Writer Identification from Non-uniformly Skewed Handwriting Images

H. E. S. Said[1], G. S. Peake[1], T. N. Tan[2] and K. D. Baker[1]
[1]Department of Computer Science, University of Reading
Reading RG6 6AY, UK.
[2]National Laboratory of Pattern Recognition, Institute of Automation,
Chinese Academy of Sciences, Beijing, China.
[H.E.S.Said|G.S.Peake|K.D.Baker]@reading.ac.uk,
TNT@prlsun2.ia.ac.cn

Abstract

Many techniques have been reported for handwriting-based writer identification. Most such techniques assume that the written text is fixed (e.g., in signature verification). In this paper we attempt to eliminate this assumption by presenting a novel algorithm for automatic text-independent writer identification from non-uniformly skewed handwriting images. Given that the handwriting of different people is often visually distinctive, we take a global approach based on texture analysis, where each writers' handwriting is regarded as a different texture. In principle this allows us to apply any standard texture recognition algorithm for the task (e.g., the multi-channel Gabor filtering technique). Results of 96.0% accuracy on the classification of 150 test documents from 10 writers are very promising. The method is shown to be robust to noise and contents.

1 Introduction

Signature verification has been an active research topic for several decades in the image processing and pattern recognition community [1]. Despite the continuous effort, signature verification remains a challenging issue. Signature verification provides a means of identifying the writer of a piece of handwriting in order to verify claimed identity in security and related applications. The writer requires to write the same fixed text. In this sense, signature verification may also be called text-dependent writer verification (which is a special case of text-dependent writer identification where more than one writer has to be considered). In practice the requirement for the use of fixed text makes writer verification prone to forgery. Furthermore text-dependent writer identification is inapplicable in many important practical applications, for example, the identification of the writers of archived handwritten documents, crime suspect identification in forensic sciences etc. In these applications, the writer of a piece of handwriting is often identified by professional handwriting examiners (graphologist). Although human intervention in text-independent writer identification has been effective, it is costly and prone to fatigue.

BMVC 1998 doi:10.5244/C.12.48

Research into writer identification has been focused on two streams, off-line and on-line writer identification. This paper focuses on the off-line identification. Off-line systems are based on the use of computer image processing and pattern recognition techniques. There are two groups of off-line approaches, text-dependent and text-independent. Our work is a text-independent approach where a texture analysis technique is introduced. Similar work has been proposed by Kuckuck [2], where a Fourier transform technique is used.

There we use multi-channel spatial filtering techniques to extract texture features from a non-uniformly skewed handwriting image. There are many available filters in the multi-channel technique. In this paper we use Gabor filters, since they have proven to be successful in extracting features for similar applications [3,4,5,6,7]. We have also used grey scale co-occurrence matrices (***GSCM***) for feature extraction (for comparison purposes). Two classification techniques are adopted here, namely the weighted Euclidean distance (***WED***) and the ***K-NN*** classifiers. The subsequent sections describe the normalisation of the handwriting images, the extraction of writer features, the experimental results and finally the conclusions.

2 The Algorithm

The algorithm is based on texture analysis and is illustrated diagrammatically in Figure 1. It consists of three main stages which are described in turn in the remainder of this section.

Figure1: Block Diagram of the Algorithm

2.1 Normalisation of Handwriting images

The normalisation of the handwriting images of different writers requires in two separate stages. Firstly, the detection and correction of the skewed words in the handwriting images has been performed. Then, the space between vertical and horizontal lines has been normalised to produce a well defined pattern for texture analysis. Details are given in the following subsections.

2.1.1 Skewed Words Normalisation

In applications such as writer identification and other document image processing systems, individual words and characters in handwriting images are not guaranteed to be upright during scanning. The inclination of different words, characters and even lines in handwriting text images may affect the identification process of writers in later stages. To solve this problem, we need to perform some preprocessing to detect the skew angle of individual words and perform skew-normalisation before identification of different writers. Many methods have been proposed to detect the skew angle of text, but this has, in general, been limited to uniformly skewed lines in typed document images. One of the most popular skew estimation techniques is based on the projection profile of the typed documents.

The horizontal/vertical projection profile [8] is the histogram of the number of black pixels along horizontal/vertical scan lines. For a script document with horizontal text lines, the horizontal projection profile has peaks at text lines positions and troughs at positions in between successive text lines. To determine the skew of a document, the projection profile is computed at a number of angles and for each angle, a measure of difference of peaks and troughs height is made. The maximum difference corresponds to the best alignment with the text line direction which, in turn, determines the skew angle.

Baird [9] proposed a modified approach for quick convergence of this approach. Akiyama [10] described an approach where documents are partitioned into vertical strips. The horizontal projection profiles are calculated for each strip, and the skew angle is determined from the correlation of the profile of the neighbouring strips. This is fast but less accurate method. In [11], Pavilidis proposed a method based on the vertical projection profile of horizontal strips, which works well if the skew angle is small. Postal [12], proposed an approach where the direction in which the density in the Fourier transform is largest gives an estimate of skew angle.

Figure2: Skewed Word Normalisation of One Handwriting Image

Another class of approaches is based on nearest neighbour clustering connected component. There has been little work that introduced in the effort to detect and correct individual skewed words in handwriting images, although Spitz [13] has described a method which may be applied to typed documents. Our work is based on detecting and correcting individual words and characters in handwriting images using line fitting on the connected components. Figure 2 shows the steps of the skewed word detection and correction. Below are the steps for the skewed words normalisation procedure:

i) Open a handwriting image.

ii) First, detect text lines and empty spaces using the horizontal projection profile (HPP) method (this is simply to demonstrate the uneven lines spacing).

iii) Perform a closing procedure on the image using a 3×3 ***structuring element (only the middle row of the element is set so as to close the image in the horizontal direction to avoid joining text lines).***

iv) Extract the connected components.

v) Then, compute the minimum, maximum and mean connected component heights.

vi) Filter out the smallest 5% (in terms of height) to eliminate small blobs (such as punctuation etc.). Then remove components with a height > $2 \times mean$ ***height to eliminate components which are already connected across more than one text line.***

ii) Then for each of the remaining connected component perform the following:

- ***Copy the component into a blank image, in which the image has the component bounding box size.***
- ***Perform the line fit on the connected component.***
- ***Rotate the area (that is defined by the bounding boxing coordinates) in the original image by the negative value of the angle computed from the gradient of the line.***

Finally, perform base line fitting (base lines are computed from the HPP of the de-skewed image) to the de-skewed handwriting image, to produce horizontal text lines.

2.1.2 Text Normalisation

Texture analysis cannot be applied directly to handwriting images (even after skewed word correction as discussed above), as texture is affected by different word spacing, varying line spacing, etc. The influence of such factors is minimised by the text normalisation stage. The input to this stage is a binary image of any handwritten document, in which the skewed words have been normalised and from which graphics and pictures have been removed (at present the writer is asked to write only text non-textual information will be considered in the future). Document image segmentation can be applied to separate text and graphics.

The handwriting may contain lines of different point sizes, and different spaces between lines, words and characters. The normalisation is done as follows: First, text lines are located using the horizontal projection profile [6]. Then, spacing between lines/words and margins are set to predefined size by means of text padding. Finally, random non-overlapping blocks (of 128x128 pixels) are extracted from the normalised image.

Texture analysis is applied to these blocks. Further details on normalisation may be found in [6]. The normalisation of the handwriting images in order to produce texture block is summarised in Figure 3.

2.2 Feature Extraction

In principle any texture analysis technique can be applied to extract features from each uniform handwriting. Here two established methods are implemented to obtain texture features, namely the multi-channel Gabor filtering technique [3] and the grey scale co-occurrence matrix (GSCM) [6]. The former is one of the most popular methods and well recognised, and the latter is often used as a benchmark in texture analysis.

2.2.1 Multi-channel Gabor Filtering

The multi-channel Gabor filtering technique is inspired by the psychophysical findings that the processing of pictorial information in the human visual cortex involves a set of parallel and quasi-independent mechanisms or cortical channels which can be modelled by bandpass filters.

A simple computational model for the cortical channels is described in [3]. Briefly stated, each cortical channel is modelled by a pair of Gabor filters $h_e(x,y;f,\theta)$ and $h_o(x,y;f,\theta)$. The two Gabor filters are of opposite symmetry and are given by

$$\begin{cases} h_e(x,y;f,\theta) = g(x,y)\cos(2\pi f(x\cos\theta + y\sin\theta)) \\ h_o(x,y;f,\theta) = g(x,y)\sin(2\pi f(x\cos\theta + y\sin\theta)) \end{cases} \quad (2)$$

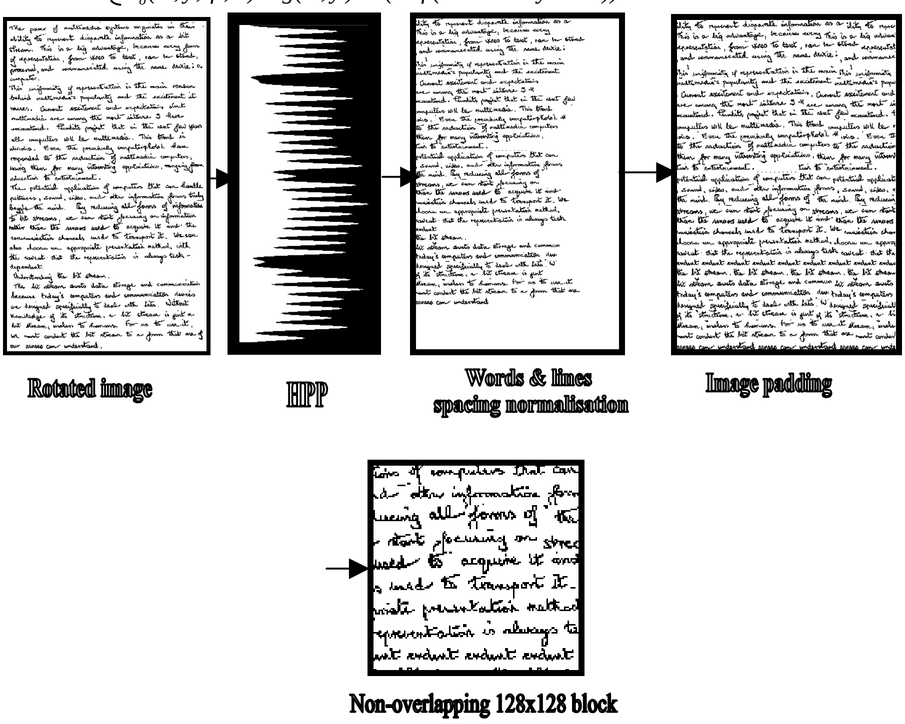

Figure3: Text Normalisation of Another Handwriting Image

where $g(x,y)$ is a 2-D Gaussian function and f and θ are the radial frequency and orientation which define the location of the channel in the frequency plane. Commonly used frequencies are of power 2. In [3] it has been shown that for any image of size $N \times N$, the important frequency components can be found within $f \le \frac{N}{4}$ cycles/degree. For this reason we use frequencies of 4, 8, 16 and 32 cycles/degree. For each central frequency f, filtering is performed at $\theta = 0°, 45°, 90° and 135°$ .This gives a total of 16 output images (4 for each frequency), from which the writer's features are extracted. These features are the mean and the standard deviation of each output image. Therefore, 32 features per input image are calculated. Testing was performed by using either all 32 features or various subsets (e.g., features associated with a particular radial frequency).

2.2.2 Grey Scale Co-occurrence Matrices (GSCM)

GSCMs are also considered. Generally speaking, GSCMs are very expensive to compute. For an image represented using $N \times N$ grey levels, each GSCM is of size $N \times N$. In the case of binary handwriting images, we have only two grey levels. Therefore, in this case, it is reasonable to use the GSCM technique. In this paper, GSCMs were constructed for five distances (d=1,2,3,4,5) and four directions $\theta = 0°, 45°, 90° and 135°$ $\alpha = 0°, 45°, 90° and 135°$. This gives each input handwriting image 20 matrices of dimension 2×2 .

When the size of the GSCM is too large to allow the direct use of matrix elements, measurements such as energy, entropy, contrast and correlation are computed from the matrix and used as features [6]. For each 2×2 GSCM derived from a binary handwriting image however, there are only 3 independent values due to the diagonal symmetry. There these 3 values are used directly as features. We then have 60 ($= 20 \times 3$) GSCM features per handwriting image.

2.2 Writer Identification

Here two classifiers were considered, namely the weighted Euclidean distance (WED) classifier and the nearest neighbour classifier (K-NN).

2.3.1 The Weighted Euclidean Distance (WED) Classifier

Representative features for each writer are determined from the features extracted from training handwriting texts of the writer. Then, for an unseen handwritten text block by an unknown writer (who has contributed training images), similar feature extraction operations are carried out. The extracted features are then compared with the representative features of a set of known writers. The writer of the handwriting is identified as writer K by the WED classifier *iff* the following distance function is a minimum at K :

$$d(k) = \sum_{n=1}^{N} \frac{(f_n - f_n^k)^2}{(v_n^k)^2} \qquad (2)$$

where f_n is the n^{th} feature of the input document, and $f_n^{(k)}$ and $v_n^{(k)}$ are the sample mean and sample standard deviation of the n^{th} feature of writer K respectively.

2.3.2 The Nearest Neighbour Classifier (K-NN)

When using the nearest neighbour classifier (K-NN), for each class K in the training set the ideal feature vectors is given as f_k. Then we detect and measure the features of the unknown writer represented as U. To determine the class K of the writer we measure the similarity with each class by computing the distance between the feature vectors f_k and U. The distance measure used here is the Euclidean distance. Then the distance computed d_k of the unknown writer from class is given by

$$d_k = \left[\sum_{j=1}^{N} \left(U_j - f_{kj} \right)^2 \right]^{\frac{1}{2}} \tag{3}$$

where $j = 1, \ldots, N$ (N is the number of the features considered).

The writer is then assigned to the class R such that:

$$d_R = min(d_k) \tag{4}$$

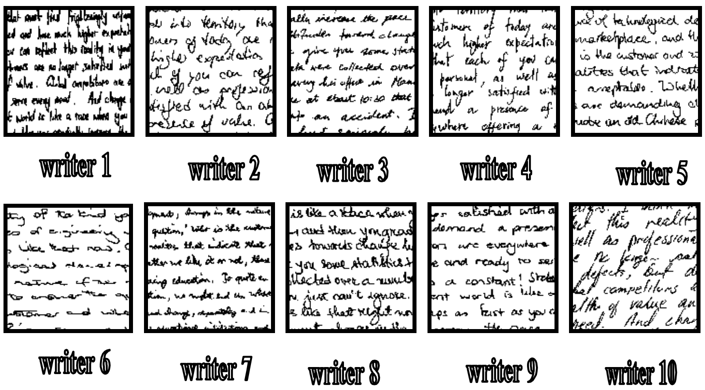

Figure 4: Examples of Handwriting of 10 Different People

Where ($k = 1, \ldots,$ no of classes). Other more sophisticated measures and classifiers such as neural network classifiers could have been used. The emphasis in this paper, however, is computational simplicity.

3 Experimental Results

A number of experiments were carried out to show the effectiveness of the proposed algorithm. Ten writers were chosen. Examples of handwriting by these people are shown in Figure 4. For the purpose of classification, 25 non-overlapping handwriting blocks were extracted for each person. Each sample was selected from an A4 page, scanned using a HP

ScanJet4c in custom mode with extra heavy lighting, at a resolution of 150dpi. Each sample block was of 128×128 pixels. The sample images were divided first into 10 training and 15 test images per writer (Set A) followed by 15 training and 10 test images (Set B). Images in the test sets did not appear in the training sets. Testing was conducted using different combinations of features under both classifiers.

3.1 Results from Gabor Filtering

The effects of the Gabor filtering on classification were investigated. Tables1-2 show the results of the multi-channel Gabor filter features using the two classifiers. It shows that features that were extracted using the channels at f = 4,8,16,32 and $\theta = 0°, 45°, 90° and 135°$ (hence there were a total of 32 features) and combinations of different frequencies and orientations were used in classification. In Table 1, the results from the weighted Euclidean distance (WED) are tabulated and in Table2 the results from the nearest neighbour classifier (KNN) are given. The accuracy of both classifiers is compared, in Figure 5 (1).

Table 1: The Identification accuracy of the Gabor filtering technique under WED

Features	***All***	***SD***	***Mean***	***Mean at f=16,32***	***All at f=32***	***All at f=16***	***All at f=8***	***All at f=4***
Set A	84.1	82.8	83.4	92.1	85.4	61.3	34.4	29.1
Set B	***96.0***	90.1	92.1	***93.0***	84.2	65.3	34.7	22.8

Table 2: The Identification accuracy of the Gabor filtering technique under K-NN

Features	***All***	***SD***	***Mean***	***Mean at f=16,32***	***All at f=32***	***All at f=16***	***All at f=8***	***All at f=4***
Set A	54.7	59.3	40.0	56.7	66.7	66.7	37.6	21.1
Set B	77.0	***78.0***	60.0	75.0	***90.0***	88.0	47.0	27.7

The horizontal axis (feature sets) in Figure 5 represents the order of the features in Table 1-2. Similar number of features were used for both classifiers. For the WED classifier higher identification rate were, in general, observed (particularly for Set B). For example, a classification rate of 96.0% was obtained when all the features were used. The result of (93.0%) was observed when f=16 and 32 were chosen. Under the K-NN classifier a classification as high as 78.0% was achieved, when only the standard deviation (SD) was used (for Set B). The best result (90.0%) under the KNN was achieved when the frequency of f=32 was used.

3.2 Results from GSCM

In Tables 3-4, features were extracted using distances at d=1,2,3,4,5 and directions (there were a total of 60 features). There were also different combinations of feature sets, e.g. features at d= 1,2,3 and at the four directions given above (i.e. there were a total of 36 features) etc. Here the results can be seen to be relatively less accurate than those obtained with the Gabor filtering method, especially those under the WED classifier. This observation is consistent with the findings in [6]. Figure 5 (2) shows the plots of the identification accuracies for the GSCM features under both classifiers.

The features sets are in the same order as in Table 3-4. The best results (in the shaded box) show that only 82.2% of the images were identified correctly when using the WED classifier. In this case, 36 texture features were required.

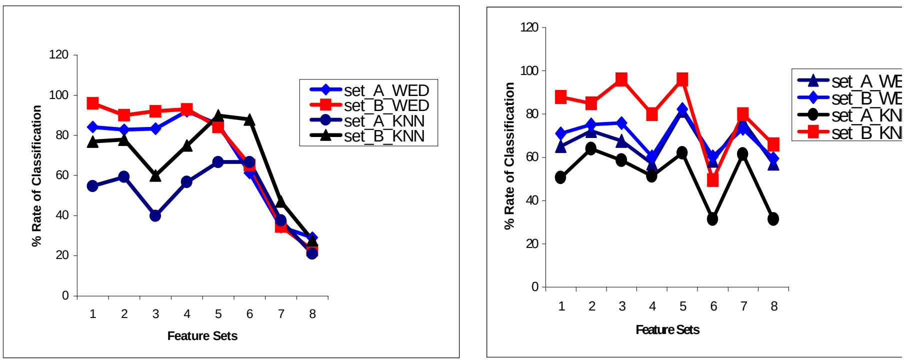

Figure 5: Results from Gabor Filter and GSCM Using the WED and K-NN Classifiers

Table 3: The Identification accuracy of the GSCM technique under WED

Features	***d=1,2, 3,4,5***	***d=1,2, 3***	***d=2,3, 4***	***d=3,4,5***	***d= 1,2***	***d=4,5***	***D=1***	***d=4***
Set A	65.1	72.2	67.6	57.0	81.5	58.3	76.8	57.0
Set B	71.0	***75.2***	75.8	60.4	***82.2***	60.4	73.3	59.4

Table 4: The Identification accuracy of the GSCM technique under K-NN

Features	***d=1,2, 3,4,5***	***d=1,2, 3***	***d=2,3, 4***	***d=3,4,5***	***d= 1,2***	***d=4,5***	***d=1***	***d=4***
Set A	50.7	64.0	58.7	51.3	62.0	31.4	61.3	31.4
Set B	***88.0***	85.0	96.0	80.0	***96.0***	49.5	80.0	66.0

4 Conclusions

We have described a new approach towards writer identification based on non-uniformly skewed handwriting images. Most existing approaches make an implicit assumption that handwritten texts are fixed. The novel approach introduced in this paper eliminates such an assumption and allows the use of any text. The algorithm is based on the observation that the handwriting of different people is visually distinctive, and a global approach based on texture analysis has been adopted. The novel approach is therefore text or content independent.

A number of experiments have been conducted. The experiments use 10 different writer classes. Features were extracted from handwriting images using the multi-channel

Gabor filtering and the grey scale co-occurrence matrix (GSCM) technique. Identification was performed using two different classifiers (the weighted Euclidean distance (WED) and the K-nearest neighbour (K-NN) classifiers). The results achieved were very promising, and an identification accuracy as high as 96.0% was obtained. The two classifiers have shown good performance, but at some stages the KNN classifier's performance is relatively poor when compared to the WED classifier.

We are currently investigating ways of reducing the impact of such factors on the performance of the proposed global approach. We will also consider local approaches which seek writer specific features to improve the recognition accuracy. In the future, both global and local approaches will be integrated as one system for a better identification accuracy.

References

[1] Plamond and G. Lorette, "Automatic Signature Verification and Writer Identification-The State of Art", *Pattern Recognition*, 1989, Vol.22, No.2, pp107-131.

[2] W. Kuckuck, "Writer Recognition by Spectra Analysis", *Proc. Int. Conf. In Security Through Science Engineering*, 1980, West Berlin, Germany, pp.1-3.

[3] T. N. Tan, "Written Language Recognition Based On Texture Analysis", *Proc. IEEE ICIP'96*, Lausanne, Switzerland, Sept. 1996, Vol.2, pp.185-188.

[4] T. N. Tan, "Texture Edge Detection by Modelling Visual Cortical Channels", *Pattern Recognition*, 1995,Vol.28, No.9, pp.1283-1298.

[5] A. K. Jain, and S. Bhattacharjee, "Text Segmentation Using Gabor Filters for Automatic Document Processing", *Machine and Vision Applications*, 1992, pp.169-184.

[6] G. S. Peake and T. N. Tan, "Script and Language Identification from Document Images", *Proc. BMVC '97*, Essex, UK, Sept. 97, Vol. 2, pp.610-619.

[7] T. R. Reed and J. M. Hans Du Buf, A Review of Recent Texture Segmentation And Feature Extraction Techniques, *CVGIP: Image Understanding*, 1993, Vol.57, pp.359-372.

[8] R. Jain, R. Kasturi, B. G. Schunck, "***Machine Vision***", McGraw-Hill, Inc., 1995.

[9] H.S Baird, "The Skew Angle of Printed Documents", *Proceedings of Society of Photographic Scientific Engineering*, 1987,Vol. 40, pp. 21-24.

[10] T. Akiyama and N. Hagita, "Automatic Entry for Printed Documents", *Pattern Recognition*, 1990,Vol.23, pp.1,141-1,154.

[11] T. Pavlidis and J. Zhou, "Page Segmentation and classification", *Computer Vision Graphics and Image Processing*, 1995, Vol.26, pp. 2,107-2,123.

[12] W. Postal, "Detection of Linear Oblique Structures and Skew in Digitized Documents", *Proceeding of Eighth International Conference on Pattern Recognition*, 1986, pp. 464-468.

[13] A. L. Spitz, "Text Characterization by Connected Component Transformations", *Proc. of SPIE, Document Recognition* (1994), Vol. 2181, pp. 97-105