A Survey on Writer Identification Schemes

Sreeraj.M
Department of Computer Science
Cochin University of Science and
Technology
Cochin, India

Sumam Mary Idicula
Department of Computer Science
Cochin University of Science and
Technology
Cochin, India

ABSTRACT

This paper presents a survey of the literature on writer identification schemes and techniques up till date. The paper outlines an overview of the writer identification schemes mainly in Chinese, English, Arabic and Persian languages. Taxonomy of different features adopted for online and offline writer identification schemes is also drawn at. The feature extraction methods adopted for the schemes are discussed in length outlining the merits and demerits of the same. In automated writer identification, text independent and text dependent methods are available which is also discussed in this paper. An evaluation of writer identification schemes under multiple languages is also analyzed by comparing the recognition rate.

General Terms

Pattern Recognition, Writer identification

Keywords

Feature extraction, online and offline schemes, text independent, text dependent, Writer identification.

1. INTRODUCTION

The growth of artificial intelligence and pattern recognition fields owes greatly to one of the highly challenged problem of handwriting identification. Identifying the handwriting of a writer is highly essential today due to the immense growth in technology and its applications in wide areas. The application of writer identification is in wide areas, such as, digital rights management in the financial sphere, to solve the expert problems in criminology by forensic expert decision-making systems, where a narrowed-down list of identified writers provided by the writer identification system. By combining with the writer verification as an authentication system this can be used to monitor and regulate the access to certain confidential sites or data where large amounts of documents, forms, notes and meeting minutes are constantly being processed and managed, knowing the identity of the writer would provide an additional value. It can also be used for historical document analysis [1], handwriting recognition system enhancement [2] and hand held and mobile devices [3]. To a certain extent its recent development and performance consider as a strong physiologic modalities of identification, such as DNA and fingerprints [4].

It is evident that the importance of writer identification has become more significant in these days. Obviously, the number of researchers involved in this challenging problem is going high as a result of these opportunities. There are numerous languages throughout the world. Each language poses a different threat to the writer identification problem depending on the characteristics of the language. So it is very clear that the identification problem varies across multiple languages.

The handwriting-based writer identification is an active research arena. As it is one of the most difficult problems encountered in the field of computer vision and pattern recognition, the handwriting-based writer identification problem faces with a number of sub problems like a) designing algorithms to identify handwritings of different individuals b) identifying relevant features of the handwriting c) basic methods for representing the features d) identifying complex features from the basic features developed and d) evaluating the performance of automatic methods.

The rest of the paper is organized as follows. The state of art in writer identification in languages like Chinese, English, Arabic and Persian is presented in detail. Also a taxonomy for online and offline writer identification depending on features is depicted. Also the performance evaluation of various writer identification schemes across multiple languages is tabulated. The paper is concluded with a discussion of the same.

2. WRITER IDENTIFICATION – CURRENT STATE OF ART

Until 1989 a comprehensive review of automatic writer identification had been given in [5]. As an extension the work from 1989 -1993 has been published in [6]. Consequently the approaches proposed in the last several years renewed the interests in this scientific community for the research topic. The following Figure 1 describes the standard framework of writer identification [7]. The necessary features from the handwritten documents are extracted as the first step. Later the features extracted are used to classify to which writer the document belongs using similarity score method. The document is classified as belonging to a writer with high similarity score.

Based on the input method of writing, automated writer identification has classifieds into on-line and off-line. The on-line writer identification task is considered to be less difficult than the offline one as it contains more information about the writing style of a person, such as speed, angle or pressure, which is not available in the off-line one [8][9]. Based on the different features associated with the writing, such as character, word, line, paragraph and the document, this has classified. The Figure 2 shows the taxonomy of the classification mentioned.

Text-dependent & text-independent are the other classification of automated writer identification. Dependent on the text content, text-dependent methods only matches the same characters and requires the writer to write the same text consequently. The text-independent methods are able to identify writers independent of the text content and it does not

require comparison of same characters. Thus it is very similar to signature verification techniques and uses the comparison between individual characters or words of known semantic (ASCII) content. This method considers as the global style of hand writing text as the metric for comparison, and also got better identification results. As it requires the same writing

content this method is not apt for many practical situations. Even though it got a wider applicability, text-independent methods do not obtain the same high accuracy as text-dependent methods do.

The following section describes the various approaches addressed for writer identification in different languages.
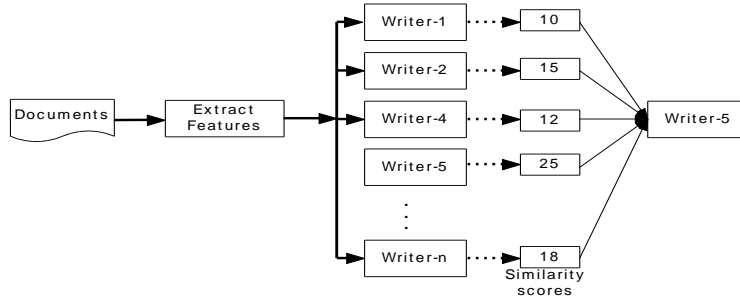
Fig 1: Writer Identification framework
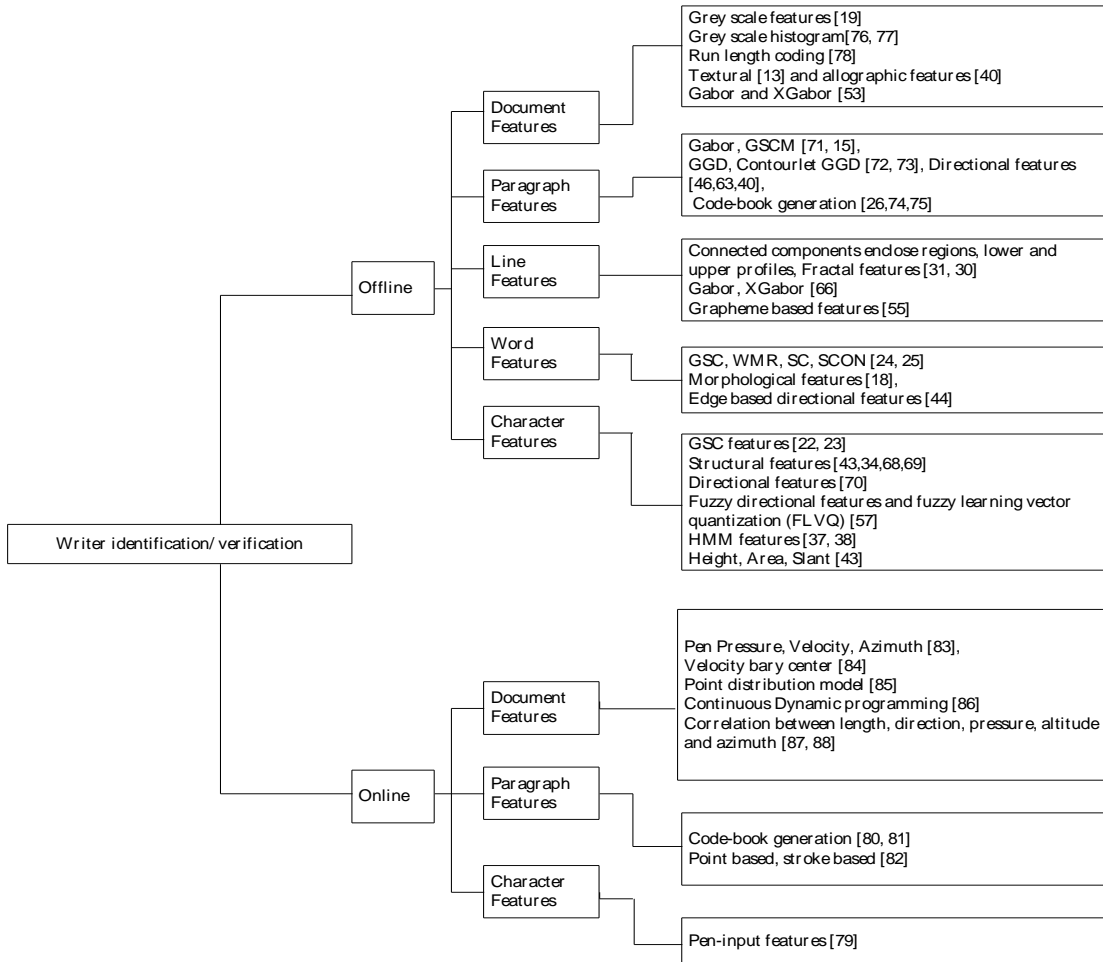
Fig 2: Taxonomy of Writer Identification

2.1 Chinese, English and other languages

In the end nineties, Said et al. [14] [15] proposed a text-independent approach for writer identification that derives writer-specific texture features using multichannel Gabor filtering and Gray-Scale Co-occurrence Matrices. The framework required uniform blocks of text that are generated by word deskewing, and also setting a predefined distance between text lines/words and text padding. Two sets of twenty writers and 25 samples per writer were used in the experimentation. Nearest centroid classification using weighted Euclidean distance and Gabor features achieved 96 percent writer identification accuracy, thus revealing that the two-dimensional Gabor model outperformed gray-scale co-occurrence matrix. A similar approach has also been used on machine print documents for script [16] and font [17] identification.

Zois and Anastassopoulos [18] implemented writer identification in 2000 and verified using single words. Experiments were performed on a data set of 50 writers. The word "characteristic" was written 45 times by each writer, both in English and in Greek. After image thresholding and curve thinning, the horizontal projection profiles were resampled, divided into 10 segments, and processed using morphological operators at two scales to obtain 20-dimensional feature vectors. Classification was performed using either a Bayesian classifier or a multilayer perceptron. The system showed an accuracy of 95% for both English and Greek words. In the writer identification scheme suggested by Marti et al. [30] and Hertel and Bunke [31], text lines was the basic input unit from which text-independent features are computed using the height of the three main writing zones, slant and character width, the distances between connected components, the blobs enclosed inside ink loops, the upper/lower contours, and the thinned trace processed using dilation operations. Using a k-nearest-neighbour classifier, identification rates exceeded 92 percent in test cases on a subset of the IAM database [33] with fifty writers and five handwritten pages per writer.

Graham Leeham et al. proposed a methodology to identify the writer of numerals [43]. The features included parameters such as height, width, area, center of gravity, slant, number of loops, etc. The system was tested among fifteen people and the accuracy was 95%. However to determine the precise accuracy it should be verified across larger databases. Srihari et al. [19], [20] proposed a large number of features for the writing which can be classified into two categories. a) Macrofeatures – They operate at document/paragraph/word level. The parameters used are gray-level entropy and threshold, number of ink pixels, number of interior/exterior contours, number of four-direction slope components, average height/slant, paragraph aspect ratio and indentation, word length, and upper/lower zone ratio. b) Microfeatures – They operate at word/character level. The parameters comprise of gradient, structural, and concavity (GSC) attributes. These features were used originally for handwritten digit recognition in [21]. Text-dependent statistical evaluations were performed on a data set containing thousand writers who copied a fixed text of 156 words (the CEDAR letter) three times. This is the largest data set ever used until now in writer identification methodologies. Microfeatures outperform macrofeatures in identification tests with an accuracy exceeding 80 percent. A multilayer perceptron or parametric distributions were used for writer verification with an accuracy of about 96 percent. Writer discrimination was also done using individual characters in [22],[23] and using words in [24], [25].

Bensefia et al. [26], [27], [28], [29] use graphemes generated by a handwriting segmentation method to encode the individual characteristics of handwriting independent of the text content. Grapheme clustering was used to define a feature space common for all documents in the data set. Experimentations were done on three data sets containing 88 writers, 39 writers (historical documents), and 150 writers, with two samples (text blocks) per writer. Writer identification was performed in an information retrieval framework, while writer verification was based on the mutual information between the grapheme distributions in the two handwritings which were used for comparison. Concatenations of graphemes are also analyzed in the mentioned papers. An accuracy of about 90 percent was reported on the different test data sets. A feature selection study is also performed in [32].

In [26, 27] Ameur Bensefia et al. have developed a probability based approach using a codebook of graphemes in the IAM and PSI databases. The system accuracy was 95% in IAM database and 86% in PSI database. Also, Laurens van der Maaten et al. have used a combination of simple directional features and codebook of graphemes [41]. The method was tested on 150 writers and the system accuracy was 97%. Vladimir Pervouchine et al. only focused on letters ''t'' and ''h'' on their English identification system. After detecting these shapes in the image, their skeletons were extracted. A cost function along the curve is then calculated and the similarity of cost functions identifies the writer [42]. It is obvious that this method cannot be extended for other languages. Schomaker et al. has presented a method based on fragmented connected-component contours (FCO3) [35, 36]. They used the method in the classification phase to calculate distance. Also, they tested it in an English data set with 150 writers. The top-1 of the method results had 72% and the top-10 had 93% accuracy. However, the top-10 results were satisfactory but its top-1 is not.

Schlapbach et al. implemented an HMM based writer identification and verification method [37, 38]. An individual HMM was designed and trained for each writer's handwriting. To determine which writer has written an unknown text, the text is given to all the HMMs. The one with biggest result is assumed to be the writer. The identification method was tested by using documents gathered from 650 writers. The identification accuracy was 97%. Also, this method was tested as a writer verification method. This was achieved by a collections writings from 100 people and twenty unskilled and twenty skilled imposters, who forged the originals. Experimentations results obtained showed about 96% overall accuracy in verification. Thus it is obvious, that this method can be extended to other languages by applying some changes on feature extraction phase. The difference between the two writer identifications schemes in [39] and [40] is that the former was used in English handwriting and got about 80% accuracy in top-1 results and about 92% in top-10 results while the latter supported Arabic handwriting and its accuracy was 88% in top-1 and 99% in top-10 results.

In 2007, Vladimir Pervouchine et al. [34] implemented a writer identification scheme based on high frequent characters. In this method, the high frequent characters ('f','d','y','th') are first identified, and then according to the similarity of those

characters, the writer is selected. The similarity is calculated with respect to the features (such as height, width, slant, etc.) associated with the characters. The number of features associated with each character is different (e.g. 'f' had 7 features while 'th' had 10 ones). A simple Manhattan distance was used in the classification phase. In order to select the best subset of the features, a GA was used which evaluated about 5000 of the subsets, out of 231 possible subsets. The system was tested in a database with 165 writers (between 15 to 30 patterns per writer), and the system accuracy was more than 95%. However, this method is simple and has good results, but the main concern of this method is that if a writer knows the procedure of method, he/she can write a text in test phase such that its characters are totally different with trained ones and so that the system cannot identify him/her.

A major contribution by Bangy Li et al. [10], again in 2007, used the feature vector of hierarchical structure in shape primitives along with the dynamic and static feature for writer identification for 242 writers using NLPR online database and attained a result of above 90% for Chinese and about 93% for English. The substantiation given is that English text contains more oriental information than Chinese text. In 2008, Zhenyu He et al. suggested an offline Chinese writer identification scheme which used Gabor filter to extract features from the text. They also incorporated a Hidden Markov Tree (HMT) in wavelet domain. The system was tested against a database containing 1000 documents written by 500 writers. Each sample contained 64 Chinese characters. The top-1, top-15, and top-30 results had 40%, 82.4%, and 100% accuracy, respectively [12]. Also, these authors have used a combination of general Gaussian model (GGD) and wavelet transform on Chinese handwriting in Ref. [13]. They tested the method on a database gathered from 500 people. This database consisted of 2 handwriting images per person. In the experiments, top-1, top-15 and top-30 results had 39.2%, 84.8% and 100% accuracy, respectively. As, the authors reported, the accuracy of proposed methods was low especially in top-1.

In 2009, YuChenYan et al. [11] presented spectral feature extraction method based on Fast Fourier Transformation which was tested on the 200 Chinese handwriting text collected from 100 writers. The methodology showed 98% accuracy for top 10 and 64% for top1 using the Euclidean and WED classifiers. This scheme has the advantage of stable feature and also it reduces the randomness in Chinese character. Another advantage is that it is feasible for large volume of dataset. However it has higher computation costs.

2.2 Arabic

Bulacu et al. proposed text-independent Arabic writer identification by combining some textural and allographic features [40, 45]. After extracting textural features (mostly relations between different angles in each written pixel) a probability distribution function was generated and the nearest neighborhood classifier using the as a distance measure was used. For the allographic features, a codebook of 400 allographs was generated from the handwritings of 61 writers and the similarity of these allographs was used as another feature. The database in experiments consisted of 350 writers with 5 samples per writer (each sample consisted of 2 lines (about 9 words)). The best accuracy seen in experiments was 88% in top-1 and

99% in top-10. Also, a simpler definition of this method was presented by M. Bulacu et al. earlier in [46].

Also, Ayman Al-Dmour et al. designed an Arabic writer identification system in 2007 [47]. Different feature extraction methods such as hybrid spectral-statistical measures (SSMs), multiple-channel (Gabor) filters, and the grey-level co¬occurrence matrix (GLCM) were verified to find the best subset of features. For the same purpose a support vector machine (SVM) was used to rank the features and then a GA (whose fitness function was a linear discriminant classifier (LDC)) chose the best one. Several classification methods such as LDC, SVM, weighted Euclidean distance (WED), and the K nearest neighbors (KNN) were also considered. The KNN-5, WED, SVM, and LDC results after feature selection per sub-images were reported as 57.0%, 47.0%, 69.0% and 90.0%, respectively. The results were better when the whole image was used, for instance the LDC result was increased to 100% (with no rotation). The database tested was gathered from 20 writers; each writer was asked to copy 2 A4 documents, one for training and the other one for testing. The used documents for each writer were different from the others and the sub-images were generated by dividing each document into 3x3 = 9 non-overlapping images. However, this method has good accuracy when LDC was used, but it seems the test database and samples per writer was small and it needs to be tested on more popular dataset. Faddaoui and Hamrouni opted for a set of 16 Gabor filters [48] for handwriting texture analysis. Gazzah and Ben Amara applied spatial-temporal textural analysis in the form of lifting scheme wavelet transforms. Angular features were considered as well in the task of Arabic writer identification [49].

Somaya Al-Ma'adeed et al. presented a text-dependent writer identification method in Arabic using only 16 words [44]. The features extracted include some edge-based directional features such as height, area, length, and three edge-direction distributions with different sizes and WED has been used as classifier. The test data was 32 000 Arabic text images from 100 people; the system was trained with 75% of the data and tested it by using 25%. They did not mention the top-1 accuracy of the method, but the best result in top-10 was 90% when 3 words were used. The main concern of this method is its dependency to text and the small dataset that were used in experiments. This method employed edge-based directional probability distributions, combined with moment invariants and structural word features, such as area, length, height, length from baseline to upper edge and length from base line to lower edge. On the other hand, Abdi et al. used stroke measurements of Arabic words, such as length, ratio and curvature, in the form of PDFs and cross-correlation transform of features [50] for the writer identification scheme.

Although, Arabic language is similar to Persian in character set and some writing styles, the Arabic methods cannot be extended to Persian language completely because of some special symbols that exists in Arabic language.

2.3 Persian

In 2006, Shahabi et al. proposed a Gabor based system for Persian writer identification and the accuracy of their work was reported about 92% in top-3 and 88% in top-1[51]. It is observed that the testing was not adequate; because in the test phase, there

was only one page per person such that 34 of it were used in training and the rest of page used in test phase. On retesting it, the method accuracy was of 60% in 80 people. In another scheme, Soleymani Baghshah et al. designed a fuzzy approach for Persian writer identification [57]. In this proposal, fuzzy directional features were used and the fuzzy learning vector quantization (FLVQ) has been trained in order to recognize the writers. The drawback of this method is that it only works on disjoint Persian characters that are not conventional in Persian language. This system was tested using 128 writers and results were around 90%–95% in different situations of test.

A Persian handwritten identification system that was based on a new generation of Gabor filter that was called XGabor filter is proposed in 2008 [52]. Feature extraction was done using Gabor and XGabor filters; in the classification phase, weighted Euclidian distance (WED) classifier was used. The proposed method in [52] got 77% accuracy using the PD100. Rafiee and Motavalli [58] introduced a new Persian writer identification method, using baseline and width structural features, and relying on a feed forward neural network for the classification.

In another recent work proposed an LCS (longest common subsequence) based classifier to classify features that are extracted by Gabor and XGabor filters [53,54]. This classifier improved the system accuracy up to 95% on PD100. However, the features extracted by XGabor filter could model the characteristic of written documents but the accuracy of these methods was not proper because of problems in data classification and representation. Therefore, in the present paper, XGabor filter was used together with Gabor filter with different data representation, classification, and identification schemes. In another research, a mixture of some different methods has been used by Sadeghi ram et al. Grapheme based features are clustered by fuzzy clustering method and after selecting some clusters, final decision is made based on gradient features. The scheme got about 90% accuracy in average on 50 people that were selected randomly from PD100 [55].They also used a three layer MLP(multi layer perceptron) to classify the gradient based features, and they got about 94% average accuracy on same data set [56]. To the best of our knowledge, there is no other reported method in Persian writer identification.

Table 1 summarizes the Writer Identification Methods on Multiple Languages. A graphical plot in Figure 3 compares the performance evaluation of different writer identification schemes across multiple languages.

Table 1. Writer Identification Methods on Multiple Languages

System	Sample Space	Features	Classification Methodology	Accuracy	Language
Text -dependent					
Srihari et al.s [19, 59]	1000 writers (CEDAR letter / paragraph / word)	Two levels of features; one at the macro level, micro level.	multi-layer perceptron	98%	English
Zois et al[18]	50 writers (45 samples of the same word)	The horizontal projection profiles are resampled into 10 segments, and processed using morphological operators	Bayesian classifiers and neural networks	95% for both English and Greek	English and Greek
Tomai et al. [25]	1000 writers	Character level, Word level features	Euclidean distances	99%	English
Zuo et al. [60]	40 writers	Offline PCA based method	Squared Euclidean distance	97.5%	Chinese
Zhang et al. [22]	*1000 writers*	Gradient (192 bits),structural (192 bits), and concavity (128 bits) features	k-nearest neighbor classification	97.71%	English
Somaya Al-Ma'adeed et	*100 writers(320 words(16diff*	Height area, length and Edge –direction	WED classifier	Top-10: 90%	Arabic

al.[44]	*erent types))*	distribution			
Schlapbach et al.[8]	200 writers(8 paragraph of about 8 lines)	Point-based (speed, acceleration, vicinity linearity, vicinity slope), stroke-based (duration, time to next stroke, number of points, number of up strokes, etc.),	Gaussian mixture model (GMM)	98.5%	English
Text-independent					
Pitak et al. [61]	81 writers	velocities of the barycenter of the pen movements	Fourier transformation approach	98.5%	Thai
Schlapbach et al. [89].	100 writers	X-Y coordinates	Hidden Markov Models	96%	English
Said et al. 15], T. Tan [16], Y. Zhu [17]	Two sets of 20 writers, 25 samples per writer (Few lines of handwritten text)	texture features using multichannel Gabor filtering and gray-scale co-occurrence matrices	Nearest centroid classification using weighted Euclidean distance	96%	English
Bensefia et al. [26], [27], [28], [29]	88 writers (French), 150 writers (English)	A textual based Information Retrieval model, local features such as graphemes extracted from the segmentation of cursive handwriting	Cosine similarity	95% on 88 writers 86% on 150 writers	French/English
S. K. Chan [62]	82 writers	x-y coordinates, direction, curvature of x-coordinates and the status of pen up or pen down.	Discrete Character prototype distribution approach (Euclidean distance)	*95%*	*Ferench*
Marti et al. [30] and Hertel and Bunke [31]	20 writers (5 samples of the same text)	Height of the three main writing zones, the distances between connected components	a k-nearest neighbor and a feed forward neural network classifier	90%	English
M. Bulacu [46],[63],[64], [65]	650 writers	Edge based directional PDFs as features (Textural and allograph prototype approach)	k-nearest neighbor and a feed forward neural network classifier	92%	English
Guo Xian Tan	120 writers	Continuous Character prototype distribution	Minimum distance	99%	French

Christian[66]		approach	classifier		
Neils et al.[67]	43 writers	Allograph prototype matching approach using the dynamic time warping (DTW) distance function	*af-iwf (allograph frequency – inverse writer frequency) measure*	60%	English
B. Helli, et al.[45], [53], [54]	100 writers (PD100 dataset), 50 writers[46]	Point-based (speed, acceleration, vicinity linearity, vicinity slope), stroke-based (duration, time to next stroke, number of points, number of up strokes, etc.).	Tey proposed an LCS (longest common subsequence) based classifier	95%	Persian
Bangy Li et al. [10]	242 writers*(NLPR online handwriting Database and 50 Chinese and English words in one page)*	Hierarchical Structure in Shape Primitives + Fusion Dynamic and Static Features	nearest neighbor classifier	Chinese accuracy>90% English accuracy>93%	English and Chinese text
YuChen Yan et al. [11]	200 handwritings from 100 writers	Spectral feature based on Fast Fourier Transformation	Euclidean and WED classifiers	98% -top 10 64%-top1	Chinese
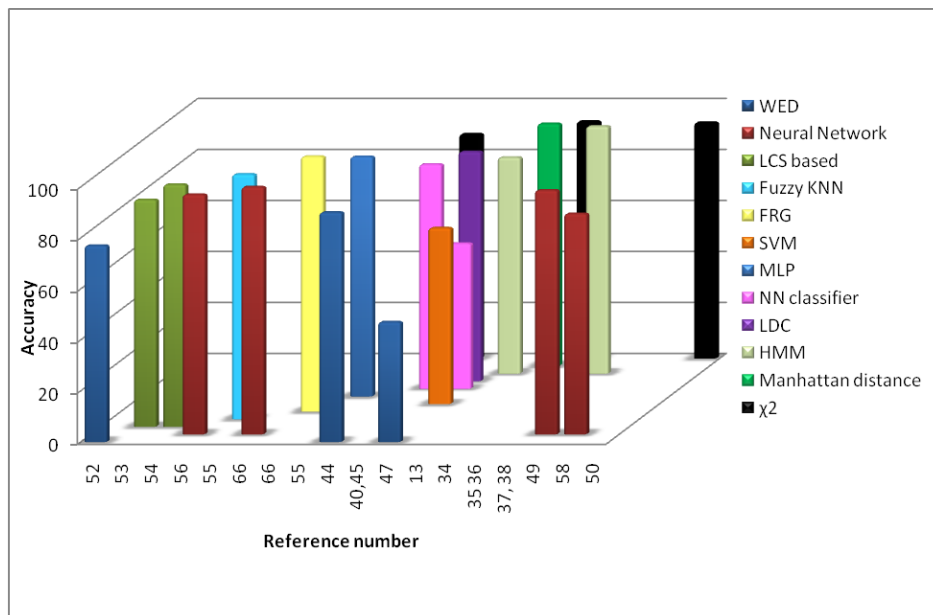
Fig 3: comparative evaluation of writer identification schemes

3. CONCLUSION

From this state of art we can see that a wide variety of features are used for writer identification. In Chinese language, writer-specific texture features using multichannel Gabor filtering and Gray-Scale Co-occurrence Matrices are common, but in English it varies from micro level features to macro level and edge distribution. Also the studies are carried out in other languages such as Arabic as well as in Persian languages. Combinations of some textural and allographic features as well as hybrid spectral-statistical measures (SSMs), multiple-channel (Gabor) filters, XGabor are carried out to obtain the individuality of the writers. Also studies are based on the features used in other languages but it could achieve only less accuracy. From this we understand that features must be selected based on the characteristic features of each language.

From the discussion of text-dependent and text-independent methods, we can conclude that in general, higher identification rates are achievable with the former type. Text- independent methods are much more useful and applicable. These methods, however, require a certain minimum amount of text to produce acceptable results. Resuming, we could say that the research on writer recognition that started with the analysis of very constrained writings and very few writers has matured really well over time. Regarding the methods developed, in addition to the structural and statistical features, codebook generation has emerged as a very popular as well as effective method for writer identification. These codebooks could be computed universally for the entire set of writers or for each of the writers separately. The methods based on a universal codebook are generally efficient in terms of computational cost, however, a new codebook is to be generated if the script changes. On the other hand, writer specific codebooks have high computational costs but they could present a generic framework independent of the alphabet under study. In the writer identification methods discussed here the features are independent of the textual content of each language.

We have been through and seem to be continuing in a rather extra-ordinary phase of experiments. Looking at the number of people, who is trying to make the most result out of these experiments, hopefully on its last legs? But of course there has to be a conscious collective efforts need to succeed on that.

4. REFERENCES

[1] Fornes, A., Llados, J., Sanchez, G., Bunke, H. (2008).Writer Identification in Old Handwritten Music Scores. In: 8th IAPR Workshop on Document Analysis Systems, 347—353.

[2] Sas, J. (2006) Handwriting Recognition Accuracy Improvement by Author Identification. In: L. Rutkowski et al. (eds.), ICAISC 2006, LNAI 4029, 682--691.Springer, Heidelberg.

[3] Chaudhry, R., Pant, S. K. (2004) Identification of authorship using lateral palm print—a new concept. *J.ForensicScience International*, volume (141), 49--57.

[4] Schomaker, L. (2007) Advances in Writer Identification and Verification. In: 9th International Conference on Document Analysis and Recognition (ICDAR'07), volume (2), 1268--1273.

[5] Plamondon, R., Lorette, G. (1989) "Automatic Signature Verification and Writer Identification—The State of the Art," Pattern Recognition, vol. 22, no. 2, pp. 107-131.

[6] Leclerc, F., Plamondon, R. (1994) Automatic signature verification: The state of the art 1989-1993. In Progress in Automatic Signature Verification edited by R. Plamandon, World Scientific Publ. Co., pp. 13-19.

[7] Gupta, S. (2008).Automatic Person Identification and Verification using Online Handwriting. Master Thesis. International Institute of Information Technology Hyderabad, India

[8] Schlapbach, A., Marcus, L. Bunke, H. (2008) A writer identification system for on-line whiteboard data, Pattern Recognition Journal 41 (2008) 23821–23897.

[9] Schomaker, L. (2007) Advances in Writer identification and verification, in: Ninth International Conference on Document Analysis and Recognition (ICDAR).

[10] Li, B., Sun, Z., Tan, T.N. (2007) Hierarchical Shape Primitive Features for Online Text-independent Writer Identification, *Proc. of 2th ICB*, pages 201–210.

[11] Yan, Y., Chen, Q., Deng, W., Yuan, F. (2009) "Chinese Handwriting Identification Based on Stable Spectral Feature of Texture Images" *International Journal of Intelligent Engineering and Systems*, Vol.2, No.1.

[12] He, Z., You, X., Tang, Y.Y. (2008) Writer identification of Chinese handwriting documents using hidden Markov tree model, Pattern Recognition Journal 41, 1295–13072008-06-15.

[13] Zhenyu, H., Xinge, Y., Tang, Y.Y. (2008)Writer Identification using global wavelet-based features, Neurocomputing 71, 1832–1841.

[14] Said, H., Tan, T., Baker, K. (2000) "Personal Identification Based on Handwriting," Pattern Recognition, vol. 33, no. 1, pp. 149-160.

[15] Said, H., Peake, G., Tan, T., Baker, K. (1998) "Writer Identification from Non-Uniformly Skewed Handwriting Images," Proc. Ninth British Machine Vision Conf., pp. 478-487.

[16] Tan, T. (1998) "Rotation Invariant Texture Features and Their Use in Automatic Script Identification," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 7, pp. 751-756.

[17] Zhu, Y., Tan, T., Wang, Y. (2001) "Font Recognition Based on Global Texture Analysis," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 23, no. 10, pp. 1192-1200.

[18] Zois, E., Anastassopoulos, V. (2000) "Morphological Waveform Coding for Writer Identification," Pattern Recognition, vol. 33, no. 3, pp. 385-398.

[19] Srihari, S., Cha, S., Arora, H., Lee, S. (2002) "Individuality of Handwriting," J. Forensic Sciences, vol. 47, no. 4, pp. 1-17.

[20] Srihari, S., Beal, M., Bandi, K., Shah, V., Krishnamurthy, P. (2005) "A Statistical Model for Writer Verification," Proc. Eighth Int'l Conf. Document Analysis and Recognition (ICDAR), pp. 1105-1109.

[21] Favata, J., Srikantan, G. (1996). "A Multiple Feature/Resolution Approach to Hand printed Digit and Character Recognition," Int'l J. Imaging Systems and Technology, vol. 7, pp. 304-311.

[22] Zhang, B., Srihari, S., Lee, S. (2003) "Individuality of Handwritten Characters," Proc. Seventh Int'l Conf. Document Analysis and Recognition (ICDAR), pp. 1086-1090.

[23] Srihari, S., Tomai, C., Zhang, B., Lee, S. (2003) "Individuality of Numerals," Proc. Seventh Int'l Conf. Document Analysis and Recognition (ICDAR), pp. 1096-1100.

[24] Zhang, B., Srihari, S. (2003) "Analysis of Handwritten Individuality Using Word Features," Proc. Seventh Int'l Conf. Document Analysis and Recognition (ICDAR), pp. 1142-1146.

[25] Tomai, C., Zhang, B., Srihari, S. (2004) "Discriminatory Power of Handwritten Words for Writer Recognition," Proc. 17th Int'l Conf. Pattern Recognition, pp. 638-641.

[26] Bensefia, A., Paquet, T., Heutte, L. (2005) "A Writer Identification and Verification System," Pattern Recognition Letters, vol. 26, no. 10, pp. 2080-2092.

[27] Bensefia, A., Paquet, T., Heutte, L. (2005) Handwritten Document Analysis for Automatic Writer Recognition," Electronic Letters on Computer Vision and Image Analysis, vol. 5, no. 2, pp. 72-86.

[28] Bensefia, A., Nosary, A., Paquet, T., Heutte, L. (2002) "Writer Identification by Writer's Invariants," Proc. Eighth Int'l Workshop Frontiers in Handwriting Recognition, pp. 274-279.

[29] Bensefia, A., Paquet, T., Heutte, L. (2003) "Information Retrieval Based Writer Identification," Proc. Seventh Int'l Conf. Document Analysis and Recognition (ICDAR), pp. 946-950.

[30] Marti, U.V., Messerli, R., Bunke, H. (2001) "Writer Identification Using Text Line Based Features," Proc. Sixth Int'l Conf. Document Analysis and Recognition (ICDAR), pp. 101-105.

[31] Hertel, C., Bunke, H. (2003) "A Set of Novel Features for Writer Identification," Proc. Fourth Int'l Conf. Audio and Video-Based Biometric Person Authentication, pp. 679-687.

[32] Schlapbach, A., Kilchherr, V., Bunke, H. (2005) "Improving Writer Identification by Means of Feature Selection and Extraction," Proc. Eighth Int'l Conf. Document Analysis and Recognition (ICDAR), pp. 131-135.

[33] Marti, U.V., Bunke, H. (2002) "The IAM-Database: An English Sentence Database for Offline Handwriting Recognition," Int'l J. Document Analysis and Recognition, vol. 5, no. 1, pp. 39-46.

[34] Pervouchine, V., Leedham, G. (2007)Extraction and analysis of forensic document examiner features used for writer identification, Pattern Recognition Journal 40, 1004–1013.

[35] Schomaker, L., Franke, K., Bulacu, M. (2007) Using codebooks of fragmented connected-component contours in forensic and historic writer identification, Pattern RecognitionLetter28, 719–727.

[36] Schomaker, M.B.L. (2004) Analysis of texture and connected-component contours for the automatic identification of writers, in: 16th Belgium–Netherland Conference on Artificial Intelligence(BNAIC).

[37] Schlapbach, A., Bunke, H. (2007) A writer identification and verification system using HMM based recognizers, Pattern Analysis Application (Springer)10,33–43, doi:10.1007/s10044-006-0047-5.

[38] Schlapbach, A., Bunke, H. (2005) Writer identification using an HMM-based hand- writing recognition system: to normalize the input or not? In: 12th Conference of the International Graphonomics Society, Salerno, Italy, June 26–29, pp.138–142.

[39] Bulacu, M., Schomaker, L. (2006) Combining multiple features for text-independent writer identification and verification, in:10th international Workshop on Frontiers in Handwriting Recognition(IWFHR).

[40] Bulacu, M., Schomaker, L., Brink, A. (2007) Text-independent writer identification and verification on offline Arabic handwriting, in: Ninth Conference on Document Analysis and Recognition(ICDAR).

[41] Vander Maaten, L., Postma, E. (2005) Improving automatic writer identification, in:17thBelgium-Netherland Conference on Artificial Intelligence.

[42] Pervouchine, V., Leedham, G., Melikhov, K. (2005). Handwritten character skeletonisation for forensic document analysis, in: ACM Symposium on Applied Computing.

[43] Leeham, G., Chachra, S. (2003) Writer identification using innovative binarised features of handwriting numerals, in: Seventh International Conference on Document Analysis and Recognition (ICDAR).

[44] Al-Ma'adeed, S., Mohammed, E., AlKassis, D., Al-Muslih, F. (2008) Writer identification using edge-based directional probability distribution features for Arabic words, in: IEEE/ACS International Conference on Computer Systems and Applications (AICCSA), 582—590.

[45] Bulacu, M., Schomaker, L.(2007) Text-independent writer identification and verification using textural and allographic features, IEEE Transactions on Pattern Analysis and Machine Intelligence(PAMI)29(4) 701–717 Special Issue—Biometrics: Progress and Directions.

[46] Bulacu, M., Schomaker, L., Vuurpijl, L. (2003) Writer identification using edge-based directional features, in: Seventh International Conference on Document Analysis and Recognition (ICDAR).

[47] Al-Dmour, A., Zitar, R.A. (2007) Arabic writer identification based on hybrid spectral- statistical measures, Journal of Experimental & Theoretical Artificial Intelligence 19(4) 307–332.

[48] Feddaoui, N., Hamrouni, K. (2007). Personal Identification based on Texture Analysis of Arabic Handwriting Text.In: IEEE International Conference on Information and Communications Technologies (ICTTA'06), vol. (1), 1302-1307.

[49] Gazzah, S., Ben Amara, N. E. (2007). Arabic Handwriting Texture Analysis for Writer Identification using the DWT-lifting Scheme. In: 9th International Conference on Document Analysis and Recognition (ICDAR'07), vol. (2), 1133-1137.

[50] Abdi, M. N., Khemakhem, M., Ben-Abdallah, H. (2009) A Novel Approach for Off-Line Arabic Writer Identification Based on Stroke Feature Combination. In: 24th IEEE International Symposium on Computer and Information Sciences, (ISCIS'09).

[51] Shahabi, F., Rahmati, M. (2006) Comparison of Gabor-based features for writer identification of Farsi/Arabic handwriting, in: 10th International Workshop on Frontiers in Handwritten Recognition (IWFHR).

[52] Helli, B., Moghaddam, M.E. (2008) Persian writer identification using extended Gabor filter, in: International Conference on Image Analysis and Recognition (ICIAR).

[53] Helli, B., Moghaddam, M.E. (2008) A text-independent Persian writer identification system using LCS based classifier, in: IEEE International Symposium on Signal Processing and Information Technology, ISSPIT 2008.

[54] Helli, B., Moghaddam, M.E. (2009) A writer identification method based on XGabor and LCS, IEICE Electronics Express 6 (10).

[55] Ram, S.S., Moghaddam, M.E. (2009) Text-independent Persian writer identi- fication using fuzzy clustering approach, in: International Con- ference on Information Management and Engineering (ICIME), Malaysia.

[56] Ram, S.S., Moghaddam, M.E. (2009) A Persian writer identification method based on gradient features and neural networks, in: Second International Conference on Image and Signal Processing (CISP), China.

[57] Soleymani Baghshah, M., Bagheri Shouraki, S. Kasaei, S. (2006) A novel fuzzy classifier using fuzzy LVQ to recognize online Persian handwriting, in: Second IEEE Conference on Information and Communication Technology (ICTTA).

[58] Rafiee, A., Motavalli, H. (2007). Off-Line Writer Recognition for Farsi text. In: 6th Mexican International Conference on Artificial Intelligence (MICAI'07), Special Session, 193--197.

[59] Srihari, S. N., Cha, S.-H., Lee, S. (2001) "Establishing Handwriting Individuality Using Pattern Recognition Techniques," in Proceedings of the Sixth International Conference on Document Analysis and Recognition, pp. 1195-1204.

[60] Long Zuo, Yunhong Wang, Tieniu Tan, (2002).Personal Handwriting Identification Based on PCA, *Proceedings of SPIE Second International Conference on Image and Graphics*, pp 766-771.

[61] Pitak, T., Matsuura, T. (2004) On-line writer recognition for Thai based on velocity of barycenter of pen-point movement, in: International Conference on Image Processing, ICIP `04, vol. 2, pp. 889–892.

[62] Chan, S. K., Viard-Gaudin, C., Tay, Y. H. (2008)"Online writer identification using character prototypes distributions," inProceedings of SPIE - The International Society for Optical Engineering.

[63] Bulacu, M. Schomaker, L. (2003) "Writer Style from Oriented Edge Fragments," Proc. 10th Int'l Conf. Computer Analysis of Images and Patterns, pp. 460-469.

[64] Schomaker, L., Bulacu, M., Franke, K. (2004) "Automatic Writer Identification Using Fragmented Connected-Component Contours," Proc. Ninth Int'l Workshop Frontiers in Handwriting Recognition (IWFHR), pp. 185-190.

[65] Bulacu, M., Schomaker, L. (2005) "A Comparison of Clustering Methods for Writer Identification and Verification," Proc. Eighth Int'l Conf. Document Analysis and Recognition, vol. II, pp. 1275-1279.

[66] Guo Xian Tan Christian. Automatic Writer Identification Framework for Online Handwritten Documents Using Character Prototypes

[67] Niels, R., Gootjen, F. Vuurpijl, L. (2008) "Writer Identification through Information Retrieval: The Allograph Weight Vector," in International Conference on Frontiers in Handwriting Recognition, pp. 481-486.

[68] Sutanto, P.,Leedham, G., Pervouchine, V. (2003) "Study of the consistency of some discriminatory features used by document examiners in the analysis of handwritten letter 'a'," in *International Conference on Document Analysis and Recognition*, pp. 1091–1095.

[69] Pervouchine, V., Leedham, G. (2006) "Extraction and analysis of document examiner features from vector skeletons of grapheme 'th'," in *Document Analysis Systems*, pp. 196–207.

[70] Xianliang, X. D., Wang, Liu, H. (2003) "Writer identification using directional element features and linear transform," in *International Conference on Document Analysis and Recognition*.

[71] Tan, T. N. (1992) "Texture feature extraction via visual cortical channel modeling," in *International Conference of Pattern Recognition*, vol. 3, pp. 607–610.

[72] Zhenyu, T. Y., He, Xinge, Y. (2005).A contourlet-based method for writer identification, International *Conference on Systems, Man and Cybernetics*, vol. 1, pp. 364–368.

[73] He, J. Y. Z., Fang, B., You, X. (2005) "A novel method for off-line handwriting-based writer identification," in *International Conference on Document Analysis and Recognition*, pp. 242–256.

[74] Schomaker, L. Bulacu. (2004)Automatic writer identification using connected component contours and edge-based features of uppercase western script, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, pp. 787–798.

[75] Seropian, A., Grimaldi, M., Vincent, N. (2003) "Writer identification based on the fractal construction of a reference base," in *International Conference of Document Analysis and Recognition*, pp. 1163–1167.

[76] Arazi, B. (1977).Handwriting identification by means of run-length measurements, *IEEE Transactions of System, Man and Cybernatics*, vol. 7, pp. 878–881.

[77] Arazi, B. (1983) "Automatic handwriting identification based on the external properties of the samples," *IEEE Transactions of System, Man and Cybernatics*, vol. 13, pp. 635–642.

[78] Zimmerman, K., Varady, M. (1985) Hand writer identification from one-bit quantized pressure patterns, *Pattern Recognition*, vol. 18, no. 1, pp. 63–72.

[79] Nakamura, Y. Kidode, M. (2005) "Individuality analysis of online kanji handwriting," in *International Conference on Document Analysis and Recognition*.

[80] Namboodiri, A.M., Gupta, S. (2006) "Text independent writer identification from online handwriting,"in *Proceedings of 10th International Workshop on Frontiers in Handwriting Recognition*, (La Baule, Centre de Congreee Atlantia, France), pp. 23–26.

[81] Jin, W., Wang, Y., Tan, T. (2005) "Text-independent writer identification based on fusion of dynamic and static features," in *International Workshop Biometric Recognition Systems*, p. 197.

[82] Schlapbach, A. L., Marcus, Bunke, H. (2008).A writer identification system for on-line whiteboard data, *Pattern Recognition Journal*, 41, 23821–23897.

[83] Seiichiro Hangai, S. Y., Hamamoto, T. (2000) On-line signature verification based on altitude and direction of pen movement, vol. 1, pp. 489–492.

[84] Thumwarin, P., Matsuura, T. (2004) "On-line writer recognition for Thai based on velocity of barycenter of pen-point movement," in *IEEE International Conference on Image Processing*, vol. Singapore, pp. 889–892, October 24-27.

[85] Tsai, L. M. Y. (2005) "Online writer identification using the point distribution model," in *International* Conference on System, Man and Cybernetics, vol. 2, pp. 1264–1268.

[86] Hiroshi Kameya, S.M., Oka, R. (2003) "Figure-based writer verification by matching between an arbitrary part of registered sequence and an input sequence extracted from on-line handwritten figures," in *International Conference on Document Analysis and Recognition*,

[87] Chapran, J. (2006).Biometric writer identification: feature analysis and classification, International JournalofPatternRecognitionandArtificialIntelligence20 (4), 483–503.

[88] Chapran, J., Fairhurst, M.C.(2006) Biometric writer identification based on the interdependency between static and dynamic features of handwriting, in: Proceedings of the 10th International Workshop on Frontiers in Handwriting Recognition, pp. 505–510.

[89] Schlapbach, A., Bunke, H. (2004) Using HMM based recognizers for writer identification and verification, in: Proceedings–International Workshop on Frontiers in Handwriting Recognition, IWFHR, Tokyo, pp. 167–172.