Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Image and Vision Computing 25 (2007) 230–239
image AND vision COMPUTING
www.elsevier.com/locate/imavis

Fuzzy technique based recognition of handwritten characters

R.M. Suresh *, S. Arumugam

[a] Department of Computer Science and Engineering, RMK Engineering College, RSM Nagar, Gummidipoondi Taluk Thiruvallur, Kavaraipettai 601206, Tamil Nadu, India

[b] Arulmigu Kalasalingam College of Engineering, Anand Nagar, Krishnankoil 626 190, Tamil Nadu, India

Received 6 May 2004; received in revised form 1 November 2005; accepted 31 January 2006

Abstract

The different methods for automatic pattern recognition are motivated by the way in which pattern classes are characterized and defined. In this paper, the handwritten characters (numerals) are preprocessed and segmented into primitives. These primitives are measured and labeled using *fuzzy logic*. Strings of a character are formed from these labeled primitives. To recognize the handwritten characters, conventional string matching is performed. However, the problem in this string matching has been avoided using the *membership value* of the string. This result is being compared with the *Modified Parser* generated from the *Error-free fuzzy context-free grammar.*

Keywords: Fuzzy logic; Fuzzy context-free grammar; Preprocessing; Polygonal approximation; Segmentation; Labeling; Handwritten numerals; Modified parsing algorithm

1. Introduction

A lot of research effort has been dedicated to handwritten character recognition. Number of schemes are available for this purpose. Some of the areas where the handwritten character recognition is being carried are fuzzy methods [2]. Knowledge-based techniques [11] and neural networks [3–7]. The different methods for automatic pattern recognition are motivated by the way in which pattern classes are characterized and defined [1]. The idea in syntactic pattern recognition is to describe a complex pattern in terms of a hierarchical composition of simple sub-patterns [25]. In syntactic pattern recognition a basic set of primitives forms the terminal set of grammar. The pattern class is the set of strings generated by the pattern grammar. But the concept of formal grammar is too rigid to be used for the representation of real-life patterns such as handwritten documents.

This rigidity can be changed if certain fuzziness is introduced which describes the vagueness of such patterns. Accordingly a fuzzy language can handle imprecise patterns when the indeterminacy is due to inherent vagueness [16]. The conventional approaches to knowledge representation usually lack the means to represent the imprecise concepts. Due to Zadeh [18], fuzzy sets offer a theoretical basis to cope with the vagueness of patterns, which we exploited in the proposed method. First the motivation for this method is given. This is followed by the inference of fuzzy context-free grammar and its Inference method. Then the case study with results has been discussed.

2. Motivation of this method

In recent years, some of the development tools in fuzzy software and hardware such as fuzzyclips [15], FUNN-Lab [12] have been introduced. These tools provide a convenient way to configure the membership functions, defining rules, input and output functions, etc. But they are not suitable for highly structured patterns recognition. The symbolic and structural description of a pattern is more useful for analysis and recognition [14]. The allograph-based method to recognize cursive handwritten words with fuzzy logic has been proposed by Parizeau et al. [13]. The drawback of this method is that, there is no direct way of generating handwriting features all graphs automatically. Malaviya et al. [11] have proposed FOHDEL a new fuzzy language to automatic generation of a pattern description in a rule-base and the representation of patterns in a linguistic form. The problem rest in this method is that, the large number of input features make the rule-base incomprehensible and consumes more time to recognize. The theory of fuzzy grammars and quantitative fuzzy semantics

* Corresponding author. Tel.: +91 4119 225338; fax: +91 4119 225193.
E-mail address: rmsuresh@hotmail.com (R.M. Suresh).

0262-8856/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2006.01.029

[17] give very interesting ideas like the connection between context-free grammar and natural grammar through transformational grammar and the derivation trees (structural descriptions or pattern markers). The idea here is to identify primitives using fuzzy logic and label them in the form of a compact fuzzy language (Strings). The labeled strings are parsed using a modified parser algorithm to recognize the pattern.

The purpose of this paper is to offer the system, which infers a complete error-handling fuzzy context-free grammar (FCFG) from samples and manipulates fuzzy languages as sets of trees and parse them.

3. Fuzzy context-free grammar and its inference

It appears that much of the existing formal grammars can be extended quite readily to fuzzy grammars. We introduce the concept of fuzzy grammar and fuzzy context-free grammar.

3.1. Fuzzy grammar

Definition 1. Informally fuzzy grammar may be viewed as a set of rules for generating the elements of a fuzzy language. A fuzzy grammar or simply a grammar is a quadruple:

$G = (V_N, V_T, P, S)$ where

V_N a set of non-terminal symbols,
V_T a set of terminal symbols, $(V_T \cap V_N = \phi)$
P a set of fuzzy production rules,
S an initial non-terminal symbol. $(S \in V_N)$

Essentially, the elements of V_N are labels for certain fuzzy subsets of V_T^* called fuzzy syntactic categories with S being the label for the syntactic category sentence. The elements of P define conditioned fuzzy sets in $(V_T \cup V_N)^*$. The concept of fuzzy grammar and fuzzy context-free grammar can be referred in [16,17].

3.2. Fuzzy context-free tree language

Definition 2. Fuzzy context-free language L(G) is defined as the language whose elements are completely accepted by the FCFG. For example

Let

$$G = (\{V_N = S, A, B\}\{V_T = a, b\}$$

$$\{P = S \to AB, A \to AB, B \to BA, A \to a, A \to b, B \to a,$$

$$B \to b\}\{S\})$$

then it can be represented in tree format as

$$\lambda \xrightarrow{0.8} \underset{A\ B}{\overset{S}{\wedge}} \quad \underset{A\,B}{\overset{A \xrightarrow{0.6} A,}{\wedge}} \quad \underset{B\ A}{\overset{B \xrightarrow{0.9} B}{\wedge}},\ A \xrightarrow{0.5} \underset{a}{\overset{A}{|}},\ A \xrightarrow{0.5} \underset{b}{\overset{A}{|}},\ B \xrightarrow{0.5} \underset{a}{\overset{B}{|}},\ B \xrightarrow{0.5} \underset{b}{\overset{B.}{|}}$$

Suppose there is a string abb in language L, then it is acceptable by the FCFG. Then the string tree transition is

$$\lambda \xrightarrow{0.8} S(A\ B) \xrightarrow{0.5} S(A(a)\ B) \xrightarrow{0.9} S(A(a)\ B(B\ A)) \xrightarrow{0.5} S(A(a)\ B(B(b)\ A)) \xrightarrow{0.5} S(A(a)\ B(B(b)\ A(b)))$$

The membership value of this tree term is predicted according to the definition 1 as 0.5.

It appears that much of the existing formal grammars can be extended quite readily to fuzzy grammars [16]. The formalization of grammatical inference includes the following phases:

(i) *Hypothesis space* is a subsets of the general rewriting systems-context-free grammars; (ii) the *measure of adequacy* is that the grammar inferred generates all of the known strings in the language L and none of the known non-strings; (iii) the *rules* by which the samples are drawn (iv) the *criterion for success* in the limit of the inference process and the *source of information* about a language L is an information sequence which is an infinite sequence of strings from the obtained set; A routine has been developed [19–22] for the automatic generation of samples. From the generated samples the strings are obtained; from the strings the grammar is inferred.

3.2.1. Method

The method employs straight lines to build-up the prototype character patterns. For the prototype, the end points of the segments are specified as the coordinates of the end point pixels in a normalized frame. The various instances are obtained by various coordinate specifications with different length. The specification for the prototype can be made as in the following function with reference to Fig. 1. The parameters represents the starting vertex (coordinate), the number of pixels in length, the next vertex (coordinate) and the direction using which the line to be drawn. The directions are horizontal (0), vertical (1), right slant (2), and left slant (3), respectively. The function to draw the picture KA is

DRAW L (1,10,2,0),(2,5,3,0), (2,6,4,1),(4,8,9,0),(9,3,10,2), (10,3,11,3), (4,6,5,3), (11,16,6,0),(6,6,7,2) (7,14,4,0),(1,7,8,1).

Using such sample characters the strings are obtained and the grammar is inferred. The algorithm for the fuzzy context-free grammar inference is given below.

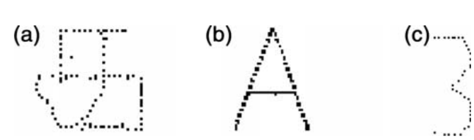

Fig. 1. Characters obtained from DRAW function (a) Tamil character KA (b) English alphabet A. (c) Numeral 3.