

Learning to Segment Cursive Words using Isolated Characters*

Jean-François Hébert, Marc Parizeau
Laboratoire de vision et systèmes numériques
Université Laval
Ste-Foy (Qc), Canada, G1K 7P4
{jfhebert,parizeau}@gel.ulaval.ca

Nadia Ghazzali
Dept. de mathématiques et de statistique
Université Laval
Ste-Foy (Qc), Canada, G1K 7P4
ghazzali@mat.ulaval.ca

Abstract

This paper presents a new strategy for isolating handwritten characters in cursive words without making an explicit a priori segmentation of the script, and without imposing any lexical or other linguistic constraints. Furthermore, this approach can be completely trained using data sets of isolated characters only.

The main idea behind this strategy is to have a window of attention moving around in the cursive word, searching for instances of known characters. If one assumes that the current window contains some significant part of a character, then the problem is to translate and scale the window of attention in such a way that it converges to the bounding box of that character. This process is implemented using both a detector network and a set of locator networks. The detector network is responsible for recognizing whole characters of any class and thus for stopping the iterative process, whereas a locator network is assigned the task of recognizing the crucial parts of a given character class and producing the corresponding transformation parameters for the window. The feasibility of this process is shown through experiments using the UNIPEN database of on-line scripts.

1 Introduction

The automatic recognition of on-line cursive script is a difficult pattern recognition problem [1, 2], mainly because of the great variations encountered in different handwriting styles, but also because of the so-called segmentation/recognition dilemma where characters need both to be segmented before they can be recognized, and recognized before they can be segmented. Many recognition systems have been designed to tackle these difficulties using various methods and techniques. Some of the more recent works include the use of neural networks [3, 4, 5], hidden markov models [6, 7] and formal grammars [8].

However, most of these systems deal with the segmentation/recognition dilemma by, on the one hand, over-segmentation of the scripts in smaller unit than characters (usually graphemes) and analysis of all possible segmentation paths, and, on the second hand, by imposing lexical constraints in order to limit combinatorial explosion. The object of this paper is to present a somewhat different strategy where no explicit a priori segmentation is made and where no lexical constraints are a priori imposed.

The main idea behind this strategy, is to be able to initialize a window of attention somewhere near a character in a cursive word, and have a segmentation process fine tune the position and size of this window in such a way that it searches for the bounding box of this character. By repeating this process on a sufficiently large set of initial points, one could then build a segmentation graph for the cursive word which could be interpreted using adjacency constraints [10], lexical knowledge, or both. It is hoped that using such a strategy, a complete and flexible cursive script recognition system can be built, although no claims to such a system is made in this paper (work on this is still under way). What is claimed in this paper, however, is the feasibility of this new segmentation strategy and that the underlying process can be trained automatically using only isolated characters.

In previous work, Suen et al. [9] have already demonstrated that handwritten alphanumeric characters are made up of *crucial parts that tend to preserve the invariant characteristics and exhibit the distinctive features of characters*. It is our objective to exploit these crucial parts in order to create a more flexible character segmentation/recognition algorithm.

The rest of this paper is organized as follows. First, Section 2 presents a global overview of our segmentation process. Then, Section 3 to Section 6 describe in details each of its four components. Finally, experimental results are presented and discussed in Section 7.

*This work was supported in part by NSERC and FCAR grants to M. Parizeau and in part by an NSERC grant to N. Ghazzali.

