8

Off-Line Roman Cursive Handwriting Recognition

Horst Bunke and Tamás Varga

8.1 Introduction

Automatic handwriting recognition has been a subject of research for more than 40 years [84]. On the one hand, the reading of human handwriting by machine has been considered an interesting and intellectually challenging problem in its own right. To approach, or even surpass, the performance of humans in text recognition has been a major driving force behind many research activities. On the other hand, the field has been quite important from the commercial and application-oriented point of view. Automatic address reading [88], bank cheque processing [43], and recognition of text filled in by hand on forms [25, 102, 105] have been major challenges in automatic handwriting recognition research. Moreover, handwritten data have often been used to validate and test the performance of new pattern classification methods.

Since the beginning of the 1990s, a significant growth of activities in handwriting recognition research has been observed. There is no doubt that enormous progress has taken place in this area. For example, for the tasks of handwritten address reading and amount recognition on bank cheques, commercial systems have become available [26]. Nevertheless, there is a clear need to further develop the field. All successful applications, for example, address and cheque reading, work in narrow domains with limited vocabularies, where task-specific knowledge and constraints are available. Examples are the relation between zip code and city name in address reading, or the redundancy of courtesy and legal amount on a cheque. However, when it comes to general word or sentence recognition where no constraints exist and one is faced with a large, possibly open lexicon, the state of the art is quite limited and recognition rates are rather low. Yet the problem of unconstrained word and sentence recognition is important in a number of

future applications, for example, the transcription of personal notes, faxes and letters, or the electronic conversion of historical handwritten archives in the context of the creation of digital libraries [4].

The field of handwriting recognition can be divided into on-line and off-line recognition. In the current chapter, we focus our attention on off-line recognition and consider only roman script. On-line handwriting recognition is the topic of Chapters 6 and 7, and for scripts other than roman see Chapters 3–5 and 10 in this book. Recent surveys on roman script recognition are [8, 72, 89, 96].

This chapter is organized as follows. In Section 8.2, the state of the art in roman cursive handwriting recognition is reviewed, including pre-processing of document images, and recognition of isolated characters or digits, individual words and unconstrained text. A few emerging topics such as databases for handwriting recognition, synthetic training data and multiple classifier systems are discussed in Section 8.3. Finally, Section 8.4 provides some conclusions and an outlook for future work.

8.2 Methodology

Roman cursive handwriting recognition can be divided into the tasks of recognizing isolated characters or digits, individual words and unconstrained text consisting of a sequence of an a priori unknown number of words. Recognition of isolated characters and digits is by far the simplest problem for which mature solutions have become available. The other two problems, word and word sequence recognition, are considerably more difficult and are still subject to research. In this section, we first review methods of text image pre-processing and normalization that are commonly found in any of the three tasks mentioned above. Subsequently, the recognition of isolated characters and digits, individual words and sequences of words is discussed.

8.2.1 Document Image Pre-processing

In the off-line mode an image of the handwriting to be recognized is captured by a sensor, for example, a scanner or a camera. Traditionally, the first processing step consists in converting the grey-level image acquired by the sensor to a binary image. A number of algorithms for this step are known from the literature (see Chapter 2 of [69], for example). However, with increasing processing speed and memory capacity of modern computers, the direct use of grey-level images is becoming increasingly more popular.

Before or after binarization, images are often filtered to remove noise. Popular methods of noise removal in both binary and grey-level images are based on image filtering theory and mathematical morphology (see [34], for example).