Fuzzy Hough Transform, Linguistic Sets and Soft Decision MLP for Character Recognition

Shamik Sural[†] and P.K.Das[‡]

[†]CMC Limited, 28 Camac Street, Calcutta 700 016, India.
[‡]Dept. of Computer Science and Engineering, Jadavpur University, Calcutta 700 032, India.

Keywords: Hough Transform, Fuzzy Sets, Multilayer Perceptron, Character Recognition

Abstract

We present a neuro-fuzzy system for character recognition from printed documents using a fuzzy Hough transform technique. For each character pattern, fuzzy Hough transform extracts information from the standard Hough transform accumulator cells into a number of fuzzy sets. These basic fuzzy sets are combined by t-norms to synthesize additional fuzzy sets whose heights form an n-dimensional feature vector for the pattern. A 3n-dimensional fuzzy linguistic vector is then generated from the n-dimensional feature vector by defining three linguistic fuzzy sets *weak*, *moderate* and *strong*. A multilayer perceptron (MLP) is trained with the fuzzy linguistic vectors by back propagation of errors. The MLP outputs represent fuzzy sets denoting belongingness of an input feature vector to different fuzzy character pattern classes. Recognition accuracy of the system is more than 98% for single font documents.

1. INTRODUCTION

Recognition of characters from scanned documents is a key step in document image processing. The two fundamental approaches to character recognition are feature classification and template matching. Multilayer perceptron and other neural networks are also often used for character recognition after they are trained with a set of standard patterns by supervised learning [1,2]. An advantage of using neural networks is that a high computation rate can be achieved through their massive parallelism, so that real time processing is possible. It may be mentioned that human beings are more efficient than computers in handling complex recognition problems including character recognition from document images. Human reasoning is somewhat fuzzy in nature which enables us to combine even visually degraded features in the brain using the millions of neurons working in parallel. Fuzzy sets have the ability to model vagueness and ambiguity in data which is encountered in character recognition. Our novel approach combines the robustness of feature extraction with the speed of operation of neural networks in a framework of fuzzy systems.

2. FUZZY HOUGH TRANSFORM

Hough transform [3,4] maps each black pixel point (x_i,y_i) in an image to the $(\rho-\theta)$ plane using the transformation $\rho = x_i\cos\theta + y_i\sin\theta$. The $(\rho-\theta)$ plane is quantized into quadruled grids forming a two-dimensional accumulator array. At the end of the transformation, each accumulator cell holds the number of points lying on a line represented by its $(\rho-\theta)$ parameter values. We define a number of fuzzy sets based on the values of ρ, θ and count of each $(\rho-\theta)$ cell for a character pattern. Two of these fuzzy sets, namely, *Long Line* (LL) and *Short Line* (SL) extract length information of the different lines in the pattern. *Nearly Horizontal* (HL), *Nearly Vertical* (VL) and *Slant Line* (TL) represent their skewness while *Near Top* (NT), *Near Bottom* (NB), *Near Vertical Centre* (NVC), *Near Right* (NR), *Near Left* (NL) and *Near Horizontal Centre* (NHC) extract the position of these lines. The 1-cuts i.e., the core of the fuzzy sets TL, NT, NB, NVC, NR, NL and NHC represent crisp features in a character pattern. The core of the fuzzy sets HL and VL denote strictly horizontal and strictly vertical lines, respectively, while the core of LL is a diagonal line. Support of the fuzzy sets LL and SL represent straight lines with all possible lengths in the pattern. The fuzzy Hough transform, thus, maps the characteristics of the different lines in an image into the properties of these fuzzy sets. Additional fuzzy sets are synthesized from the basic fuzzy sets using t-norms to represent each line as a combination of its length, position and orientation. These fuzzy sets are defined as *Long Slant Line* (LSL) $\equiv$ *i*(TL,LL), *Short Slant Line* (SSL) $\equiv$ *i*(TL,SL), *Nearly Vertical Long Line near Left* (VLL) $\equiv$ *i*(VL,LL,NL), etc., where *i* denotes a t-norm. A non-null support of a synthesized fuzzy set implies the presence of the corresponding feature in a pattern. We, therefore, choose the height of each synthesized fuzzy set to define a feature element and the set of 'n' such feature elements constitute an n-dimensional feature vector for a character. Thus, if $\overline{F}_1, \overline{F}_2, \ldots, \overline{F}_P$ denote P feature vectors derived from P input characters using fuzzy Hough transform, the feature elements of each pattern $\overline{F}_i$ are defined as $F_{i1} = h(LSL) = \underset{(\rho\text{-}\theta)}{\text{Sup}}\ LSL$, $F_{i2} = h(SSL) = \underset{(\rho\text{-}\theta)}{\text{Sup}}\ SSL$, etc., where Sup denotes Supremum of a fuzzy set.

3. MLP WITH FUZZY INPUT/OUTPUT

When an n-dimensional feature vector, as mentioned above, is extracted from a degraded character pattern for recognition, the strength of the features in the vector may vary due to the presence of noise. To combat the effect of noise, we generate membership values in three linguistic fuzzy sets, namely, *weak*, *moderate* and *strong* from the individual feature elements. The linguistic set membership functions are derived from the Butterworth filter transfer functions [5] as shown below.

$$\mu_{weak}(x) = \left[1 + \left(\frac{x}{a}\right)^{2n}\right]^{-\frac{1}{2}}$$

$$\mu_{moderate}(x) = \left(\left[1 + \left(\frac{x}{a_1}\right)^{2n}\right]\left[1 + \left(\frac{a_2}{x}\right)^{2n}\right]\right)^{-\frac{1}{2}}$$

$$\mu_{strong}(x) = \left[1 + \left(\frac{a}{x}\right)^{2n}\right]^{-\frac{1}{2}}$$

The membership value is 0.7 for x = a (a_1, a_2 for $\mu_{moderate}$), the cut-off point, for all values of 'n' where 'n' controls the slope of the functions. The 3n-dimensional vectors form the MLP input both during training and recognition. The advantage of using linguistic features is that, for small variations in the extracted feature values, the linguistic set memberships remain unchanged. The system can then recognize even degraded character patterns.

In a conventional MLP, an input pattern belongs only to a particular output pattern class. We, however, use fuzzy character pattern classes as outputs and the MLP is trained to learn the degree by which a feature vector belongs to each of these classes. For a P-class problem domain with P nodes in the output layer of the MLP, the Euclidean distance between each input vector $\overline{F}_i$ and other feature vectors is calculated as follows.

$$d_{ik} = \sqrt{\sum_j \left(F_{ij} - F_{kj}\right)^2} \qquad k = 1,2,\ldots, P.$$

Here the summation is done over all the feature elements subscripted by j. The membership of the ith character pattern to the kth fuzzy pattern class, and hence the value of the kth expected output of the MLP for the input vector $\overline{F}_i$ is then determined using the following relation.

$$O^i_{k(exp)} = \mu_k\left(\overline{F}_i\right) = \left[1 + \left(\frac{d_{ik}}{f_{den}}\right)^{f_{pow}}\right]^{-1}$$

Here 'f_{den}' and 'f_{pow}' control the membership grades of the different output fuzzy sets for each input pattern. For all the fuzzy class membership functions, $\mu_k\left(\overline{F}_i\right) \in [0,1]$, $\mu_k\left(\overline{F}_i\right) = \mu_i\left(\overline{F}_k\right)$, $\mu_k\left(\overline{F}_k\right) = 1$ and $d_{ik} \geq d_{il} \Rightarrow \mu_k\left(\overline{F}_i\right) \leq \mu_l\left(\overline{F}_i\right)$. Distance measures other than the Euclidean distance may also be considered in a similar manner. The MLP is trained with the input fuzzy feature vectors and fuzzy expected outputs by the standard back propagation algorithm [6]. The error between the actual output and the expected output is minimized through updation of weights, initially set with random values.

The recognition decision of the MLP is based on the α-cuts of the output fuzzy sets for α = th_opt, a threshold value for the outputs. For a fixed value of the parameters 'f_{den}' and 'f_{pow}' if th_opt is low, the α-cuts contain more than one element while a high value of th_opt results in null α-cuts for some of the outputs. Conversely, in the presence of noise in a test character, one of the MLP outputs goes above the threshold and the others take on low values. In this case, the highest value output is considered to be the unknown character. If, however, the membership value is above the threshold for more than one output, indicating a possibility of misclassification, a dictionary search is used to uniquely identify the character. The advantage of using fuzzy set membership functions at the MLP output is that, only outputs with high membership values need to be considered for the search. Since the MLP outputs denote their similarities to the different pattern classes, this decision making process is justified. The fuzzy feature based MLP has been found to be stable for different initial random values of the inter-layer connector weights. The feature set has been chosen after performing a sensitivity analysis of the MLP output for different fuzzy features using the neural network based technique proposed in [7]. A block diagram of the complete fuzzy character recognition system is shown in Fig. 1.

4. RESULTS

The fuzzy OCR system has been implemented for character recognition using a 300 dpi HP scanner. The linguistic set parameter 'n' and the MLP output parameter 'f_{pow}' are typically chosen as 4 and 0.87, respectively. The ratio d_{ik} / f_{den} is kept less than unity. The character recognition efficiency is more than 98% for single font documents when the system is trained with a particular font. The recognition rate is expected to be higher if other features including curves are also extracted from the character patterns to form the fuzzy feature vectors. When trained with more than one font, the recognition efficiency of the system is in the range of 96-98% for multifont documents.

5. CONCLUSIONS

The advantage of fuzzy sets is harnessed in two stages in the proposed fuzzy character recognition system. The fuzzy Hough transform does not reject any feature since thresholding is not done on the Hough transform accumulators cells. All the pattern features are, therefore, retained for decision making at a higher level. The output fuzzy sets of the MLP enable the system to consider characters which are suitable candidates for final selection from word level knowledge. Here also, we do not use the *winner-take-all* logic of crisp perceptrons. Character features other than those mentioned here, can be extracted using the Hough transform for curve extraction. The analysis for the generalized fuzzy Hough transform will be similar in nature with new fuzzy sets defined for the extra features. The linguistic set definitions and the fuzzy MLP can be used with fuzzy feature extraction techniques other than the fuzzy Hough transform proposed here. Besides character recognition, the fuzzy Hough transform technique is useful for a variety of other pattern recognition problems also. Fuzzy features like *thick* lines, *thin* lines, *nearly parallel* lines, lines *slightly above* or *slightly below* other fixed lines similar to those presented in [8] may be extracted from document as well as non-document images using the fuzzy Hough transform. The present work can be extended to include fuzzy rule based systems to combine features extracted by fuzzy Hough transform instead of using fuzzy MLP.

REFERENCES

[1] B. Hussain and M.R.Kabuka, "A novel feature recognition neural network and its application to character recognition", IEEE Trans. Pattern Anal. Mach. Intell., Vol. 16, 98-106, 1994.

[2] S. Sural and P.K.Das, "A document image analysis system on parallel processors", Proc. Fourth International Conference on High Performance Computing, Bangalore, India, 527-532, 1997.

[3] J. Illingworth and J.Kittler, "A survey of the Hough transform", Computer Vision, Graphics and Image Processing, Vol. 44, 87-116, 1988.

[4] S.Y.Yuen and C.H.Ma, "An investigation of the nature of parametrization for the Hough transform", Pattern Recognition, Vol. 30, 1009-1040, 1997.

[5] J.Millman and C.C.Halkias, "Integrated Electronics: Analog and Digital Circuits and Systems", McGrawhill, Singapore, 1972.

[6] D.E.Rumelhart, G.E.Hinton and R.J.Williams, "Learning internal representation by error propagation", in: D.E.Rumelhart and J.L.McClelland, Eds., "Parallel Distributed Processing : Explorations in the microstructure of cognition, Vol. 1 : Foundations", Chapter 8, MIT Press, 1986.

[7] R.K.De, N.R.Pal and S.K.Pal, "Feature analysis : Neural network and fuzzy set theoretic approaches", Pattern Recognition, Vol. 30, 1579-1590, 1997.

[8] R.Krishnapuram, J.M.Keller and Y.Ma, "Quantitative analysis of properties and spatial relations of fuzzy image regions", IEEE Trans. Fuzzy Systems, Vol. 1, 222-233, 1993.