Simple Distances Between Handwritten Signatures

J.R. Parker
Laboratory for Computer Vision
Department of Computer Science
University of Calgary

Abstract

When analyzing handwritten signatures using a computer, a certain amount of variation within any particular class is to be expected. Successful recognition demands that this variation be less than that between two different signatures. This paper describes three simple ways to compare signatures that do not use any complicated or derivative feature measurements, each of which defines a distance between signatures that allows individual variation.

1. Introduction

The use of handwritten signatures to verify and individual's identity is older than photography. Common wisdom and practice has it that signatures are unique, and that human experts can match two signatures well enough for most practical purposes. A difficulty with automatic signature verification, the comparison of an input signature with a known file sample signature using computer algorithms, is that two signatures are never exactly the same, not even for the same person under identical circumstances. Thus, the signatures are never compared directly.

Signature verification is frequently performed by making various measurements on a digitized image[4,5,10] and subjecting these to a classifier, like a neural net or support vector machine[1,2,3]; the measurements are called *features*, and often a great many are used, creating a feature vector. The features are numerical values, and will vary somewhat between instances created by the same writer. The hope is that the features will vary more for two different writers (*between-class* variance) than for the same writer (*in-class* variance).

It has been claimed that signatures are written with a great deal of consistency, especially with respect to orientation and scale (for example [11]). This may be true if the writer is constrained to use a particular form, but is not sufficiently true in unconstrained cases to yield consistent sets of features for recognition. Figure 1 shows a set of 50 signatures drawn freehand on a Wacom data tablet. The variation can be seen clearly, but is difficult

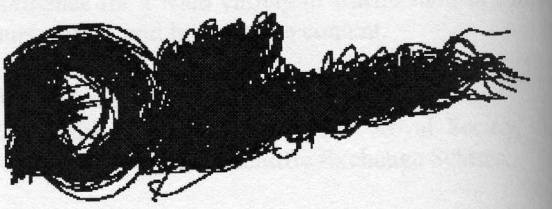

FIGURE 1. 50 signatures overlaid. The variation in position of related pixels is over 10%.

to measure in any really meaningful way. For example, the maximum distance between temporally related points (points the same time from the time of the first pen down) is 76 pixels, or 19% of the signature width. The maximum distance between any two closest points is 22 pixels (5.5% of width). Its not clear what any of this really means. or how it could be used to verify the identity of the writer.

What is clear is that if, as is claimed, in-class signatures are in some sense more similar to each other than they are to out-class signatures, then there is some obvious (at least to the human visual system) measure that minimizes the small variations normal to signatures and uses larger ones to make distinctions.

2. Measuring Distance

How can it be decided automatically that some variation from a template signature is natural, while another variation indicates that two distinct signatures are involved? Usually by looking at what 'distances' between signatures written by the same writer are typical, and comparing to the distances obtained between classes. By 'distance', we could mean any number of things, usually involving feature vectors and somewhat complicated measurements. The work described here is an effort to determine whether simple, relatively obvious distances can yield good results.

Again, many different things may be meant by 'relatively obvious'. What is meant here is distance will be