Off-Line Handwritten Signature Retrieval

M. S. Shirdhonkar
Dept. of Computer Science & Engineering,
B.L.D.E.A's College of Engineering & Technology, Bijapur, India
+919900394341
ms_shirdhonkar@rediffmail.com

M. B. Kokare
Dept. of Electronics & Telecommunication
S.G.G.S Institute of Engineering & Technology, Nanded, India
+919850770829
mbkokare@yahoo.com

ABSTRACT
This paper presents a new method for off-line handwritten signature retrieval using rotated complex wavelet filters (RCWF) and dual tree complex wavelet transform (DT-CWT) jointly for extracting details in twelve different directions of the signature image. This paper addresses issue in the context of a database of handwritten signature images and describes a system for similarity retrieval. The performance of the system has been tested with a signature image database of 192 signatures. From experimental results, we achieved a recall of 73.95% when considering the top 12 results. The results indicate that the proposed system is able to retrieve signatures with high accuracy even when a part of a signature is missing.

Categories and Subject Descriptors
H.3.3 [**Information Storage and Retrieval**]:–Information search and retrieval.

General Terms
Algorithm, Performance, Experimentation.

Keywords
Handwritten recognition, Similarity retrieval, Content based image retrieval, Rotated complex wavelet filters.

1. INTRODUCTION
1.1 Motivation
A presence of signature on different types of documents such as bank cheques in daily life and credit slips etc. Hence signature has a great importance in a person's life. Automatic bank cheque processing is an active topic in the field of document analysis and processing. Signature validity confirmation of different document is one of the important problems in automatic document processing. Now a days, person identification and verification are very important in security and resource access control. For this purpose the first and simple way is to use Personal Identification Number (PIN), but PIN code may be forgotten. Now an interesting method to identification and verification is biometric approach [11]. Biometric is a measure for identification that is unique for each person. Always biometric is together with person and cannot be forgotten. In addition biometric usually cannot be misused.

Handwritten signature retrieval is still a challenging work in the situations of a large database. Unlike fingerprint, palm print and iris signatures have significant amount of intra class variations making the research even more compelling. This approach with the potential applications of signature recognition/verification system optimized with efficient signature retrieval mechanism.

1.2 Related Works
Signature verification contains two areas: off-line signature verification, where signed papers or documents are scanned into image representation and on-line signature verification, where signature samples are collected from a digitizing tablet which is capable of pen movements during the writing. In our work, we survey the offline signature identification and retrieval. In 2009, Ghandali and Moghaddam have proposed an off-line Persians signature identification and verification based on Image registration, DWT (Discrete Wavelet Transform) and fusion. They used DWT for feature extraction and Euclidean distance for comparing features. It is language dependent method [11]. In 2008, Larkins and Mayo have introduced a person dependent off-line signature verification method that is based on Adaptive Feature Threshold (AFT) [8]. AFT enhances the method of converting a simple feature of signature to binary feature vector to improve its representative similarity with training signatures. They have used combination of spatial pyramid and equimass sampling grids to improve representation of a signature based on gradient direction. In classification phase, they used DWT and graph matching methods. In another work, Ramachandra et al [10], have proposed cross-validation for graph matching based off-line signature verification (CSMOSV) algorithm in which graph matching compares signatures and the Euclidean distance measures the dissimilarity between signatures.

In 2007, Kovari et. al presented an approach for off-line signature verification, which was able to preserve and take usage of semantic information[7]. They used position and direction of endpoints in features extraction phase. Porwik [9] introduced a three stages method for offline signature recognition. In this approach the hough transform, center of gravity and horizontal-vertical signature histogram have been employed, using both static and dynamic features that were processed by DWT [12]. The verification phase of this method is based on fuzzy net using the enhanced version of the MDF(Modified Direction feature) extractor [1]. The different neural classifier such as Resilient Back Propagation (RBP), Neural network and Radial Basis Function (RBF) network have been used in verification phase of this method. In 1995, Han and Sethi [5], described offline signature retrieval and use a set of geometrical and topological features to map a signature onto 2D strings.

The main contribution of this paper is that, we have proposed off-line handwritten signature retrieval using RCWF and DT-CWT. In retrieval phase, Canberra distance measure is used for comparing features. The experimental results of proposed method were satisfactory and found that it had better results compare with earlier approach. The rest of paper is organized as follows: In section 2, we discuss the signature retrieval phase. In section 3, the experimental results have been discussed and finally section 4 concludes the work.

2. SIGNATURE RETRIEVAL PHASE

The steps involved in signature retrieval are as follows:

2.1 Preprocessing

For degraded and poor quality document, a preprocessing stage of the gray scale source I(x, y) is essential for the elimination of noisy areas and smoothing of background areas. In image processing, it is usually necessary to perform high degree of noise reduction in an image before performing higher level processing steps. The median filter is a filtering techniques often used to remove noise from image. This filtering procedure is used to examine a sample of the input.

2.2 Feature Extraction

The main objective is to extract the signature region from the original image and transform it into a new compact features vector that supports measuring the similarity among signatures for retrieval purposes. In signature retrieval, edge information is very important in characterizing signature properties. We proposed the use of DT-CWT and DT-RCWF jointly, which extracts details in twelve different directions [6].

2.2.1 DT-CWT and DT-RCWF

The real DWT has two main disadvantages, firstly it gives poor directionality, and secondly it is shift sensitivity. These problems of real wavelet can be overcome by using complex wavelet transform, which gives signature information in six different directions. The proposed method uses the complex wavelet, which can be defined as

$$\psi(x) = \psi_h(x) + j\,\psi_g(x) \tag{1}$$

where $\psi_h(x)$ and $\psi_g(x)$ are both real-valued wavelets. The scaling and directional complex wavelets are given by (2-8)

$$\phi_1(x,y) = \phi_h(x)\,\phi_h(y) \quad and \quad \phi_2(x,y) = \phi_g(x) \tag{2}$$

$$\psi_{1,1}(x,y) = \psi^{+15^\circ}(x,y) = \phi_h(x)\,\psi_h(y) \tag{3}$$

$$\psi_{2,1}(x,y) = \psi^{-15^\circ}(x,y) = \phi_g(x)\,\psi_g(y) \tag{4}$$

$$\psi_{1,2}(x,y) = \psi^{+75^\circ}(x,y) = \psi_h(x)\,\phi_h(y) \tag{5}$$

$$\psi_{2,2}(x,y) = \psi^{-75^\circ}(x,y) = \psi_g(x)\,\phi_g(y) \tag{6}$$

$$\psi_{1,3}(x,y) = \psi^{+45^\circ}(x,y) = \psi_h(x)\,\psi_h(y) \tag{7}$$

$$\psi_{2,3}(x,y) = \psi^{-45^\circ}(x,y) = \psi_g(x)\,\psi_g(y) \tag{8}$$

Form the above equations the six directional wavelets can be defined by (9 and 10)

$$\psi_i(x,y) = \frac{1}{\sqrt{2}}\left(\psi_{1,i}(x,y) + \psi_{2,i}(x,y)\right) \tag{9}$$

$$\psi_{i+3}(x,y) = \frac{1}{\sqrt{2}}\left(\psi_{1,i}(x,y) - \psi_{2,i}(x,y)\right) \tag{10}$$

where $1 \le i \le 3$. A wavelet transform based on these six wavelets can be implemented by taking sum and difference of two separable 2-D DWTs. The resulting directional wavelet transform is two-times redundant. We used new rotated complex wavelet filters that are obtained by rotating nonseparable wavelet filters obtained using equations (2-8) by 45° so that the decomposition is performed along the new direction, which are 45° apart from decomposition directions of CWT. The six subbands of the 2D DT-RCWF gives information strongly oriented at $\left\{-30^0,\ 0^0,\ 30^0,\ 60^0,\ 90^0,\ 120^0\right\}$. This characteristic of RCWF sets provides important complementary information to the CWT filter set in extracting signature features in twelve different directions by considering them jointly

2.2.2 Feature Database Creation

To conduct the experiments, we have computed features using combined DT-CWT and DT-RCWF. To construct the feature vectors of each signature in the database, we decomposed each signature using DT-CWT and DT-RCWF up to 6th level to obtain better retrieval performance. The energy and standard deviation were computed separately on each sub band and the feature vector was formed using these two parameter values. The Energy E_k and Standard deviation σ_k of kth sub band is computed as follows

$$E_k = \frac{1}{M \times N}\sum_{i=1}^{M}\sum_{j=1}^{N}\left|W_k(i,j)\right| \tag{11}$$

$$\sigma_k = \left[\frac{1}{M \times N}\sum_{i=1}^{N}\sum_{j=1}^{M}\left(W_k(i,j)-\mu_k\right)^2\right]^{\frac{1}{2}} \tag{12}$$

where $W_k(i,j)$ is the k^{th} wavelet-decomposed sub band, *MxN* is the size of wavelet decomposed sub band, and μ_k is the mean of the k^{th} sub band. The resulting feature vector using energy and standard deviation are

$\bar{f}_E = \left[E_1 \quad E_2 \quad \dots \quad E_n\right]$ and $\bar{f}_\sigma = \left[\sigma_1 \quad \sigma_2 \quad \dots \quad \sigma_n\right]$ respectively. So combined feature vector is

$$\bar{f}_{\sigma\mu} = \left[\sigma_1 \quad \sigma_2 \quad \dots \quad \sigma_n \quad E_1 \quad E_2 \quad \dots \quad E_n\right] \qquad (13)$$

2.2.3 *Signature Retrieval*

There are several ways to work out the distance between two points in multidimensional space. We have used Canberra distance metric as distance measure. If *x* and *y* are the feature vectors of the database and query signature, respectively, and have dimension d, then the Canberra distance is given by

$$\text{Canb}(x, y) = \sum_{i=1}^{d} \frac{|x_i - y_i|}{|x_i| + |y_i|} \qquad (14)$$

The step by step procedure of signature retrieval is discussed in Algorithm 1.

Algorithm 1: Handwritten Signature Retrieval

Input: Test signature: St
Feature database: FV
Output: Distance vector: Dist
Handwritten signature retrieval
Begin
Calculate feature vector of test signature using RCWF and DT-CWT
For each fv in FV **do**
Dist= Calculate distance between test signature and fv using (14)
Sort Distance Vector Dist.
End for
Display the top N signature from distance vector.
End

---

3. EXPERIMENTAL RESULTS

3.1 Image Database

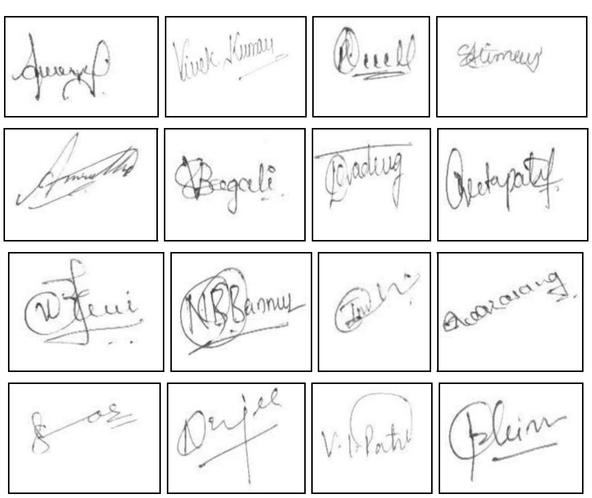

Figure 1. Sample Signature Images Database

The signatures were collected using either black or blue ink (No pen brands were taken into consideration), on a white A4 sheet of paper, with eight signature per page. A scanner subsequently digitized the eight signatures, contained on each page, with a resolution in 256 grey levels. Afterwards the signature images were created in rectangular shape of size 256x256 pixels. Sample signature database for 16 persons are shown in Fig. 1. A group of 16 persons were selected for 12 specimen signatures which make the total of 16x12=192 signature database.

3.2 Retrieval Performance

For each experiment, one image was selected at random as the query image from each writer and thus retrieved images were obtained. For performance evaluation of the signature image retrieval system, it is significant to define a suitable metric. Two metrics are employed in our experiments as follows.

$$Recall = \frac{Number\ \ of\ relevant\ \ signatures\ \ retrieved}{Number\ \ of\ relevant\ \ signatures} \qquad (15)$$

$$Precision = \frac{Number of\ \ relevant signatures retrieved}{Number of\ signatures retrieved} \qquad (16)$$

Results correspond to precision and recall rate for a Top 2, Top 5, Top 8, Top 10, and Top 12. The comparative retrieval performance of the proposed system is shown in Table 1.

Table 1: Retrieval Performance based on Precision and Recall

	Ref.[10] method		Proposed method	
Number of Top matches	Precision %	Recall %	Precision %	Recall %
Top 2	70.66	12.77	100	16.66
Top 5	61.66	25.55	96.25	40.10
Top 8	52.50	35	84.37	56.25
Top 10	51.66	42.22	81.87	68.22
Top 12	45.55	46.11	**69.83**	**73.95**

Retrieval performance of the proposed method is compared with earlier approach based on Gradient Structural Concavity (GSC) [4] features. We evaluated the performance in terms of precision and recall images as function of the number of top retrieved images [1][3]. Fig. 2 shows graph illustrating this comparison between GSC and proposed method. From Fig. 2, it is clear that the proposed method is superior to earlier approach based on precision and recall.

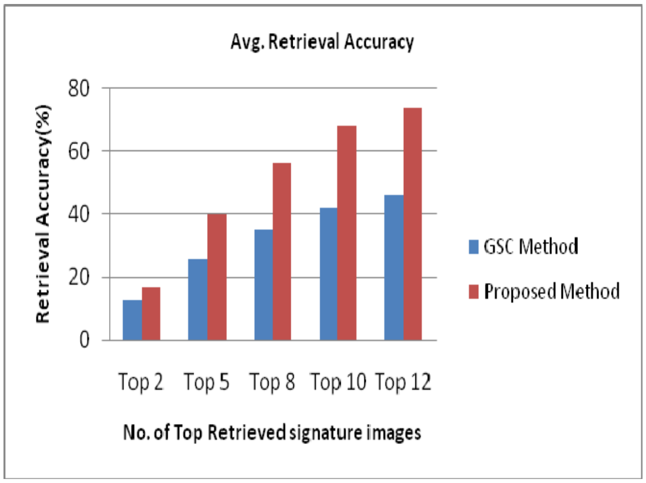

Figure 2. Comparison between GSC [3] and proposed method

Fig. 3 shows retrieval example for one of the query image from the database using proposed method. Images are displayed from top left to right bottom in the increasing order of distance from the query image. There are 16 ground truth images of each class in the database. The new method retrieves perceptually very similar signatures from the database.

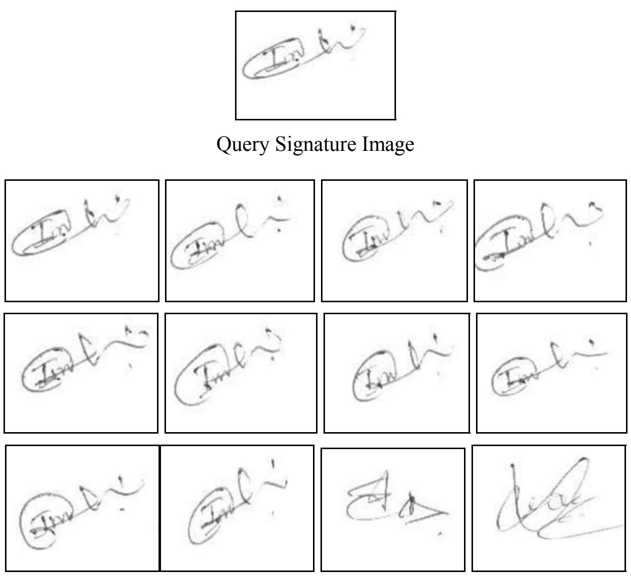

Figure 3. Sample handwritten signature retrieval example

4. CONCLUSION

Experimental were conducted for quick retrieval of offline signature and result are presented. The proposed approach uses RCWF and DT-CWT jointly for extracting signatures features in twelve different directions and Canberra distance measures for comparing the features. With this approach we could improve a recall from 46.11% to 73.95% when considering the top 12 results compared with earlier approach.

5. ACKNOWLEDGMENTS

The authors would like appreciate all writers who gave permission to use their handwritten signature in this study.

6. REFERENCES

[1] Armand S., Blumenstein, M., Muthukkumarasamy V. 2006. Off-Line Signature and Neural Based Classification, *IJCNN*, 684-691.

[2] A.K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall,India

[3] D. Rafael C. Gonzales, Richard E. Words and Steven L, *Digital Image Processing Using MATLAB*, Low Price Edition, India.

[4] G. Srikantan, S. Lam, ND s. Srihari, 1996. Gradient based contour encoding for character recognition, *Pattern Recognition*, vol. 29, no. 7, 1147-1160.

[5] Han Ke and Sethi I. K.. 1995. Handwritten signature retrieval and identification, *Pattern Recognition Letter*, vol.17, 83-90.

[6] Kokare M., P.K. Biswas, and B.N. Chatterji, 2005. Texture Image retrieval using New Rotated Complex Wavelet Filters, *IEEE Trans. on systems, man, and Cybernetics-Part B: Cybernetics*, vol. 35, no.6, 1168-1178.

[7] Kovari, B. Kertesz, Z . and M a j o r ,a., 2007.Off-Line Signature Verification Based on Feature Matching, *In Intelligent Engineering Systems*, 93-97.

[8] Larkins, R. Mayo, M., 2008. Adaptive Feature Thresholding for Off-Line Signature Verification, *Int. Image and vision computing* New Zealand, 1-6.

[9] Porwik P., 2007. The Compact Three Stages Method of the Signatures Recognition, *6th International Conference on Computer Information Systems and Industrial Management Applications*, 282-287.

[10] Ramachandra, A.C. Pavitra, K.and Yashasvini, K. and Raja, K.B. and Venugopal, K.R. and Patnaik, L.M., 2008. Cross-Validation for Graph Matching based Off-Line Signature Verification, *In IDICON,* India, 17-22.

[11] Samanesh Ghandali and Mohsen Ebrahimi Moghaddam, 2009. Off- Line Persian Signature Identification and Verification based on Image Registration and Fusion, *Int. Journal of Multimedia*, volume 4, 137-144.

[12] Wei Tian Yizheng Qiao Zhiqiang Ma, 2007. A New Scheme for Off-Line Signature Verification uses DWT and Fuzzy net, *Int. Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*, 3 0 - 35.