

Handwritten Thai Character Recognition Using Fourier Descriptors and Robust C-Prototype

Olarik Surinta
 Management Information System
 Faculty of Informatics
 Mahasarakham University
 Vittayaborikan B Building
 Tambon Kamreang,
 Amphur Kantarawichai, Mahasarakham
 44150, THAILAND
 Phone: +66-4375-4359 Ext 2453
 olarik.s@msu.ac.th

Supot Nitsuwat
 Multimedia Technology Center (MTC)
 Department of Computer Education
 Faculty of Technical Education
 King Mongkut's Institute of Technology
 North Bangkok
 1518 Piboolsongkram Rd., Bangsue,
 Bangkok 10800, THAILAND
 Phone: +66-2913-2500 Ext. 2703
 sns@kmitnb.ac.th

Abstract

This research paper proposes the statistical-based method for Thai handwritten character recognition using Fourier descriptors and robust C-Prototype clustering. Recognition scheme is based on features extracted from Fourier transform of the edge of character-image. Therefore, the character-image is described by a group of descriptors. We train the system using the RCP training scheme to find the centroid of the prototype (44 Prototypes) and membership function. Finally, the FD of unknown character-image is used to perform recognition step. In this way the experimental results of recognition, RCP can perform with accuracy up to 91.5%.

Keywords

Fourier Descriptors, Robust C-Prototypes, Handwritten Thai Character Recognition

1. Introduction

Pattern recognition scheme has numerously become a suffice tool utilized in character recognition. Generally, computer will be programmed to provide the functionality in order to classify each of character's property separately, defined as input character. These inputs will be determined for matching with the provided character patterns consequently. This paper proposes the offline processing, which the input data, gray scale of 256 levels and 300dpi, is retrieved through scanner. The processing is based on Thai characters on which preprocessing have been conducted. There are 44 Thai characters:

ก ข ฃ ค ฅ ฆ ง	จ ฉ ช ซ ฌ ญ
ฎ ฏ ฐ ฑ ฒ ณ	ด ต ถ ท ธ น
บ ป ผ ฝ พ ฟ	ภ ม ย ร ล ว
ศ ส ฮ ห อ อ	

2. Data Preprocessing

Character-images are images of Thai handwritten characters on which are written down on non-line paper. The output will be stored in the term of digital data by scanning. One bitmap file with gray scale pattern and 256 levels specifics one character.



Figure 1 A prototype character-image.

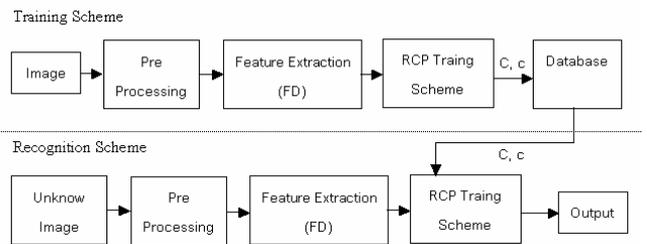


Figure 2 Thai handwritten recognition scheme flow diagrams.

3. Image Processing

3.1 Edge Detection

Edge detection is one of an important image processing phases. Detecting edge in any object has 2 important conventions, the object (image) must be continuous image and the image must be scaled as white and black tone. This paper uses chaining technique to detect the image's edge. More general, this technique implies that motivation of moving along the edge of the image and stop at the beginning position in order to get the image's edge.

3.2 Binarization

Binarization converts gray-level image to black-white image. Generally, image will be separated into 2 components. The first component is the object, while another is background. Generally, object component is smaller, in size, than background in any image. To exacting the object component from background, this scheme will check on every point of pixel value to one intermediate value, called Thresholding that can be calculated as the following formula:

$$g(x, y) = \begin{cases} 0 & \text{if } g(x, y) < T \\ 1 & \text{if } g(x, y) \geq T \end{cases} \quad (1)$$



(a)



(b)

Figure 3 The example of binarization scheme.

In this paper, data, the image file of Thai handwritten, is stored as bitmap pattern. The individual bit bares 2 possible values:

- 1 refers to background and
- 0 refers to object or content

Figure 4 show the diagram of extracting the object from the background component in the image.

3.3 The extraction of outer edge from the object component

From the previous step, extracting the object from its background, the result is all bits contain 1. Therefore, this will be used to detect the outer rim, laying on the object's edge, of the object and separate those from the resulting object in order to get it actual edge. There are many methodologies used to manipulate on this particular case, i.e. chaining code and morphology, etc. Moreover, this paper used chaining code convention.

The resulting image is actually a structural character; comprising many actually points laying on the image. Therefore, detecting the direction of those points has been applied in order to simplify the processing. This implementation is based on chaining code technique, changing the points to the form of numeric representation. Eventually, in this paper, the direction has been classified by 8 categories:

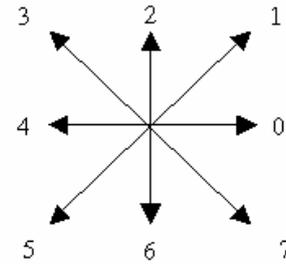


Figure 5 chaining with 8 directions

Chaining application

1. Get the starting point of the code by detecting the points whose content is 0 and specify that point to be the beginning point for the matrix consequently, defined as the current point.
2. From (1), the system will check for any neighbor point and the current pointer will move along a closed point which its values is 0, regarding to the 8-direction. If there exists more that 1 point closed to the current point whose value contains 0, the rotation will be forced in the clockwise pattern. This routine work recursively until it reach to the beginning point.

Once the edge of image has discovered, shown in figure 4, the process needs to find the character line. The coordinate (x_k, y_k) is represented by complex number as the formula:

$$u_k = x_k + iy_k \quad (2)$$

6.0000 + 4.0000i	5.0000 + 5.0000i	4.0000 + 6.0000i
4.0000 + 10.0000i	4.0000 + 11.0000i	4.0000 + 12.0000i
4.0000 + 16.0000i	4.0000 + 17.0000i	4.0000 + 18.0000i
4.0000 + 22.0000i	4.0000 + 23.0000i	4.0000 + 24.0000i
4.0000 + 28.0000i	4.0000 + 29.0000i	4.0000 + 30.0000i
8.0000 + 33.0000i	9.0000 + 33.0000i	10.0000 + 33.0000i
14.0000 + 33.0000i	15.0000 + 33.0000i	16.0000 + 33.0000i
20.0000 + 33.0000i	21.0000 + 33.0000i	22.0000 + 33.0000i
26.0000 + 33.0000i	27.0000 + 33.0000i

Figure 6 coordinate (x_k, y_k) represented in character image

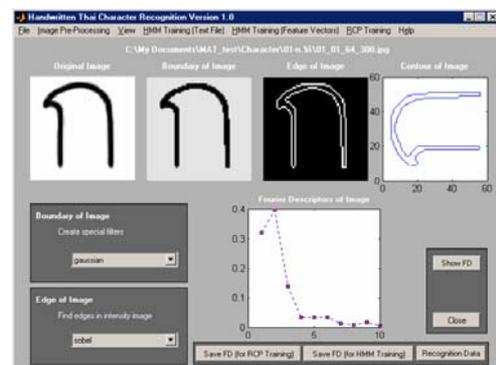


Figure 7 the character images the adjustment scheme

4. Fourier Descriptors

Generally, signals or functions existing in general physical problem and engineering are recognized as a period function, which can be represented in the term of sine and cosine function. Moreover, there two types of signal are the basis signal used to determine all other signals by replace each term of series with sine and cosine function.

4.1 Fourier Transform

The main objective of signal analysis is to be able transform any signal to periodic function. Furthermore, in the frequency analysis of any linear system, input function or deriving function will particularly transformed to the period function, either discrete or continuous, as well.

Implicitly, Fourier features of Fourier transform is generally said to be to the tool used to show or transform any function in them of frequency. If any function can be represented by Fourier Transform $F(\omega)$, on another hand, the Fourier Transform $F(\omega)$ can, also, transformed to function $f(t)$ as the relation shown below :

$$f(t) \leftrightarrow F(\omega) \tag{3}$$

Particularly, some special properties concealed in the image must be detected in the phases. Using Fourier conversion, the Fourier descriptors has been consistent to pixel alliterating, image rotating and image resizing of the pattern image; simply the robustness has been identified to the pattern images.

4.2 Fourier Features

Fourier features used to describe the edge of the object works by identify coordinate (x_k, y_k) ; $k = 0, 1, \dots, N-1$ where N is any other area in the image. All point

(x_k, y_k) will be represents as complex number. That is:

$$u_k = x_k + iy_k \quad ; i = \sqrt{-1} \tag{4}$$

Therefore, the DFT (Discrete Fourier Transform) f_l can be derived as below:

$$f_l = \sum_{k=0}^{N-1} u_k \exp\left(-j \frac{2\pi}{N} lk\right) \quad l = 0, 1, \dots, N-1 \tag{5}$$

From the above formula, coefficient vector will be automatically calculated. This vector fits as 1 dimension with the size of 1x10 or 1xn

5. Fuzzy C-Mean (FCM) Algorithm

FCM algorithm minimizes the following objective function

$$J_F(B, U; X) = \sum_{i=1}^C \sum_{j=1}^N (u_{ij})^m d_{ij}^2 \tag{2}$$

Where $u_{ij} \in [0,1]$ and $\sum_{i=1}^C u_{ij} = 1 \forall j$

x = the set of Feature Vector \bar{x}_j

$j = 1, 2, \dots, N$: N represents the image pixel

C = the group of images

$m \in [1, \alpha)$ = Fuzzifier

d_{ij}^2 = the distance from Feature Vector to the group of pattern

u_{ij} = member indicator of x_j in β_i

$B = (\beta_1, \dots, \beta_C)$ = C-tuple indicating

C - Cluster

$U = [u_{ij}]$ = matrix, $C \times N$ in size, representing the condition

(Constrained Fuzzy C - Partitions)

5.1 Object-Function Minimization in Robust C-Prototypes (RCP)

RCP can be determined in grouping phase in order to estimate C-Prototypes spontaneously, utilizing loss function (ρ) and square distance to reduce some noise. The definition can be expressed as:

$$J_F(B, U; X) = \sum_{i=1}^C \sum_{j=1}^N (u_{ij})^m \rho_i d_{ij}^2 \tag{3}$$

where $u_{ij} \in [0,1]$ and $\sum_{i=1}^C u_{ij} = 1 \quad \forall j$

The highest efficiency can be calculated as follow:

$$u_{ij} = \frac{1}{\sum_{k=1}^C \left[\frac{\rho_i(d_{ij}^2)}{\rho_i(d_{ik}^2)} \right]^{\frac{1}{m-1}}} \quad (4)$$

The best portion of weight function is that function with significant impact and efficiency in RCP-Algorithm. This weight function used for general estimation has been designed as symmetric distribution, which its center is the original point that differs from RCP-Algorithm used in this experiment. Therefore, Weight Function must be rebuilt:

$$\begin{aligned} w : \mathbb{R}^+ &\rightarrow [0,1] \text{ Subject to} \\ \mathfrak{T}_1 : w(d^2) &: \text{Monotonically function} \\ \mathfrak{T}_1 : w(d^2) &= 0 \text{ for } d^2 > T + \alpha S \end{aligned}$$

T and S are robust estimation of the average of square of distance and standard deviation, while α is any constant.

$$p_1 : w(0) = 1, w(T) = 0.5, w'(0) = 0$$

T and S are truly important for weighting, constructing the weight function. Therefore, the process needs the efficiency estimation. Obviously, Med-Median and Median of Absolute Deviation have been identified to be used in estimation, which is:

$$T_i = \underset{x_i \in X_j}{\text{Med}}(d_{ij}^2) \text{ and} \quad (5)$$

$$S_i = 1.418 \times \underset{x_i \in X_j}{\text{MAD}}(d_{ij}^2)$$

$$\text{where } X_i = \left\{ x_j \mid d_{ij}^2 \leq d_{kj}^2 \quad \forall k \neq i \right\} \quad (6)$$

the weight function $w : \mathbb{R}^+ \rightarrow [0,1]$ is defined as $\mathfrak{T}_1 - \mathfrak{T}_3$, which is :

$$w_i(d^2) = \begin{cases} 1 - \frac{d^4}{2T_i^2} & \text{where } d^2 \in [0, T] \\ \frac{[d^2 - (T_i + \alpha S_i)]^2}{2\alpha^2 S_i^2} & \text{where } d^2 \in [T_i, T_i + \alpha S_i] \\ 0 & \text{where } d^2 > T_i + \alpha S_i \end{cases} \quad (7)$$

loss function derived from weight function can be calculated regarding to (7)

(8)

$$\rho_i(d^2) = \begin{cases} d^4 - \frac{d^4}{6T_i^2} & \text{where } d^2 \in [0, T] \\ \frac{[d^2 - (T_i + \alpha S_i)]^2}{6\alpha^2 S_i^2} + \frac{5T_i + \alpha S_i}{6} & \text{where } d^2 \in [T_i, T_i + \alpha S_i] \\ \frac{5T_i + \alpha S_i}{6} + K_i & \text{where } d^2 > T_i + \alpha S_i \end{cases}$$

K_i is a constant;

$$K_i = \max_{1 \leq j \leq C} \left\{ \frac{5T_j + \alpha S_j}{6} \right\} - \frac{5T_i + \alpha S_i}{6}$$

when $i = 1, \dots, C$

in the (8), it has to add constant K_i in order to impede any noise. This will force the total values of loss function to be at least the average value of normal data, which every point of noise has the same member value.

5.2 The calculating of the distance between groups of images and relation among them

The required conditions to adjust the pattern characters for Mahalanobis Distance is

$$d_{ij}^2 = (\bar{x}_j - \bar{c}_i)^T M (\bar{x}_j - \bar{c}_i) \quad (9)$$

\bar{x}_j ; the feature vector if group of data

\bar{c}_i ; center vector of each cluster

M_i ; symmetric vector, which is positive definite matrix derived from

$$\bar{c}_i = \sum_{j=1}^N (u_{ij})^m w_{ij} \bar{x}_j / \sum_{j=1}^N (u_{ij})^m w_{ij} \quad (10)$$

and

$$M_i = |R_i|^{1/n} R_i^{-1} \quad (11)$$

where

$$C_i = \sum_{j=1}^N (u_{ij})^m w_{ij} (\bar{x}_j - \bar{c}_i)(\bar{x}_j - \bar{c}_i)^T / \sum_{j=1}^N (u_{ij})^m w_{ij}$$

C_i is ‘‘Robust Fuzzy Covarian Matrix’’

Therefore, the diagram of solving the problem by Robust-C Type is shown below:

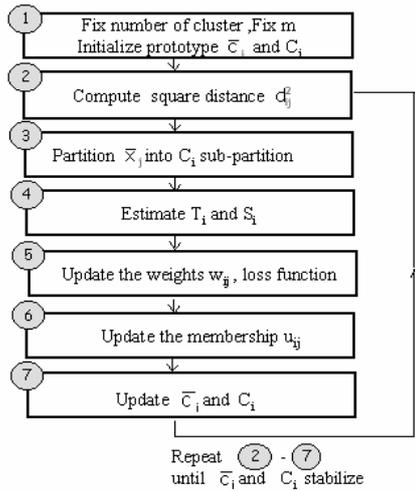


Figure 8 RCP Algorithm

of Informatics. University of Athens, Academic Press 24-28 Oval Road, London NW1 7DX, UK, 1998.
 [5] Yerin Yoo. Tutorial on Fourier Theory. n.p. , 2001.

6. Experimental Result

The experiment of Handwritten Thai Character Recognition using Fourier descriptors and robust C-Prototype has been divided into 2 steps to estimate accuracy. Step 1 and 2 will get 100 random characters from the first set of data.

The accuracy was computed, regarding to the results of the experiments. The accuracy of recognition RCP is 91.5 respectively. Figure 9.

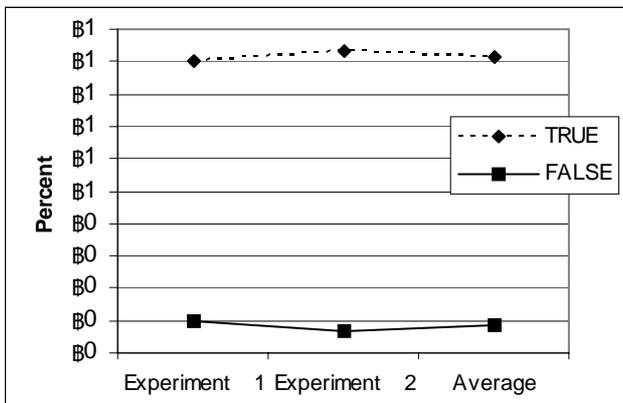


Figure 9 Graph of the accuracy of the experiment

7. References

[1] Castleman, K. R. Digital Image Processing. Prentice Hall. NJ, n.p. , 1995.
 [2] Hichem Frigui and Raghu Krishapuram. A Robust Algorithm for Automatic Extraction of an Unknown Number of Clusters from Noisy Data. Pattern Recognition Letters(17). 1223-1232, n.p. , 1996.
 [3] Lawrence R. Rabiner. A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. IEEE, n.p. , 1989.
 [4] Sergios Theodoridis and Konstantinos Koutroumbas. Pattern Recognition. Department