WAVELET PACKET TRANSFORM AND FUZZY LOGIC APPROACH FOR HANDWRITTEN CHARACTER RECOGNITION

Sreela Sasi and Loren Schwiebert and Jatinder Bedi

Department of Electrical & Computer Engineering
Wayne State University, Detroit, MI 48202
{ssasi,loren,jbedi}@ece.eng.wayne.edu

ABSTRACT

This paper presents a novel method for handwritten character recognition using wavelet packet transform and fuzzy logic. This method exploits the time-frequency localization and compression capability of wavelet packet transform, using the best basis algorithm to enhance the accuracy of recognition at the pixel level and the computational capability of fuzzy logic with linguistic variables, which is a universal approximator if it uses enough rules. The best basis algorithm automatically adapts the transform to best match the characteristics of the signal, as well as minimize the additive cost function. The wavelet packet transform of the handwritten characters are taken using the best basis algorithm. The standard deviation of the spread of the coefficients in each multi-resolution level are computed, which forms the characteristic features for the characters. These features are given as input to the fuzzy logic character recognition system, where these are fuzzified, analyzed, and the corresponding characters are given as output using IF ... THEN rules. This method is more efficient for handwritten character recognition than energy sorted wavelet transform of character images, since it contains only a few edges in the image. Simulation of four multi-resolution levels for each character is done using symmlet and results show that they have better accuracy than the methods using only fuzzy logic.

1. INTRODUCTION

Handwriting identification and recognition are of great practical interest in the extraction of discriminating and invariant information from a handwritten text [1]. Generally, handwriting identification and recognition systems can be divided into off-line and on-line systems. In off-line systems, only static information can be obtained while in on-line systems dynamic information can also be considered. So, in order to recognize a cursive script, off-line systems use critical points like end points, crossings, branch points, local extrema in the X and Y dimensions and so on, while in on-line systems a stroke of a cursive script occurs between a pen-down and the successive pen-up. One of the greatest advantage of on-line devices is that the acquisition process itself yields a natural segmentation of a cursive script into strokes. On-line systems require the use of special equipment for the writer such as electromagnetic/electrostatic pens and pressure sensitive digitizing tablets, which are not as simple, natural, and cheap as pen and paper. Besides a hard copy must be available to keep a physical trace for justification. Off-line handwriting identification and recognition are more difficult, not only because of the loss of dynamic information, but also because the camera reading process of a cursive script adds additional noise to the remaining information. It is practically impossible to restore the real trajectory of a penball. To recognize a static handwritten text, one needs to determine the discriminating, robust, and perceptually salient features of handwriting that reflect the fundamental factors of shapes and curves that make up a cursive script. Machine recognition of general handwritten text faces a number of challenges. The dynamic nature of handwriting styles requires development methods that are adaptive to local variations. There is no perfect mathematical model that can describe such extreme variations, and hence it is impossible to find characteristic features that are invariant with different writing styles. So, one may extract characteristic features that are reasonably insensitive to variations caused by individual writing styles, while maintaining the ability to separate between characters. In this paper, wavelet packet transform (WPT) which is good for time-frequency localization at different multi-resolution levels is used for feature extraction and fuzzy logic technique is used for classification and recognition.

Wavelet packet transform is used for sub-band coding by researchers [2, 3, 4]. Image coding based on energy sorted wavelet packets is discussed by Kong [5]. In the energy sorted wavelet packet decomposition, all the sub-images in the packet are sorted according to their energies and the most important sub-images as measured by the energy are preserved and coded. Character recognition using fuzzy logic is discussed in [6, 7, 8].

2. WPT AS FEATURE EXTRACTOR

Feature extraction is a fundamental problem in image processing. Feature extraction in off-line text analysis can be classified into statistical and structural methods [9]. In statistical techniques, features extracted from handwritten text images can be global or local, although a successful system may combine both of them. Generally the global features are easily extracted, insensitive to noise, and can tolerate minor distortions and stylistic variations commonly associated with writing of the same words by the same person. On the other hand, the local features extracted from handwritten text based on local information are usually stable and have a high tolerance to distortions and stylistic variations. Wavelet transforms of signals are well known for their time and frequency localization property. The wavelet transform decomposes the original image into waveforms at different scales and locations. The large wavelet coefficient corresponds to important features in the image, such as edges. Also, orthonormal wavelets with finite support exhibit a very powerful mathematical tool for decomposing a function into a multi-resolution hierarchy of different localized frequency components. Such a decomposition allows us to analyze a function simultaneously at different levels of resolution, and supports coarse to fine representation strategies similar to what exist in human visual systems [10]. Wavelet approximation can compact the energy of a signal into a relatively small number of wavelet functions. This data compression feature of wavelets is valuable for applications such as non-parametric statistical estimation and classification. This property is used in our approach for the classification of handwritten characters. Wavelets are fundamental building block functions, similar to the trigonometric sine and cosine functions oscillating about zero. Wavelet packet approximators are based on translated and scaled wavelet packet functions $W_{j,b,k}$. These are generated from a base function of the form $W_{j,b,k}(t) = 2^{-j/2} W_b(2^{-j}(t-k))$ [11], where j is the resolution level, b is the number of oscillations, and k is the translation shift. In wavelet packet analysis, a signal $f(t)$ is represented as a sum of orthogonal wavelet packet functions $W_{j,b,k}(t)$ at different scales, oscillations, and locations:

$$f(t) \approx \sum_j \sum_b \sum_k w_{j,b,k} W_{j,b,k}(t)$$

where $w_{j,b,k}$ is the wavelet packet coefficient. The range of summations for the levels j and the oscillations b is chosen so that the wavelet packet functions are orthogonal. A fast splitting algorithm [12] which is an adaptation of the pyramid algorithm [13] for discrete wavelet transform is used for finding the wavelet packet table. The splitting algorithm differs from the pyramid algorithm in that low-pass and high-pass filters are applied to the detailed coefficients in addition to the smooth coefficients at each stage in the algorithm. Also, all the coefficients are retained, including those at intermediate filtering stages. The best basis algorithm [14] is used for selecting optimal bases (transforms) from wavelet packet tables. The best basis algorithm automatically adapts the transform to best match the characteristics of the image. The best basis algorithm finds the wavelet packet transform W that minimizes the additive cost function, $E(W) = \sum_{j,b} E(w_{j,b})$, where $(j,b) \in I$ with I as the set of index pairs *(j,b)* of the components in the transform W. At the heart of the best basis algorithm is the wavelet packet cost table which is the table of costs $E(w_{j,b})$. Minimizing the default cost function is equivalent to finding the minimum *entropy* function. Feature extraction is done by taking the wavelet packet transform of the character image using the best basis algorithm for a desired number of multi-resolution levels.

3. FUZZY RECOGNITION SYSTEM

Common sense knowledge can be characterized as pieces of knowledge that are usually true, but not necessarily always true. A serious drawback of conventional approaches to knowledge representation is their inability to represent uncertainty and imprecision. Hence the conventional approaches are inadequate for representing human reasoning which are approximate rather than exact. Fuzzy logic, which can be viewed as an extension of classical logical systems, provides an effective conceptual framework for dealing with the problem of knowledge representation in an environment of uncertainty and imprecision. Zadeh's theory of incompatibility states that as the complexity of a system increases, our ability to make precise yet significant statements about its behavior diminishes until a threshold is reached beyond which precision and significance or relevance become almost mutually exclusive characteristics [15]. When the brain receives complex data through the senses, it tries to summarize data, reducing massive detail to chunks of perception. For instance, the million light-sensitive cells of the human retina take in far too much information for the brain to decode. But it has tactics for sifting and reducing this flow.

In other words, it summarizes the information. That is, we perceive the precise information in a fuzzy way. Fuzzy logic could handle complexity in a similar way. By summing up words mathematically, fuzzy sets could help bring complex systems, like the visual information, under control. Fuzzy systems are universal approximators if they use enough rules [16]. In this sense, any continuous function or system can be modeled by them and the quality of the fuzzy approximation depends on the quality of the rules. According to Zadeh [17], exploiting the tolerance of imprecision is an important issue in Computing with Words (CW). A key aspect of CW is that it does computation using fuzzy variables in natural languages. Humans can express their mental perceptions by formulating reasons from premises which help them to reach conclusions using words in natural languages. Due to the impreciseness, lack of structure, tremendous variations in writing styles of the same person and between persons, etc. associated with handwritten characters, a fuzzy logic approach is used for character recognition purposes. Here a multi-input single-output fuzzy system is used which is given by $f : U \subset R^n \mapsto V \subset R$, where $U = U_1 \times U_2 \times ...U_n$ is the input space and $V \subset R$ is the output space in U_j. A fuzzy inference engine can be formed by a set of linguistic rules in the form of "IF <*a set of conditions are satisfied*> THEN <*a consequence is inferred*>". In our handwritten character recognition system, the statistical features from the standard deviation of the coefficients of the wavelet packet components are given as input to the fuzzy recognition system. Fuzzy sets are formed for each wavelet component with the name of the component, the membership value in the range, and the name of the range. The fuzzy logic system consists of a fuzzifier, a fuzzy inference engine, and a defuzzifier.

4. SIMULATION

The handwritten text is segmented, characters are enlarged to fit into a 4 x 4 matrix cell, and scanned to get the character image which is sampled in our case at $128x128$ pixels and given to the Feature Extractor. WPT of each character image is taken at enough multi-resolution levels (four levels in our case) for better precision, using best basis algorithm. The wavelet symmlet s8 is used for simulation. Exploratory data analysis is done on the coefficients and standard deviation of coefficients for all the four levels computed for each character. These values form the characteristic feature of the character image and are given to the fuzzy logic recognition system. Here, fuzzy sets are formed for each wavelet component by the fuzzification process, and the fuzzy rule-based inference engine will classify and recognize the character and the character is given as output. As an example, consider the character image for "A". Standard deviations of wavelet coefficients for slanting towards right "A" (RA), normal straight "A" (NA), and slanting towards left "A" (LA) are computed after WPT is taken for four levels.

The WPT of character image for "A" is shown in Table 1. The coefficient values are fuzzified and rules are formed for the recognition system. The membership overlaps 'LOW', 'MEDIUM' and 'HIGH' are assigned for the range 0-5. Range span is chosen as a multiple of 5, the higher end being a multiple of 5 higher than the largest standard deviation value of coefficients, and the lower end is zero. Each span of 5 starting from zero is given a range name R1, R2,...,Rn. So the value 21 will be 'LOW' in the range span R5 with an exact membership of 0.2 in R5. The rule for character "A" can be as follows: IF w1.0-w1.1 is LOW OR MEDIUM OR HIGH in R8 AND w1.1-w1.1 is HIGH in R5 OR LOW in R6 AND ... AND w4.13-w4.1 is HIGH in R3 OR LOW in R6 THEN character is "A". The feature extraction is done using S-PLUS software from Mathsoft and the fuzzy logic recognition system is simulated using LISP, since we have to process a list of fuzzy sets for each character.

5. CONCLUSION

The best basis algorithm for taking WPT reduces cost over other transforms. Also it has lowest mean square energy uniformly across all compression ratios. Since the energy concentrates on two or three wavelet components, the energy sorted approach fails for character recognition. The fuzzy rules can be tuned to incorporate more variations in writing.

6. REFERENCES

[1] R. Plamondon and G. Lorette. *Designing an automatic signature verifier: problem definition and system description.* World Scientific, 1990.

[2] W. Kim and Y. Hu. Wavelet packet-based subband coder. In *ICASSP*, volume v4, pages 2225–2228, 1995.

[3] M.T. Orchad Z. Xiong, K. Ramachandran and K. Asai. Wavelet packet-based image coding using joint space frequency quantization. In *IEEE International Conf. On Image Processing*, volume v3, pages 324–328, 1994.

[4] P. Onno and C. Guillemot. Wavelet packet coding with jointly optimized lattice vector quantization

and data rate allocation. In *IEEE International Conf. On Image Processing*, volume v3, pages 329–333, 1994.

[5] L. Kong and K. Lay. Image coding based on energy-sorted wavelet packets. In *SPIE - The International Society for Optical Engineering*, volume v2501/2, pages 1279–1287, 1995.

[6] J.M. Keller P. Gader and J. Cai. A Fuzzy Logic system for the detection and recognition of handwritten street numbers. *IEEE Transactions on Fuzzy Systems*, 3(1):83–95, February 1995.

[7] C. Zheru and H. Yan. ID3 Derived Fuzzy Rules and Optimized defuzzification for Handwritten numeral recognition. *IEEE Transaction on Fuzzy Systems*, 4(1):24–31, February 1996.

[8] S.Sasi and J.S. Bedi. Handwritten Character Recognition using Fuzzy Logic. In *37th MWSCS-94, Lafayette, USA*, pages 1399–1402, 1994.

[9] R. Sabourin and R. Plamondon. *Off-line identification with handwritten signature images: Survey and Perspectives.* Springer Verlag, 1992.

[10] D. Marr. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information.* San Francisco: W. H. Freedman, 1982.

[11] I. Daubechies. *Ten Lectures on Wavelets.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.

[12] A. Bruce and H. Gao. *Applied Wavelet Analysis with S-Plus.* Springer-Verlag, New York, 1996.

[13] S.G. Mallat. A Theory for multiresolution signal decomposition: wavelet representation. *IEEE Transactions On Pattern Analysis and Machine Intelligence*, 11(7):674–693, July 1989.

[14] R. Coifman and V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Transactions on Information Theory*, 38(2):713–718, 1992.

[15] L.A. Zadeh. Outline of a New Approach to the Analysis of Complex Systems and Decision Processes. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3(1):28–44, January 1973.

[16] B. Kosko. *Fuzzy Engineering.* Prentice Hall, New Jersey, 1997.

[17] L.A. Zadeh. Fuzzy Logic = Computing with Words. *IEEE Transaction On Fuzzy Systems*, 4(2):103–111, May 1996.

Table 1: WPT coefficients for character "A"

WPT COMP.	NA	RA	LA
w1.0-w1.1	35.93	37.29	38.08
w1.1-w1.1	19.79	19.69	20.38
w2.0-w2.1	64.09	63.53	63.04
w2.1-w2.1	—	36.79	36.59
w2.2-w2.1	14.76	15.97	16.29
w2.3-w2.1	17.60	16.68	17.63
w3.0-w3.1	95.34	97.82	95.61
w3.1-w3.1	—	58.79	54.39
w3.2-w3.0	117.92	—	112.16
w3.2-w3.1	24.25	24.48	23.65
w3.2-w3.2	34.87	—	—
w3.2-w3.3	59.47	—	—
w3.3-w3.0	79.09	78.90	69.00
w3.3-w3.1	15.23	19.06	18.58
w3.3-w3.2	15.38	—	—
w3.3-w3.3	12.21	—	—
w3.4-w3.1	12.10	14.56	14.36
w3.5-w3.1	14.38	22.97	18.70
w3.6-w3.0	—	—	56.12
w3.6-w3.1	19.98	19.02	23.51
w3.7-w3.0	64.08	55.50	40.53
w3.7-w3.1	27.38	19.54	21.55
w4.0-w4.0	232.31	231.86	239.74
w4.0-w4.1	139.18	126.55	141.06
w4.1-w4.0	162.14	157.02	163.49
w4.1-w4.1	70.40	94.82	96.46
w4.2-w4.0	94.37	93.97	87.34
w4.2-w4.1	24.21	27.34	28.34
w4.2-w4.2	59.98	—	—
w4.2-w4.3	96.63	—	—
w4.3-w4.0	163.70	165.21	173.34
w4.3-w4.1	20.33	21.22	29.38
w4.3-w4.2	20.98	—	—
w4.3-w4.3	14.90	—	—
w4.4-w4.0	—	132.45	—
w4.4-w4.1	—	17.25	—
w4.5-w4.0	—	191.72	—
w4.5-w4.1	—	20.94	—
w4.8-w4.0	142.91	143.41	155.38
w4.8-w4.1	8.83	—	23.23
w4.9-w4.0	89.62	86.53	91.91
w4.9-w4.1	10.38	11.22	16.75
w4.10-w4.0	152.26	149.43	160.93
w4.10-w4.1	13.82	23.22	22.34
w4.11-w4.0	79.41	87.67	86.57
w4.11-w4.1	11.90	23.99	24.71
w4.12-w4.0	96.44	97.65	—
w4.12-w4.1	12.03	24.84	—
w4.13-w4.0	71.83	73.43	—
w4.13-w4.1	13.97	25.57	—

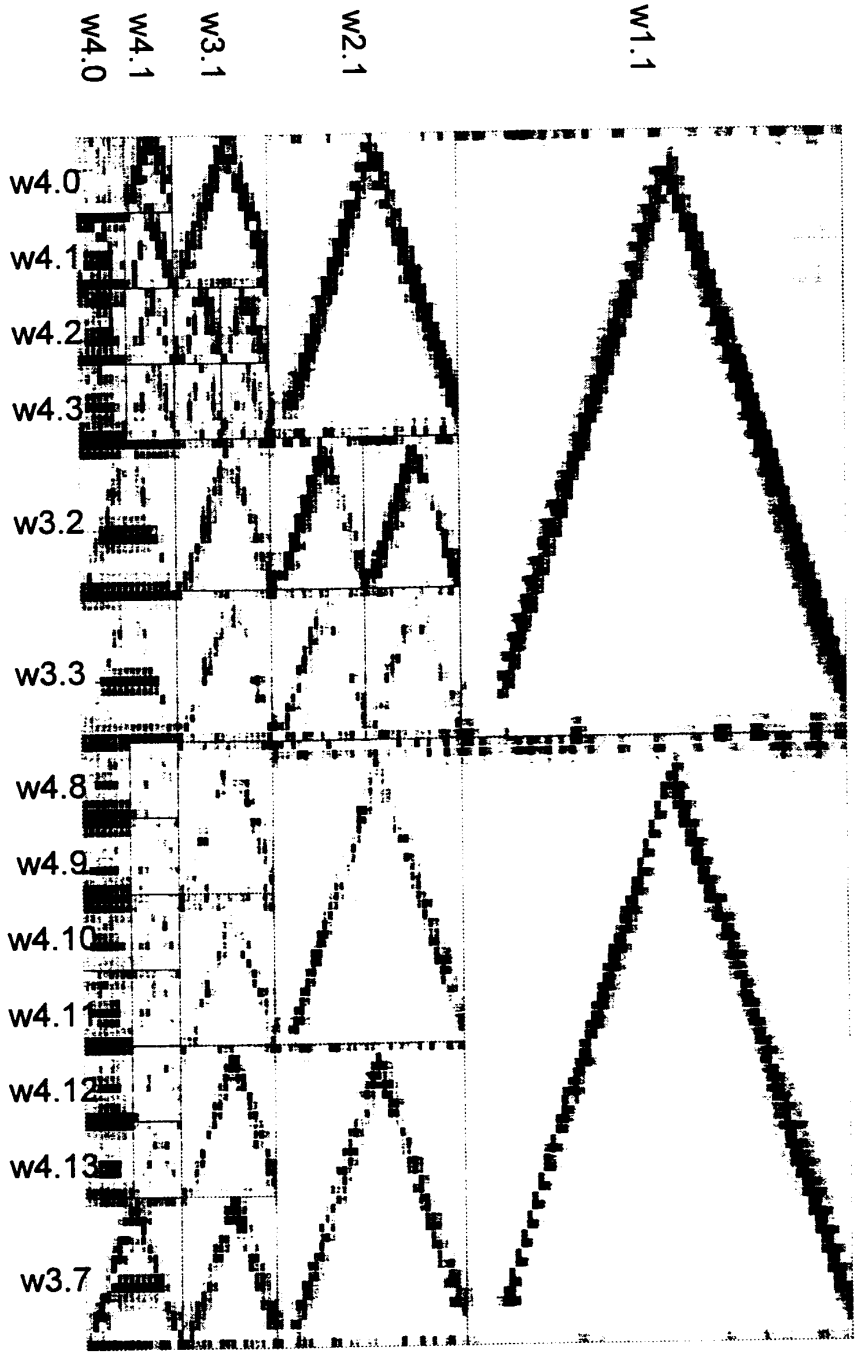

Fig. 1 WPT of Character image for "A" using Best-Basis algorithm