ภาคผนวก

ภาคผนวก ก.

ผลงานวิจัยที่ได้รับการตีพิมพ์เผยแพร่

1. Arit Thammano and Pruegsa Duangphasuk, **“Printed Thai Character Recognition Using the Hierarchical Cross-correlation ARTMAP”**, The 17^{th} IEEE International Conference on Tool with Artificial Intelligence (ICTAI05), Hong Kong, November 14-16 2005, pp.695-698

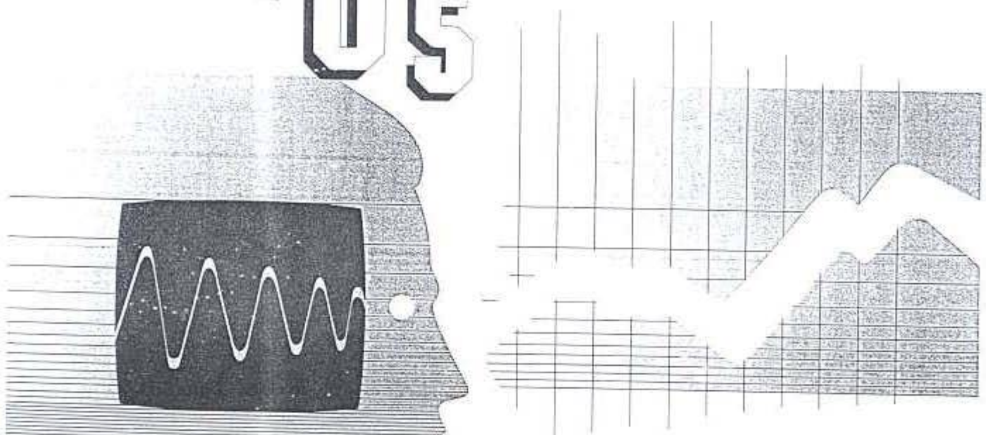
Proceedings

Seventeenth International Conference on

TOOLS WITH ARTIFICIAL INTELLIGENCE

ICTAI 05

14–16 November 2005
Hong Kong, China



Sponsored by
IEEE Computer Society
Co-Sponsored by
Information Technology Research Institute,
Wright State University
The Hong Kong University of Science and Technology

Edited by
Andrew Lim

IEEE COMPUTER SOCIETY

IEEE

Printed Thai Character Recognition Using the Hierarchical Cross-correlation ARTMAP

Arit Thammano and Pruegsa Duangphasuk
Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang
Bangkok, 10520 Thailand
Phone: (662)737-2551 to 4 ext. 529, Fax: (662)326-4332
E-mail: arit@it.kmitl.ac.th and pruegsa@mut.ac.th

Abstract

Traditionally, Thai characters are composed of circle, zigzag line, curve, and head. However, many new Thai fonts, which are now gaining in popularity, do not follow the traditional writing rule; the head has been omitted from the characters. Without the head, it is very difficult to segregate the characters. Even the best commercial Thai OCR software has difficulty in recognizing this kind of character. Therefore, the hierarchical cross-correlation ARTMAP is proposed in this paper to recognize the no-head Thai characters.

1. Introduction

Thai character recognition is a very complex problem. Even though many researches have been conducted on the recognition of Thai characters since the past decade, the success in recognizing Thai characters is still limited, compared to English characters. This is due to the following reasons: (a) there are a number of characters in Thai language, each of which has a complicated structure; (b) many Thai characters look very alike, e. g. ด-ต-ค, ข-ช-ซ, ม-ฆ, ฎ-ฏ, ก-ถ-ภ, ผ-พ; (c) many new Thai fonts violate the traditional writing rule. Traditionally, Thai characters are composed of circle, zigzag line, curve, and head. However, many new fonts omit the head of the characters for the sake of its beauty. The head is the loop at the beginning of the characters. It is one of the important features used to differentiate one character from another. Without the head, the characters are even more similar as shown in Figure 1. The situation gets even worse when these font styles are growing in popularity, but researchers do not pay much attention to them.

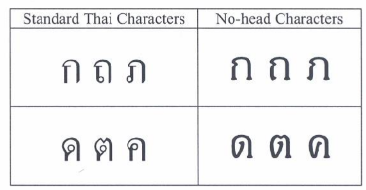


Figure 1. The similarity of Thai characters when the head of the characters are left out

Since the past decade, there has been a considerable growth in the Thai character recognition area. However, most of the researches put a lot of effort into the recognition of standard Thai characters. The methods widely used are the statistical approach [1, 2], the neural network based approach [3, 4, 5], and the hybrid approach [6, 7]. In this paper, the hierarchical cross-correlation ARTMAP is proposed to recognize the no-head Thai characters.

Following this introduction, section 2 describes the architecture of the proposed model and its learning algorithm. The preprocessing of the character images is described in section 3. In section 4, the experimental results are demonstrated and discussed. Finally, section 5 is the conclusion.

2. The proposed model

The proposed model is a four-layer feedforward neural network as shown in Figure 2. The first layer is the input layer, which consists of 9 segments. Each segment contains N input nodes. The number of input nodes in a segment is associated with the number of feature components extracted from each segment of the

Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'05)
1082-3409/05 $20.00 © 2005 IEEE

character image. The second layer is the hidden layer. Each node in the s^{th} segment of the second layer is fully connected to the input nodes of the same segment via the connections w_{ji}^s. The weight vector w_j^s of dimension N represents the reference pattern of the j^{th} node in the s^{th} segment of the hidden layer. The third layer is the cluster layer. The nodes in the second and the third layers are constructed during the training process. The fourth layer is the output layer. Each node in the output layer represents a class of character. During the supervised learning, the input vector is presented to the model, together with its respective target output vector. The input vector is denoted by $I = \{I^1,...,I^s,...,I^9\} = \{(I_1^1,...,I_N^1),...,(I_1^9,...,I_N^9)\}$, where s is the s^{th} segment in the input layer, and N is the number of feature components in each segment. Once the model receives the input and its associated target output (I, O), the maximum of the normalized cross-correlation between the input vector of the s^{th} segment (I^s) and each weight vector of the same segment (w_j^s) is computed and the outputs of the j^{th} node, $T_j^s(I^s)$, in the s^{th} segment of the hidden layer are then determined.

$$T_j^s(I^s) = f(I^s, w_j^s) \tag{1}$$

$$f(x, y) = \max_p \left\{ \frac{\sum_i x_i y_{i-p}}{\sqrt{\left(\sum_i x_i^2\right)\left(\sum_i y_i^2\right)}} \right\} \tag{2}$$

where $i = 1, \ldots, N$.
$j = 1, \ldots, M^s$.
$s = 1, \ldots, 9$.

N is the number of feature components in each segment.

M^s is the number of hidden nodes in the s^{th} segment.

f(x, y) is the maximum of the normalized cross-correlation at any shifted period p.

p is the lag variable, which is between -(N-1) to (N-1).

For each segment, the hidden node with the highest output value, $T_j^s(I^s)$, is selected as the winning node. Next, the vigilance criterion is evaluated to check whether the degree of matching between the input pattern and the reference pattern of the chosen node is within an acceptable level.

$$T_j^s(I^s) \geq \rho_{hidden} \tag{3}$$

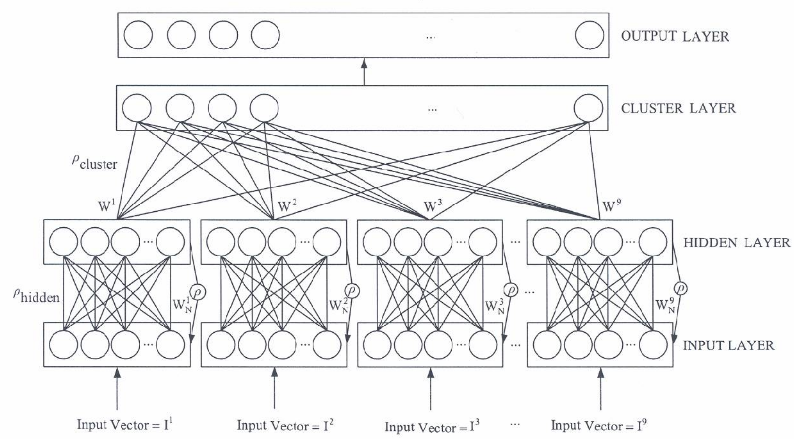


Figure 2. Architecture of the hierarchical cross-correlation ARTMAP

where ρ_{hidden} is the vigilance parameter at the hidden layer. It has a value between 0 and 1. If the winning node meets the above criterion, the weight vector of the winning node (w_J^s) will be updated according to the following equation:

$$w_J^s(t+1) = \beta(I^s \wedge w_J^s(t)) + (1-\beta)w_J^s(t) \quad (4)$$

However, if the condition in (3) is not satisfied, a new hidden node is recruited to code the input pattern. The weight of this new node is initialized to be equal to the input pattern.

$$w_J^s = I^s \quad (5)$$

Next, the weight vector of the winning node is transmitted to the next layer. The choice function of each k^{th} node in the cluster layer is then evaluated as follows:

$$C_k = f(w_J, w_k) \quad (6)$$

where $w_J = \{w_J^1, w_J^2, ..., w_J^s, ..., w_J^9\}$.

$w_k = \{w_{k1}, w_{k2}, ..., w_{ks}, ..., w_{k9}\}$.

$f(w_J, w_k)$ is the maximum of the normalized cross-correlation between the weight vector transmitted from the hidden layer and the weights of the connections from the hidden layer to the k^{th} node of the cluster layer.

The system then makes a cluster choice by selecting the winning node K with maximum choice function value, among all the nodes k in the cluster layer.

$$C_K = \max_k\{C_k\} \quad (7)$$

Next, C_K is compared to the vigilance parameter ($\rho_{cluster}$). $\rho_{cluster}$ is the vigilance parameter at the cluster layer. It has a value between 0 and 1. If C_K is greater than or equal to $\rho_{cluster}$ and the winning cluster node K belongs to the correct class defined by the target output vector, the weight vector w_{ks} will be updated according to where the weight vector w_J^s is originated:

(a) The weight vector w_J^s is transmitted from the newly recruited J node,

$$w_{ks}(t+1) = \beta(w_J^s \wedge w_{ks}(t)) + (1-\beta)w_{ks}(t) \quad (8)$$

(b) Otherwise,

$$w_{ks}(t+1) = w_{ks}(t) \quad (9)$$

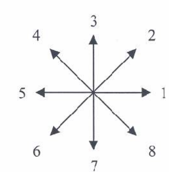
However, if C_K is less than $\rho_{cluster}$ or the winning cluster node K does not belong to the correct class defined by the target output vector, a new cluster node is recruited and its weight is initialized to be equal to the weight vector transmitted from the hidden layer (w_J). Then the connection between a new cluster node and the target output is created.

During testing, each testing vector is applied in turn and its class is predicted. The class whose cluster node returns the maximum output value is the result of the prediction.

3. Preprocessing

An optical scanner scans a printed document and converts it into an image. The segmentation process then segments the entire image into individual character blocks. Once the image is turned into character blocks, the thinning algorithm is applied to each character image to reduce the thickness of the character image to its skeleton. Next, each character block is divided into 9 equal segments and the feature is extracted from the thinned character image. In this paper, the merely feature is a list of directional codes. There are 8 directional codes, labelled "1" to "8". "1" represents the line whose angle is between $15\pi/8$ and $\pi/8$, "2" represents the line whose angle is between $\pi/8$ and $3\pi/8$, "3" represents the line whose angle is between $3\pi/8$ and $5\pi/8$ and so on.

In the feature extraction process, the starting point of a stroke must be identified first, then the directional code is used to traverse along the contour of the character. Figure 4 shows the directional codes for the character "ə."

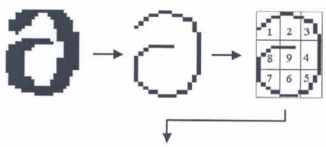


"1" → 1.000
"2" → 0.125
"3" → 0.250
"4" → 0.375
"5" → 0.500
"6" → 0.625
"7" → 0.750
"8" → 0.875

Figure 3. Eight directional codes

4. Experimental results

The performance of the proposed network is evaluated against the fuzzy ARTMAP neural network and two of the best commercial software available in the market. The commercial products used in this research are "ArnThai version 2.5" and "Thai OCR version 1.5b." ArnThai is a commercial software developed by National Electronic and Computer Technology Center (NECTEC). It is the most widely used and the most precise character recognition software in Thailand.



[0.125 1 0.125 1 0.125 1 1 1 0.125],[1 1 1 1 0.875 1 1 1], [0.875 1 1 1 1 0.875 0.75 0.75 0.75 0.875 0.75 0.75 0.75], [0.75 0.75 0.75 0.75 0.75 0.75 0.75 0.75 0.75 0.75], [0.75 0.75 0.625 0.75 0.625 0.75 0.625 0.5 0.5 0.5 0.5 0.5], [0.625 0.5 0.5 0.5 0.5 0.5 0.5 0.375],[0.5 0.5 0.5 0.375 0.5 0.375 0.5 0.375 0.25 0.25 0.25 0.25 0.25],[0.25 0.25 0.25 0.25 0.125 1 0.125 1 0.125 1 1 1 1],[1 1 1 1 1]

Figure 4. Directional codes for the character "ə"

To compare the performance of the selected systems, eight experiments have been conducted on 12 no-head fonts. The twelve no-head fonts used in this research are Boaboon, Chalit, Chanok, DB Private, Jasmine, Js 75 Pumpu, Kodchiang, Lily, Pat Ex, Patpong, Sathorn, and Silom. In each experiment, six fonts are randomly chosen to use as the training data, while the remaining fonts are used as the testing data.

The experimental results (Table 1) demonstrate a very strong performance of the proposed approach. It outperforms fuzzy ARTMAP and two of the best commercial products by a wide margin. However, it should be noted here that the inability of the available commercial products to retrain the models with new data sets is probably the cause of their failure to recognize the above no-head fonts.

Table 1. The experimental results

	Recognition Rate (%)			
	Proposed Model	Fuzzy ARTMAP	ArnThai	Thai OCR
1	76.74	55.04	27.13	22.87
2	86.11	59.52	21.83	19.05
3	89.29	59.92	23.02	25.00
4	81.40	54.26	31.01	29.46
5	82.56	55.04	24.42	20.16
6	85.71	58.33	19.05	21.83
7	85.94	56.64	22.27	23.44
8	82.42	55.86	25.00	22.66

5. Conclusion

The experiments show that the proposed approach can be used successfully to solve the printed Thai character recognition problem. It can recognize the no-head Thai characters, which are problematic even for the best commercial Thai OCR software. With its ability to separately train each segment of the network, the hierarchical cross-correlation ARTMAP can achieve much higher performance on the experimental data in comparison to the fuzzy ARTMAP neural network and two of the best commercial software available in the market.

6. References

[1] R. Nopsuwanchai and D. Povey, "Discriminative Training for HMM-based Offline Handwritten Character Recognition," *Proceedings of the Seventh International Conference on Document Analysis and Recognition*, Vol. 1, August 2003, pp. 114-118.

[2] T. Theeramunkong and C. Wongtapan, "Off-line Isolated Handwritten Thai OCR Using Island-based Projection with N-gram Model and Hidden Markov Models," *Information Processing and Management*, Vol. 41, Issue 1, January 2005, pp. 139-160.

[3] C. Tanprasert and T. Koanantakool, "Thai OCR: A Neural Network Application," *Proceedings of the 1996 IEEE TENCON*, Vol. 1, 1996, pp. 90-95.

[4] B. Kijsirikul, S. Sinthupinyo, and A. Supanwansa, "Thai Printed Character Recognition by Combining Inductive Logic Programming with Backpropagation Neural Network," *Proceedings of the 1998 IEEE Asia-Pacific Conference on Circuits and Systems*, 1998, pp. 539-542.

[5] P. Phokharatkul and C. Kimpan, "Recognition of Handprinted Thai Characters Using the Cavity Features of Character Based on Neural Network," *Proceedings of the 1998 IEEE Asia-Pacific Conference on Circuits and Systems*, 1998, pp. 149-152.

[6] N. Premchaiswadi, "A Study on Printed Thai Character Recognition," *Doctoral Dissertation*, Waseda University, Tokyo, March 2001.

[7] A. Thammano and P. Ruxpakawong, "Printed Thai Character Recognition Using the Hybrid Approach," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, Vol. E85-A, No. 6, June 2002, pp. 1236-1241.

Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence (ICTAI'05)