M-Band Packet Wavelet Farsi Handwriting Word Recognition
Farsi Script Segmentation Based on New Wavelet Function

J. Shanbehzadeh[1,2], A. Broumandnia[1]
[1]Oloum Va Tahghighat IAU, [2]Tarbiat Moalem University
Teheran, Iran

Abstract
Farsi Handwriting Word Recognition (FHWR), especially those with different orientation and scale changes as well as different handwriting style, is a challenging and important problem in document image analysis. This paper proposes a holistic FHWR scheme using local features of M-Band packet wavelet. The rotation and scale invariant local feature extraction for a given word image involves applying a polar transform to eliminate the rotation and scale effects, but at same time produce a row shifted polar image, which is then passed to an M-Band wavelet transform with row shift invariant to eliminate the row shift effects. So, the output M-Band wavelet coefficients are rotation and scale invariant. A local feature vector extracted from each subband of M-Band wavelet coefficients is constructed for rotation and scale invariant FHWR. In the experiments, we employed a Mahalanobis classifier to recognition a set of 224 distinct Farsi words selected from different persons with different style of writing. The experimental results, based on different handwriting style testing data sets for images with different orientations and scales, show that the proposed classification scheme using M-Band wavelet signatures outperforms other holistic handwriting word recognition methods, its overall accuracy rate for joint rotation and scale invariance being 95.8 percent, demonstrating that the extracted local energy features are effective rotation and scale invariant features.
Keywords: Holistic Word Recognition, Farsi Handwritings, M-Band Packet wavelet, row shift invariant.

1. Introduction
The holistic approach or word shape recognition for FHWR treats the word as a whole and features vector are extracted from the M-Band packet wavelet coefficients of Farsi handwriting word image. In this paper we implement this approach to cursive Farsi handwriting words. Each handwriting word image is decomposed in several sub bands with M-Band packet wavelet and from each band we extracted suitable feature which used FHWR with holistic paradigm. Word recognition plays a very important role in computer vision and pattern recognition, because used to automatic recognize text documents. There are many different applications involving word recognition, including license plate recognition, document recognition, recognition of handwritten check amounts, interpretation of handwritten addresses on pieces of mail, reading handwritten responses on forms, and automatic filing of faxes.Word recognition has been studied widely for over three decades [1] [2] [3] [4] [5] [7] [8] [9] [10] [11]. Despite so many different methods proposed, most of them assumed that the word images have the same orientation and scale. However, this assumption is not realistic for most practical applications. For example, if document images are obtained from scanning photographs, they are usually subject to certain random skew angles. Also, taking photographs of the same objects in different distances will cause different scales of objects.
We propose an effective scheme for rotation and scale invariant handwriting word recognition using M-band packet wavelet transform. The rotation and scale invariant feature extraction for a given word image involves applying a polar transform [5] [6] to eliminate the rotation and scale effects, but at same time produces M row shifted polar image, which is then passed to an row shift invariant M-band wavelet packet transform to eliminate the row shift effects. So, the output wavelet coefficients are rotation and scale invariant. Then for each subband of these wavelet coefficients a set of local energy features are computed. The local energy features, only computed for end subband of wavelet coefficients. We extracted features vectors form subbands of wavelet coefficients for M=2, 3 and 4, then we have three independent features vector for handwriting word classification and compare classification result for M=2, 3 and 4 at this paper. The proposed polar M-band wavelet feature has been well tested using a Mahalanobis classifier to classify a set of 224 distinct natural handwriting Farsi words provided from different person with different handwriting style.

2 Rotation and scale invariant M-band packet wavelet transform
Discrete M-band Wavelet Transforms (DMWT) has been shown to be useful for image analysis in literature [13], [15] due to wavelets having finite duration which provides both the frequency and spatial locality and efficient implementation. The hierarchical one-dimensional discrete M-band wavelet transform uses a set of M filters derived from wavelet functions to decompose the original signal into M sub bands: details and approximation. The decomposition process is recursively applied to the approximation and details sub band to generate the next level of the hierarchy. If an orthonormal wavelet basis has been chosen, the coefficients computed are independent and possess a distinct feature of the original signal. However, it is well known that one of the major drawbacks of DMWT is their lack of invariance to the shifting of the input signal due to the M down sampling structure of the wavelet expansions. And, for two dimensional input signals/images, 2D-DMWT is also sensitive to orientation/rotational changes; that is, the same images with different orientations may have different wavelet coefficients. The main reason is that the efficient implementation of 2D-DMWT requires applying a filter bank along the rows and columns of an image [15]. Due to separability of the filters, the separable 2D-DMWT is strongly oriented in the horizontal and vertical directions. This makes it hardly possible to extract rotation-

invariant features from the wavelet coefficients. In this section, we define a transform to convert Farsi word image into new image namely polar image with size $M^q \times M^q$ .With rotate Farsi word image we have row shift in related polar image. For this reason, we applying row shift invariant M-band wavelet packet transform. Then, for each sub band of the output rotation and scale invariant wavelet coefficients, an energy signature is computed according to some energy measures. The set of sub band energy signature are constructed feature vector and this vector used for rotation and scale invariant holistic Farsi word recognition. The details are as follows.

2.1 Polar transform

The polar transform map Farsi word image with size $k \times l$ in to new polar image with size $M^q \times M^q$ which this transform is used to eliminate the rotation and scale effects in the input Farsi word image. The fig.1 displays graphically a sample Farsi word image in different rotation angles and their corresponding polar images. As is shown in fig.1, if the word image is rotated then corresponding polar image will be row-shifted. However, the resultant polar image is row-shifted. The polar transform algorithm is described as follows.

Step 1: For a given $k \times l$ word image, finding bounding circle of word image with radius R and point center (X_m, Y_m).

Step 2: With help of radius R and equation (23) computed polar image with size $S \times R$ .The radius R is used as a scan line to sample S times from 0^o to 360^o for produce its equivalent $S \times R$ polar image.

$$p(a,r) = f(\left\lfloor R + r\cos(\frac{2\pi a}{S})\right\rfloor, \left\lfloor R - r\sin(\frac{2\pi a}{S})\right\rfloor) \qquad (23)$$

$$a = 0,1,.....,S-1 \qquad r = 0,1,.....,R$$

Where f(x,y) , p(a,r) are given the Farsi word image and polar image.

Step 3: Resize polar image p(x,y) to square polar image lp(x,y) with size $M^q \times M^q$. Where $q = ceil(\log_M^R)$.

For resizing image, we used interpolation methods. Three interpolation methods: Nearest neighbor interpolation, bilinear interpolation and Bicubic interpolation are used for image resizing. Interpolation is the process by which we estimate an image value at a location in between image pixels. For example, if you resize an image so it contains more pixels than it did originally, the software obtains values for the additional pixels through interpolation.

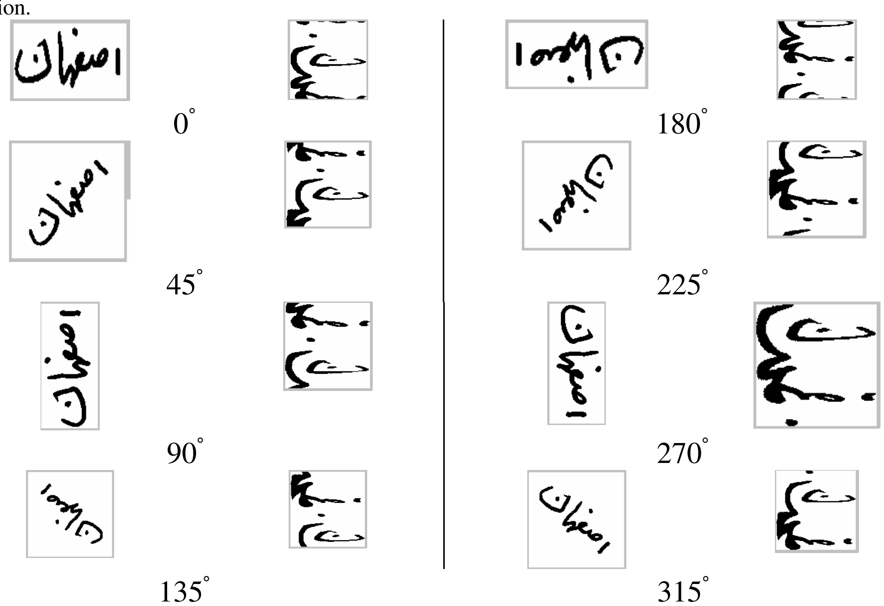


Fig1: The polar for a sample Farsi word image in rotation angles 0 , 45, 90, 135,180,225,270 and 315

2.2 Row Shift Invariant M-Band Wavelet Packet Transform

After applying the polar transform operation in Farsi word image, a rotated word image would be converted into a corresponding polar image with size $M^q \times M^q$ which is rotation invariant and scale invariant. However, any orientation changes would cause a row shifting in the polar image. The general approach of 2D-M-band wavelet packet transform or decomposition, as introduced in Section 2.4 , computes the M-band wavelet packet coefficients efficiently with a complexity of $O(n)$ (where n is the number pixels in the given image). However, the M-band wavelet packet coefficients thus obtained are not shift invariant. Many shift-invariant 2-band wavelet decomposition algorithms [19], [20], [21 have been proposed and are shift invariant in both rows and columns. These algorithms are worked for 2-band wavelet and are not suitable for M-band wavelet. Nevertheless, these algorithms generating more redundant wavelet coefficients are not suitable for the row-shifted output image produced by our polar transform. We need an M-band wavelet packet the decomposition, which should be invariant to row shifts only, without generating redundant wavelet coefficients. Also, we propose here a row shift invariant M-band wavelet packet decomposition to generate row shift invariant and no redundant wavelet coefficients. The block diagram fig.2 is display proposed row shift-invariant M-Band wavelet packet transform graphically.

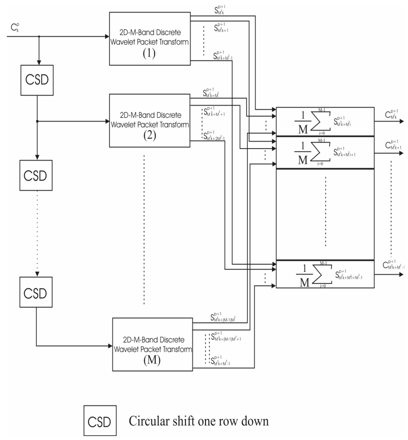


Fig.2: Rotation and scale invariant M-band packet wavelet transform

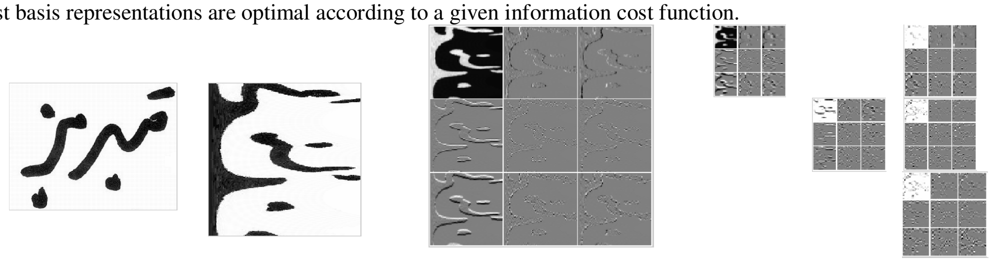
2.3 Best basis selection

In order to have a more effective and concise representation and, at the same time, having the feature of row shift invariance, we construct the best basis representation for the image during our row shift invariant M-band wavelet packet decomposition. Best basis selection approach proposed by Coifman and Wickerhauser for 1D signal. For best basis selection of subbands image, we can adaptively select some subbands to decompose further, instead of decomposing every subband. The basic idea is to compute the information cost of each subband and compare it with that of the sum of all next level subbands. If the information cost of the current subband is less than that of the sum of all next level subbands, then the current subband will not be decomposed; otherwise, we decompose the current subband further and do comparison again until maximum level is reached. Hence, the best basis representation can be obtained by an efficient recursive selection process, which determines the best decomposition of the image based exclusively on the local minimization of the information cost function. The best basis for image x can be computed recursively by:

$$A_k^p = \begin{cases} C_k^p & if \quad M(C_k^p) \le \sum_{i=0}^{M^2-1} M(A^{p+1}_{M^2k+i}) \quad or \quad p \ge J \\ \bigoplus_{i=0}^{M^2-1} A^{p+1}_{M^2k+i} & \end{cases} \qquad (32)$$

Where A_0^0 is word image x, A_k^j is basis coarse resolution at level j and J is maximum level.

Fig.3 show a example of best basis selection for polar images of a Farsi word image in different band M. Their best basis representations are optimal according to a given information cost function.



Farsi Word Polar Transform 3-Band Packet wavelet Best selected sub bands

Fig.3:The best basis selection for Three-Band Packet Wavelet

3. Local Features Extraction and Classification

For FHWR, we need a set of features to classification; these features are extracted from subband wavelet coefficients. The stages of feature-extraction of FHWR are shown in Fig.4. As shown in fig.4, in the first stage with M-band shift invariant packet wavelet, word image decomposed into a number of sub band images, in the second stage local or global features are extracted form sub band images with one of energy measurement, in the third stage feature vector are sorted in descending order and in the final stage feature vector is smoothed by desirable filter.

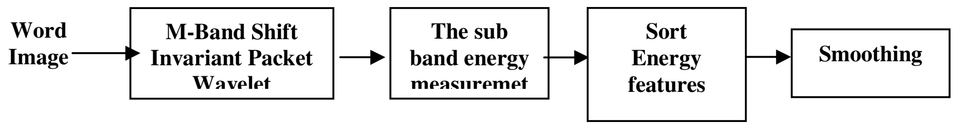


Fig.4: The stages of features extraction for FHWR

4. EXPERIMENTAL RESULTS

The effectiveness of the proposed M-Band shift invariant packet wavelet for rotation and scale invariant FHWR has been well tested via several experiments using a set of 224 distinct Farsi Handwriting words selected from different persons with different style of writing. Three major experiments were carried out with the objectives: 1) to investigate the FHWR performance based on the proposed M-Band shift invariant wavelet feature composed of different energy measures, 2) to investigate the FHWR performance of the proposed method on the different of M-Band wavelet for M=2, 3, 4 and 3) to investigate the FHWR performance of the proposed method on Farsi word images with rotation or scale changes only.

For the above experiments, we prepared ten different manuscript testing data sets and each data set has 224 distinct Farsi handwriting words. To rotation experiment; for each word of data sets, we provided seven words with different orientation (45^0 to 315° with 45° intervals, such as Fig.2). Then for each data set we have 1792 words.

5. References

[1] S.Madhvanath and V.Govindaraju, “The Role of Holistic Paradigms in Handwritten Word Recognition” IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 23, No.2, February 2001.

[2] J.T.Favata “Offline General Handwritten Word RecognitionUsing an Approximate BEAM Matching Algorithm” IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 23, No.9, September 2001.

[3] V.Govindaraju, P.Slavk and H.Xue, “Use of Lexicon Density in Evaluating Word Recognizers” IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 24, No.6, June 2002.

[4] G.Kim and V.Govindaraju “A Lexicon Driven Approach to Handwritten Word Recognition for Real-Time Applications” IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 19, No.4, April 1997.

[5] M.S.khorsheed and W.Clocksin,“Spectral Features For Arabic Word Recognition” IEEE 2000.

[6] T.Rauber and A.Steiger-Garcao, “2-d form descriptor based on a normalized parametric polar transform” In Proceeding of The IAPR Workshop on Machine Vision Applications, MVA’92, Tokyo,1992.

[7] R.Bozinovic and S.Srihari, “Off-Line Cursive Script Word Recognition” IEEE Transaction on Pattern Analysis and Machine Intelligence, VOL.11, NO. 1, Jan 1989.

[8] M.-Y. Chen, A. Kundu, and J. Zhou, “Off-Line Handwritten Word Recognition Using a Hidden Markov Model Type Stochastic Network” IEEE Transaction on Pattern Analysis and Machine Intelligence, VOL.16, NO.5, Jan 1994.

[9] P.D. Gader, M. Mohammad, and J.H. Chiang, “Handwritten Word Recognition with Character and Inter Character Neural Networks” IEEE Transaction System, Man, and Cybernetics, VOL. 27, NO. 1, 1997

[10] M. Gilloux, M. Leroux, and J.M. Bertille, “Strategies for Cursive Script Recognition using Hidden Markov Models” J. Machine Vision and Application , VOL.8, pp. 197-205,1995.

[11] F.Kimura, M.Shridhar, and N.Narasimhamurhi, “Lexicon Directed Segmentation-Recognition Procedure for Unconstrained Handwritten Words” Proc. Third Int’l workshop Frontiers in handwritteng recognition, pp. 122-130, 1993

[12]Y.Y.Tang, L.H.Yang, J.Liu, H.Ma "Wavelet Theory and Its Application to Pattern Recognition” World Scientific Publishing Co. 2000

[13]A. Primer "Introduction to wavelets and wavelet Transformations" Prentice Hall Press 1998

[14]Stephane Mallat, “A Wavelet Tour of Signal Processing, Second Edition”, Academic Press, 14 September 1999

[15] Stephane G.Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation" IEEE Transaction on Pattern Analysis and Machine Intelligence, VOL.11, NO.7, July 1989.
[16] O. Alkin and H. Caglar, "Design of efficient M-band coders with linear phase and perfect reconstruction properties," *IEEE Trans. Signal Processing*, vol. 43, pp. 1579–1590, July 1995.
[17] P. Steffen, P. N. Heller, R. A. Gopinath, and C. S. Burrus, "Theory of regular M-band wavelet bases," *IEEE Trans. Signal Processing*, vol. 41, pp. 3497–3510, Dec. 1993.
[18] S.R. Fountain and T.N. Tan, "Extraction of Noise Robust Invariant Texture Features via Multichannel Filtering," Proc. Int'l Conf. Image Processing '97, vol. 3, pp. 197-200, 1997.
[19] G. Beylkin, "On the Representation of Operators in Bases of Compactly Supported Wavelets," SIAM J. Numerical Analysis, vol. 6, no. 6, pp. 1716-1740, Dec. 1992.
[20] R.R. Coifman and D.L. Donoho, "Translation-Invariant De-Noising," Wavelet and Statistics, Lecture Notes in Statistics,A. Antoniadis and G. Oppenheim, ed., Springer-Verlag, pp. 125-150, 1995.
[21] J. Liang and T.W. Parks, "A Translation-Invariant Wavelet Representation Algorithm with Applications," IEEE Trans. Signal Processing, vol. 44, Feb. 1996.
[22] J.C. Pesquet, H. Hrim, and H. Carfantan, "Time-Invariant Orthonormal Wavelet Representations," IEEE Trans. Signal Processing, vol. 44, Aug. 1996.