x} \tag{3}$$

$$magnitude = \sqrt{\left(g_x^2 + g_y^2\right)} \tag{4}$$

$$\begin{aligned} dim = &\left(\frac{Image\ width}{Cell\ size} - Block\ size + 1\right) \\ &\times \left(\frac{Image\ width}{Cell\ size} - Block\ size + 1\right) (Block\ size)^2 \times Orientation \end{aligned} \tag{5}$$

2.4.2. *BOVW feature.* In order to achieve well recognition, we need to consider an addition feature such as bag of visual words. The general concept of BOVW is to represent an image as a set of visual words composed of key points and descriptors. Then these points and descriptors will form vocabularies. The histogram of those vocabularies for an image will represent the associated image. From the frequency histogram, later, we can find similar images or predict the category of the image. Next, we make clusters from the descriptors. In this case we use K-means. Finally, for each image, we make frequency histogram from the vocabularies and the frequency of the vocabularies in the image. Those histograms are our Bag of Visual Words (BOVW). This method divides the training image into grids, finds characteristic blocks based on SURF (Speeded-Up Robust Features), treats one block as one word, and generates visual vocabulary, K-means. Use categories to classify categories. After that, the degree of coincidence with the image is compared based on the generated data, and the category is searched [8,9]. Figure 5 shows an example of category classification by BOVW.

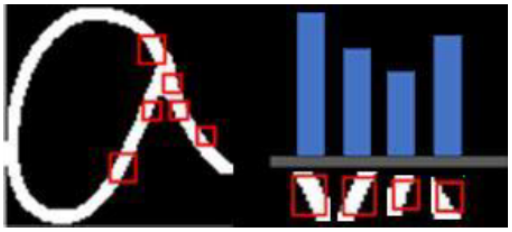

FIGURE 5. Example of category classification by BOVW

2.5. **Character classification recognition using support vector machine.** In this process, we employ support vector machine classifier for the character recognition process. SVM is one of the pattern recognition models that can apply models such as classification and regression [10]. A straight line (decision boundary) that divides the data into two is determined, and classification is performed based on the straight line. At that time, it is selected by a straight line that maximizes the minimum distance to the learning data, and is expressed by Equation (6). Here, a, b, c are constants, and x_1, x_2 represent variables.

$$f(x) = ax_1 + bx_2 + c \tag{6}$$

Equation (6) is a hyper-plane or a set of hyper-planes in a high dimensional space to perform the SVM classification tasks. SVM requires a training dataset and a testing dataset. After preprocessing and feature extraction training is done by taking the feature vectors which are stored in matrices form. SVM approach has some advantages compared to other classification techniques. SVM technique shows greater ability to generate good classifiers. Even when number of training samples is small, the result of SVM technique is robust, accurate and effective.

3. Some Experimental Results.

3.1. **Experimental design.** The experiments are conducted on the training images from the EMNIST data set [11]. The EMNIST data set is learning data that is publicly available on the Internet and can be easily acquired. An improved version of the MNIST data set, in addition to numbers, large and small English is available. The number of characters is respectively 45,699 large and small letters and 280,000 numbers. Note that both resolutions are 28×28. In this paper, 5,000 pieces were used for each category.

We also conducted the experiments on the self-collected hand written images on two tablets. Table 1 and Table 2 show a list of acquired words and numbers. English words were excerpted from the words learned in primary education [12]. We have used all 34

TABLE 1. List of used numbers

No.	Digit
1	18
2	24
3	35
4	60
5	97

TABLE 2. List of alphabets

No.	Character	No.	Character	No.	Character
1	apple	13	face	25	milk
2	banana	14	head	26	quarter
3	lemon	15	mouth	27	queen
4	mango	16	classroom	28	avocado
5	orange	17	homework	29	vegetable
6	bird	18	school	30	zoo
7	cat	19	student	31	zebra
8	cow	20	teacher	32	box
9	dog	21	coffee	33	fox
10	fish	22	tea	34	yard
11	ear	23	juice		
12	eye	24	water		

letters in lower case. The numbers were selected to be two-digit numbers so that all numbers from 0 to 9 could be included. Then, 15 subjects were asked to write twice, and 30 sheets were collected per word, and a total of 1,170 sheets were prepared. Each of the collected sample images was subjected to segmentation processing, and feature value extraction and classification were performed. The images extracted from the collected images are 5,070 letters and 300 letters.

3.2. **Performance evaluation method.** In this experiment, the two types of images are considered: alphabetic and numeric. The processes of image segmentation, feature extraction and classification are performed. Therefore, the evaluation of the experiment was performed by calculating the HOG feature quantity and the BOVW feature quantity for each segmented letter and number. SVM classifier is employed. At the time of creation of the classifier, three patterns were used for learning data for accuracy comparison. First, we use learning data of EMNIST only. The second is learning data in which the images collected from EMNIST and the tablet are mixed, and the third is learning data of only the image collected from the tablet. The images collected from the tablet use 4,000 letters and 200 numbers. The image used for learning and the test image used for input used different images, and 1,000 sheets of alphabets and 100 sheets of numbers were used.

4. **Comparative Analysis.** Table 3 represents the percentage accuracy obtained by SVM classifier using our experimental data set for digits, and alphabets. The results of the proposed method are comparable with the other state-of-the-art methods for off line. As can be seen from the table, as for the recognition rate in the case of using the HOG feature quantity, the recognition rate was improved when using the collected image as learning data for both alphabetic and numeric characters. In particular, the discrimination rate using collected images for training data was 96.0% in English and 90.0% in numerical results, and high recognition rates were obtained for all of a to z and

TABLE 3. (a) Identification result for alphabetic character and (b) identification result for numbers

(a)

Feature	EMNIST	EMNIST and Tablet	Tablet
HOG	60.1%	94.0%	96.0%
BOVW	4.0%	7.7%	9.6%

(b)

Feature	EMNIST	EMNIST and Tablet	Tablet
HOG	55.8%	70.8%	90.0%
BOVW	12.9%	25.0%	26.3%

0 to 9. On the other hand, when BOVW feature quantities were used, the recognition rate was improved as a whole when using the collected image as learning data, but the recognition rate was lower as a whole. Finally, when the HOG feature and the BOVW feature were compared, the HOG feature clearly had a higher recognition rate.

5. **Conclusion and Future Prospects.** In this paper, we had presented an off-line character recognition method based on HOG and BOVW features. The experimental results show that the proposed method achieves the recognition outcomes with high accuracy. However, we realized that some more works need to be done to get almost sure recognition rates.

REFERENCES

[1] *https://unesdoc.unesco.org/ark:/48223/pf0000232565*, Accessed on January 21 2019.

[2] *https://www.mofa.go.jp/mofaj/gaiko/oda/sdgs/pdf/about_sdgs_summary.pdf*, Accessed on January 21 2019.

[3] D. C. Shubhangi and P. S. Hiremath, Handwritten English character recognition by combining SVM classifier, *International Journal of Computer Science and Applications*, vol.2, no.2, pp.122-126, 2009.

[4] D. Mohamad, M. F. Zafar and R. M. Othman, On-line cursive handwriting recognition: A survey of methods and performances, *The 4th International Conference on Computer Science and Information Technology (CSIT2006)*, Amman, Jordan, 2006.

[5] S. A. Vaidya and B. R. Bombade, A novel approach of handwritten character recognition using positional feature extraction, *International Journal of Computer Science and Mobile Computing*, vol.2, no.6, pp.179-186, 2013.

[6] V. L. Sahu and B. Kubde, Offline handwritten character recognition techniques using neural network: A review, *International Journal of Science and Research (IJSR)*, vol.2, no.1, pp.87-94, 2013.

[7] N. Dalal and B. Triggs, Histograms of oriented gradients for human detection, *International Conference on Computer Vision & Pattern Recognition (CVPR)*, vol.1, pp.886-893, 2005.

[8] G. Katiyar, A. Katiyar and S. Mehfuz, Off-line handwritten character recognition system using support vector machine, *American Journal of Neural Networks and Applications*, vol.3, no.2, pp.22-28, 2017.

[9] H. Bay, A. Ess, T. Tuytelaars and L. V. Gool, Speeded-up robust features (SURF), *Computer Vision and Image Understanding*, vol.110, no.3, pp.346-359, 2008.

[10] M. Y. Chen, A. Kundu and J. Zhou, Off-line handwritten word recognition using a hidden Markov model type stochastic network, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.16, pp.481-496, 1994.

[11] A. Choudhary, R. Rishi and S. Ahlawat, A new character segmentation approach for off-line cursive handwritten words, *Procedia Computer Science*, vol.17, pp.88-95, 2013.

[12] N. Arica and F. T. Yarman-Vural, An overview of character recognition focused on off-line handwriting, *IEEE Trans. Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol.31, no.2, pp.216-233, 2001.