Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT®

PERGAMON

Pattern Recognition 38 (2005) 323–339

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Directed connection measurement for evaluating reconstructed stroke sequence in handwriting images

K.K. Lau*, Pong C. Yuen, Yuan Y. Tang

Department of Computer Science, Hong Kong Baptist University, Hong Kong

Received 28 May 2003; received in revised form 15 July 2004; accepted 15 July 2004

Abstract

Generally speaking, the performance of on-line handwriting recognition systems is better than that of off-line handwriting recognition systems. Many researchers then started to extract/estimate on-line/temporal information from a static handwritten image to improve the performance of off-line recognition systems. More specifically, various algorithms have been developed to extract the stroke sequence, which is an important temporal information, from a static handwritten image in the last decade. However, to the best of our knowledge, there are no methods to evaluate the performance of these algorithms. In view of this limitation, this paper presents a directed connection measurement for evaluating the performance of these methods. The measurement is designed in terms of direction and connection. The direction measurement is based on the consecutive arrangements of the items, while the connection measurement is based on both the number of disconnections and the Feigin and Cohen model (FCM) in ranking analysis. We replace the Kendall's distance, the key metric component in FCM, by a new connection metric. We have also proved that the new direction and connection metrics satisfy all axioms for being a valid metric. The two correlations are combined to form a directed connection correlation, which is applicable in comparing the performance of different stroke sequence recovery algorithms.

Keywords: Ranking analysis; Handwriting recognition; Stroke sequence estimation

1. Introduction

Handwriting recognition has been an active and challenging research area since the 1980s and attracts a lot of researchers working on this area. Basically, handwriting recognition can be divided into two problem domains, namely on-line and off-line. In off-line handwriting recognition, what we have is a static image. The typical applications include bank check processing and postal address recognition. Over the two decades of development, the performance has greatly improved. However, due to the large variations on writing images, the performance is still not fully satisfied. In on-line handwriting recognition, in addition to the static handwritten image, we have on-line writing information, such as stroke sequence, pressure, velocity and acceleration. Thus, on-line handwriting recognition system usually performs better than that of an off-line one. A very good review can be found in Ref. [1].

Govindaraju and Krishnamurthy [2] suggested that although each person has his/her own personal handwriting style, the basic pen-tip movements for two different persons writing the same script are close. Besides, Plamondon and Srihari [1] stated that the recovered temporal information may be helpful in improving the recognition performance. In view of this, some researchers started to think about the

* Corresponding author.
E-mail addresses: kklau@comp.hkbu.edu.hk (K.K. Lau), pcyuen@comp.hkbu.edu.hk (P.C. Yuen), yytang@comp.hkbu.edu.hk (Y.Y. Tang).

0031-3203/$30.00 © 2004 Published by Elsevier Ltd on behalf of Pattern Recognition Society.
doi:10.1016/j.patcog.2004.07.006

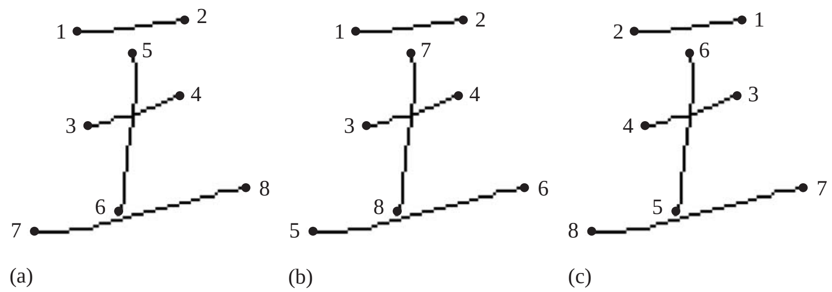

Fig. 1. The stroke sequences of a sample handwritten script. (a) Original stroke sequence. (b) Stroke sequence recovered by Algorithm 1. (c) Stroke sequence recovered by Algorithm 2.

recovery of the on-line information from a static handwritten image. Along this direction, more and more algorithms have been developed for recovering temporal information from a single-stroke or multiple-stroke handwritten images [3–10]. However, to the best of our knowledge, there are no algorithms to evaluate "how good" these algorithms perform. Of course, there are no problems in evaluating the algorithm if the estimated stroke sequence is the same as the original. However, in practice, the estimated stroke sequence is only partially matched with the original.

Take a simple handwritten script in Fig. 1 as an example. Fig. 1(a) shows a handwritten script with the labeled sequence, which refers to the original stroke sequence and is used as reference. It can be seen that the image consists of four strokes (represented by 8 points). Suppose there are two recovery algorithms, namely, Algorithm 1 and Algorithm 2. Their recovered sequences are shown in Fig. 1(b) and (c), respectively. This paper addresses the performance evaluation issue of these two algorithms by defining a new metric. The new metric will truly reflect the similarity between two stroke sequences, which is in line with human judgment.

The proposed new measurement for evaluating the stroke sequence estimation algorithm from a handwritten image is based on ranking analysis. More specifically, a new measurement is developed based on the classical Feigin and Cohen model (FCM) [11]. In FCM, the core metric component is the Kendall's distance [12]. However, Kendall's distance cannot be directly applied to our problem because it has not considered the direction and connection attributes, which are very important properties in handwritten scripts. Thus, we propose and develop a directed connection measurement, which is discussed in detail in Section 3. The preliminary version of this paper has been published in Ref. [13].

The rest of this paper is organized as follows. Section 2 briefly reviews ranking analysis. Section 3 reports our proposed directed connection measurement. Section 4 shows how the directed connection measurement is applied on evaluating the performance of stroke sequence recovery algorithms. Moreover, two well-known algorithms, namely Doermann and Rosenfeld's algorithm [5] and Plamondon and Privitera's algorithm [10], are selected for the evaluation. Doermann and Rosenfeld's stroke sequence recovery algorithm [5] includes local glues and global glues. The local glues include striations, end-point intensity variations, feathering, hooking, tramlines, and relative width and intensity. The global clues involve "hypotheses as to the ordering of strokes (left–right); the general top to bottom flow, energy conservation, and global consistency are applied to the stroke graph segments and junctions attributed with partial direction and ordering information". On the other hand, Plamondon and Privitera's stroke sequence recovery algorithm [10] "initially starts at the top-left corner of the image and continues down, scanning the picture by means of a series of 45° sloping trajectories". The conclusions are drawn in Section 5.

2. A brief review on ranking analysis

Ranking is a classical research topic and several well-known and popular ranking models have been developed. Daniels [14] classified the various types of ranking comparisons, and Feigin and Cohen [11] reorganized them into three classes: (I) parametric, (II) paired comparison and (III) sampling.

Class (II) from Ref. [11] is defined as follows: "each ranking is the result of the relevant judge making all of the $n(n-1)/2$ possible paired comparisons in a consistent manner." In this paper, we assume that a set of handwritten strokes is extracted, and we want to evaluate the performance of stroke sequence recovery algorithm against the correct sequence of the set of extracted strokes. Hence, this paper focuses on this class, but the analysis is concentrated in the direction and the connectivity between each pair of consecutive items.

The study of metrics and ranking models used in this class is classical. The earliest papers in this area can be dated back to Cayley's work in 1849 [15]. Many textbooks [16–18] provided detailed descriptions of various methods. This section gives a brief review of the works with the notations used by Critchlow [16].