*Int. J. Appl. Math. Comput. Sci., 2010, Vol. 20, No. 4, 727–738*
DOI: 10.2478/v10006-010-0055-x

amcs

# EFFICIENT ONLINE HANDWRITTEN CHINESE CHARACTER RECOGNITION SYSTEM USING A TWO–DIMENSIONAL FUNCTIONAL RELATIONSHIP MODEL

YUN FAH CHANG *, JIA CHII LEE *, OMAR MOHD RIJAL **,
SYED ABDUL RAHMAN SYED ABU BAKAR ***

* Department of Mathematical and Actuarial Sciences
Universiti Tunku Abdul Rahman, Petaling Jaya, Malaysia
e-mail: changyf@utar.edu.my, sharon_jclee@hotmail.com

** Institute of Mathematical Sciences
University of Malaya, Kuala Lumpur, Malaysia
e-mail: omarrija@um.edu.my

*** Faculty of Electrical Engineering
Universiti Teknologi Malaysia, Skudai, Malaysia
e-mail: syed@fke.utm.my

This paper presents novel feature extraction and classification methods for online handwritten Chinese character recognition (HCCR). The $X$-graph and $Y$-graph transformation is proposed for deriving a feature, which shows useful properties such as invariance to different writing styles. Central to the proposed method is the idea of capturing the geometrical and topological information from the trajectory of the handwritten character using the $X$-graph and the $Y$-graph. For feature size reduction, the Haar wavelet transformation was applied on the graphs. For classification, the coefficient of determination ($R_p^2$) from the two-dimensional unreplicated linear functional relationship model is proposed as a similarity measure. The proposed methods show strong discrimination power when handling problems related to size, position and slant variation, stroke shape deformation, close resemblance of characters, and non-normalization. The proposed recognition system is applied to a database with 3000 frequently used Chinese characters, yielding a high recognition rate of 97.4% with reduced processing time of 75.31%, 73.05%, 58.27% and 40.69% when compared with recognition systems using the city block distance with deviation (CBDD), the minimum distance (MD), the compound Mahalanobis function (CMF) and the modified quadratic discriminant function (MQDF), respectively. High precision rates were also achieved.

**Keywords:** 2D functional classifier, coefficient of determination, handwritten Chinese character recognition, Haar wavelet, multidimensional functional relationship model.

## 1. Introduction

The problem of handwritten Chinese character recognition (HCCR) has received considerable attention from the research community due to its wide-ranging applications, which include text entry for form filling, message composition in mobile, computer-aided education, personal digital assistants (PDA) and handwritten document retrieval. However, HCCR is different from handwriting recognition of other languages. It poses a special challenge due to a complex structure, a large shape variation, a large character set and many instances of highly similar characters in Chinese words. Furthermore, even for the same character, there is a large variation in its graphical pattern because of the variability in writing styles of different writers. The area of handwriting recognition can be divided into two different categories: offline and online approaches. This study will concentrate on the latter.

Feature extraction (Michalak and Kwaśnicka, 2006; Świniarski, 2001) and classification (Miquelez *et al.*, 2004; Fujarewicz and Wiench, 2003) are crucial to a recognition system. However, defining a feature vector for handwritten Chinese character recognition is not trivial. A high recognition rate (accuracy and precision)

can be achieved by investigating the use of various features. Feature extraction schemes can be classed into three categories: structural, statistical and hybrid statistical-structural (Liu *et al*., 2004). Structural approaches such as the attributed relational graph (ARG) (Liu *et al*., 1996) and the fuzzy attributed relational graph (FARG) (Zheng *et al*., 1997) are amongst the earliest and most widely used methods. The increase in character complexity, for example, the increase in the number of strokes, will increase the size of the feature vector and subsequently influence recognition rates. Numerical measurements that describe the structure can be used to develop statistical approaches due to its computational efficiency. An alternative to the number of strokes, the number of occurrences for stroke directions of each character can be used to describe the structure (Kawamura *et al*., 1992; Kimura *et al*., 1997). The structural approach and statistical ideas have been combined, (for example, the hidden Markov model (HMM) (Shimodaira *et al*., 2003; Takahashi *et al*., 1997)), as an efficient way of temporal modeling.

Many distance measures for online character recognition have been developed which include modified quadratic discriminant functions (MQDFs) (Kimura *et al*., 1987), support vector machines (SVMs) (Gao *et al*., 2002), neural networks (NNs) (Romero *et al*., 1995), compound Mahalanobis functions (CMF) (Suzuki *et al*., 1997) and the city block distance with deviation (CBDD) (Kato *et al*., 1999). These distance measures achieve a high recognition rate in HCCR. Applications for embedded systems employ simple distance measures such as the minimum distance (MD) classifier (Gonzalez and Woods, 1993) for limited storage consideration (Liu *et al*., 2005) but may face problems with character shape deformation. These methods require normalization, which uses methods that may incur costs in terms of time.

This paper proposes a novel HCCR system using features extracted from the $X$-graph and the $Y$-graph. The sequence of points $(x_t,y_t)$, $1 \leq t \leq N$, where $t, N \in \mathbb{Z}^+$, obtained from the trajectory of handwritten Chinese character, is transformed into two separated graphs, firstly the graph of the $x$-coordinate versus the time sequence (called the $X$-graph), and secondly the graph of the $y$-coordinate versus the time sequence (called the $Y$-graph). The $X$-graph (or the $Y$-graph) may be treated as a discrete signal, $\boldsymbol{x} = (x_1, x_2, \ldots, x_N)$. The Haar wavelet transform was applied to handle the dimensionality problem. The coefficient of determination ($R_p^2$) for the two-dimensional unreplicated linear functional relationship model is proposed as a similarity measure and used for estimating recognition rates.

The main experiment in this study involved comparing a hand-written version of Chinese characters from the database CL2009. This study will describe the database or library of characters in the next section. Section 3 explains how the online handwritten characters were obtained, and the process of normalization and feature extraction is described. This is followed by the classification procedure in Section 4. Results and discussions are detailed in Section 5.

## 2. Database

HCL2000, ETL9B and CASIA are commonly used databases in Chinese character recognition, with some details given in Table 1. The number of classes is defined as the number of different characters, whereas a sample is defined as the character reproduced by different writers for each class (or character). The HCL2000 database was used by Long and Jin (2008) as well as Liu and Ding (2005), the ETL9B database by Dong *et al*. (2005) as well as Gao and Liu (2008), and the CASIA database by Gao and Liu (2008), respectively. The existing databases (Table 1) store many samples of different writing styles for each character, in order to cope with the problem of handwriting variation of different writers. This technique is faced with the problem of a large storage space. This study considers situations for a limited storage space only. In particular, only one sample of each character is available in the database.

A new database with only a single sample for each character will be created, namely, CL2009. This database is based on the Jun Das modern Chinese character frequency list (Dan, 2004), collected from a large corpus of Chinese texts obtained from online sources. 3000 frequently used simplified Chinese characters are selected and reproduced individually by an expert in *songti*, thus creating the characters in our database. A subset of CL2009 is given in Fig. 1.

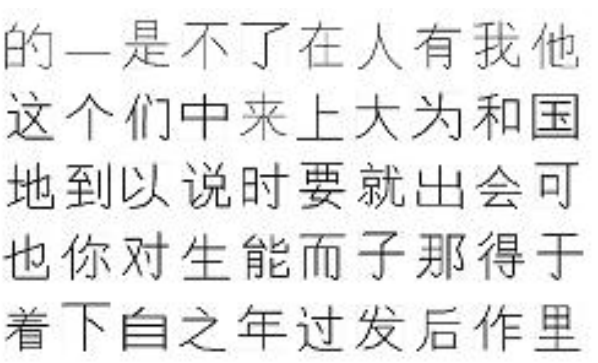



Fig. 1. Examples of 50 normalized Chinese characters in the *songti* written style.

## 3. Experiment

The recognition system was developed on a Dell Vostro 1400 N-Series notebook with an Intel(R) Core(TM)2 Duo

Table 1. Three commonly used databases for Chinese character recognition and a database newly created for this research.

| Database | Origin | No. of classes | No. of samples |
|---|---|---|---|
| HCL2000 | Collected by the Beijing University of Posts and Telecommunications for China 863 project . | 3755 | 1000 |
| ETL9B | Collected by the Electro-Technical Laboratory (ETL) of Japan | 2965 | 200 |
| CASIA | Collected by the Institute of Automation of the Chinese Academy of Sciences | 3755 | 300 |
| CL2009 | Based on the Jun Das modern Chinese character frequency list (Dan, 2004) | 3000 | 1 |

Processor T5470 and 1GB ($2 \times 512$ MB) 667MHz Dual Channel DDR2 SDRAM. The programming environment used is MATLAB and the implementations only utilize one CPU core.

Two writers reproduce each of the 3000 characters in CL2009 only once. Each writer was given one week to complete the reproduction process under similar conditions. The first writer (A) has more than 15 years of experience with Chinese characters, whereas the second writer (B) has six years of experience. The Wacom Intuos 3 pen tablet was used by each writer for the reproduction process. For each character, 128 points are used to represent each stroke. Thus, a $w$-stroke character, for example, will have a total of $128 \times w$ points.

Samples of 20 characters from Writer A and Writer B illustrate the size, slant and position variation, as well as character deformation as given in Figs. 2 and 3, respectively.

Once a character is written, it is cropped and normalized (or otherwise). This is followed by extracting the feature vector from the $X$-graph and $Y$-graph, which include using the Haar wavelet transformation for reducing the dimension. The derived feature vector is subjected to two-stage classification procedures: firstly, the rough classification, and secondly, the fine classification.

The proposed recognition system (Fig. 4) is compared with four other systems using the city block distance with deviation (CBDD), the minimum distance (MD), the compound Mahalanobis function (CMF) and the modified quadratic discriminant function (MQDF). The experiment was carried out for both normalized and non-normalized characters.

Fig. 2. Sample of 20 Chinese characters taken from Writer A.

Fig. 3. Sample of 20 Chinese characters taken from Writer B.

**3.1. Cropping.** Given the sequence of points for an input character, the maximum $x$ and $y$ coordinates and also their corresponding minimum were determined. Then, a subset of the original image, labeled as the subarea, $(y_{\min} : y_{\max}, x_{\min} : x_{\max})$ (from row $y_{\min}$ to row $y_{\max}$ and column $x_{\min}$ to column $x_{\max}$ ) was cropped (Fig. 5).

**3.2. Normalization.** The sequence of points $[x_{?t,y_{?t}}]$, which ranges within the cropped subarea is normalized to the size of $128 \times 128$:

$$x_t^* = 127\Big(\frac{x_t - x_{\min}}{x_{\max} - x_{\min}}\Big) + 1, \tag{1}$$

$$y_t^* = 127\Big(\frac{y_t - y_{\min}}{y_{\max} - y_{\min}}\Big) + 1. \tag{2}$$

An illustration of normalization is given in Figs. 5 and 6. The linear normalization (LN) method (Saeed, 2000; Deepu *et al.*, 2004) is preferred in our experiment over non-linear normalization (NLN) (Casey, 1970, Liu *et al.*; 2003, Liu and Marukawa, 2004;2005 ; Horiuchi *et al.*, 1997).

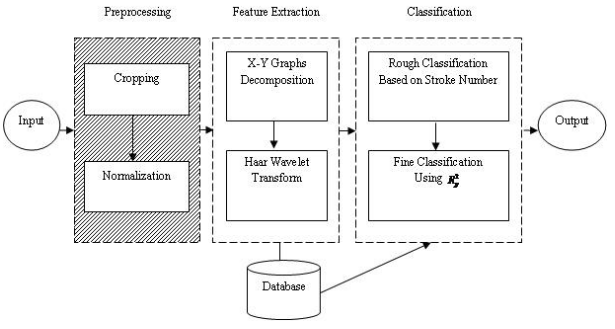



Fig. 4. Complete recognition system with preprocessing.

**3.3. *X*-graph and *Y*-graph.** The $X$-graph is defined as $\{t, x_t\}$, $1 \le t \le N = 128 \times w$, where $x_t$ is the value of the $x$-coordinate of a point on the character in position (space) $t$. The $Y$-graph $\{t, y_t\}$ is similarly defined. The subscript $t$ in $\{t, x_t\}$ and $\{t, y_t\}$ depends on the stroke direction, order and number, thus preventing the possibility of different patterns for the same character. In the example shown in Fig. 7, the values in the $X$-graph drop while in the $Y$-graph the values rise when the first stroke is written, whereas in the case of the second (horizontal) stroke (this order is also fixed in the Chinese character) the values in the $X$-graph rise while the corresponding values for the $Y$-graph remain unchanged. The process was repeated until the 7-th stroke. Consequently, the feature vectors obtained are $[x_1, \ldots, x_N]^T$ and $[y_1, \ldots, y_N]^T$, $N = 7 \times 128$ representing the $X$-graph and $Y$-graph, respectively.

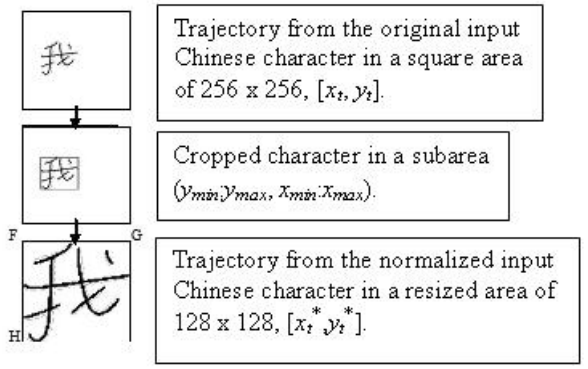



Fig. 5. Diagram of the whole preprocessing procedure for the Chinese character 'I' or 'me'. Point F is defined as the origin, FG is the $x$-axis and FH is the $y$-axis.

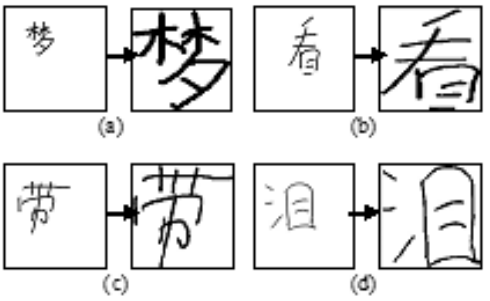



Fig. 6. Examples of non-normalized (left) and normalized (right) Chinese characters: character 'Dream' (a), character 'See' or 'look' (b), character 'Bring' (c) and character 'Tear' (d).

**3.3.1. Properties of the *X* and *Y*-graphs.** The motivation to use the $X$ and $Y$-graphs lies in properties. Only one character can be represented by a pair of the $X$-graph and $Y$-graph. Different writing styles will still yield the same pair of the $X$-graph and $Y$-graph. Finally, it is simple to use the $X$-graph and $Y$-graph. Henceforth the properties of uniqueness, invariance and simplicity are illustrated by examples in the following subsections.

(i) *Uniqueness*. For similarly shaped characters, Fig. 8, the $X$-graph and $Y$-graph are of different shapes. Hence, both the $X$-graph and $Y$-graph can be considered as useful features for discrimination.

(ii) *Invariance to different writing styles*. Different writing styles pose the main problem faced in developing a recognition system. The character, obtained from the database with its corresponding $X$-graph and $Y$-graph is given in Fig. 9(a). The same character written by two different writers is given in Figs. 9(b) and (c). Figures 9(b) and 9(c) show clear differences in writing styles, yet their $X$-graph and $Y$-graph are similar in appearance. Similar remarks can be made about the other characters in CL2009.

(iii) *Simplicity*. Transforming the character coordinates separately into the $X$-graph and the $Y$-graph and obtaining the feature vectors from the corresponding graphs are the only tasks required. This simple approach should boost the efficiency and speed of the recognition system.

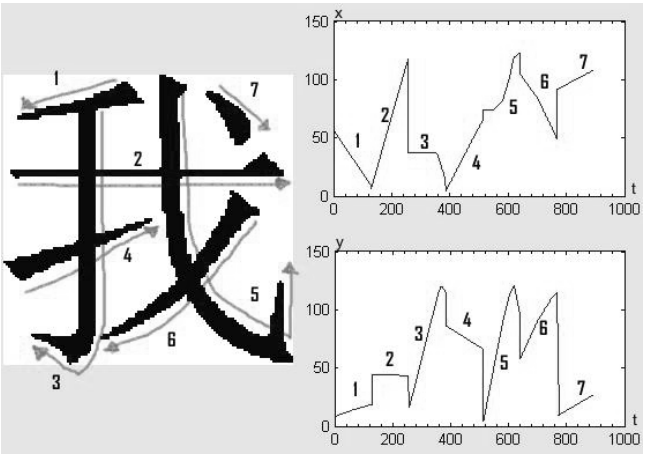
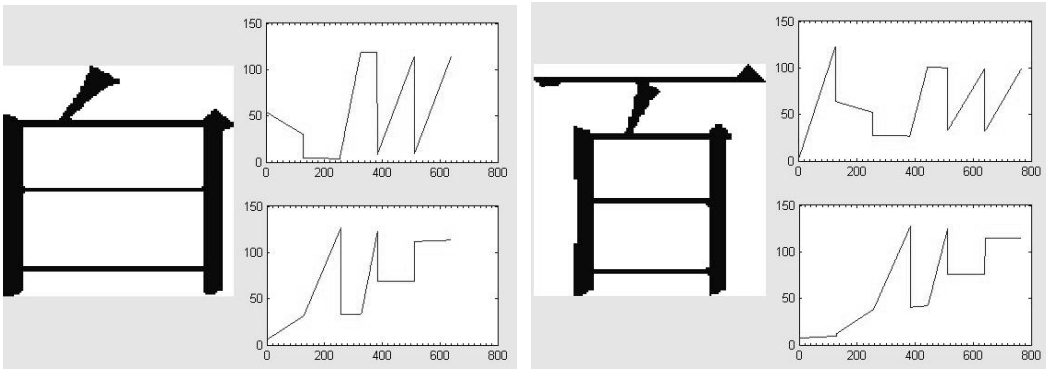


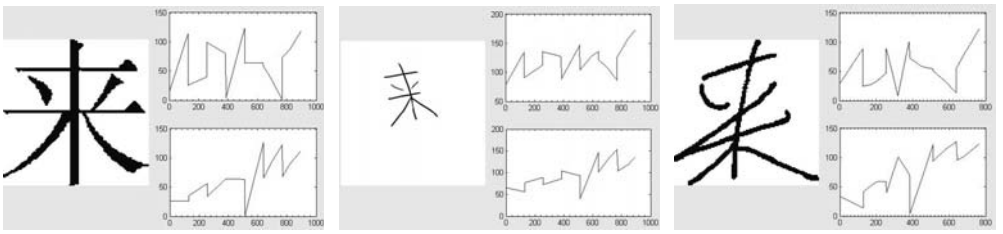

Fig. 7. $X$-graph (above) and $Y$-graph (below) of Chinese character 'I' or 'me'.

**3.4. Haar wavelet transform.** Wavelet transformation (Shioyama *et al.*, 1998; Huang and Huang, 2001) was used to reduce the dimension of the feature vector. Although there are many different wavelet families, such as Haar, Daubechies, Coiflet, Symmlet and Mallat, only the Haar wavelet transformation will be considered due to its simplicity. The Haar wavelet reduces the size of the



(a) (b)

Fig. 8. *X*-graphs (above) and *Y*-graphs (below) plotted for two similar characters: character 'White' (a) and character 'Hundred' (b).



(a) (b) (c)

Fig. 9. *X*-graph (above) and *Y*-graph (below) plotted for the character 'Come': regular character in the database (a), written by Writer A (b) and written by Writer B (c).

feature vector by creating two new sequences of points $\mathbf{a}_j = [a_{xj}, a_{yj}]$ and $\mathbf{d}_j = [d_{xj}, d_{yj}]$, $1 \le j \le D$, $2^5 \le D < 2^6$, which are known as approximation and detailed coefficients, respectively. Only the approximation vector $\mathbf{a}_j$ will be used. In particular,

$$a_{xj} = \frac{x^*_{2j-1} + x^*_{2j}}{\sqrt{2}} \tag{3}$$

represents the *X*-graph and

$$a_{yj} = \frac{y^*_{2j-1} + y^*_{2j}}{\sqrt{2}} \tag{4}$$

represents the *Y*-graph. The new extracted feature $\mathbf{a}_j$ is then used for classification.

Let $\mathbf{b}_j = [b_{xj}, b_{yj}]$ represent the approximation vector of the corresponding character in the database obtained in the same manner as the vector $\mathbf{a}_j$ . The notations in Eqns. (3) and (4) follow Ritter and Wilson (2001).

## 4. Classification

Compared with numerals and alphabets, the number of Chinese character sets is extremely large. Hence, in order to speed up the recognition system, the classification process is separated into two stages: rough classification and fine classification.

**4.1. Rough classification.** The number of strokes in each character, $w$ for all characters in CL2009, is determined as follows:

$$w = \frac{N}{128}, \tag{5}$$

where $N$ is a multiple of 128 (see Section 3). Characters with the same number of strokes are gathered in one group. Rough classification takes into account the number of strokes only. Defining a character by the number of strokes may create fewer classes of characters that need to be considered and possibly help to speed up the recognition system.

**4.2. Fine classification.** After the rough classification stage, the process of fine classification is applied sequentially to each class obtained. Within each class, a given feature vector $\mathbf{a}_j$ (from a given writer) is compared with all the $\mathbf{b}_j$ vectors (from CL2009) within the given class according to some predefined distance measures described in Sections 4.2.1–4.2.4. Generally, a small distance value indicates that the character represented by $\mathbf{a}_j$ belongs to the same class as $\mathbf{b}_j$.

**4.2.1. City block distance with deviation (CBDD).** Define the D-dimensional handwritten character as $\mathbf{a}$ and the $M$ character classes in the database as $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_M$, where $\mathbf{b}_i = (b_{i1}, b_{i2}, \ldots, b_{iD})^T$ for $i = 1, 2, \ldots, M$. The

city block distance with deviation (CBDD) measure (Kato *et al*., 1999) is defined as

$$d_i^{CBDD}(\mathbf{a}) = \sum_{j=1}^{D} \max\{0, |a_j - b_{ij}| - \theta \cdot s_{ij}\}, \quad (6)$$

for a class $\mathbf{b}_i$, where $s_j$ denotes the standard deviation of the *j*-th element, and $\theta$ is a constant. Kato *et al*., (1999) claim that variations of handwritten characters be taken into account in the city block distance measure.

**4.2.2. Minimum distance (MD).** One way to determine the class membership of an unknown input character $\mathbf{a}$ is to assign it to the character class of its closest prototype. To determine the closeness, the Euclidean distance is used:

$$D_i(\mathbf{a}) = \|\mathbf{a} - \mathbf{b}_i\|, \quad (7)$$

for class $\mathbf{b}_i$, $i = 1, 2, \ldots, M$, where $\|\mathbf{h}\| = (\mathbf{h}^T\mathbf{h})^{1/2}$ is the Euclidean norm. Without loss of generality, it is equivalent to evaluating the functions

$$d_i^{MD}(\mathbf{a}) = \mathbf{a}^T\mathbf{b}_i - \frac{1}{2}\mathbf{b}_i^T\mathbf{b}_i. \quad (8)$$

Equation (8) is, therefore, the discriminant function of the minimum distance (MD), as mentioned by Gonzalez and Woods (1993). The MD classifier is also known as a 1-nearest neighbor (NN) rule.

**4.2.3. Modified quadratic discriminant function (MQDF).** The modified quadratic discriminant function (MQDF) was originally proposed by Kimura *et al*. (1987). It is a modified version of the ordinary QDF, in which the complexity of the QDF is reduced by replacing the minor eigenvalues of the covariance matrix of each class with a constant. The MQDF is defined as

$$\begin{aligned} d_i^{MQDF}(\mathbf{a}) &= \sum_{j=1}^{K} \frac{1}{\lambda_{ij}}[\phi_{ij}^T(\mathbf{a} - \mathbf{u}_i)]^2 + \sum_{j=K+1}^{D} \frac{1}{\delta_i}[\phi_{ij}^T(\mathbf{a} - \mathbf{u}_i)]^2 \\ &\quad + \sum_{j=1}^{K} \log \lambda_{ij} + (D - K)\log \delta_i \\ &= \frac{1}{\delta_i}(\|\mathbf{a} - \mathbf{u}_i\|^2 - \sum_{j=1}^{K}(1 - \frac{\delta_i}{\lambda_{ij}})[\phi_{ij}^T(\mathbf{a} - \mathbf{u}_i]^2) \\ &\quad + \sum_{j=1}^{K} \log \lambda_{ij} + (D - K)\log \delta_i, \end{aligned} \quad (9)$$

where $\mathbf{u}_i$ is the mean of the *i*-th class for $i = 1, 2, \ldots, M$, $\lambda_{ij}$ denotes the eigenvalues (in descending order) of $\Sigma_i$ (the covariance matrix of the *i*-th class) for $j = 1, \ldots, D$, $\phi_{ij}$ are the ordered eigenvectors, $\delta_i$ is a constant replacing the minor eigenvalues and $K$ denotes the number of dominant eigenvectors.

**4.2.4. Compound Mahalanobis function (CMF).** The compound Mahalanobis function (CMF) (Suzuki *et al*., 1997) is a discriminant function which improves the discriminant performance of the ordinary Mahalanobis distance (MD), by projecting the difference class-mean feature vectors of two similar classes onto a certain subspace such that the two similar classes can be clearly differentiated.

Let $\mathbf{b}_1$ and $\mathbf{b}_2$ be two similar classes. We denote by $\lambda_1, \lambda_2, \ldots, \lambda_D$ the eigenvalues of the covariance matrix of classes $\mathbf{b}_1$, where $\lambda_j \geq \lambda_{j+1}, j = 1, 2, \ldots, D-1$ and by $\phi_j$ the eigenvectors which correspond to $\lambda_j$. The difference between classes $\mathbf{b}_1$ and $\mathbf{b}_2$ is considered to be clearly demonstrated in a subspace constructed by the eigenvectors $\phi_{K+1}, \phi_{K+2}, \ldots, \phi_D$ and, hence, the CMF for class $\mathbf{b}_1$ is defined as

$$\begin{aligned} &CMF_1(\mathbf{a}, \mathbf{u}) \\ &\quad = \sum_{j=1}^{p} \frac{\{\phi_j^T(\mathbf{a} - \mathbf{u})\}^2}{\lambda_j + Q} \\ &\qquad + \frac{1}{Q}\{\|\mathbf{a} - \mathbf{u}\|^2 - \sum_{j=1}^{p}\{\phi_j^T(\mathbf{a} - \mathbf{u})\}^2\} \\ &\qquad + \mu\Big\{ \sum_{j=K+1}^{p} \frac{\{\phi_j^T(\delta_1)\}^2}{\lambda_j + Q} \\ &\qquad + \frac{1}{Q}\{\|\delta_1\|^2 - \sum_{j=1}^{p}\{\phi_j^T(\delta)\}^2\}\Big\}, \end{aligned} \quad (10)$$

where $\mathbf{u}$ is the class-mean vector of class $\mathbf{b}_1$, $Q$ is a bias, $\mu$ is the weighting parameter and $\delta_1$ is a projective vectors for $\mathbf{b}_1$ as shown in the following:

$$\begin{aligned} \delta_1 &= \{\psi^T(\mathbf{a} - \mathbf{u})\}\psi, \\ \psi &= \frac{(\mathbf{u} - \mathbf{v}) - \sum_{j=1}^{K}\{\phi_j^T(\mathbf{u} - \mathbf{v})\}\phi_j}{\sqrt{\|\mathbf{u} - \mathbf{v}\|^2 - \sum_{j=1}^{K}\{\phi_j^T(\mathbf{u} - \mathbf{v})\}^2}}, \end{aligned} \quad (11)$$

where $\mathbf{v}$ is the mean vector of $\mathbf{b}_2$ and $\psi$ is a unit vector obtained by projecting the difference class-mean vectors $(\mathbf{u} - \mathbf{v})$ onto a subspace constructed by $\phi_{K+1}, \phi_{K+2}, \ldots, \phi_D$ and normalizing the length to 1. In the same way, we can calculate $CMF_2(\mathbf{a}, \mathbf{v})$ for class $\mathbf{b}_2$. Finally, the two classes are differentiated by comparing $CMF_1(\mathbf{a}, \mathbf{u})$ and $CMF_2(\mathbf{a}, \mathbf{v})$.

**4.3. Proposed $R_p^2$ for the two-dimensional ULFR model.** The two-dimensional functional relationship model is derived from the multidimensional unreplicated linear functional relationship (MULFR) model developed

by Chang *et al*. (2009). Consider the extracted feature for the input character trajectory, **a**, where $\mathbf{a}_j = [a_{xj}, a_{yj}]$, $1 \le j \le D$, and the extracted feature for the trajectory of a character in the database, **b** , where $\mathbf{b}_j = [b_{xj}, b_{yj}]$, $1 \le j \le D$. Suppose that **a** and **b** are observed with errors $\boldsymbol{\delta}_\mathbf{j} = [\delta_{xj}, \delta_{yj}]$ and $\boldsymbol{\varepsilon}_\mathbf{j} = [\varepsilon_{xj}, \varepsilon_{yj}]$, respectively, as

$$\mathbf{a}_j = \mathbf{A}_j + \boldsymbol{\delta}_\mathbf{j}, \tag{12}$$

$$\mathbf{b}_j = \mathbf{B}_j + \boldsymbol{\varepsilon}_\mathbf{j}, \tag{13}$$

where $\mathbf{A}_j = [A_{xj}, A_{yj}]$ and $\mathbf{B}_j = [B_{xj}, B_{yj}]$ are two linearly related unobservable true values of $\mathbf{a}_j$ and $\mathbf{b}_j$ such that

$$\mathbf{B}_j = \boldsymbol{\alpha} + \beta \mathbf{A}_j, \tag{14}$$

where $\boldsymbol{\alpha} = [\alpha_1, \alpha_2]$ is the intercept vector and $\beta$ is the slope of the functional model.

Assume that $\boldsymbol{\varepsilon}_\mathbf{j}$ and $\boldsymbol{\delta}_\mathbf{j}$ are mutually independent and normally distributed random variables with mean $\mathbf{0}$ but different variances $\boldsymbol{\Omega}_{11} = \tau^2\mathbf{I}$ and $\boldsymbol{\Omega}_{22} = \sigma^2\mathbf{I}$, i.e., $\boldsymbol{\varepsilon}_\mathbf{j} \sim N(\mathbf{0}, \boldsymbol{\Omega}_{11})$ and $\boldsymbol{\delta}_\mathbf{j} \sim N(\mathbf{0}, \boldsymbol{\Omega}_{22})$, where $\mathbf{I}$ is the identity matrix. Equations (12)–(14) are called the multidimensional unreplicated linear functional relationship (MULFR) model. Note that there is only one relationship between the two variables $\mathbf{A}_j$ and $\mathbf{B}_j$ . It can be shown that if the ratio of the error variances is equal to one, i.e., $\sigma/\tau = \lambda = 1$ , then the maximum likelihood estimators for $\boldsymbol{\alpha}$ , $\beta$ , $\sigma^2$ and $\mathbf{A}_j$ are given below:

$$\hat{\boldsymbol{\alpha}} = \bar{\mathbf{b}} - \hat{\beta}\bar{\mathbf{a}}, \tag{15}$$

$$\hat{\beta} = \frac{(S_{bb} - \lambda S_{aa}) + \sqrt{(S_{bb} - \lambda S_{aa})^2 + 4\lambda S_{ab}^2}}{2S_{ab}}, \tag{16}$$

$$\hat{\sigma}^2 = \frac{1}{p(n-2)}\Big[\sum_{j=1}^{D}(\mathbf{a}_j - \hat{\mathbf{A}}_j)^T(\mathbf{a}_j - \hat{\mathbf{A}}_j) + \frac{1}{\lambda}\sum_{j=1}^{D}(\mathbf{b}_j - \hat{\boldsymbol{\alpha}} - \hat{\beta}\hat{\mathbf{A}}_j)^T(\mathbf{b}_j - \hat{\boldsymbol{\alpha}} - \hat{\beta}\hat{\mathbf{A}}_j)\Big], \tag{17}$$

$$\hat{\mathbf{A}}_j = \frac{1}{(\lambda + \hat{\beta}^2)}[\lambda\mathbf{a}_j + \hat{\beta}(\mathbf{b}_j - \hat{\boldsymbol{\alpha}})], \tag{18}$$

where

$$\bar{\mathbf{a}} = \frac{\sum_{j=1}^{D}\mathbf{a}_j}{n}, \qquad \bar{\mathbf{b}} = \frac{\sum_{j=1}^{D}\mathbf{b}_j}{n},$$

$$S_{aa} = \sum_{j=1}^{D}(\mathbf{a}_j - \bar{\mathbf{a}})^2, \qquad S_{bb} = \sum_{j=1}^{D}(\mathbf{b}_j - \bar{\mathbf{b}})^2$$

and

$$S_{ab} = \sum_{j=1}^{D}(\mathbf{a}_j - \bar{\mathbf{b}})^2.$$

According to Chang *et al*. (2009), the coefficient of determination for the MULFR can be defined as

$$R_p^2 = \frac{SS_R}{S_{bb}} = \frac{\hat{\beta}S_{ab}}{S_{bb}}. \tag{19}$$

The residual sum of squares ($SS_E$) and the regression sum of squares ($SS_R$) are given as follows:

$$SS_E = \frac{S_{bb} - 2\hat{\beta}S_{ab} + \hat{\beta}^2 S_{aa}}{(1 + \hat{\beta}^2)}, \tag{20}$$

$$SS_R = S_{bb} - SS_E = \frac{\hat{\beta}^2(S_{bb} - S_{aa}) + 2\hat{\beta}S_{ab}}{1 + \hat{\beta}^2}. \tag{21}$$

The quantity $R_p^2$ measures the proportion of variation on a character in the database explained by the input character. We would like the similarity measure to satisfy the following conditions.

*Property 1*: Boundedness

To date, there is no standard range for a similarity or dissimilarity measure between characters. For example, the peak signal to noise ratio (PSNR) (Battaglia, 1996) is defined on $[0, \infty)$. Wang *et al*. (2004) defined a good similarity measure in the range $(-\infty, 1]$, but Van de Weken *et al*. (2002) defined it on $[0, 1]$. Among others, $[0, 1]$ is the most commonly and well accepted dynamic range for a good quality measure. The proposed measure $R_p^2$ is also defined on $[0, 1]$, and this can be easily shown from the regression sum of squares

$$0 \le SS_R = S_{bb} - SS_E \le S_{bb},$$

$$0 \le \frac{SS_R}{S_{bb}} \le \frac{S_{bb}}{S_{bb}} = 1,$$

$$\therefore 0 \le R_p^2 \le 1. \tag{22}$$

*Property 2*: Reflexivity

Note from Eqns. (16) and (19) that when a character in the database and an input character are identical or $\mathbf{a} = \mathbf{b}$, this implies $R_p^2 = 1$ and vice versa. On the other hand, $R_p^2 = 0$ indicates that the two characters are dissimilar. $R_p^2$ decreases when the degree of similarity between the two characters decreases.

*Property 3*: Symmetry

Given $S_{bb} = kS_{aa}$ where $k > 0$. Let $R_p^2$ and $\breve{R}_p^2$ be the coefficients of determination for the MULFR model with $\lambda = 1$ as defined by $\mathbf{B}_j = \boldsymbol{\alpha} + \beta\mathbf{A}_j$ and $\mathbf{A}_j = \boldsymbol{\alpha}^* + \beta^*\mathbf{B}_j$, respectively. Then we have

$$R_p^2 = \frac{1}{k}(\breve{R}_p^2 - 1) + 1. \tag{23}$$

This indicates that $R_p^2$ is non-symmetric when $k \ne 1$. In order to solve this problem, we regard the character in

the database or the input character, whichever has larger variance, as **A**, and the other character as **B**. With this arrangement, the full range $0 \leq R_p^2 \leq 1$ will be granted. Otherwise, Eqn. (23) provides a simple way to convert $\breve{R}_p^2$ to $R_p^2$.

## 5. Results and discussion

Some parameters used in distance measures need to be defined before performing experiments. Selected parameter values are as follows:

(i) CBDD: $\theta = 1.2$, as stated by Kato *et al.* (1999).

(ii) MQDF: $K = 1$ since there is only one sample for each character in the database and hence it results in only one dominant eigenvalue.

$\delta_i = 1.3641$, in which it is made class independent and equals the average of all eigenvalues of all classes. Note that, as stated by Long and Jin (2008), the performance of the classifier is superior when setting, the constant class independent rather than class dependent.

(iii) CMF: $p, K = 1$ (refer to (ii).
$Q = 1.3641$, which is the average of all eigenvalues of all classes.
$\mu = 2.8$, as stated by Suzuki *et al.* (1997).

**5.1. Recognition rate (accuracy) and precision.** The recognition rate or accuracy is defined as follows:

$$\text{RecognitionRate} = \frac{\text{Number of Test Samples with Correct Matching}}{\text{Total Number of Test Samples}} \times 100\%. \quad (24)$$

Every character recognized correctly is given the unit weight.

Tables 2(a) and (b) show that $R_p^2$ performs consistently for both normalized and non-normalized characters as well as for both writers. Recognition rates using $R_p^2$ tend to be higher than those using the other four distance measures.

When similar recognition rates are achieved for different levels of character complexity (variation in the number of strokes), the recognition system is said to be precise. Three categories of character complexity are defined, namely

(i) character of low complexity with less than 6 strokes,

(ii) character of medium complexity with 6 to 12 strokes,

(iii) very complex character with more than 12 strokes.

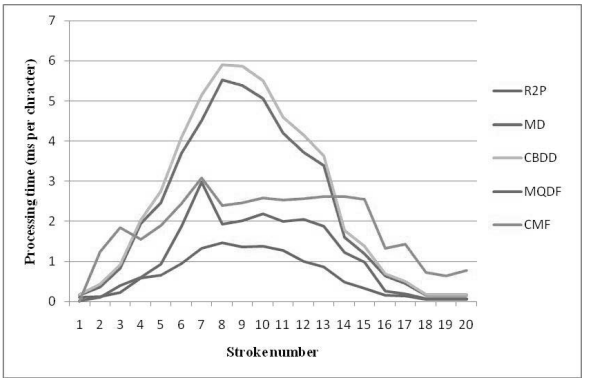



Fig. 10. Processing time for recognizing characters with varying stroke numbers using the MD, CBDD, MQDF and CMF classifiers.

Writer A with greater experience shows higher precision rates regardless of normalization or otherwise when using $R_p^2$ (Table 3(a)). Writer B is less precise than Writer A with respect to $R_p^2$, although low precision is seen when CBDD, MD, MQDF and CMF are used (Tables 3(a) and (b)).

Characters in the high complexity category are most likely to be recognized accurately. This is probably because the larger number of strokes makes recognition easier.

**5.2. Processing time.** The speed of the recognition system is an important factor for an efficient recognition system, in which the reduction in the processing time is crucial. The processing times are listed in Table 4 for feature extraction, rough classification and fine classification.

**5.2.1. Comparison between different feature extraction methods.** Compared with other feature extraction methods such as the attributed relational graph (ARG), whole character-based hidden Markov model (HMM) and directional feature densities (DFDs), $X$-graph and $Y$-graph transformation may be regarded as the simplest method in terms of computation. This is illustrated in Table 5.

**5.2.2. Comparison between distance measures.** Figure 10 shows the processing time for the recognition of characters with varying numbers of strokes by using $R_p^2$, the MD, CBDD, MQDF and CMF. Characters withs number of strokes between 6 and 12 consume longer processing time. The frequencies of these characters are highest, suggesting that the processing time taken for fine classification will be greater. It is also seen in Fig. 10 that the recognition system with the proposed $R_p^2$ results in the least processing time, followed by the MQDF, CMF, MD and CBDD (see Table 6).

Table 2. Experimental results for different writers: with (a) and without normalization (b). Each writer writes all 3000 different Chinese characters.

(a)

| Distance measures | | CBDD | MD | MQDF | CMF | $R_p^2$ |
|---|---|---|---|---|---|---|
| Recognition rate (%) with normalization | Writer A | 96.6 | 98.2 | 98.2 | 97.6 | 98.0 |
| | Writer B | 93.1 | 94.1 | 94.1 | 94.1 | 96.1 |

(b)

| Distance measures | | CBDD | MD | MQDF | CMF | $R_p^2$ |
|---|---|---|---|---|---|---|
| Recognition rate (%) without normalization | Writer A | 70.0 | 79.6 | 81.6 | 75.0 | 98.2 |
| | Writer B | 55.9 | 66.7 | 66.7 | 54.9 | 97.4 |

Table 3. Results for different number of strokes: Writer A (a) and Writer B (b). Number of strokes is grouped into three categories: less than six strokes, between six to 12 strokes, and more than 12 strokes.

(a)

| Distance measures | | CBDD | MD | MQDF | CMF | $R_p^2$ |
|---|---|---|---|---|---|---|
| Recognition rate (%) with normalization | Low complexity | 95.3 | 96.6 | 97.3 | 97.3 | 94.6 |
| | Medium complexity | 97.0 | 98.8 | 98.5 | 97.6 | 99.4 |
| | High complexity | 100 | 100 | 100 | 100 | 100 |
| Recognition rate (%) without normalization | Low complexity | 58.1 | 67.6 | 73.6 | 64.9 | 96.6 |
| | Medium complexity | 73.8 | 83.7 | 84.0 | 78.0 | 98.8 |
| | High complexity | 95.0 | 100 | 100 | 100 | 100 |

(b)

| Distance measures | | CBDD | MD | MQDF | CMF | $R_p^2$ |
|---|---|---|---|---|---|---|
| Recognition rate (%) with normalization | Low complexity | 86.1 | 86.1 | 86.1 | 86.1 | 88.9 |
| | Medium complexity | 96.9 | 98.5 | 98.5 | 98.5 | 100 |
| | High complexity | 100 | 100 | 100 | 100 | 100 |
| Recognition rate (%) without normalization | Low complexity | 58.3 | 69.4 | 63.9 | 52.8 | 94.4 |
| | Medium complexity | 53.8 | 64.6 | 67.7 | 55.4 | 98.5 |
| | High complexity | 100 | 100 | 100 | 100 | 100 |

Table 4. Result of timing analysis.

| | Average processing time (millisecond (ms) per character) |
|---|---|
| Feature extraction | 0.913 |
| Rough classification | 1.176 |
| Fine classification | 0.957 |

**5.3. Feature size.** A comparison of the size of features with other feature extraction schemes such as (i) the ARG (ii) the HMM and (iii) the DFDs is illustrated in Table 7, where the feature derived from the $X$-graph and $Y$-graph has the lowest dimension.

## 6. Conclusion

A novel online HCCR system was proposed in this paper and was found to be efficient and simple to apply. This was made possible by the uniqueness and invariance of the $X$-graph and $Y$-graph which involved only simple computations. High accuracy and precision rates were achieved, strongly suggesting that the system is robust to non-normalization of characters as well as for characters in a wide range of complexity (number of strokes). The recognition system saved between 40.36% to 75.31% of processing time and this may be regarded as an added advantage of using the system.

Central to the development of the proposed recognition system is the coefficient $R_p^2$ used, whose properties of boundedness, reflexivity and symmetry motivate its application as a similarity measure. The experiments showed $R_p^2$ being robust to size, slant and position variation. In general, $R_p^2$ is to be preferred over the use of the CBDD, MQDF, MD and CMF when accuracy and precision are considered.

Table 5. Processing steps for four different feature extraction methods.

| Methods | Processing steps |
|---|---|
| Attributed relational graph (ARG) (Liu *et al*., 1996) | Step 1: Perform strokes identification.<br>Step 2: Fit the strokes with straight lines.<br>Step 3: Determine geometric centres for each stroke.<br>Step 4: Construct a complete ARG with nodes and arcs.<br>Step 5: Convert the ARG to a generalized relation matrix. |
| Whole character-based hidden Markov model (HMM) (Takahashi *et al*., 1997) | Step 1: Estimate parameters, such as state transition probabilities, output emission probabilities and initial state probability through learning process, in which it is time-consuming.<br>Step 2: Determine HMM of the character. |
| Directional feature densities (DFDs) (Kawamura *et al*., 1992) | Step 1: Define vectors for each consecutive point on the strokes.<br>Step 2: Compute directional feature vectors.<br>Step 3: Define vector for square areas.<br>Step 4: Perform dimension condensation. |
| *X*-*Y* graphs decomposition with the haar wavelet | Step 1: Define feature vectors from *X*-*Y* graphs.<br>Step 2: Implement the Haar wavelet transform. |

Table 6. Average reduced time rates in comparing the algorithm of $R_p^2$ with CBDD, MD, MQDF and CMF classifiers.

| Distance measures | | CBDD | MD | MQDF | CMF |
|---|---|---|---|---|---|
| Reduced time rate (%) using $R_p^2$ | with preprocessing | 74.57 | 72.28 | 40.36 | 58.09 |
| | without preprocessing | 75.31 | 73.05 | 40.69 | 58.27 |

Table 7. Feature sizes for four different feature extraction methods.

| Methods | Feature size (dimension) |
|---|---|
| Attributed Relational graph (ARG) (Liu *et al*., 1996) | (stoke-number)$^2$, increase in the number of strokes will increase the feature size massively. For characters with the number of strokes between six and 12, which is typical in Chinese characters, dimension $= [6^2, 12^2] = [36, 144]$ . |
| Whole character-based hidden Markov model (HMM) (Takahashi *et al*., 1997) | Sum of the sizes of parameters $\{a_{ij}\}$, $\{b_{ik}^1\}$, $\{b_{il}^2\}$ and $\{\pi_i\}$, which amounts to $2(N-1)+N(L+2)+N$, where *N* is the number of states and *L* is the number of quantized directions. In the work of Takahashi *et al*. (1997), it results in the feature size of 187 with $N = 9$ and $L = 16$. |
| Directional feature densities (DFDs) (Kawamura *et al*., 1992) | $8 \times 8 \times 4 = 256$. |
| *X*-*Y* graphs decomposition with the Haar wavelet | Between $2^5 \times 2 = 64$ (inclusive) and $2^6 \times 2 = 128$ (exclusive). |

More work needs to be done to address two limitations of the proposed recognition system. Firstly, it is dependent on the number of strokes and order. In particular, the problem encountered with characters of the number of strokes between 6 and 12 suggests investigating different recognition approaches. Secondly, only two writers using the CL2009 database were considered and there is a need to compare it with other databases.

## Acknowledgment

We wish to thank H.V. Chen and W.S. Ng at UTAR for their useful comments. This work was supported by an E-Science Grant (No. 01-02-11-SF0053) of the Ministry of Science, Technology and Innovative of Malaysia.

## References

Battaglia, G.J. (1996). Mean square error, *AMP Journal of Technology* **5**(1): 31–36.

Casey, R.G. (1970). Moment normalization of handprinted character, *IBM Journal of Research and Development* **14**(5): 548–557.

Chang, Y.F., Rijal, O.M. and Abu Bakar, S.A.R. (2010). Multidimensional unreplicated linear functional relationship model with single slope and its coefficient of determination, *WSEAS Transactions on Mathematics* **9**(5): 295–C313.

Dan, J. (2004). Modern Chinese Character Frequency List, `http://lingua.mtsu.edu/chinesecomputing/statistics/char/list.php?Which=MO`.

Deepu, V., Sriganesh, M. and Ramakrishnan, A.G. (2004). Principal component analysis for online handwritten character recognition, *Proceedings of the 17th International Conference on Pattern Recognition, Cambridge, UK*, Vol. 2, pp. 327–330.

Dong, J.X., Krzyżak, A. and Suen, C.Y. (2005). An improved handwritten Chinese character recognition system using support vector machine, *Pattern Recognition Letters* **26**(12): 1849–1856.

Fujarewicz, K. and Wiench, M. (2003). Selecting differentially expressed genes for colon tumor classification, *International Journal of Applied Mathematics and Computer Science* **13**(3): 327–335.

Gao, T.F. and Liu, C.L. (2008). High accuracy handwritten Chinese character recognition using LDA-based compound distances, *Pattern Recognition* **41**(11): 3442–3451.

Gao, X., Jin, L.W., Yin, J.X. and Huang, J.C. (2002). SVM-based handwritten Chinese character recognition, *Chinese Journal of Electronics* **30**(5): 651–654.

Gonzalez, R.C. and Woods, R.E. (1993). *Digital Image Processing*, Addison-Wesley Publishing Co., New York, NJ, pp. 580–583.

Horiuchi, T., Haruki, R., Yamada, H. and Yamamoto, K. (1997). Two-dimensional extension of nonlinear normalization method using line density for character recognition, *Proceedings of the 4th International Conference on Document Analysis and Recognition, Ulm, Germany*, pp. 511–514.

Huang, L. and Huang, X. (2001). Multiresolution recognition of offline handwritten Chinese characters with wavelet transform, *Proceedings of the 6th International Conference on Document Analysis and Recognition, Washington, DC, USA*, pp. 631–634.

Kato, N., Suzuki, M., Omachi, S.I., Aso, H. and Nemoto, Y. (1999). A handwritten character recognition system using directional element feature and asymmetric Mahalanobis distance, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**(3): 258–262.

Kawamura, A., Yura, K., Hayama, T., Hidai, Y., Minamikawa, T., Tanaka, A. and Masuda, S. (1992). On-line recognition of freely handwritten Japanese characters using directional feature densities, *Proceedings of the 11th International Conference on Pattern Recognition, The Hague, the Netherlands*, Vol. 2, pp. 183–186.

Kimura, F., Takashina, K., Tsuruoka, S. and Miyake, Y. (1987). Modified quadratic discriminant functions and its application to Chinese character recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **9**(1): 149-153.

Kimura, F., Wakabayashi, T., Tsuruoka, S. and Mayake, Y. (1997). Improvement of handwritten Japanese character recognition using weighted direction code histogram, *Pattern Recognition* **30**(8): 1329–1337.

Liu, C.L. and Marukawa, K. (2004). Global shape normalization for handwritten Chinese character recognition: A new method, *Proceedings of the 9th International Workshop on Frontiers of Handwriting Recognition, Tokyo, Japan,* pp. 300–305.

Liu, C.L. and Marukawa, K. (2005). Pseudo two-dimensional shape normalization methods for handwritten Chinese character recognition, *Pattern Recognition* **38**(12): 2242–2255.

Liu, C.L., Jaeger, S. and Nakagawa, M. (2004). Online recognition of Chinese characters: The-state-of-the art, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**(2): 198–213.

Liu, C.L., Mine, R. and Koga, M. (2005). Building compact classifier for large character set recognition using discriminative feature extraction, *Proceedings of the 8th ICDAR, Seoul, Korea,* pp. 846–850.

Liu, C.L., Sako, H. and Fujisawa, H. (2003). Handwritten Chinese character recognition: Alternatives to nonlinear normalization, *Proceedings of the 7th International Conference on Document Analysis and Recognition, Edinburgh, UK*, pp. 524–528.

Liu, H. and Ding, X. (2005). Handwritten character recognition using gradient feature and quadratic classifiers with multiple discrimination schemes, *Proceedings of the 8th ICDAR, Seoul, Korea*, pp. 19–23.

Liu, J.Z., Cham, W.K. and Chang, M.M.Y. (1996). Online Chinese character recognition using attributed relational graph matching, *IEE Proceedings: Vision, Image, Signal Processing* **143**(2): 125–131.

Long, T. and Jin, L.W. (2008). Building compact MQDF classifier for large character set recognition by subspace distribution sharing, *Pattern Recognition* **41**(9): 2916–2925.

Michalak, K. and Kwaśnicka, H. (2006). Correlation-based feature selection strategy in classification problems, *International Journal of Applied Mathematics and Computer Science* **16**(4): 503–511.

Miquelez, T., Bengoetxea, E. and Larranaga, P. (2004). Evolutionary computation based on Bayesian classifiers, *International Journal of Applied Mathematics and Computer Science* **14**(3): 335–349.

Ritter, G.X. and Wilson, J.N. (2001). *Handbook of Computer Vision Algorithms in Image Algebra*, CRC Press LLC, Boca Raton, FL, pp. 225–228.

Romero, R., Berger, R., Thibadeau, R. and Touretsky, D. (1995). Neural network classifiers for optical Chinese character recognition, *Proceedings of the 4th Annual Symposium on Document Analysis and Information Retrieval, Las Vegas, NV, USA*, pp. 385–398.

Saeed, K. (2000). A projection approach for Arabic handwritten characters recognition, *in* P. Sincak and J. Vascak (Eds.), *Quo Vadis Computational Intelligence? New Trends and Approaches in Computational Intelligence*, Physica-Verlag, Berlin, pp. 106–111.

Shimodaira, H., Sudo, T., Nakai, M. and Sagayama, S. (2003). On-line overlaid-handwriting recognition based on substroke HMMs, *Proceedings of the 7th International Conference on Document Analysis and Recognition, Edinburgh, UK*, Vol. 2, p. 1043.

Shioyama, T., Wu, H.Y. and Nojima, T. (1998). Recognition algorithm based on wavelet transform for handprinted Chinese characters, *Proceedings of the 14th International Conference on Pattern Recognition, Hong Kong, China*, Vol. 1, pp. 229–232.

Suzuki, M., Ohmachi, S., Kato, N., Aso, H. and Nemoto, Y. (1997). A discrimination method of similar characters using compound Mahalanobis function, *IEICE Transactions on Information and Systems* **J80-D**(10): 2752–2760.

Świniarski, R.W. (2001). Rough sets methods in feature reduction and classification, *International Journal of Applied Mathematics and Computer Science* **11**(3): 565–582.

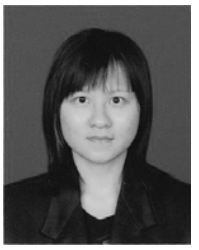
Takahashi, K., Yasuda, H. and Matsumoto, T. (1997). A fast HMM algorithm for on-line handwritten character recognition, *Proceedings of the 4th International Conference on Document Analysis and Recognition, Ulm, Germany*, pp. 369–375.

Van der Weken, D., Nachtegael, M. and Kerre, E.E. (2002). Image quality evaluation, *Proceedings of the 6th International Conference on Signal Processing, Beijing, China*, Vol. 1, pp. 711–714.

Wang, Z., Bovik, A.C., Sheikh, H.R. and Simoncelli, E.P. (2004). Image quality assessment: From error visibility to structural similarity, *IEEE Transactions on Image Processing* **13**(4): 600–612.

Zheng, J., Ding, X. and Wu, Y. (1997). Recognizing online handwritten Chinese character via FARG matching, *Proceedings of the 4th International Conference on Document Analysis and Recognition, Ulm, Germany*, Vol. 2, pp. 621–624.

**Yun Fah Chang** received the B.A. degree in mathematics and the M.Sc. degree in applied statistics from the University of Malaya, Malaysia, in 1998 and 2002, respectively. He is currently pursuing the Ph.D. degree in the area of statistical image analysis at the same university. His research interests are in image and video quality and compression, handwritten character recognition, multivariate analysis and functional linear models. From 2001 until 2007, he was a lecturer at the Faculty of Engineering of Multimedia University. In 2007, he joined the Department of Mathematical and Actuarial Sciences, Universiti Tunku Abdul Rahman, Malaysia.



**Jia Chii Lee** received the B. Sc. degree in applied mathematics with computing from Universiti Tunku Abdul Rahman (UTAR), Malaysia, in 2008. Currently, she is pursuing her M.Sc. degree in handwritten Chinese character recognition. Her research interests include pattern recognition, online and offline character recognition, speech recognition and signal processing.

**Omar Mohd Rijal** is an associate professor of applied statistics at the University of Malaya, Malaysia. He received his B.Sc. degree in mathematics (operational research) from the New University of Ulster, Ireland (1979) and the Ph.D. (applied statistics) from the University of Glasgow (1984). His research interest is in applied statistics, image and data analysis for medical, industrial and remote sensing applications.

**Syed Abdul Rahman Syed Abu Bakar** received the B.Sc. degree from Clarkson University in Potsdam (USA) in 1990, the M.S.E.E. degree from Georgia Tech (USA) in 1991, and the Ph.D. degree from the University of Bradford (UK) in 1997. In 1992, he joined the Faculty of Electrical Engineering of Universiti Teknologi Malaysia as a lecturer and currently he is an associate professor at the same faculty as well as the head for the Computer Vision, Video and Image Processing Research Lab. His current research interests include image processing with application in medical imaging, biometrics, agricultural and industrial applications, and computer vision, especially in the area of security and surveillance. He has published more than 90 scientific papers both at the national and the international level. He is also a senior member of the IEEE.

Received: 28 November 2009
Revised: 23 June 2010