Presented at the 44th Annual Conference of the
Society for Imaging Science and Technology
May 12-17, 1991, St. Paul, Minnesota pp. 475-478

050391

DECISION BOUNDARY FEATURE SELECTION FOR NON-PARAMETRIC CLASSIFIER[1]

Chulhee Lee and David A. Landgrebe

School of Electrical Engineering
Purdue University, W. Lafayette, IN
Tel:(317)494-3846, FAX:(317)494-6440

ABSTRACT

Feature selection has been one of the most important topics in pattern recognition. Although many authors have studied feature selection for parametric classifiers, few algorithms are available for feature selection for non-parametric classifiers. In this paper we propose a new feature selection algorithm based on decision boundaries for non-parametric classifiers. We first note that feature selection for pattern recognition is equivalent to retaining "discriminantly informative features" and a discriminantly informative feature is related to the decision boundary. Next a procedure to extract discriminantly informative features based on a decision boundary for non-parametric classification is proposed. Experiments show that the proposed algorithm finds effective features for the non-parametric classifier with Parzen density estimation.

I. INTRODUCTION

Linear feature selection can be viewed as finding a set of vectors which effectively represent an observation while reducing the dimensionality. The performance criteria for feature selection could be different depending on the application. In signal representation, a widely used criteria is mean square error. The Karhunen-Loeve transformation is one of the techniques under such a criterion and is optimum in the sense that the mean square error is minimum for a given number of features. In pattern recognition, however, it is desirable to extract features which are focused on discriminating between classes.

Although many authors studied feature selection for parametric classifiers, not many algorithms are available for non-parametric classifiers. Depending on the characteristics of data, there are some cases in which the use of non-parametric classifiers is desirable.

In this paper, we address this problem and propose a new algorithm for feature selection based on the decision boundary for non-parametric classifiers. The proposed algorithm predicts the minimum number of features to achieve the same classification accuracy as in the original space while at the same time finding the needed feature vectors.

---

1 This work was funded in part by NASA under grant NAGW-925.

II. DECISION BOUNDARY FEATURE SELECTION

In Lee and Landgrebe [1] we showed that discriminantly informative features and discriminantly redundant features are related to the decision boundary and can be extracted from the decision boundary. We also defined the decision boundary feature matrix as follows:

Definition 1. The decision boundary feature matrix (DBFM): Let $\mathbf{N}_i$ be the normal vector to the decision boundary at a point on the decision boundary for a given pattern classification problem. Let $\mathcal{R}_i$ be a collection of points on the decision boundary which have the same normal vector $\mathbf{N}_i$. Then the decision boundary feature matrix is defined as

$$\Sigma_{DBFM} = \sum_i \mathbf{N}_i \mathbf{N}_i^t W(\mathcal{R}_i)$$

$$\text{where } W(\mathcal{R}_i) = \frac{\text{area of } \mathcal{R}_i}{\text{total area of decision boundary}}$$

It was shown that the rank of the decision boundary feature matrix is the intrinsic discriminant dimension and the eigenvectors of the decision boundary feature matrix of a pattern recognition problem corresponding to non-zero eigenvalues are the necessary feature vectors to achieve the same classification accuracy as in the original space for the pattern recognition problem [1].

In general the decision boundary for non-parametric classifiers can not be expressed in analytic forms. Therefore the decision boundary for non-parametric classifiers must be calculated numerically. We propose the following procedure to calculate the decision boundary feature matrix for non-parametric classifiers. If there are more than two classes, the procedure can be repeated for each pair of classes. Fig. 1 shows an illustration of the proposed procedure.

Procedure for Feature Selection for Non-Parametric Classifier Utilizing Decision Boundary

(2 pattern class case)

STEP 1: Classify the training data

STEP 2: For each sample classified as class Ω_1, find the closest sample classified as class Ω_2. Repeat the same procedure for the samples classified as class Ω_2.

STEP 3: Connect the pair of samples found in STEP 2. Then the line connecting the pair of samples must pass through the decision boundary. By moving along the line, find the point on the decision boundary or near the decision boundary within a threshold.

STEP 4: At each point found in STEP 3, estimate the normal vector $\mathbf{N}_i$ by

$$\nabla h(\mathbf{X}) = \frac{\partial h}{\partial x_1}\mathbf{x}_1 + \frac{\partial h}{\partial x_2}\mathbf{x}_2 + \cdots + \frac{\partial h}{\partial x_n}\mathbf{x}_n$$

$$\approx \frac{\Delta h}{\Delta x_1}\mathbf{x}_1 + \frac{\Delta h}{\Delta x_2}\mathbf{x}_2 + \cdots + \frac{\Delta h}{\Delta x_n}\mathbf{x}_n$$

STEP 5: Calculate the decision boundary feature matrix using the normal vectors found in STEP 4.

$$\Sigma_{EDBFM} = \sum_i N_i N_i^t$$

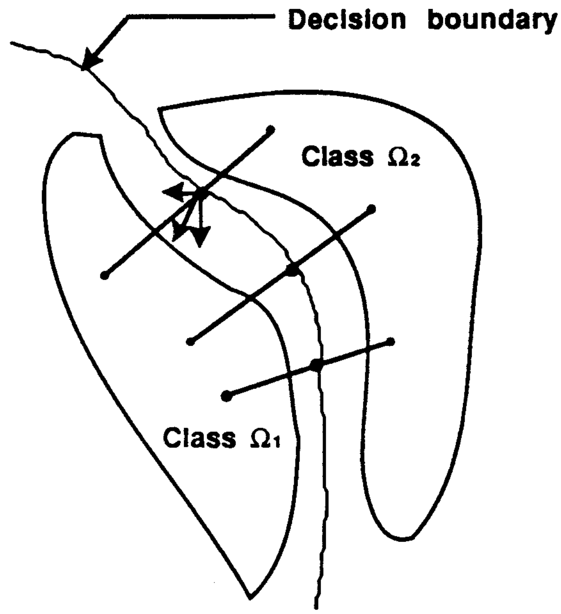

Fig. 1 Illustration of the procedure feature selection for non-parametric classifier utilizing decision boundary

III. EXPERIMENT AND RESULT

Four classes are chosen from the data collected at HAND CO. S. DAK. in July 26, 1978. TABLE I shows the description of the four classes. The data was collected as a part of the LACIE program [2].

TABLE I. Class Description

SPECIES	No. of Samples
SPRING WHEAT	518
NATIVE GRASS PAS	217
OATS	177
SUMMER FALLOW	204

The non-parametric classifier is modeled by Parzen density estimation. Two feature selection algorithms are tested and compared. The first one (Uniform Parzen) is a simple band combination procedure. For example, if the number of features is to be reduced in half, each two consecutive bands are combined to form a new feature. The second one (DBFM Parzen) is the proposed feature selection algorithm. 60 randomly chosen samples from each class are used for training and the rest of data are used for test. Fig.2 shows the performance comparison. The classification accuracy using all 10 features is about 70 percent. The proposed feature selection algorithm (DBFM Parzen) successfully achieves about the same classification accuracy with just one feature while the band combination procedure (Uniform Parzen) needs about 6 features to achieve the same classification accuracy.

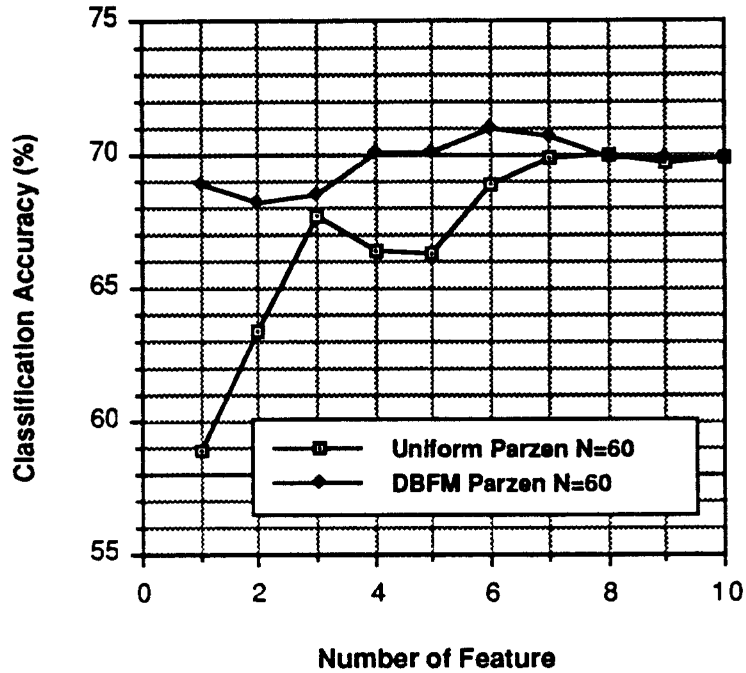

Fig. 2 Performance comparison of the non-parametric classifiers with Parzen density estimation

IV. CONCLUSION

The decision boundary feature selection is a powerful feature selection technique which can be used both for parametric and non-parametric classifiers. The method is robust and effective. In this paper we proposed an algorithm for non-parametric classifiers. Although the proposed algorithm was applied to the non-parametric classifier with Parzen density estimation, the algorithm can be used for any other classifier including parametric classifiers. The experiment shows that the performance of the proposed algorithm is promising.

With more use of non-parametric classifiers such as neural networks, the proposed algorithm can be successfully used for feature selection for high dimensional data and multi-source data.

REFERENCES

[1] C. Lee and D. A. Landgrebe, "Classification Algorithm for High Dimensional Data," Ph.D. dissertation in preparation.

[2] L. L. Biehl, M. E. Bauer, B. F. Robinson, C. S. T. Daughtry, L. F. Silva and D. E. Pitts, "A Crops and Soils Data Base For Scene Radiation Research," Proceedings of the Machine Processing of Remotely Sensed Data Symposium, West Lafayette, IN 1982, pp 169-177.