N-tuple Classifier Stacking

Robert M. Haralick

Computer Science, Graduate Center
City University of New York

Stacking Subspace Classifiers

- Stacking means having a subspace classifier composed of layers of subspace classifiers
- Each layer has fewer inputs
- The last layer has a small enough outputs that the classification can be made by a Bayesian Classifier
- This is analogous to neural net hidden layers

Stacking Example: Layer 1

Number of Dimensions	96
Number of Index Sets/Class	12
Size of Each Index Set	8
Number of Classes	3
Number of Scores Calculated	$12 \times 3 = 36$

Stacking Example: Layer 2

Number of Dimensions	36
Number of Index Sets/Class	9
Size of Each Index Set	4
Number of Classes	3
Number of Scores Calculated	$3 \times 9 = 27$

Stacking Example: Layer 3

Number of Dimensions	27
Number of Index Sets/Class	6
Size of Index Sets	$4, 4, 4, 5, 5, 5$
Number of Classes	3
Number of Scores Calculated	18

Stacking Example: Layer 4

Number of Dimensions	18
Number of Index Sets/Class	3
Size of Index Sets	6
Number of Classes	3
Number of Scores Calculated	9

Stacking Example: Layer 5

Number of Dimensions	9
Number of Index Sets/Class	2
Size of Index Sets	4, 5
Number of Classes	3
Number of Scores Calculated	6

Stacking Example: Layer 6

Number of Dimensions	6
Number of Possibilities per Dimension	6
Size of Measurement Space	$6^6 = 46,656$
Use Bayesian Classifier	

Requisite Variety Question

- In the Requisite Variety Analysis we did
  - We utilized mutually exclusive index sets
  - Small Overlapping of index sets should not change the analysis by much
  - Stacking may change the analysis
- For the Unstacked N-tuple Classifier We Wanted
  - For each class
  - The size of the training
  - to be more than 10 times
  - The size of all arrays storing class conditional probabilities

Requisite Variety Question

- For the case that all arrays store class conditional probabilities
  - Meaning of subsets of feature or scores
  - Where there is no optimization of values in the arrays
- Does the requisite variety criterion
  - Just apply
  - Class by class
  - To the arrays only in the first layer?
  - Or does it apply to all arrays in the stack?

Experiment

- What kind of experiment can be done
- To determine the requisite variety criterion for each class
- On whether it is
  - The size of the arrays for each class for the first layer
  - Or the size of all arrays for each class in all stack layers