Multiple/Parallel Handprinted Digit Recognition

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

Abstract

The use of multiple classification algorithms applied to the same input can give higher accuracy classifications than using a single algorithm. This paper discusses three new schemes for classifying handprinted digits, and shows how to merge the results from the individual classifiers into a coherent single classification.

1. Introduction

There are two ways to use features to classify objects. In *statistical* approaches many features are combined into a large feature vector. Because features are measurements, the same object can correspond to a wide variety of feature vectors just through the errors in the measurements. However, these measurements will be clustered in some region of N-space, where N is the dimension of the feature vector. Hence, a statistical recognizer will construct a feature vector from a data object and classify it based on how far (Euclidean distance) it is in N-space from the feature vectors for known objects.

The basic idea behind *structural* pattern recognition is that objects are constructed from smaller components using a set of rules. Characterizing an object in an image is a matter of locating the components, which at the lowest level are features, and constructing some representation for storing the relationships between them. This representation is then checked against known patterns to see if a match can be identified. Structural pattern recognition is, in fact, a sophisticated variation on template matching, one which must match relationships between objects as well as the objects themselves.The problems involved in structural pattern recognition are two: locating the components, and finding a good representation for storing the relationships between the components.

Both methods have their strengths and weaknesses, and within each class of recognizer there are many variations of the basic theme. It is hoped that by combining many classifiers, some statistical and some structural, the resulting system will be more robust in the sense of yielding a correct classification in a wide variety of circumstances. To this end, we will examine briefly five methods for recognizing handprinted digits, and will look at methods for combining the results of all five into a single classification.

The author would like to acknowledge the support provided by the Natural Sciences and Engineering Research Council of Canada

2. Properties of the Character Outline

In a couple of interesting articles, Shridhar et al [4,10,11] describe a collection of topological features that can be used to classify hand printed numerals. Most of these features are properties of the outline, or *profile*, of the numeral. For instance, a digit '8' might be described as having a smooth profile on both the left and the right sides, and as having the width a minimum in the center region. Not all hand printed '8' digits would be recognized by this description, and certainly some other digits might also have this description; the idea is to provide a sufficient number of descriptions for each digit that a high recognition rate can be achieved.

Figure 1 shows how the left and right profiles are defined and calculated for a sample digit '9'. After the digit is isolated and thresholded, the number of background pixels between the left side of the character's bounding box and the first black pixel is counted and saved for each row in the bounding box. This gives a sampled version of the left profile (LP), which is then scaled to a standard size, in this case 52 pixels. A similar process gives the right profile (RP); the difference is that the last black pixel on each row is saved.

Having the profiles, the next step is to measure some of their properties. For example, one important property is the location of the extrema: L_{min} is the location of the minimum value on the left profile, and L_{max} is the location of the maximum value. of the actual peaks in LDIF and RDIF and their values happens to be quite important in characterizing numerals, and the peaks are located using a range rather than a single position. Thus, a digit '5' would have the RDIF peak near the top of the digit, and the peak would have a relatively large value. This set of features is not comprehensive. In all, 48 features are used and others could be defined. (See [10] for a complete list and description)

In the training phase all 48 features are computed for