Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition 41 (2008) 1295–1307
PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/pr

Writer identification of Chinese handwriting documents using hidden Markov tree model

Zhenyu He[a], Xinge You[b], Yuan Yan Tang[a,b,c,*]

[a]Department of Computer Science, Hong Kong Baptist University, Hong Kong
[b]Department of Electronics and Information Engineering, Huazhong University of Science and Technology, China
[c]Department of Computer Science, Chongqing University, PR China

Received 24 April 2006; received in revised form 23 August 2007; accepted 28 August 2007

Abstract

Handwriting-based writer identification, a branch of biometrics, is an active research topic in pattern recognition. Since most existing methods and models aim to on-line and/or text-dependent writer identification, it is necessary to propose new methods for off-line, text-independent writer identification. At present, two-dimensional Gabor model is widely acknowledged as an effective and classic method for off-line, text-independent handwriting identification, while it still suffers from some inherent shortcomings, such as the excessive calculational cost. In this paper, we present a novel method based on hidden Markov tree (HMT) model in wavelet domain for off-line, text-independent writer identification of Chinese handwriting documents. Our experiments show this HMT method, compared with two-dimensional Gabor model, not only achieves better identification results but also greatly reduces the elapsed time on computation.

Keywords: Writer identification; Chinese document; Wavelet; Hidden Markov tree model; Two-dimensional Gabor model

1. Introduction

Since the beginning of civilization, identifying the statuses of uncertain persons has been crucial to the human society. Consequently, personal identification is widely used in diverse commercial and governmental sections such as financial access, health care, security control, border control and communication. In particular, personal identification is in highly increasing demand after the $\frac{9}{11}$ terroristic attack. Traditionally, the ways for personal identification are identification cards (ID cards) and passwords, while they cannot provide us an unique, secure and consistent personal identification. For example, passwords and ID cards can be shared by others and therefore are not unique. Furthermore, it is possible that we forget to bring the ID cards along with us or forget the passwords and thus they are not consistent. A nationwide survey in USA showed that heavy web users have an average of twenty-one passwords, and they often make confusion on their varied passwords [1]. So we need better solutions to personal identification.

Writer identification, which, speaking in a simple way, is to determine the writer from his/her handwritings (including signatures, letters, notes, etc.), is such a technique that satisfies four requirements of personal identification: accessible, cheap, reliable and acceptable. Therefore, in spite of the existence and development of other techniques on personal identification based on DNA [2,3], iris [4], fingerprint [5], etc., it appears that the writer identification still remains an attractive application. As a result, writer identification enjoys a huge interest from both industry and academia [6–8].

Writer identification can be classified in several ways, however the most straightforward one is to classify it into on-line and off-line writer identifications [7,9]. The former assumes that a transducer device is connected to the computer, which can convert writing movement into a sequence of signals and then send these signals to the connected computer. The most common form of the transducer is a tablet digitizer, which consists of a plastic or electronic pen and a pressure or

* Corresponding author. Department of Computer Science, Hong Kong Baptist University, Hong Kong. Tel.: +852 3411 7092; fax: +852 3411 7892.
E-mail addresses: zyhe@comp.hkbu.edu.hk (Z. He), xyou@comp.hkbu.edu.hk (X. You), yytang@comp.hkbu.edu.hk (Y.Y. Tang).

0031-3203/$30.00 © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2007.08.017

electrostatic-sensitive writing surface on which the user writes down his/her handwritings with the electronic pen. Since dynamic information of the writing process captured by the transducer device contains many useful writing features, on-line writer identification, compared to the off-line writer identification, is easier to achieve a high identification accuracy. On the other hand, off-line writer identification deals with handwritings scanned into a computer file in two-dimensional image representation. Despite continuous effort, off-line writer identification still remains as a challenging issue [7]. In fact, off-line systems rely on more sophisticated architectures to accomplish the same identification task, and their identification results are still lower than those of on-line identification systems under same testing conditions. Unfortunately, on-line systems are inapplicable in many cases. For example, on-line systems cannot help us to find out the writer of an existing handwriting document. Therefore, developing effective techniques on off-line writer identification is an urgent task.

Further, off-line writer identification can be divided into two types: text-dependent and text-independent writer identification [7,9]. Text-dependent identification matches one or a small group of same characters/words and consequently requires the writer to write the same fixed text in the handwriting documents. For example, signature identification, which is well known to us, is a special case of text-dependent writer identification. Commonly, the geometry or structure features of those given characters/words are extracted as the writing features in text-dependent writer identification. While in many applications, it is impossible to find out the same text from different handwriting documents and therefore text-dependent identification is unavailable. In this case, we need text-independent identification. Text-independent identification does not use the writing features of some specific characters/words, while instead considers handwriting document layout features, text line features, etc. Generally speaking, text-dependent identification, compared to text-independent identification, has a better identification performance. However, as mentioned above, its applicability is lower than text-independent identification because of its strict requirement on same characters/words. In this paper, we focus on the research on off-line, text-independent writer identification, which still is a challenging research topic and comparatively less touched by researchers. In the following, we briefly review the previous researches on off-line, text-independent writer identification.

In Ref. [9], R. Plamondon made a summary of the early researches on the handwriting-based writer identification. He pointed out that two general approaches had been proposed for the off-line, text-independent writer identification: transform techniques and histogram descriptions. In transform techniques, Duvernoy et al. had reported that the most important variations of the writer's transfer function are reflected in the low frequency bands of the Fourier spectrum of handwritten pages [10,11]. Durvernoy et al. had also designed a hybrid optical-digital image processing system to extract features from Fourier spectra of handwritten texts. The dominant eigenvectors of the data covariance matrix of the Fourier description of the difference between spectras were used as features [11]. Similarly Kuckuck et al. had used Fourier transform techniques to preprocess handwritten text as a texture. The feature sets extracted in this study were either composed of a sequence of spectrum mean values per bandwidth, or polynomial fitting coefficients or linear transform of these coefficients [12]. In the second technology, frequency distributions of different global or local properties were used. In most cases, the properties were extracted by measuring global run length, the handwritten line's curvature [12], length and direction of straight pixel chains [13,14], relative number of connected segments, distance between special points [9,15].

In view of that the handwritings of different people usually are visually different, and at the same time inspired by the idea of multi-channel spatial filtering technique, Said et al. proposed a texture analysis approach [7]. In this method, they regarded the handwriting as an image containing some special textures and applied a well-established two-dimensional Gabor model to extract features of such textures. In their paper, Said et al. also compared the two-dimensional Gabor model with the grey-scale co-occurrence matrix, and concluded the two-dimensional Gabor model outperformed grey-scale co-occurrence matrix.

Except the global style of handwriting, some researchers found valuable features from single word or text line. In Ref. [16], Zois et al. morphologically processed horizontal projection profiles of single words. To increase the identification efficiency, the projections were derived and processed in segments. Bayesian classifier or neural network was used for classification. In Ref. [17], Hertel et al. designed a system for writer identification base on the text line features. They segmented a given text into individual text lines and then extracted a set of features from each text line. The text line features were regarded as the writing features. And then the k-nearest neighbor classifier was adopted for classification. In Ref. [18], Schlapbach et al. proposed the hidden Markov model (HMM) based recognizer which also extracted the text line features. They trained an HMM recognizer on text line for each writer in the database. The query handwriting was presented to each of these recognizers, and thus a series of log-likelihood score results were obtained. They ranked log-likelihood score results and regarded the recognizer with the highest log-likelihood score belonged to the writer of query handwriting.

Structure features and geometrical features also came into the researchers' views. In Ref. [19], Bulacu et al. used the edge-based directional probability distributions as features for writer identification. They found out that the joint probability distribution of the angle combination of two "hinged" edge fragments outperformed all individual features. Limitation of their method is that a large number of handwritten materials were needed to obtain the reliable distribution estimates.

In Refs. [20–22], researchers gave a definition of writer invariants. They suggested that each handwriting could be characterized by a set of writer invariant features and these invariant features could be detected by using an automatic grapheme clustering procedure.

In Ref. [23], G. Leedham et al. presented eleven innovative features, which could be extracted from handwritten digits. All these features were binarized to form a binary feature vectors of constant lengths. Then the formed binary feature vectors were measured by Hamming distance for writer discrimination.

Naturally, some papers integrated multiple features to writer identification. In Ref. [24], Cha et al. integrated several distance measures for many feature types: element, histogram, string, convex hull, etc. A satisfied result was achieved on their experiment database.

The rest of the paper is organized as follows. In Section 2, two-dimensional Gabor model, a classic method for off-line, text-independent writer identification, is briefly introduced, which is used as a benchmark to be compared with our method in this paper. In Section 3, we propose our method on off-line, text-independent writer identification. The experiments for writer identification using our method and relative discussions are offered in Section 4. Finally, a short conclusion is made in Section 5.

2. A classic method for writer identification: two-dimensional Gabor model

Gabor function is the name given to a Gaussian weighted sinusoid. The function was named after Dennis Gabor who used this function in the 1940s. Later, Daugman proposed the function to describe the spatial response of cells in visual stimuli experiments [25]. Gabor function is chosen for image processing because of its biological relevance and technical properties: (1) Gabor function is of similar shape as the receptive fields of simple cells in the primary visual cortex. (2) Gabor function is localized in both space and frequency domains and has the shape of plane waves restricted by a Gaussian function. Generally, the Gabor function is a representative of time-frequency analysis and multi-channel filtering technology, and is used in a wide range of image processing applications.

In Ref. [7], Said et al. firstly applied the two-dimensional Gabor model on English off-line, text-independent writer identification. Later, in Ref. [8], Zhu et al. also applied the same two-dimensional Gabor model on Chinese off-line, text-independent writer identification. Both of the two papers said two-dimensional Gabor model achieved good results and outperformed the co-occurrence matrix in their experiments. And nowadays, the academia also widely acknowledges this two-dimensional Gabor model is one of the best methods for off-line, text-independent writer identification.

The computational model of the two-dimensional Gabor filters proposed in Ref. [7,8] is given as follows:

$$h_e(x, y) = g(x, y)\cos[2\pi f(x\cos\theta + y\sin\theta)], \tag{1}$$

$$h_o(x, y) = g(x, y)\sin[2\pi f(x\cos\theta + y\sin\theta)], \tag{2}$$

where h_e and h_o denote the so-called even- and odd- symmetric Gabor filters, and $g(x, y)$ is an isotropic Gaussian function.

The spatial frequency responses of the Gabor filters are

$$H_e(u, v) = \frac{[H_1(u, v) + H_2(u, v)]}{2}, \tag{3}$$

$$H_o(u, v) = \frac{[H_1(u, v) - H_2(u, v)]}{2j}, \tag{4}$$

where $j = \sqrt{-1}$ and

$$H_1(u, v) = \exp\{-2\pi^2\sigma^2[(u - f\cos\theta)^2 + (v - f\sin\theta)^2]\},$$

$$H_2(u, v) = \exp\{-2\pi^2\sigma^2[(u + f\cos\theta)^2 + (v + f\cos\theta)^2]\}.$$

Here, f, θ, σ are the spatial frequency, orientation, and space constant of the Gabor envelope, respectively. $h_e(x, y)$ and $h_o(x, y)$ will combine to provide multi-channel outputs of the input image with different f, θ and σ. An example of multi-channel output of Gabor filters are shown in Fig. 1.

The mean (M) and standard derivation (σ) of the multi-channel outputs are selected as features to represent writer global features for writer identification. Weighted Euclidean Distance (WED) is applied for feature matching after extracting the writing features,

$$WED(k) = \sum_{i=1}^{N} \frac{(M_i - M_i^k)^2}{\sigma_i^k}, \tag{5}$$

where M_i denotes the ith mean value of the query handwriting, M_i^k and σ_i^k denote the ith mean and standard derivation of the training handwriting of writer K separately, and N denotes the total number of mean values.

3. Our algorithm for writer identification

Our algorithm for writer identification, which can be regarded as a problem of pattern recognition to some extent, contains three main steps. They are

(1) *Preprocessing*: Removing the image noises and other detrimental factors which would disturb the later processings.
(2) *Feature extraction* (*FE*): Extracting features to fully represent the given handwriting image.
(3) *Similarity measurement* (SM): Using a certain measurement function to calculate the similarity between extracted features of the query handwriting image and the training handwriting images.

The whole procedure of our algorithm is described in Fig. 2.

3.1. Preprocessing

As we know, the original handwriting image contains characters of different sizes, spaces between text lines and even noises. So prior to FE, the original image should be preprocessed firstly. In the whole identification procedure, preprocessing plays an important role and inevitably influences the later processes and even the identification result.

In Refs. [7,8,26], preprocessing steps are as follows: firstly, removing the noises in the handwriting image; secondly, locating the text line and separating the single character using projection; thirdly, normalizing each character into a same size; finally, creating the preprocessed handwriting image (PHI) by text padding. However this method only aims to handle the handwriting documents with a regular layout. Admittedly, automatic localization and segmentation of irregular handwriting

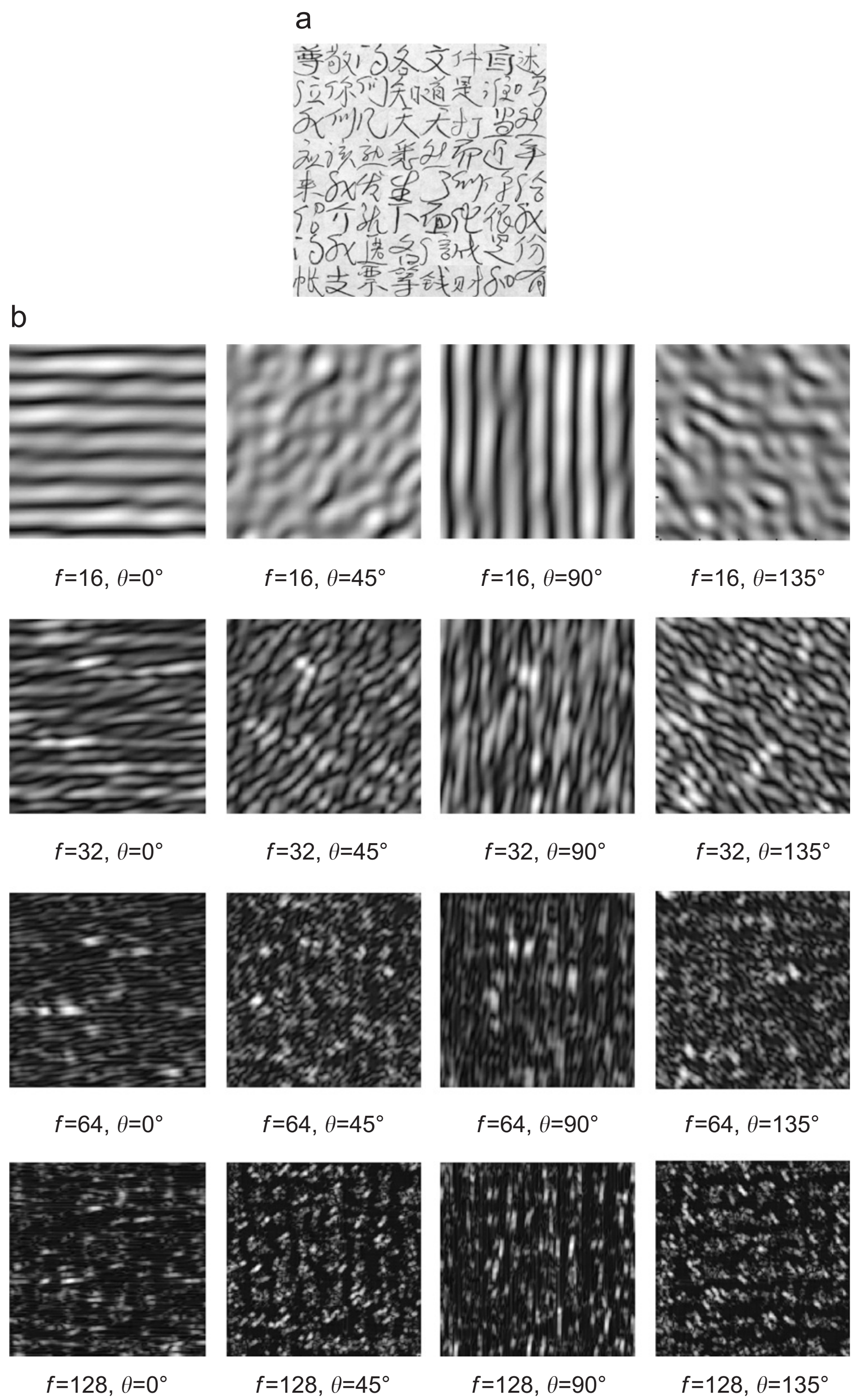

Fig. 1. (a) One preprocessed handwriting image (we will introduce how to obtain the preprocessed handwriting image from the original handwriting image in Section 3.1), (b) Multi-channel outputs of this preprocessed handwriting image by two-dimensional Gabor filters at different orientations and frequencies.

documents are far from being successfully solved [9], and evasive in nearly all relevant papers. While we cannot guarantee that all involved handwritings are written in a regular layout in practical applications. Therefore, we must find out an effective method to deal with the irregular handwriting documents. For this, in our research, we develop a software which can interactively localize and segment the characters manually from the irregular part of handwritings and generates the preprocessed handwriting images (PHIs) with high-quality. In the following, we offer an example to show how our software works.

Our software provides a rubber-like tool to remove the noises and needless marks. Fig. 3(a) is the original situation of the Chinese characters we want to process. Obviously, in this figure, the right character is surrounded by a circle, which may be a

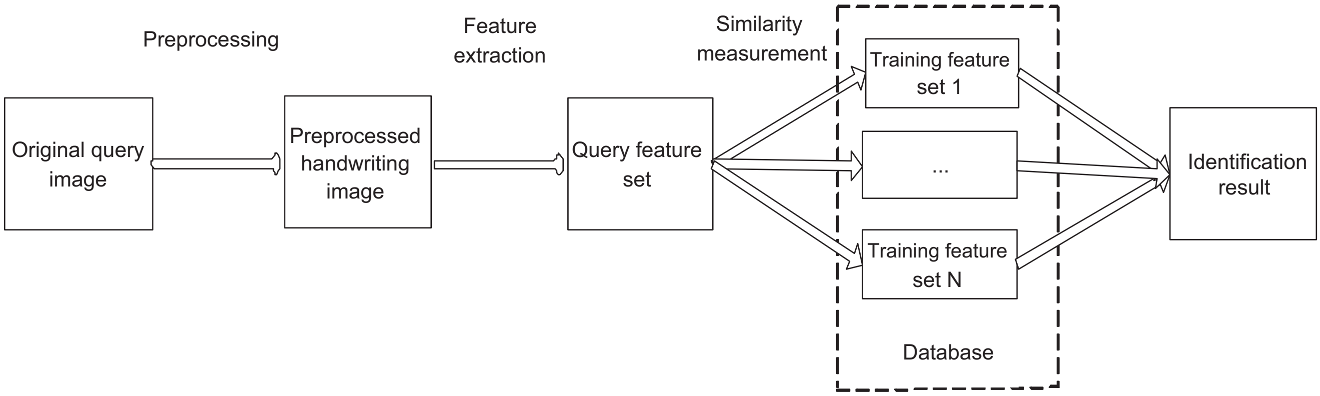

Fig. 2. The flow chart of our algorithm for writer identification.

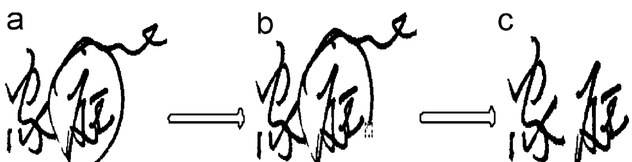

Fig. 3. Removing the noise.

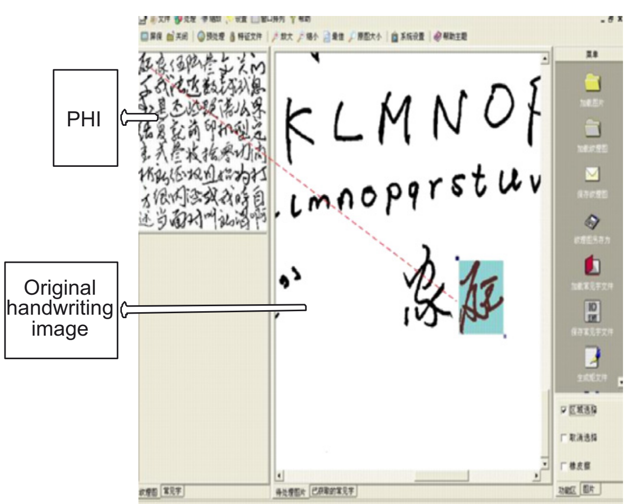

Fig. 4. Segmenting the character.

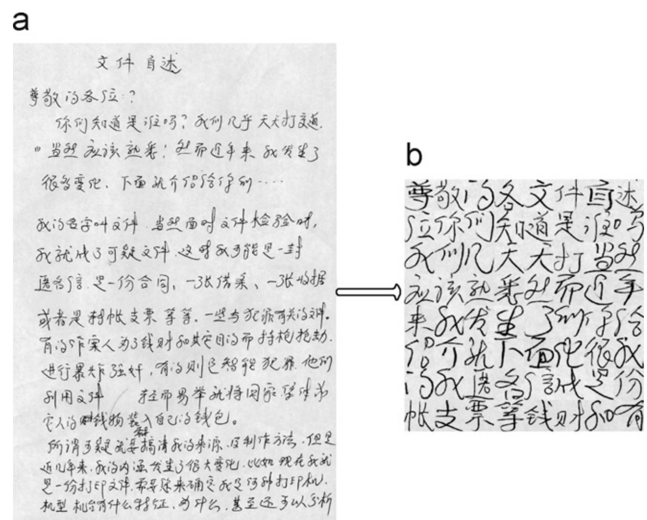

Fig. 5. An example of preprocessing. (a) One original handwriting image. (b) PHI obtained from this original handwriting image via our software. To offer a good visual effect, the size of PHI is enlarged in this figure.

revising mark. Fig. 3(b) shows how to remove the outside circle, where the dotted box is the area the rubber-like tool is working in. The size of this tool can be adjusted. Therefore, one user can select large size when dealing with a large area of noise and small size when carefully dealing with the overlapping. Fig. 3(c) is the result after removing the needless mark.

Our software also provides a segmenting tool, which is a rectangle box with two blue "ears" at the left upper corner and right bottom corner separately. By manipulating these two "ears", the user can segment any rectangle area from the image. Fig. 4 shows the Chinese character we want is segmented and then padded into the PHI. The normalization of the character is implemented automatically by our software when the character is padded into the PHI.

Fig. 5 shows one original Chinese handwriting and the PHI generated from it using our software. Ref. [27] provides more details about the preprocessing.

3.2. FE based on the hidden Markov tree model in wavelet domain

Though two-dimensional Gabor model is effective to extract the global writing style of handwriting images, this method still suffers from some inherent disadvantages, which greatly limit its practicability. One of the most serious disadvantages is its intensively computational cost, because the two-dimensional Gabor filters have to convolute the whole handwriting image for each orientation and each frequency. As a better multi-channel analysis tool, wavelet transform can decompose the image into a series of wavelet subbands with different resolutions, and afterward only subbands with interest are taken into account. Another disadvantage of the Gabor method is that it does not consider relation among the Gabor coefficients in each subband and only use the mean and standard derivation to represent a whole subband. As we know, mean and standard derivation are not accurate statistical description of one subband. To well capture the relation among wavelet coefficients, hidden Markov tree (HMT) is an ideal model.

3.2.1. Two-dimensional wavelet decomposition

Here we assume that point (x, y) is a pixel in an image. It has a gray function $f(x, y)$, which indicates the gray level at this pixel. Then, the wavelet transform of function $f(x, y)$ is defined as [29]

$$W_f(x, y) = f(x, y) * \psi(x, y), \tag{6}$$

where, '$*$' stands for the two-dimensional convolution operator, and $\psi(x, y)$ is a two-dimensional wavelet function, which satisfies the "admissibility" condition,

$$c_\psi := \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (\hat{\psi}(\omega_x, \omega_y))^2/(\omega_x^2+\omega_y^2)\mathrm{d}\omega_x \mathrm{d}\omega_y < \infty, \tag{7}$$

where, $\hat{\psi}(\omega_x, \omega_y)$ is Fourier transform of the function $\psi(x, y)$. In order to extract features from an image at different resolutions, the multi-scale wavelet function can be written explicitly as

$$\psi_s(x, y) = (1/s^2)\psi(x/s, y/s), \tag{8}$$

where s is the scale. Wavelet transform of $f(x, y)$ at scale s is,

$$W_{s,f}(x, y) = f(x, y) * \psi_s(x, y). \tag{9}$$

Furthermore, some constraints are forced into the "mother" wavelet function so as to guarantee the transform to be non-redundant, complete, and form a multi-resolution representation for the original image. A well-known example is Daubechies wavelet, which is the orthonormal bases of compactly supported wavelet [28], from which the pyramid algorithm of wavelet decomposition can be drawn out.

After we decompose the handwriting image into a series of wavelet subbands, what we should do next is to is find the feature information contained in these wavelet subbands, based on which we can discriminate one handwriting image from others.

3.2.2. Training HMT model for the handwriting image

In Refs. [29–31], researchers found the wavelet coefficients satisfy two properties: clustering, persistence across scale. Clustering means the large/small values of wavelet coefficients tend to propagate to their neighbors. Persistence across scale means the large/small values of wavelet coefficients tend to propagate across scales.

To match the non-Gaussian nature of the wavelet coefficients, M-state Gaussian mixture density is used as a probabilistic models for an individual wavelet coefficient,

$$f_W(w) = \sum_{s=1}^{M} P_S(s) f_{W|S}(\omega|s), \tag{10}$$

where, $P_S(s)_{s\in 1,2,\dots,M}$ is the probability massive function (pmf) of state variable S with value s, and $f_{W|S}(\omega|s)$ is the Gaussian conditional probability density function (pdf).

The persistence property of wavelet coefficients suggests an across-scale dependency between a wavelet coefficient w_i at a coarse resolution and its corresponding coefficients at the next resolution, which are also called the children of w_i. For example, in Fig. 6, tree nodes 2, 3, 4, 5 are children nodes of node 1. In other words, node 1 is the parent node of tree nodes 2, 3, 4, 5. The cluster property of wavelet coefficients reveals a strong inter-scale dependency between a wavelet coefficient w_i and its neighbors within the same scale.

Fig. 6 is a vivid example graph to describe the dependency among wavelet coefficients. In Fig. 6, each white node refers to the observed value of a continuous wavelet coefficient ω_i, each black node represents the mixture state value S_i for ω_i. Connecting the hidden state value nodes vertically across scale (solid links) yields the HMT model [32].

Using the M-state Gaussian mixture model mentioned above for each wavelet coefficient ω_i, the HMT model can be completely defined by the following parameters.

(1) $P_{S_1}(m)$, the pmf of state value of the root node 1.
(2) $\varepsilon_{m,n}^{i,P(i)} = P_{S_i|S_{P(i)}}(m|n)$, the *parent* $\rightarrow$ *children* link between hidden states.
(3) μ_{im}, σ_{im}, the mean and standard derivation of Gaussian pdf of wavelet coefficient ω_i given state $S_i = m$.

For simplicity, we usually assume $M = 2$ because in this case the state value has a clear physical meaning. Wavelet coefficients with large value contain significant contributions of signal energy, wavelet coefficients with small value contain little signal energy. And since the wavelet coefficients are generated by the wavelet filters with zero sum, they can be considered to be zero-mean.

The basic idea of using HMT model to characterize the features of a handwriting image is to train an HMT model for a handwriting image, and then use the trained HMT parameter set $\theta = \{P_{S_1}(m), \varepsilon_{m,n}^{i,P(i)}, \sigma_{im}\}$ as the features of handwriting.

On training HMT model, we look for the parameters that best fit a given set of wavelet coefficients. Maximum likelihood estimation is an effective principle for parameter estimation. That is, we choose the model parameters that maximize the probability of the observed wavelet data. Therefore, we adopt the EM algorithm used in Ref. [32] for training. The basic steps are given as follows:

(1) *Initialization*: Set an initial model estimate θ^0.
(2) *E step*: Estimate $P\{S|W, \theta^t\}$, the probability for the hidden state variables of the wavelet coefficients.

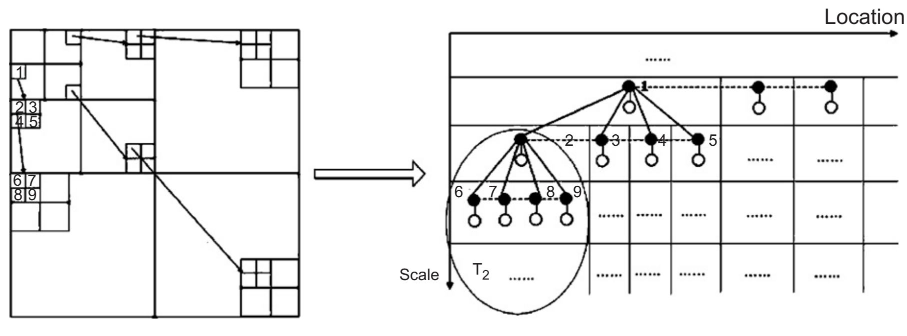

Fig. 6. The HMT model in wavelet domain.

(a) *Upward step*: Propagate hidden state information up along the HMT.
(b) *Downward step*: Propagate hidden state information down along the HMT.

(3) *M step*: Update θ^t to maximize the expected likelihood function.
(4) *Iteration step*: Set $t = t + 1$. If not converged, then return to E step; else stop the whole procedure.

3.3. Similarity measurement

After obtaining the parameter set of HMT model, we can regard a handwriting image is completely represented by its corresponding parameter set θ. In other words, the similarity between two handwriting images can be considered as the similarity between the two corresponding HMT parameter sets.

From Ref. [33], we know the Kullback–Leibler distance (KLD) is an effective measurement of the similarity between two HMT models. KLD between two pmfs is given as

$$K(\omega\|\bar{\omega}) = \sum_i \omega_i \log \frac{\omega_i}{\bar{\omega}_i}, \tag{11}$$

and KLD between two pdfs is defined below,

$$K(P(X|\theta_i)\|P(X|\theta_j)) = \int P(x|\theta_i) \log \frac{P(x|\theta_i)}{P(x|\theta_j)}. \tag{12}$$

In HMT model, the probability function is very complex which can be viewed as a mixture of large number of pdfs, and we do not have a simple, direct expression for the KLD. Monte–Carlo method is a traditional approximation of the KLD, while it needs high computational cost [34]. To save the computational cost, we use the method to compute the upper boundary of the KLD [35] instead of the Monte–Carlo method, as is based on such a lemma proposed in Ref. [33] that the KLD between two mixture densities $\sum_i \omega_i f_i$ and $\sum_i \bar{\omega}_i \bar{f}_i$ is upper bounded by

$$K\left(\sum_i \omega_i f_i \,\middle\|\, \sum_i \bar{\omega}_i \bar{f}_i\right) \leqslant K(\omega\|\bar{\omega}) + \sum_i \omega_i\, K(f_i\|\bar{f}_i) \tag{13}$$

with equality if and only if $\omega_i f_i / \bar{\omega}_i \bar{f}_i = const$.

For a tree node i, its conditional probability density of its observation value given that its state is m is defined as

$$P(O_i = o|S_i = m) = b_m^i(o). \tag{14}$$

Generally, $b_m^i(o)$ is a Gaussian function.

Then define β_m^i is the conditional probability density of the observation value of the subtree of the node i given its state is m, $P(i)$ is the parent of the tree node i, $C(i)$ is the set of children node of tree node i, T_i is the subtree of all nodes with root at i (for example, in Fig. 6, the tree inside circle is T_2), and O_{T_i} is the wavelet values (observed value) of this subtree. For a tree node i, based on the hidden Markov rule, we can get

$$\beta_m^i(O_{T_i}) = b_m^i(o_i) \prod_{c \in C(i)} \sum_{m=1}^{2} \varepsilon_{mr}^{i,P(i)} \beta_m^c(O_{T_c}). \tag{15}$$

For a leaf node i without any children node,

$$\beta_m^i(O_{T_i}) = b_m^i(o_i). \tag{16}$$

For the root node 1, the probability of the whole observation tree is defined as

$$P(O_{T_1} = o_{T_1}|\theta) = \sum_{m=1}^{2} P_{S_1}(m)\beta_m^1(o_{T_1}). \tag{17}$$

Then the upper boundary of the KLD between two HMTs is

$$K(P(O_{T_1} = o_{T_1}|\theta)\|P(\bar{O}_{T_1} = \bar{o}_{T_1}|\bar{\theta})) \leqslant K(P_{S_1}\|\bar{P}_{S_1}) + \sum_{m=1}^{2} P_{S_1}(m) K(\beta_m^1\|\bar{\beta}_m^1). \tag{18}$$

$K(\beta_m^1\|\bar{\beta}_m^1)$ can be calculated in the following way,

$$K(\beta_m^1\|\bar{\beta}_m^1) \leqslant K(b_m^1\|\bar{b}_m^1) + \sum_{c \in C(1)} (K(\varepsilon_m^1\|\bar{\varepsilon}_m^1) + \sum_{n=1}^{2} \varepsilon_{m,n}^{c,1} K(\beta_n^c\|\bar{\beta}_n^c)). \tag{19}$$

Fig. 7. A part of PHIs (20 PHIs belonging to 10 persons) carried out in our experiment. The meaning of image index '*XY*' is defined as follows: the frontal half '*X*' means this handwriting is written by the writer whose index is '*X*' (the value of *X* is from 001 to 500); the latter half '*Y*' means the purpose of this handwriting. If $Y = 01$, then this handwriting is a training handwriting. For a query handwriting, $Y = 02$.

This induction can be iteratively operated till the leaf node. As for the leaf node i of the HMT, the KLD can be written as

$$K(\beta_m^i \| \bar{\beta}_m^i) = K(b_m^i \| \bar{b}_m^i). \tag{20}$$

$K(b_m^i \| \bar{b}_m^i)$ is a KLD between two Gaussian pdfs and $K(\varepsilon_m^1 \| \bar{\varepsilon}_m^1)$ is a KLD between two pmfs. The following expression for KLD between two d-dimensional Gaussian pdfs is used [36],

$$D(N(.|\mu, C) \| D(N(.|\bar{\mu}, \bar{C}) = \frac{1}{2}\left[\log \frac{|\bar{C}|}{C} - d + trace(\bar{C}^{-1}C) + (\mu - \bar{\mu})^T \bar{C}^{-1}(\mu - \bar{\mu})\right]. \tag{21}$$

As we mentioned above, the pdf of wavelet coefficients is a Gaussian with zero-mean, so Eq. (21) can be simplified as

$$D(N(.|\mu, C) \| D(N(.|\bar{\mu}, \bar{C}) = \frac{1}{2}\left[\log \frac{|\bar{C}|}{C} - d + trace(\bar{C}^{-1}C)\right]. \tag{22}$$

In sum, the KLD of two HMTs can be expressed by Eq. (18), and $D(\beta_m^1 \| \bar{\beta}_m^1)$ in Eq. (18) can be calculated out by Eq. (19) iteratively till leaf node. The KLD between two leaf nodes is expressed by Eq. (20).

Next, we can generate the identification result according to the KLD values. The smaller KLD value is, the more similar it is. We only consider the k-nearest neighbor classifier since it is a simple while efficient scheme [37]. That is, identification result is the list of top M handwriting images which are most similar to the query handwriting image. We can know the corresponding writers from the indexes of these top ranked handwritings.

4. Experiments

4.1. Database used in our experiments

One thousand Chinese handwritings written by 500 persons have been carried out in our experiments, with one training handwriting and one query handwriting for each person. All handwritings are scanned into computer with a resolution of 300 dpis. We produce one PHI from each original handwriting, and thus totally 1000 PHIs are obtained. The training and query PHIs both consist of 64 Chinese characters with size 64×64 (the unit is pixel), arranged in an 8×8 array, shown in Fig. 7.

4.2. Identification performance evaluation 1

In our experiments, we compare our method with two-dimensional Gabor model on not only the identification accuracy but also the computational efficiency. For two-dimensional Gabor model, Tan et al. pointed out that for an image of size $N \times N$, the most significant Gabor frequency components were equal to or smaller than N/4 [38]. Therefore, for a PHI of size 512×512, the Gabor frequency should not be larger than 128. Several combinations of different Gabor frequencies are tested, ranging from 16 to 128. For each spatial frequency, we select 0, 45, 90 and 135 degree as orientations. In our method, we decompose the handwriting image via traditional discrete wavelet transform (DWT) using Daubechies orthogonal wavelets. Of course, different wavelet filters may lead to different results. While testing all possible wavelet filters and finding out which one is the best is out of the scope of this paper.

The evaluation criterion of identification is defined as follows: for each query handwriting, if the training handwriting belonging to the same writer is ranked within the top S matches, we say that this is a correct identification, otherwise a failure identification. The identification rate is the percentage of the correct identification. The identification results of our experiments are offered in the Table 1. Fig. 8 shows a graph, which more directly illustrates the comparison of our method and two-dimensional Gabor model on identification rate. Obviously, the identification rate changes at varied number of top matches considered.

An example of writer identification using our method is shown in Fig. 9. The training handwriting written by the same writer of the query handwriting is ranked at the top 1, so certainly this is a successful identification.

And KLD values between the query handwriting and top matches are much less than those between the query handwriting and undermost matches. In this example, the KLD values of the top 9 matches are $\{1.32\times10^4, 3.18\times10^4, 3.44\times10^4, 5.41\times10^4, 5.75\times10^4, 5.79\times10^4, 6.77\times10^4, 6.98\times10^4, 7.41\times10^4\}$, while the KLD values of the undermost 5 matches are $\{4.45\times10^6, 4.60\times10^6, 4.83\times10^6, 4.87\times10^6, 4.93\times10^6\}$. To some extent, the difference between the similarity distances of correct matches ("01601" in this example) and uncorrect matches can indicate one method's discriminating ability. The larger the difference is, the better the discriminating ability is.

Our program also records the elapsed time of our method and two-dimensional Gabor model. Our program is implemented in PC computer. The software environment of our computer is: Window XP, Matlab 7.0; and the hardware environment is: Intel Pentium IV 2.4 GHZ CPU, 512 MB RAM. The record of average elapsed time is given in Table 2.

From Table 1, it is clear that in two-dimensional Gabor model, the more frequencies are combined, the higher identification rate is achieved; unfortunately at the same time, the elapsed time also increases greatly. The identification rate of two-dimensional Gabor model combing four frequencies $f = 16, 32, 64, 128$ is nearly same to that of our method, while its elapsed time is four times of that used by our method. The elapsed time of two-dimensional Gabor model with $f = 16$ is close to our method, however its identification rate is large lower. Comprehensively, our method outperforms

Table 1
Writer identification rate 1(%)

Number of top matches	Our method	Gabor $f = 16$	Gabor $f = 16, 32$	Gabor $f = 16, 32, 64, 128$
1	36.4	13.4	18.2	32.8
2	43.6	24.6	31.8	39.0
3	52.2	33.8	43.2	49.4
5	60.4	41.4	51.6	56.2
7	67.8	47.2	58.4	64.8
10	74.6	55.0	64.2	71.4
15	82.4	64.6	71.6	79.8
20	89.8	70.8	79.4	85.2
25	95.4	76.2	84.2	91.2
30	100	80.6	87.8	95.6
40	100	86.6	92.8	100

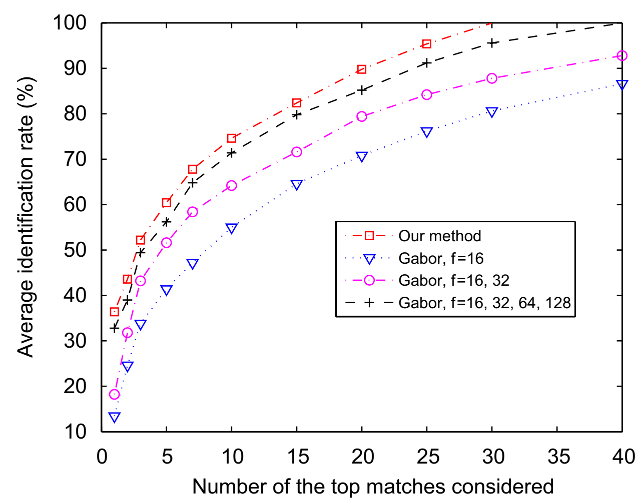

Fig. 8. Identification rate according to the number of top matches considered.

Fig. 9. One identification example using our method. The query handwriting "01602" is on the top left corner. The upper two row (except the query handwriting "01602") are the top nine matches, and the last row is the undermost five matches. The handwriting matches are ranked from left to right, from top to bottom.

Table 2
Average elapsed time 1 for writer identification(second)

Method	Our method	Gabor $f = 16$	Gabor $f = 16, 32$	Gabor $f = 16, 32, 64, 128$
Elapsed time	50.35	53.17	107.03	213.87

two-dimensional Gabor model on both identification performance and the computational efficiency.

4.3. Identification performance evaluation 2

To increase the writing samples for one writer in order to enhance the persuasion of the performance evaluation, we divide one PHI of 512×512 into four non-overlapped sub-PHIs of 256×256. A figure illustrating this division is given in Fig. 10. In this way, we can obtain eight sub-PHIs from one writer. Inspired by the evaluation criterion in Ref. [39], for each query sub-PHI, only the top $S \geqslant 7$ matches are considered since, for each query sub-PHI, there are seven sub-PHIs of the same writer. The identification percentage is the ratio of the number of correct matches within the top S identification results to

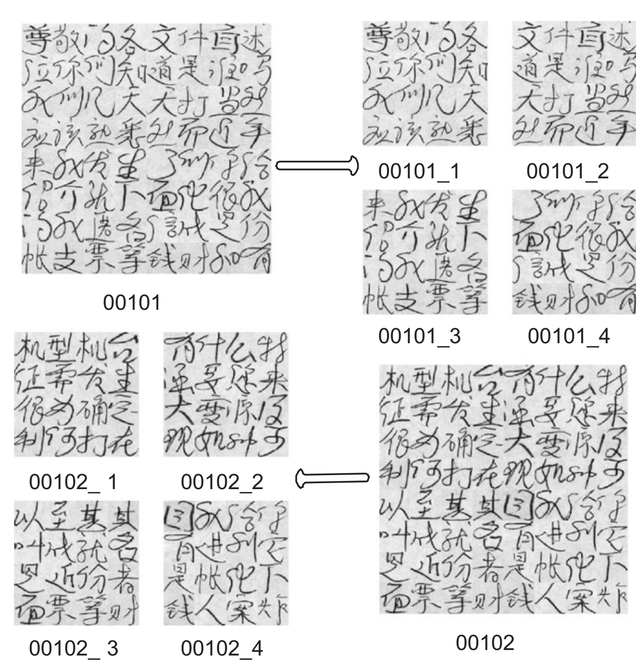

Fig. 10. Dividing one handwriting image into four non-overlapped subimages.

Table 3
Writer identification rate 2(%)

Number of top matches	Our method	Gabor $f=16$	Gabor $f=16, 32$	Gabor $f=16, 32, 64, 128$
7	26.84	10.17	15.32	22.37
10	31.94	16.65	22.48	27.56
20	41.77	25.82	31.63	37.03
30	49.39	35.73	40.07	44.89
50	60.43	44.25	49.11	54.47
70	69.25	52.19	60.38	66.72
100	77.81	60.39	68.65	74.58
150	84.57	68.48	74.61	80.25
200	91.35	73.54	79.17	88.69
300	97.83	80.26	87.53	93.71

seven. For instance, in the case of $S=10$, the identification rate is $5/7 \times 100\% = 71.43\%$ if five correct matches are at the top 10 matches. In this experiment, we do not classify the sub-PHIs into training group and query group. All sub-PHIs are used as a query handwriting, and at the same time other sub-PHIs except for the query one play the role as the training handwritings. The identification rates of our method and two-dimensional Gabor model are offered in Table 3 and Fig. 11. We also provide an example in this case, as is shown in Fig. 12.

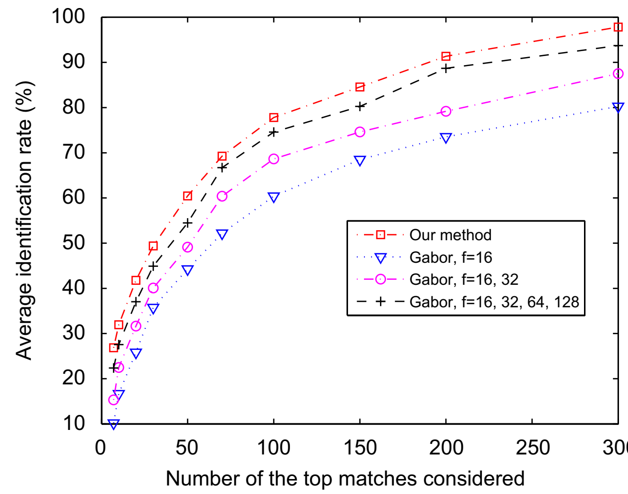

Fig. 11. Identification rate according to the number of top matches considered.

Table 4
Average elapsed time 2 for writer identification (second)

Method	Our method	Gabor $f=16$	Gabor $f=16, 32$	Gabor $f=16, 32, 64, 128$
Elapsed time	10.48	5.01	9.10	17.42

Fig. 12. Another identification example using our method. The query handwriting '01801_1' is on the top left corner. The upper two row (except the query handwriting '01801_1' are the top 7 matches, the last row is the undermost 4 matches. The handwriting matches are ranked from left to right, from top to bottom. The KLD values of the top 7 matches are $\{1.19 \times 10^4, 1.27 \times 10^4, 1.32 \times 10^4, 1.38 \times 10^4, 1.41 \times 10^4, 1.44 \times 10^4, 1.46 \times 10^4\}$, the KLD values of the undermost 4 matches are $\{1.17 \times 10^6, 1.22 \times 10^6, 1.24 \times 10^6, 1.31 \times 10^6\}$.

The average elapsed time of our method and two-dimensional Gabor model in this experiments is offered in Table 4. Combining the results in Table 3 and Table 4, our method still outperform the two-dimensional Gabor model in this experiment.

5. Conclusions

In this paper, we presented a new method using HMT model in wavelet domain for off-line, text-independent handwriting identification. Experiments on 1000 Chinese handwritings and 4000 sub-handwritings provided by 500 persons indicate that our new method is satisfactory and outperforms the two-dimensional Gabor model, one representative of the existing methods for off-line, text-independent writer identification, on both identification accuracy and computational efficiency. This is consistent with our expectation. At first, in our method, the handwritings are decomposed into a series of wavelet subbands at different resolutions via wavelet transform, and only the wavelet coefficients within subbands of interest are considered. While in two-dimensional Gabor model, the whole handwritings are convoluted with the two-dimensional Gabor filters for each frequency and each orientation and the convolution must redo when either frequency or orientation changes, as greatly decreases the computational efficiency. In addition, mean and standard derivation, two statistical parameters used in two-dimensional Gabor model, certainly cannot well describe the statistical properties of Gabor coefficients within each Gabor subband. On the contrary, the HMT model provides an accurate description for the statistical distribution of wavelet coefficients. The accurate model naturally brings about a better identification result. By the way, the methods discussed in this paper are also applicable to English, Korean, Japanese and Latin Language, etc., since text-independent methods do not care about the writing content.

Acknowledgments

This work was supported by research grants received from Research Grant Council (RGC) of Hong Kong, and Faculty Research Grant (FRG) of Hong Kong Baptist University. This work is supported by Grants 60403011 and 60773187 from the NSFC, NCET2007 and Grants 2006ABA023, 2007ABA036, 2007CA011 from the Department of Science and Technology in Hubei province, China. In addition, the authors would like to thank the anonymous reviewers for their perceptive comments, which have significantly improved the paper.

References

[1] A.K. Jain, Recent development on biometric authentication, Proceeding of Advanced Study Institute (ASI), Hong Kong Baptist University, Hong Kong, December 2004.
[2] M. Benecke, DNA typing in forensic medicine and in criminal investigations: a current survey, Natur Wissenschaften 84 (5) (1997) 181–188.
[3] B. Devlin, N. Risch, K. Roeder, Forensic inference from DNA fingerprints, J. Am. Stat. Assoc. 87 (418) (1992) 337–350.
[4] J. Daugman, The importance of being random: statistical principles of iris recognition, Pattern Recognition 36 (2) (2003) 279–291.
[5] A. Jain, L. Hong, R. Bolle, On-line fingerprint verification, IEEE Trans. Pattern Anal. Mach. Intell. 19 (4) (1997) 302–314.
[6] S. Srihari, S. Cha, H. Arora, S. Lee, Individuality of handwriting, J. Forensic Sci. 47 (4) (2002) 1–17.
[7] H.E.S. Said, T. Tan, K. Baker, Writer identification based on handwriting, Pattern Recognition 33 (1) (2000) 133–148.
[8] Y. Zhu, T. Tan, Y. Wang, Biometric personal identification based on handwriting, in: Proceedings of 15th International Conference Pattern Recognition, 2000, pp. 801–804.
[9] R. Plamondon, G. Lorrtte, Automatic signature verification and writer identification—the state of the art, Pattern Recognition 37 (12) (1989) 107–131.
[10] J. Duvernoy, Handwriting synthesis and classification by means of space-variant transform and Karhunenloeve analysis, J. Opt. Soc. Am Cybern. 65 (1975) 1331–1336.
[11] J. Duvernoy, D. Charraut, P.Y. Baures, Hybrid optical digital image processing applied to handwriting recognition and aerial photograph clustering, Opt. Acta, Univ. de Franche-comte, Besancon, France, 1977, pp. 795–810.
[12] W. Kuckuck, Writer recognition by spectra analysis, in: Proceedings of the 1980 International Conference Security Through Science Engineering, 1980, pp. 1–3.
[13] W. Kuckuck, B. Rieger, K. Steinke, Automatic writer recognition, in: Proceedings of the 1979 Carnahan Conference on Crime Countermeasures, University of Kentucky, Lexington, vol. 23(1), 1979, pp. 57–64.
[14] K. Steinke, Recognition of writers by handwriting images, Pattern Recognition 14 (1981) 357–364.
[15] V. Klement, K. Steinke, R. Naske, The application of image processing and pattern recognition techniques to the forensic analysis of handwriting, in: Proceedings of the International Conference Security Through Science Engineering 1980, pp. 155–172.
[16] E.N. Zois, V. Anastassopoulos, Morphological wavelform coding for writer identification, Pattern Recognition 33 (2000) 385–398.
[17] C. Hertel, H. Bunke, A set of novel features for writer identification, Audio-and-Video Based Biometric Person Authentication, 2003, pp. 679–687.
[18] A. Schlapbach, H. Bunke, Off-line handwriting identification using HMM based recognizers, in: Proceedings of 17th International Conference on Pattern Recognition, August 2004, pp. 654–658.
[19] M. Bulacu, L. Schomarker, L. Vuurpijl, Writer identification using edge-based directional features, in: Proceedings of 7th International Conference on Document Analysis and Recognition, 2003, pp. 937–941.
[20] A. Nosary, L. Heutte, T. Paquet, Defining writer's invariants to adapt the recognition task, in: Proceedings of 5th International Conference on Document Analysis and Recognition, 1999, pp. 765–768.
[21] A. Bensefia, A. Nosary, T. Paquet, L. Heutte, Writer identification by writer's invariants, Proc. International Workshop on Frontiers in Handwriting Recognition, 2002, pp. 274–279.
[22] A. Bensefia, T. Paquet, L. Heutte, Information retrieval based writer identification, in: Proceedings of 7th International Conference on Document Analysis and Recognition, 2003, pp. 946–950.
[23] G. Leedham, S. Chachra, Writer identification using innovative binarised features of handwritten numerals, in: Proceedings of 7th International Conference on Document Analysis and Recognition, 2003, pp. 413–417.
[24] S. Cha, S.N. Srihari, Multiple feature integration for writer verification, in: Proceedings 7th IWFHR2000, September 2000, pp. 333–342.
[25] J.G. Daugman, Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters, J. Opt. Soc. Am. 1985, 1160–1169.
[26] C. Shen, X.G. Ruan, T.L. Mao, Writer identification using Gabor wavelet, in: Proceedings of 4th World Congress on Intelligient Control and Automation, June 2002, pp. 2061–2064.
[27] Z.Y. He, Writer Identification Using Wavelet, Contourlet and Statistical Models, Ph.D Thesis, Hong Kong Baptist University, 2006.
[28] I. Daubechies, Ten Lectures on Wavelets, SIAM, Philadelphia, PA, 1992.
[29] S. Mallat, S. Zhong, Characterization of signals from multiscale edges, IEEE Trans. Pattern Anal. Mach. Intell. 14 (1) (1992) 710–732.

[30] S. Mallat, W. Hwang, Singularity detection and processing with wavelets, IEEE Trans. Inf. Theory 38 (2) (1992) 617–643.
[31] M.T. Orchard, K. Ramchandran, An investigation of wavelet-based image coding using an entropy constrained quantization framework, in: Proceedings of Data Compression Conference '94, 1994, pp. 341–350.
[32] M. Crouse, R.D. Nowak, R.G. Baraniuk, Wavelet-based signal processing using hidden Markov model, IEEE Trans. Signal Process. (Special Issue on Wavelets and Filter banks) (1998) 886–902.
[33] T.M. Cover, J.A. Thomas, Elements of Information Theory, Wiley Interscience, New York, NY, 1991.
[34] B.H. Juang, L.R. Rabiner, A probabilistic distance measure for hidden Markov models, AT&T Tech. J. 64 (2) (1985) 391–408.
[35] M.N. Do, Fast approximation of Kullback-Leibler distance for dependence trees and hidden Markov models, IEEE Signal Process. Lett. 10 (2003) 115–118.
[36] Y. Singer, M.K. Warmuth, Batch and on-line parameter estimation of Gaussian mixtures based on the joint entropy, in: Advances in Neural Information Processing Systems 11 (NIPS'98), 1998, pp. 578–584.
[37] A.R. Webb, Statistical Pattern Recognition, second ed., Wiley, New York, 2002.
[38] T.N. Tan, Written language recognition based on texture analysis, in: Proceedings of IEEE ICIP'96, September 1996, pp. 185–188.
[39] B.S. Manjunath, Texture features for browsing and retrieval of image data, IEEE Trans. Image Process. 18 (8) (1996) 837–842.

About the author—ZHENYU HE received the B.S. and M.S. degrees from Wuhan University of Technology, Wuhan, China and the Ph.D. degree in Department of Computer Science from Hong Kong Baptist University, Hong Kong, in 2000, 2003 and 2006, respectively. He is presently a post doctoral researcher in Department of Computer Science and Engineering, The Hong Kong University of Science and Technology. His research interests include wavelet, pattern recognition, biometrics, medical image analysis.

About the author—XINGE YOU received the B.S. and M.S. degrees in mathematics from the University of Hubei, Wuhan, China and the Ph.D. degree in computer science from the Hong Kong Baptist University, Hong Kong, in 1990, 2000, and 2004 respectively. He is presently a Professor in the Department of Electronics and Information Engineering at Huazhong University of Science and Technology, China. And he is currently work as postdoctoral fellow in the Department of Computer Science at Hong Kong Baptist University, Hong Kong. His current research interests include wavelets and its application, signal and image processing, pattern recognition, and computer vision.

About the author—YUAN YAN TANG received the B.S. degree in electrical and computer engineering from Chongqing University, Chongqing, China, the M.Eng. degree in electrical engineering from the Graduate School of Post and Telecommunications, Beijing, China, and the Ph.D. degree in computer science from Concordia University, Montreal, Canada. He is presently an Adjunct Professor in the Faculty of Mathematics Computer Science at Hubei University, a Chair Professor in the Department of Computer Science at Hong Kong Baptist University and Adjunct Professor in the Computer Science at Concordia University. He is an Honorary Lecturer at the University of Hong Kong, an Advisory Professor at many institutes in China.
His current interests include wavelet theory and applications, pattern recognition, image processing, document processing, artificial intelligence, parallel processing, Chinese computing and VLSI architecture. Professor Tang has published more than 250 technical papers and is the author/co-author of 21 books/book-chapters on subjects ranging from electrical engineering to computer science.
He has serviced as General Chair, Program Chair and Committee Member for many international conferences. Professor Tang will be the General Chair of the 19th International Conference on Pattern Recognition (ICPR'06). He is the Founder and Editor-in-Chief of International Journal on Wavelets, Multiresolution, and Information Processing (IJWMIP) and Associate Editors of several international journals related to Pattern Recognition and Artificial Intelligence. Professor Y.Y. Tang is an IEEE Fellow and IAPR Fellow.