PARALLEL ALGORITHMS FOR EXTRACTING SHAPE FEATURES OF HANDWRITTEN CHARACTERS

H.F. Li, M. Youssef and R. Jayakumar

Department of Computer Science
Concordia University
Montreal, Quebec, Canada H3G 1M8

Abstract

Parallel algorithms to extract shape features of handwritten characters based on edge classification are presented. The start and end points of the edges defining the boundary of a character are first detected in parallel. These edges are sent into the second stage of a pipeline to form edge chains for characterizing the corresponding shapes of a character pattern. The edge chains are identified in parallel using a two phase routing (routing update and data routing) of messages. These parallel feature extraction algorithms are simulated on a sequential computer and a classification scheme for recognizing handwritten numerals is developed. It is noted that the characters can be recognized with very high accuracy even if the characters are rotated by an angle between -15 and +15 degrees.

1. Introduction

Recent advances in hardware technology has made possible the development of special-purpose hardware based character recognition systems. The first step in developing such systems is to design parallel algorithms suitable for VLSI implementation. In this paper some parallel algorithms for extracting features from handwritten characters are proposed.

Various techniques for recognizing handwritten characters have been reported in the literature. These techniques recognize characters by (1) extracting global and statistical features [1,2] or (2) extracting geometrical and structural features [3-6]. However, the latter approach has the advantage of less sensitivity to rotation and are more amenable to natural variations in handwritten characters. Hence, this paper considers extracting the shape features based upon the approach proposed by Ahmed and Suen [5,6].

2. The Approach for Extracting Shape Features

Consider a binary pattern having M rows and N columns of pixels each of them can be either white or dark. Such a pattern is surrounded by a sequence of edges where an <u>edge</u> is a transition from a white region to a dark one or vice-versa, while scanning the pattern. For example, Fig. 1 shows a binary pattern having 11 rows and 12 columns and the edges surrounding this pattern are shown in Fig. 2. In the following it is assumed that the pattern is scanned horizontally.

An edge has a <u>start point</u> and an <u>end point</u>. These start and end points can be ranked such that points occurring in i-th row of the pattern have lower rank compared to points occurring in j-th row if and only if $i < j$. There are basically four types of edges, depending upon whether they occur along the inner periphery or the outer periphery of a pattern: (1) outer left (e^1), (2) outer right (e^2), (3) inner left (e^3) and (4)inner right (e^4). These edge types are exemplified in Fig. 2. The edges in a pattern can be ranked such that an edge having a start point with lower rank than that of another edge is assigned a lower rank than the latter one. A relation can be defined between a pair of edges meeting at a start or end point. The relations between the edges in Fig. 2 are also shown in the figure. Ahmed and Suen [5,6] has derived 14 different relations between a pair of edges meeting at a start or end points. Using the edges and the relations, simple holes and cavities can be detected in a pattern [5,6]. Furthermore, the edges form closed <u>chains</u> in and around the pattern (see Fig. 2). These chains can be used to detect complex holes and cavities in a pattern [5,6]. Ahmed and Suen proposed an approach [5,6] in which the edges along with the relation between pairs of them and the chains in a pattern can be used as shape features in recognizing handwritten characters.

In this paper, systolic algorithms to extract these shape features from a