An Evaluation of the WPE Algorithm Using Tangent Distance.

R. Paredes, E. Vidal and D. Keysers
Instituto Tecnológico de Informática,
Universidad Politécnica de Valencia, 46071 Valencia, Spain.
rparedes@iti.upv.es, evidal@iti.upv.es
keysers@I6.Informatik.RWTH-Aachen.DE

Abstract

Weighting Prototype Editing (WPE) is a novel approach to edit a given set of prototypes so that the resulting set can outperform the original one in terms of the Nearest Neighbor (NN) classification accuracy. This technique is applied in this work along with an interesting dissimilarity measure between pixel maps, known as Tangent Distance (TD). Experiments on the USPS handwriting digits benchmark corpus are presented, with results showing the capability of the WPE to improve the already good results based on TD NN classification.

Keywords: Editing, Condensing, Nearest Neighbour, Weighted Prototypes, Tangent Distance.

1 Introduction

The Nearest neighbor (NN) rule is a very common and successful approach for many pattern recognition applications. While the asymptotic optimality of this rule is well know [1], when the number of prototypes is not large enough performance can degrade dramatically. Unfortunately, this is quite often the case in real applications. One idea to circumvent this problem is the use of *Editing Techniques* [11, 10, 8, 2, 6, 3] which attempt to "clean" inter-class overlap regions, thereby leading to smoother NN-based decision boundaries between classes and hopefully increasing classification accuracy.

In [7] a new editing technique called *"Weighting Prototype Editing (WPE)"* was introduced [1]. Rather than aiming at asymptotically good performance as most editing techniques do, the WPE tries to obtain a good editing rule for *each given prototype set*. This is achieved by first learning an adequate assignment of a weight to each prototype and then pruning out those prototypes having large weights. As a result, WPE was expected to outperform other traditional editing techniques when the number of available prototypes is small. Moreover, since the prototype weights are explicitly optimized for each prototype set, performance was expected to be uniformly good for varying sizes and/or dimensionalities of the training sets of prototypes.

These expectations could be successfully confirmed in [7] throughout a series of experiments using common benchmark synthetic data sets. Moreover, as compared with *Wilson, MultiEdit* and *Cross-Validation Editing*, only WPE was actually able to achieve error rates consistently close to the corresponding Bayes bounds, despite significantly decreasing the number of prototypes and increasing the data dimension.

An interesting feature of WPE, observed in these experiments, is that the optimization algorithm tends to assign large weights not only to the prototypes laying on the inter-class Bayes confusion regions (as required for the editing mechanism), but also to prototypes which are deeply embedded into their corresponding Bayes acceptance regions. Correspondingly, by pruning prototypes with large weights, a certain degree of prototype *Condensing* is achieved along with the *Editing* effect initially aimed at.

We should emphasize that this combined Editing/Condensing effect is achieved by WPE with *complete independence of the* metric *adopted*. Therefore, it can be generally used to improve the results of many Pattern Recognition tasks for which good, may be sophisticated classification techniques are already available. If these techniques can be seen under a NN-based classification scheme then, no matter how complex (even *non-vectorial*) data representation is used, or how elaborate the metric to compare these representations is, WPE is easily applied. If the available training data contains confusing and/or redundant prototypes, the WPE can take care of removing the required prototypes such that the expected test-set error rates become lower.

In the present work WPE is applied to a real task for which good results have already been achieved using appro-

[1] C source code is available at: "http:\\www.iti.upv.es\ ~rparedes\wpedit"