International Journal of Innovative
Computing, Information and Control
Volume **7**, Number **8**, August **2011**

 ISSN 1349-4198
pp. **4955–4964**

CURSIVE SCRIPT SEGMENTATION WITH NEURAL CONFIDENCE

Tanzila Saba, Amjad Rehman and Ghazali Sulong

Department of Computer Graphics and Multimedia
Faculty of Computer Science and Information Systems
Universiti Technologi Malaysia
Skudai 81310, Johor Malaysia
stanzila2@live.utm.my; { rkamjad; ghazalis }@utm.my

Received March 2010; revised July 2010

Abstract. *This paper presents a new, simple and fast approach for character segmentation of unconstrained handwritten words. The proposed approach first seeks the possible character boundaries based on characters geometric features analysis. However, due to inherited ambiguity and a lack of context, few characters are over-segmented. To increase the efficiency of the proposed approach, an Artificial Neural Network is trained with significant number of valid segmentation points for cursive handwritten words. Trained neural network extracts incorrect segmented points efficiently with high speed. For fair comparison, benchmark database CEDAR is used. The experimental results are promising from complexity and accuracy points of view.*
Keywords: Handwriting recognition, Character segmentation, Feature extraction, Character recognition, Back propagation learning

1. **Introduction.** An extensive research has been done in the field of handwriting recognition in the last few decades [1]. It seems that the research has been reached to its maturity for the recognition of isolated characters, hand printed words recognition, automatic address processing and bank check reading (holistic approaches) [2-4]. On the other hand, the recognition results for unconstrained cursive handwriting are still low due to the poor character segmentation. These segmentation errors mislead classifier during character recognition [5-7]. In fact, the segmentation problem has persisted for nearly as long as handwriting recognition problem itself. The recognition rate is highly dependent on the segmentation accuracy [11]. Hence, the segmentation is the backbone of the recognition process and still active research topic. Researchers have acknowledged the important role of segmentation that plays in handwriting recognition process [7,12,13]. That is why more innovative, accurate and fast methods are needed to be employed and compared with the work of other researchers using benchmark databases.

Generally, segmentation algorithms for unconstrained handwritten words can be classified into two categories: explicit and implicit segmentation [7]. In the first approach, letter boundaries are determined prior to recognition, while, in the latter, characters are recognized first dynamically by semantic analysis and then segmented [8-10]. However, character segmentation is byproduct.

In most of the existing segmentation algorithms, human writing is evaluated empirically to derive rules [15]. Although, the derived rules are satisfactory, and however, there is no guarantee for their optimum results in all writing style. This is due to the human writing style that varies from person to person and even for the same person depending on mood, speed, environment, etc. Hence, rules alone are not enough to segment characters in unconstrained handwriting. To overcome these limitations, researchers have employed