WAVELET PACKET TRANSFORM AND NEURO-FUZZY APPROACH TO HANDWRITTEN CHARACTER RECOGNITION

SREELA SASI
Dept. of Electrical Engineering
University of Idaho
Moscow, ID 83843

sreela@orizaba.ee.uidaho.edu

LOREN SCHWIEBERT & JATINDER SINGH BEDI
Dept. of Electrical and Computer Engineering
Wayne State University
Detroit, MI 48202

{loren, jbedi} @ece.eng.wayne.edu

ABSTRACT

This paper presents a novel method for automatic handwritten character recognition by combining wavelet packet transform with neuro-fuzzy approach. The time-frequency localization and compression capability of wavelet packet transform using best-basis algorithm is used for feature extraction, enhancing the accuracy of recognition at pixel level. The best-basis algorithm automatically adapts the transform to best match the characteristics of the signal, minimizing the additive cost function. Since fuzzy sets and fuzzy logic remains as a means for representing, manipulating and utilizing uncertain
information and to provide a framework for handling uncertainties and imprecision associated with real world problems, a fuzzy logic system is used for classification purpose. A neural network system is used for recognition purposes since they provide computational power, fault tolerance, and learning capability to the systems. Characteristic features are extracted by taking wavelet packet transform using best-basis algorithm and are given as input to the fuzzy classifier where they are fuzzified and classified using IF ... THEN rules, and given to a neural network recognition system. This method is more efficient for handwritten character recognition as well as personal identification compared to energy sorted wavelet transform of character images, since characters contain very few edges in the images. Simulation of characters is done for 3 multiresolution levels using symmlet and results show that this method is more efficient than the methods using only fuzzy logic.

1. Introduction

The advances in the field of handwriting identification and recognition result from a better understanding of the fundamental biophysical and psychophysical processes involved in handwriting generation, and the application of this knowledge to various types of specific systems. From the view point of biophysics and psychophysics, handwriting can be represented as the output of a space-time variant system equivalent to the writer, where the input is a learned motor program, and described by the curve-linear displacement, the angular displacement, and the torsion of the trajectory according to the intrinsic properties of curves [1], [2]. According to the scheme given by Plamondon and Lorette [3], certain neurons within the brain fire with a predetermined intensity and duration, and this nerve network activate the proper muscles in a predetermined order. The motion of pen on the paper resulting from muscle contraction and relaxation leaves a partial trace of the trajectory of the pen tip. How the motor program is constructed, used and influenced by other biophysical and psychophysical mechanisms is still an open question, but the design of a handwriting identification or recognition system is based on the fact that people do not write according to a standard penmanship, and the deviation from the norm is individual dependent.

In handwritten character recognition there are two general approaches [4]. The first approach involves segmenting a word into individual characters and recognizing each character separately.The second approach involves recognizing a word as a complete entity using some global features of the word. Both approaches have their strengths and weaknesses. The former has the advantage of being generalized to a large vocabulary with limited training, but seems more suitable to good handwriting because segmentation is ambiguous and different interpretations are possible at letter level. This make the recognition process of the whole word ambiguous or erroneous. The latter is nearer to the human way of reading, and more favorable in the case of bad handwriting, but it is limited in its discrimination capability, and suitable only for limited vocabulary applications. The task of grouping characters and words into lines of text is relatively straight forward for machine-printed documents. In handwritten text, the lines might modulate up and down, and ascenders and descenders frequently intersect characters from neighboring lines making the task more difficult. After the line isolation and extraction, the next challenge is the location of word boundaries in the line. Since words tend to flow together, it is necessary to develop algorithms to search for the likely boundaries of the words. Such techniques are somewhat tied to the word recognition algorithms employed, and there are a number of trade-offs to be considered.

For describing the nature of handwriting, Nishida and Mori proposed a clear, rigorous and powerful method for structural description of character shapes in terms of quasi-topologic features such as convexity and concavity, directional features, and singular points such as branch points and crossings [5]. Chhabra et.al extracted four types of raw features namely, printed line, high convex curvature, convex hull, and hole [6]. Pettier and Camillerapp described a new decomposition method, where a regular line is characterized by a fairly constant stroke and the remaining parts consists of intersections, overlappings, and discontinuities [7].

Handwriting identification and recognition are of great practical interest in the extraction of discriminating and invariant information from a handwritten specimen. There are two different techniques namely, structural and statistical, are used for handwriting identification and recognition problems. Both techniques have their merits and weaknesses. The structural information about the interconnections in a cursive handwriting cannot be handled well by statistical techniques, whereas the use of formal handwriting models to represent a cursive script is the main drawback of the structural approach. Handwriting is a natural entity which cannot strictly obey the mathematical constraints set by formal theory, where the intraclass variations are enormous. To recognize a static handwritten specimen, one needs to determine the discriminating, robust, and perceptually salient features of handwriting that reflect the fundamental factors of shapes and curves that makes up a cursive script. Kim and Govindaraju gives an efficient image handling technique for handwritten document recognition by providing image manipulation procedures for fast handwritten word recognition, including pre-processing, segmentation, and feature extraction [8]. One of the major difficulties in off-line word recognition originates from the great variations observed in different samples of script from the same writer over time or from different scriptors. There is no perfect mathematical model that can describe such extreme variations, and hence it is impossible to find characteristc features that are invariant with different writing styles. Hence, feature extraction is a fundamental problem in handwritten character recognition due to the dynamic nature of handwriting styles that require adaptive methods to local variations. In this paper, wavelet packet transform (WPT) which is good for time-frequency localization at different multiresolution levels is used for feature extraction, fuzzy logic (FL) is used for analysis and classification, and neural network is used for recognition purpose.

2. WPT as Feature Extractor

Wavlet packet analysis is is an important generalization of wavelet analysis [9], [10], [11], [12]. Wavelet packet functions are also localized in time, but offer more flexibility than wavelets in representing different types of signals. Wavelet packet approximators are based on translated and scaled wavelet packet functions $W_{j,b,k}$. These are generated from the base function [13], as $W_{jbk}(t) = 2^{j/2} W_b(2^{-j}(t-k))$, 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 $W_{jbk}(t)$ at different scales, oscillations and locations:

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

$w_{j,b,k}$ is the wavelet packet coefficient. The range of summation for the levels 'j' and the oscillations 'b' is chosen so that the wavelet packet functions are orthogonal. A fast splitting algorithm [14] which is an adaptation of the pyramid algorithm [15] 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. Best basis algorithm [9] 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, where 'I' is 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 best basis algorithm for a desired number of multiresolution levels. Wavelet packet transform is used for sub-band coding by researchers [16], [17], [18]. Image coding based on energy sorted wavelet packets is discussed by Kong [19]. 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.

3. Fuzzy Logic Classifier

Even though a tremendous amount of information is presented to the human senses in a given situation, somehow the human mind has the ability to discard most of the information and to concentrate only on task relevant information. This ability of the human mind to deal only with task relevant information is its ability to process fuzzy information. Fuzzy logic helps finding solutions to complex problems. It is a tool that enhances our ability to deal with problems that are too complex, and too ill-defined to be succeptible to solution by conventional means. 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. By summing up words mathematically, fuzzy sets could bring complex systems like visual systems under control. 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 classification purpose. Here a multi-input single output fuzzy system is used which is given by $f: U \subset R^n \rightarrow V \subset R^m$, where

$U = U_1 \times U_2 \times \ldots \times U_n$ is the input space, and $V \subset R^m$ is the output space in Vj . 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>". Fuzzy systems can be considered as universal approximators [20], if they use enough rules. In this sense, they can model any continuous function or system, and the quality of fuzzy approximation depends on the quality of the rules. 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 classification system. Fuzzy sets are formed for each wavelet component with name of the component, the membership value in the range, and the name of the range. The fuzzy logic system consists of a fuzzifier, fuzzy inference engine and a defuzzifier. Character recognition using fuzzy logic is discussed in [21], [22], [23].

4. Artificial Neural Network Recognizer

Artificial Neural network ANN) are highly parallel information processing systems resembling that of human brain configured in regular architectures. The collective behavior demonstrates the ability to learn, recall and generalize from training patterns or data [24], [25]. A neural network learns about its environment through an iterative process of adjustments applied to its synaptic weights and thresholds. Ideally, the network becomes more knowledgeable about its environment after each iteration of the learning process. Learning can be defined as a process by which the free parameters of a neural network are adapted through a continuing process of simulation by the environment in which the network is embedded. Because of their learning and memorizing capability, a neural network recognizer is used for our handwritten character recognition system. The output from the fuzzy classifier is used for training the neural network which will memorize these patterns.

5. Neuro-Fuzzy Recognition System

Fuzzy systems and neural networks have the ability to improve the intelligence of systems working in uncertain, imprecise and noisy environments. They estimate a function without requiring a mathematical description of how the output functionally depends on the inputs. Neuro-fuzzy networks combine the learning and the knowledge representational capabilities of neural networks and fuzzy sets. Here, fuzzy logic and neural network are integrated for exploiting the advantage of both technologies for the character recognition purpose. Neuro-fuzzy systems are used for pattern recognition purposes in [26], [27].

6. Simulation

The segmented and normalized handwritten text is scanned to get the character image which is sampled in our case at 128 x 128 pixels and given to the Feature Extractor. WPT of each character image is taken for enough multiresolution levels (three levels in our case), 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 three levels are computed for each character. These values form the characteristic feature of the character image and are given to the Fuzzy Logic Classification system. Here, fuzzy sets are formed for each wavelet component by the fuzzification process, and the fuzzy rule-based inference engine will analyze and does the classification. As an example, consider the character image for "B". Standard deviations of wavelet coefficients for normal straight "B" (NB), slanting towards left "B" (LB), and slanting towards right "B" (RB) are computed after WPT is taken for three levels.

The WPT of character image for "B" is shown in Fig. 1. The standard deviation of coefficients after taking WPT for three levels using best-basis algorithm for Character "B" is given 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 amultiple of 5, the higher end being multiple of 5 higher than the largest standard deviation value of the coefficients for the character, and the lower end is zero. Each span of 5 starting from zero is given a range name R_1, R_2 ,........, R_n. So, the value 26 will be 'LOW' in the range span R_6 with an exact membership of 0.2 in R_6. The rule for character "B" can be as follows: IF w1.0-w1.1 is HIGH in R_7 AND w1.1-w1.1 is LOW OR MEDIUM in R_5 AND .... AND w3.7-w3.1 is LOW in R_7 OR MEDIUM in R_9 THEN w1.1-w1.1 is 20.28 AND w2.0-w2.1 is 75.98 AND ..... AND w3.7-w3.1 is 35.10. Similar rules can be formed for each character. The consequent part is given as input to the neural network recognizer where it will be recognized. An architecture for the recognition system is shown in Fig. 2

WAV. COMP.	SD NB	SD LB	SD RB
w1.0-w1.1	-----	33.59	33.31
w1.1-w1.1	20.28	21.86	22.19
w2.0-w2.1	75.98	72.85	69.19
w2.0-w2.3	28.40	------	------
w2.1-w2.1	38.37	40.45	39.12
w2.1-w2.2	25.02	------	------
w2.1-w2.3	18.09	------	------
w2.2-w2.0	81.29	70.16	77.28
w2.2-w2.1	32.44	36.27	35.19
w2.3-w2.1	28.77	33.30	31.92
w3.0-w3.0	202.42	191.69	180.42
w3.0-w3.1	161.11	116.49	117.83
w3.0-w3.4	81.38	------	-------
w3.0-w3.5	39.64	------	-------
w3.1-w3.0	142.88	116.46	125.44
w3.1-w3.1	6 5.64	69.11	59.05
w3.1-w3.4	25.80	------	------
w3.1-w3.5	18.35	------	------
w3.2-w3.0	145.26	130.58	129.83
w3.2-w3.1	68.65	62.30	54.90
w3.3-w3.0	80.92	60.86	72.16
w3.3-w3.1	52.49	50.31	48.44
w3.6-w3.0	69.34	57.11	61.47
w3.6-w3.1	32.16	32.83	38.27
w3.7-w3.0	38.60	41.74	41.37
w3.7-w3.1	35.10	30.54	41.95

Table 1 Standard Deviation of coefficients after WPT of character image "B"

7. Conclusion

The best-basis algorithm reduces the 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, energy-sorted approach fails for character recognition. The result shows that WPT and neuro-fuzzy approach has more accurate recognition at pixel level than the one using only fuzzy logic.

8. REFERENCES

[1] R. Plamondon, “What does Differential Geometry tell us about Handwriting Generation?”, Proc. 3rd International Symposium on Handwriting and Computer Applications, Montreal, pp. 11-13, 1987.

[2] R. Plamondon, “A Handwriting Model based on Differential Geometry” *Computer Recognition and Human Production of Handwriting,*R. Plamondon, C. Y. Suen, M. L. Simner (eds.), Word Scientific Publishing, pp. 179-192, 1989.

[3] R. Plamondon and G. Lorette, “Designing an Automatic Signature Verifier: Problem Definition and System Description”, *Computer Processing of Handwriting*, R. Plamondon and C.G. Leedham (eds.), World Scientific, pp. 3-20, 1990.

[4] J. T. Favata and S. N. Srihari, “Recognition of Handwritten Words for Address Reading”, Proc. 4th U.S. Postal Service Advanced Technology Conference, Washington DC, pp. 191-205, 1990.

[5] H. Nishida and S. Mori, “Algebraic Description of Curve Structure”, IEEE PAMI-14, pp. 516-533, 1992.

[6] A. K. Chhabra et.al, “High-order Statistically Derived Combinations of Geometric Features for Hand-printed Character Recognition”, Proc. 2nd International Conference on Document Analysis and Recognition, Tsukuba Science City, pp.397-401, 1993.

[7] J. C. Pettier, J. Camillerapp, “Script Representation by a Generalized Skeleton”, Proc. 2nd International Conference on Document Analysis and Recognition, Tsukuba Science City, pp. 850-853, 1993.

[8] G. Kim and V. Govindaraju, “Efficient Chain Code based Image Manipulation for Handwritten Word Recognition”, SPIE, vol. 2660, pp. 262-272, 1996.

[9] R. Coifman and V. Wickerhauser, “Entropy-based Algorithms for Best basis Selection”, IEEE Transactions on Information Theory, 38(2):713-718, 1992.

[10] R. Coifman, Y. Meyer, S. Quake, and V. Wickerhauser, Signal Processing and Compression with Wavelet Packets, Technical Report, Yale University, 1990.

[11] R. Coifman, Y. Meyer, and V. Wickerhauser, “Wavelet Analysis and Signal Processing”, Wavelets and Their Applications, Jones and Bartlett Publishers, Boston, 1992.

[12] M. V. Wickerhauser, Adapted Wavelet Analysis - From Theory to Software, A.K. Peters Ltd., 1994.

[13] I. Daubechies, Ten Lectures on Wavelets, Society For Industrial and Applied Mathematics, Philadelphia, PA, 1992.

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

[15] S.G. Mallat “A Theory for Multiresolution signal Decomposition: The Wavelet Representation”, IEEE Trans. On Pattern Analysis and Machine Intelligence, 11(7): 674-693, July 1989.

[16] W. Kim and Y. Hu, “Wavelet Packet-based Sub-band Coder” ICASSP, vol. V4, pp. 2225-2228, 1995.

[17] Z. Xiong, K. Ramachandran, M. T. Orchard and K. Asai, "Wavelet packet-based image coding using joint space-frequency quantization", IEEE Intl. Conf. On Image Processing, v3, 1994, IEEE, Los Alamos, CA, USA 944CH355708, pp. 324-328.

[18] P. Onno and C. Guillemot, "Wavelet Packet Coding with Jointly Optimized Lattice VectorQuantization and Data Rate Allocatio", IEEE International Conference on Image Processsing, V3, pp. 329-333, 1994.

[19] L. Kong and K. Lay, "Image Coding based on Energy-sorted Wavelet Packets", SPIE, vol. 2501/2, pp. 1279-1287, 1995.

[20] B. Kosko, *Fuzzy Engineering*, Prentice Hall, New Jersey 07458, 1997.

[21] P. Gader, J.M. Kellerand J. Cai, "A Fuzzy Logic System for the Detection and Recognition of Handwritten Street Numbers", IEEE Transaction On Fuzzy Systems, vol. 3, no. 1, pp. 24-31, Feb. 1996.

[23] S. Sasi and J.S. Bedi, "Character Recognition using Fuzzy Logic", Proc. 37th MWSCS-94, Lafayette, Lousiana, USA, Aug. 3-6, pp. 1399-1402, 1994.

[24] B. Kosko, *Neural Networks and Fuzzy Systems-A Dynamical System Approach to Machine Intelligence*, Prentice Hall, Englewood Cliffs, NJ, 1992.[25] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Macmillan College Publishing Company, New York, USA,1994.[26] S. Sasi, *Handwritten Character Recognition using Fuzzy, Neuro-fuzzy and Wavelet Transform Approaches*, Ph. D Dissertation, Wayne State University, 1997.

[27] A. Joshi et.al, “On Neurobiological, Neuro-fuzzy, Machine Learning and Pattern Recognition Techniques”, IEEE Transaction on Neural Networks, vol. 8, no.1, pp. 18-31, Jan.1997.