The 5th International Conference on Information & Communication Technology and Systems (ICTS) 2009

Proceedings

August 4th 2009, Surabaya - Indonesia

SECTION 3

DEPARTMENT OF INFORMATICS
FACULTY OF INFORMATION TECHNOLOGY
INSTITUT TEKNOLOGI SEPULUH NOPEMBER

ICTS

ITS Institut Teknologi Sepuluh Nopember

IEEE INDONESIA SECTION

JICA

Faculty of Engineering Kumamoto University

Proceeding of

the 5th International Conference on Information & Communication Technology and Systems

ISSN 2085-1944

ORGANIZING COMMITTEE

Chairman

Anny Yuniarti, S.Kom, M.Comp Sc.
Informatics Department, Faculty of Information Technology, Sepuluh Nopember Institute of Technology

Secretary

Radityo Anggoro, S.Kom, M.Eng.Sc

Website and Design

Ridho Rahman Hariadi, S.Kom
Hadziq Fabroyir, S.Kom

Proceeding and Registration

Chastine Fatichah, S.Kom,M.Kom.
Umi Laili Yuhana, S.Kom, M.Sc.
Dini Adni Navastara, S.Kom
Wijayanti Nurul Khotimah, S.Kom
Ratih Nur Esti Anggraini, S.Kom
Erina Letivina Anggraini, S.Kom

Tour and Promotion

Victor Hariadi, S.Si.,M.Kom
Bagus Jati Santoso, S.Kom
Arya Yudhi Wijaya, S.Kom

Seminar Kit and Accommodation

Bilqis Amaliah, S.Kom.M.Kom
Nurul Fajrin Ariyani, S.Kom

Contact Address

Informatics Department, Faculty of Information Technology,
Sepuluh Nopember Institute of Technology Surabaya
Gedung Teknik Informatika ITS,
Jl. Raya ITS Keputih Sukolilo Surabaya Indonesia
Tel. + 62-31-5939214
Fax. +62-31-5913804
Homepage: http://icts.if.its.ac.id

Proceeding of

the 5th International Conference on Information & Communication Technology and Systems

TABLE OF CONTENT

ON THE PERFORMANCE OF BLURRING AND BLOCK AVERAGING FEATURE EXTRACTION BASED ON 2D GAUSSIAN FILTER

Linggo Sumarno
Electrical Engineering Department, Science and Engineering Faculty,
Sanata Dharma University
Kampus 3, Paingan, Maguwoharjo, Yogyakarta, 55282, Indonesia.
lingsum@sttaff.usd.ac.id

ABSTRACT

Feature extraction has an important role in the field of handwritten recognition, especially in reducing the number of data to be processed. Block averaging is one of feature extraction that based on multiresolution of images. 2D Gaussian filter is one of low-pass filter that can be able to blur the image. By performing block averaging on a blurring image, a set of feature extraction values can be resulted. Based on the experiment, it was shown that 2D Gaussian filter 14x14 with standard deviation 10, can blur 64x64 pixels binary image optimally. In this case, it can show basic shape of the image clearly, which is not too detail nor too blur, where quantitatively it give highest recognition rate. Meanwhile, based on the other experiments it was shown that feature extraction by using block averaging can give insignificant difference in terms of recognition rate performance, if it is compared with feature extraction by using wavelet and DCT.

Keywords: Feature extraction, blurring, block averaging, 2D Gaussian filter.

1 INTRODUCTION

In most of image recognition system, a compact image representation is needed in order to avoid extra complexity and also increase the recognition algorithm [1]. In order to get a compact image representation a set of feature is extracted from the image. One approach in getting compact representation is by using multiresolution approach. One method in multitresolution approach is blurring and block averaging, as inspired from Ethem [2]. (However Ethem did not describe in detail, how block averaging need to be performed). Another method in multiresolution approach is by using wavelet transform. This transform is a kind of transformation to represent image at different resolution levels. Representation coefficient called wavelet coefficients, can be used as feature of an image [4], [5], [6], [10].

Besides multiresolution approach, there is a feature extraction by using lossy compression approach e.g. Discrete Cosine Transform (DCT). By using DCT, an image can be compressed at different compression levels. Representation coefficients called DCT coefficients, can be used as feature of an image [12], [15].

In this paper, will be explored the effects of blurring on the performance of block averaging feature extraction. Besides, it will be compared block averaging feature extractions with wavelet and DCT, in terms of recognition rate performance.

2 MODEL, ANALYSIS, DESIGN, AND IMPLEMENTATION

2.1 Theory

The objective of feature extraction is to get a compact representation of an image. There are two properties of it [7].

a. It must be variant to characteristic differences among classes, so that it will help to differentiate images that have different classes.
b. It must be invariant to characteristic differences in a class, so that it will help to group images that have the same class.

A simple method in getting the compact representation is blurring and block averaging that inspired by Ethem [2]. A low-pass filter can be used to get a blur image. A kind of low-pass filter that can be used to get a blur image is 2D Gaussian filter, that formulated as follows [13].

$$h(x,y) = \frac{e^{-(x^2+y^2)/2\sigma^2}}{s_{xy}} \tag{1}$$

where

$$s_{xy} = \sum_x \sum_y e^{-(x^2+y^2)/2\sigma^2} \tag{2}$$

ISSN 2085-1944

and σ is standard deviation. Whereas $x = y = \left(-\frac{N-1}{2}, \cdots, \frac{N-1}{2}\right)$, where N is filter order.

Block averaging (see Algorithm 1) is performed to get a more compact image representation (lower image resolution), but it still has main features. If the image resolution too high, the computational load on the next process (classification process) will increase. On the contrary, if the image resolution too low, main features of the image will be too little, so that it will be difficult to differentiate images that have different classes.

Algorithm 1: Partition and Block Averaging

1. For image $p(m,n)$ which has size $2^k \times 2^k$ pixels (where $k = 1, 2, 3, \dots$), determine the size of block partition $s(u,v)$ which has size $2^\ell \times 2^\ell$ pixels (where $\ell = 0, 1, 2, \dots, 2^{k-1}$).
2. Partition image $p(m,n)$ by using partition block $s(u,v)$.
3. Compute average value in each partition block by using the following formula.

$$s_{avg}(u,v) = \frac{1}{2^\ell \times 2^\ell} \sum_{u=1}^{2^\ell} \sum_{v=1}^{2^\ell} s(u,v) \qquad (3)$$

4. A set of average values from partition blocks is feature extraction of image $p(m,n)$.

Figure 1 shows an example of blurring and block averaging of binary image of letter “a” 64x64 pixels. Blurring give basic shape of letter “a”, whereas block averaging gives a more compact representation of the image.

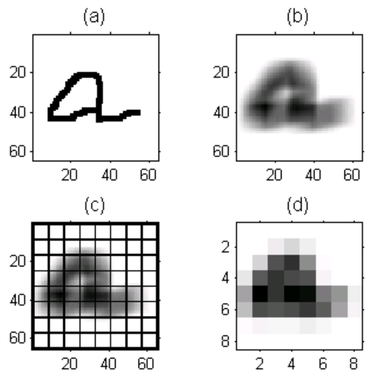

Figure 1. Example of blurring and block averaging; (a) Original binary image 64x64 pixels; (b) Blurred by using 2D Gaussian filter 14x14 with standard deviation 10; (c) Partition by using block partition 8x8 pixel; (d) The result of block averaging which gives a more compact representation of the image.

2.2 Materials and Equipments

Research materials are a set of isolated handwritten character in binary format, that come from data acquisition sheet scanned at 300 dpi. Data are got from 100 persons from various levels of age (10-70 years) and gender. In data acquisition sheet, every person write 3 times 26 letter ‘a’ – ‘z’. In this case, every person can write in cursive or hand printed style which depend on their own styles. Equipment research is a set of computer based on AMD Athlon 64 3500 + and RAM 1GB, that equipped with MATLAB 7.0.4.

2.3 Research Steps

By using materials and equipments above, the following steps has been performed.

2.3.1 Feature Extraction

The steps of blurring and block averaging feature extraction is shown in Figure 2 as follow.

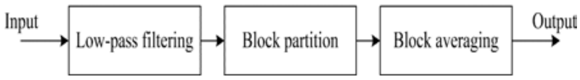

Figure 2. Feature extraction steps.

In Figure 2 above, low-pass filtering will blur the input image. The blurring level of the image depends on the size and standard deviation of 2D Gaussian filter used. See Figure 6.

The aim of block partition, in Figure 2, is to partition an image into block of images, for block averaging needs. The aim of block averaging is to get a set of values that represents input image. According to Algorithm 1, the size of partition block and the size of corresponding multiresolution images shown in Table 1.

Table 1 The size partition block and the size of corresponding multiresolution images.

Size of block partition (pixels)	Size of multiresolution image (piksel)
1	64x64
2x2	32x32
4x4	16x16
8x8	**8x8**
16x16	4x4
32x32	2x2
64x64	1

The size of multiresolution images in Table 1 is the size of feature extraction. By referring to other researcher that using multiresolution approach [6], from 64x64 pixels image, the optimal size of multiresolution image is 8x8 pixels. Therefore, according to Table 1, 8x8 pixels of multiresolution images corresponds with 8x8 pixels of block partition. In this research, 8x8 pixels of block partition was used.

2.3.2 Letter Recognition System

In order to be able to observe the performance of the above feature extraction, a system of handwritten letter recognition has been developed (see Figure 3). In that system, the input is an isolated letter image in binary format, whereas the output is a letter in text format.

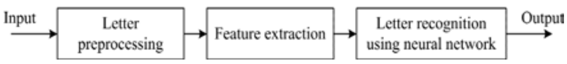

Figure 3 Letter recognition system.

2.3.2.1 Letter Preprocessing

Letter preprocessing in Fig 3 was performed in order to correct problems of slant, size, shift, and stroke-width. Figure 4 shows some steps in correcting that problems.

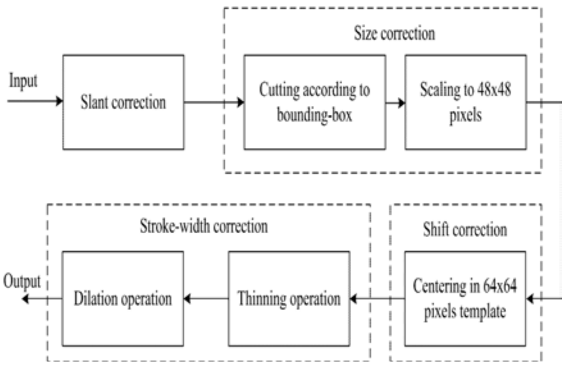

Figure 4. Steps in letter preprocessing.

Slant correction in Figure 4, was performed to straighten handwritten letter that slant to the left or right. For that purpose, it was performed evaluation of vertical projection histogram of handwritten letters that had been undergone shearing operation by using shearing coefficient {-0.4, -0.35, … , 0.4}. (In this case it was assumed that the slant of handwritten letter is in the range of shearing coefficient -0.4…0.4). Based on the observation, when the histogram has the highest variance, the letter seemed straight. Therefore, the straightness of the letter correspond with shearing coefficient that give highest variance.

The template size, in Figure 4, is 64x64 pixels. This size refer to other researchers that had been published their research. The minimum size of the template was 16x16 pixels [8], whereas the maximum one was 64x64 pixels [6]. In this research, the template size 64x64 pixels was used.

Letter scaling in Figure 4 was set to 48x48 pixels. This scaling was needed in order to avoid data cutting at the edge of filtered image. Based on the observation (by using 2D Gaussian filter 14x14 with standard deviation 10), adding 8 pixels around letter bounding box, has been adequate in order to avoid data cutting at the edge of filtered image. See Figure 6.

Thinning operation in Figure 4 used thinning algorithm from Zhang-Suen [15]. Whereas dilation operation used square structure-element 3x3. Based on simplicity, that dilation operation can be performed by using look-up table. Therefore, the use of square structure element 2x2, 3x3, and 4x4 will need 2^4= 16, 2^9=512, and 2^{16}= 65536 elements respectively. Therefore, optimal dilation that will not need too many elements is square structure-element 3x3.

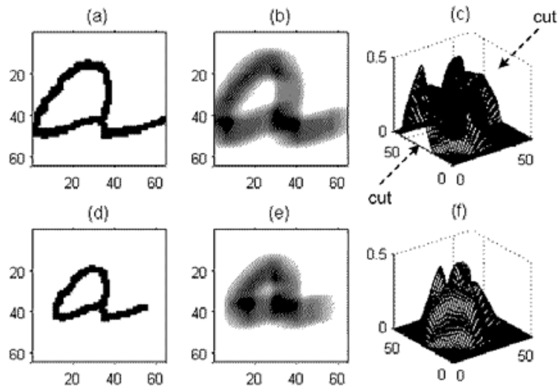

Figure 5. Example of letter scaling in 64x64 pixels template; (a) and (d) letter scaling to 64x64 and 48x48 pixels; (b) and (e) filtering result of (a) and (d) by using 2D Gaussian filter 14x14 with standard deviation 10; (c) and (f) mesh representation of (b) and (e).

2.3.2.2 Feature extraction

Feature extraction make use of feature extraction steps in Figure 2.

2.3.2.3 Neural Network

Neural network that used in letter recognition was backpropagation neural network. It is described in detail as follow.

a. Input layer has 64 neurons which correspond with the number of feature extraction elements.

b. Output layer has 26 neurons which correspond with the number of alphabet letter 'a' – 'z'. Transfer function in this layer is unipolar sigmoid, that correspond with the network output i.e. in the range of 0…1.
c. Neural network has 2 hidden layers i.e. hidden layer 1 and 2 which have 64 and 312 neurons respectively. The number of neurons in each hidden layer is got from the experiment, where by using 64 and 312 neurons in hidden layer 1 and 2, it gave the highest recognition rate. Transfer functions in each hidden layer is bipolar sigmoid, that correspond with internal data processing in neural network which in the range -1 … 1.

Notes:

a. In case of pattern recognition that based on multiresolution, Suhardi [11] found that backpropagation neural network with two hidden layers, could give better recognition rate compare with one hidden layer.
b. Sigmoid function is a function that commonly used in backpropagation neural network [3].
c. Training of backpropagation neural network can be more effective by using bipolar data processing in the range of -1…1 [11].

Training and testing of neural network

The neural network was trained by using resilient backpropagation algorithm [9]. This algorithm is the fastest algorithm for pattern recognition [14]. Stopping criterion in training make use of validation, in order to avoid under-training or over-training. Patterns that used in training and testing are images of isolated handwritten letter, that come from 100 persons which further processed become three pattern sets as follows.

1. Training set
   This set is used in training (in updating the neuron's weights). This set consist of 13,000 patterns as follows.
   a. There are 2,600 corrected patterns from group 1.
   b. There are 5,200 corrected patterns from group 2, that come from corrected patterns from group 2 which rotated -5^0 and 5^0.
   c. There are 5,200 corrected patterns from group 3, that come from corrected patterns from group 3 which rotated -10^0 and 10^0.

   Note:
   a. Every group consist of 2,600 patterns.
   b. Corrected patterns are original patterns that have undergone slant, size, shift, and stroke width corrections.
   c. It was assumed that the rotation in input patters is in the range of -10^0 .. 10^0.
2. Validation set
   This set is also used in training (in stopping the training process). This set consist of 2,600 corrected patterns from group 2.
3. Test set
   This set is used in testing the neural network whis has been trained. This set consist of 2,600 corrected patterns from group 3.

3 RESULT

3.1 The Effect of 2D Gaussian Filter

Table 4. The effect of 2D Gaussian filter sizes and standard deviation in letter recognition rate

Size of 2D Gaussian filter	Standard deviation of 2D Gaussian filter			
	2	6	**10**	14
6x6	83.38 %	83.17 %	83.72 %	83.26 %
10x10	84.23 %	85.31 %	85.38 %	85.74 %
14x14	85.08 %	86.34 %	**86.88 %**	86.42 %
18x18	85.29 %	86.55 %	85.93 %	86.18 %

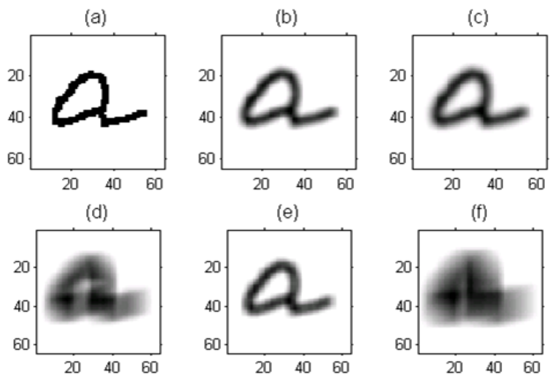

Figure 6. Example of blurring effect by using 2D Gaussian filter; (a) Binary image 'a'; (b) Blurring by using filter size 6x6 with standard deviation 2; (c) Blurring by using filter size 18x18 with standard deviation 2; (d) Blurring by using filter size 14x14 with standard deviation 10; (e) Blurring by using filter size 6x6 with standard deviation 14; (f) Blurring by using filter size 18x18 with standard deviation 14.

The effects of 2D Gaussian filter sizes in letter recognition rate is shown in Table 4. Table 4 indicates that reducing filter size and standard deviation have tendency to reduce recognition rate. If filter size and its standard deviation too small, it will give a very sharp image, so that, basic shape of the image become very clear. In this case, a class of image can have more than one basic shape, that finally it will be difficult to group images that have the same class. On the contrary, if filter size and its standard deviation too big, it will give a very blur image, so that, basic shape of the image become unclear. In this case, some class of images can have one basic shape, that finally it will be difficult to differentiate images that have different classes. See Figure 6.

Table 4 also indicates that 2D Gaussian filter 14x14 with standard deviation 10 is the optimal filter in blurring binary image 64x64 pixel, since it has the highest recognition rate. In this case optimal also has a meaning that basic shape of the image still clear, i.e. not too detail nor too blur, where quantitatively it give highest recognition rate. See Figure 6.

3.2 Performance Comparison with Other Feature Extractions

Performance comparison in terms of recognition rate with other feature extractions i.e. wavelet (by using Haar wavelet) and DCT (Discrete Cosine Transform) is shown in Table 4.

Table 4. Performance comparison in terms of recognition rate

Feature extraction	Block averaging	Wavelet	DCT
Average recognition rate	86.88 %	87.34 %	87.55 %

Notes:

a. Average recognition rate values are got from 5 neural networks.
b. There are 2,600 test patterns for testing.
c. Low-pass filter for blurring is 2D Gaussian filter 14x14 with standard deviation 10.
d. Feature 8x8 is extracted from image 64x64 pixels.

Table 4 shows insignificant differences in terms of recognition rate performance, among block averaging, wavelet, and DCT. Therefore it can be said that recognition rate performance of block averaging is similar with wavelet and DCT.

4 CONCLUSION AND DISCUSSION

Blurring by using 2D Gaussian filter which has 14x14 in size and 10 in standard deviation is the optimal filter in blurring binary image 64x64 pixel, since it has the highest recognition rate. In This case optimal also has meaning that basic shape of the image still clear, i.e. not too detail nor too blur, where quantitatively it give highest recognition rate.

Feature extraction by using block averaging can give insignificant difference in terms of recognition rate performance, if it compared to feature extraction by using wavelet and DCT.

Further exploration on the performance of block averaging feature extraction, on the different template sizes and also different block sizes will give more complete picture about block averaging.

ACKNOWLEDGEMENTS

The author would like to thank Prof. Adhi Susanto M.Sc. Ph.D. from Electrical Engineering of Gadjah Mada University for helpful discussion.

REFERENCES

[1] Arica, N. and F.T. Yarman-Vural (2001) An Overview of Character Recognition Focused On Off-line Handwriting. IEEE Trans. Systems, Man, and Cybernetics – Part C: Application and Rev. Vol. 31. No. 2, pp 216-233.

[2] Ethem, A. (1998) Techniques for Combining Multiple Learners, Proceedings of Engineering of Intelligent Systems '98 Conference. Vol.2, pp. 6-12.

[3] Fausett, L. (1994) Fundamentals of Neural Networks. Prentice Hall International, Inc. New Jersey.

[4] Lee, S. W. and Y. J. Kim (1995) Multiresolutional Recognition of Handwritten Numerals with Wavelet Transform and Multilayer Cluster Neural Network. Proceedings of 3rd International Conference Document Analysis and Recognition, Montreal, Canada, pp 1010-1014.

[5] Mozzafari, S., K. Faez and H. R. Kanan (2004) Feature Comparison between Fractal Codes and Wavelet Transform in Handwritten Alphanumeric Recognition Using SVM Classifier. Proceedings of 17th International Conference on Pattern Recognition, Cambridge, Inggris, vol 2, pp 331-334.

[6] Mozaffari, S., K. Faez, H. R. Kanan and M. Ziyaratban (2005) Farsi/Arabic Handwritten Digit Recognition Using Fractal, Wavelet Nearest Neighbor Classifiers and Eigen image Method. Proceedings of the First International Conference on Modeling,

Simulation and Applied Optimization, Sharjah, Uni Emirat Arab.
[7] Oh, S. and C. Y. Suen (1998) Distance Features for Neural Network based Recognition of Handwritten Characters. International Journal on Document Analysis and Recognition, 1(2), pp 73-88.
[8] Paterson, D. W. (1996) Artificial Neural Networks. Prentice Hall International, Inc., New Jersey.
[9] Riedmiller, M. and H. Braun (1993) A Direct Adaptive Method for Faster Backpropagation Learning: The RPROP Algorithm. Proceedings of the IEEE International Conference on Neural Networks, pp 586-591.
[10] Shioyama, T., H. Y. Wu and T. Nojima (1998) Recognition Algorithm Based On Wavelet Transform for Hand printed Chinese Characters. Proceedings of 14th International Conference on Pattern Recognition, vol. 1, pp 339-232.
[11] Suhardi, I. (2003) Evaluation of Artificial Neural Network for Handwritten Character Recognition Handprinted Style. Master Thesis. Electrical Engineering, Gadjah Mada University, Yogyakarta.
[12] Sumarno, L. (2007) Recognition of Noisy and Scaled Handwritten Recognition Based on DCT Feature Extraction by Using Probabilistic Neural Network, Sigma: Science and Technology Journal, Vol. 10, No. 2, pp 185-197.
[13] The Mathworks Inc. (a) (2005) Image Processing Toolbox: For Use with MATLAB. Version 5, The Mathworks Inc. Massachussets.
[14] The Mathworks Inc. (b) (2005) Neural Network Toolbox: For Use with MATLAB. Version 5, The Mathworks Inc. Massachussets.
[15] Trier, O. D., A. K. Jain, and T. Taxt (1996) Feature Extraction Methods for Character Recognition – A Survey. Pattern Recognition. Vol. 29, pp 641 – 662.
[16] Zhang, T.Y. and C. Y. Suen (1984) A Fast Parallel Algorithm for Thinning Digital Patterns. Comm. ACM, vol. 27(3), pp 236-239.