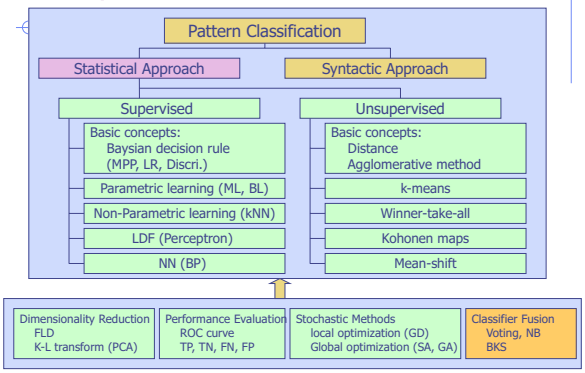
ECE 471/571 – Lecture 11

Classifier Fusion

Recap



Three heads are better than one.

Motivation

- Combining classifiers to achieve higher accuracy
  - Combination of multiple classifiers
  - Classifier fusion
  - Mixture of experts
  - Committees of neural networks
  - Consensus aggregation
  - …
- Reference:
  - L. I. Kuncheva, J. C. Bezdek, R. P. W. Duin, "Decision templates for multiple classifier fusion: an experimental comparison," *Pattern Recognition*, 34: 299-314, 2001.

Popular Approaches

- Majority voting
- Naive Bayes combination (NB)
- Behavior-knowledge space (BKS)
- Interval-based integration

4

Consensus patterns

- Unanimity (100%)
- Simple majority (50%+1)
- Plurality (most votes)

5

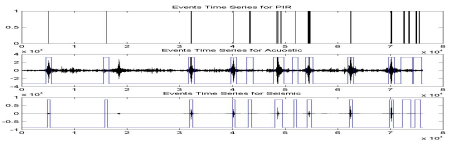
Example of Majority Voting - Temporal Fusion

- Fuse all the 1-sec sub-interval local processing results corresponding to the same event (usually lasts about 10-sec)
- Majority voting

$$\overline{\varphi_i^j} = \arg\max_{c} \omega_c, \quad c \in [1, C]$$

number of local output c occurrence

number of possible local processing results



NB (the independence assumption)

The real class is DW, the classifier says it's HMV

Confusion matrix

s

k

C1	AAV	DW	HMV
AAV	894	329	143
DW	99	411	274
HMV	98	42	713

C2	AAV	DW	HMV
AAV	1304	156	77
DW	114	437	83
HMV	13	107	450

i = 1, 2 (classifiers)

L1	AAV	DW	HMV
AAV			
DW			
HMV			

L2	AAV	DW	HMV
AAV			
DW			
HMV			

Probability that the true class is k given that C_i assigns it to s

Probability multiplication

7

NB – Derivation

- Assume the classifiers are mutually independent
- Bayes combination - Naïve Bayes, simple Bayes, idiot's Bayes
- Assume
  - L classifiers, i=1,..,L
  - c classes, k=1,...,c
  - s_i: class label given by the i^{th} classifier, i=1,...,L, s={s_1,...,s_L}

$$P(\omega_k \mid \mathbf{s}) = \frac{p(\mathbf{s} \mid \omega_k) P(\omega_k)}{P(\mathbf{s})} = \frac{P(\omega_k) \prod_{i=1}^{L} p(s_i \mid \omega_k)}{P(\mathbf{s})}$$

$$P(\omega_k) = N_k / N$$

$$p(s_i \mid \omega_k) = cm_{k,s_i} / N_k$$

$$P(\omega_k \mid \mathbf{s}) \propto \frac{1}{N_k^{L-1}} \prod_{i=1}^{L} cm^i_{k,s_i}$$

8

BKS

- Majority voting won't work
- Behavior-Knowledge Space algorithm (Huang&Suen)

Assumption:
- 2 classifiers
- 3 classes
- 100 samples in the training set

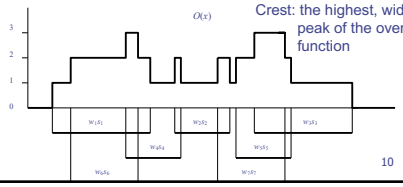
Then:
- 9 possible classification combinations

C_1, C_2	samples from each class	fused result
1,1	10/3/3	1
1,2	3/0/6	3
1,3	5/4/5	1,3
	...	
3,3	0/0/6	3

9

Value-based vs. Interval-based Fusion

- Interval-based fusion can provide fault tolerance
- Interval integration – overlap function
  - Assume each sensor in a cluster measures the same parameters, the integration algorithm is to construct a simple function (overlap function) from the outputs of the sensors in a cluster and can resolve it at different resolutions as required



Crest: the highest, widest peak of the overlap function

10

A Variant of kNN

- Generation of local confidence ranges (For example, at each node i, use kNN for each k∈{5,…,15})

	Class 1	Class 2	…	Class n
k=5	3/5	2/5	…	0
k=6	2/6	3/6	…	1/6
…	…	…	…	…
k=15	10/15	4/15	…	1/15
	{2/6, 10/15}	**{4/15, 3/6}**	…	**{0, 1/6}**

confidence level

confidence range

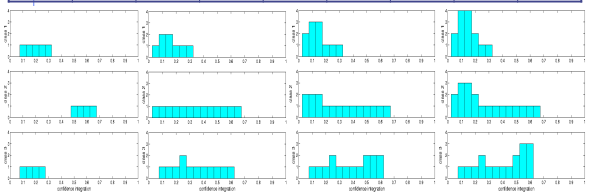
smallest largest in this column

- Apply the integration algorithm on the confidence ranges generated from each node to construct an overlapping function

11

Example of Interval-based Fusion

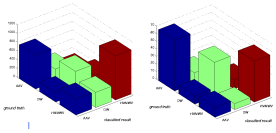
	stop 1		stop 2		stop 3		stop 4	
	c	acc	c	acc	c	acc	c	acc
class 1	1	0.2	0.5	0.125	0.75	0.125	1	0.125
class 2	2.3	0.575	4.55	0.35	0.6	0.1	0.75	0.125
class 3	0.7	0.175	0.5	0.25	3.3	0.55	3.45	0.575



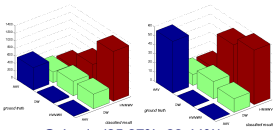
12

Confusion Matrices of Classification on Military Targets

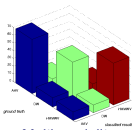
	AAV	DW	HMV
AAV	29	2	1
DW	0	18	8
HMV	0	2	23



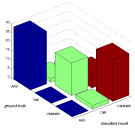
Acoustic (75.47%, 81.78%)



Seismic (85.37%, 89.44%)



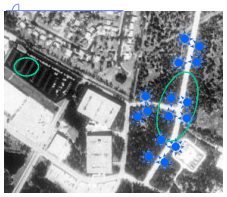
Multi-modality fusion (84.34%)

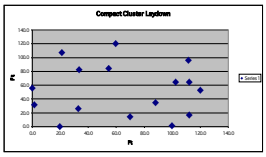


Multi-sensor fusion (96.44%)

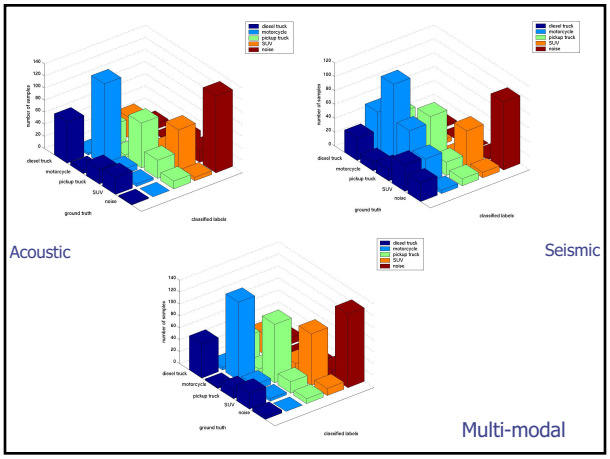
13

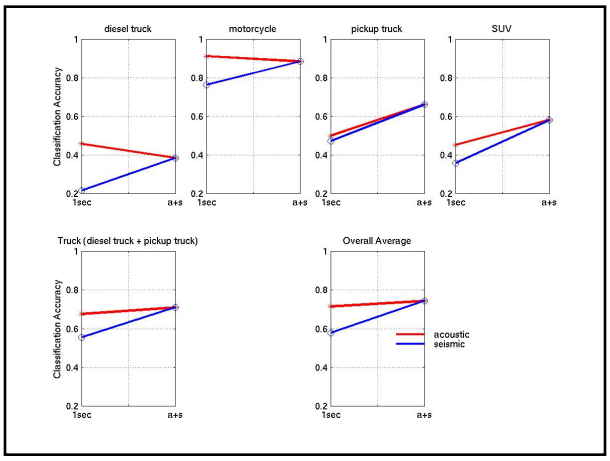
Civilian Target Recognition





Ford 250 Harley Motocycle Ford 350 Suzuki Vitara





Reference

- For details regarding majority voting and Naïve Bayes, see

http://www.cs.rit.edu/~nan2563/combining_classifiers_notes.pdf

17