



Automatic recognition of Bird Species by Their Sound

Seppo Fagerlund

Masters Thesis Presentation

3.11.2004





Outline

- Introduction
 - Bird sounds
 - Recognition of bird species
 - Segmentation
 - Feature extraction
 - Simulation results
 - Conclusion
- 



Introduction – motivation

- Automatic recognition of bird species would probably have significant impact to development of research methodology in biology
 - Research in many areas in biology requires currently lots of work in the field by researchers and volunteers
 - Automatic monitoring would reduce this need
 - There is probably also commercial interest for such system because bird watching is popular hobby in many countries
- 

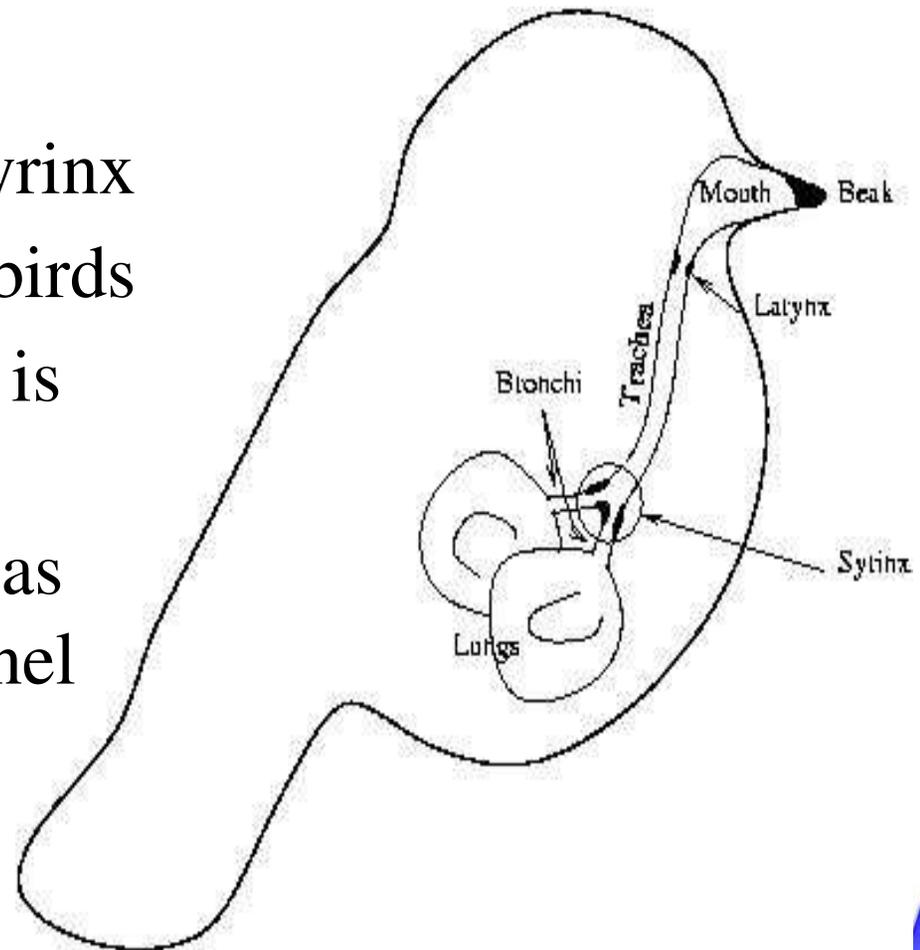


Introduction

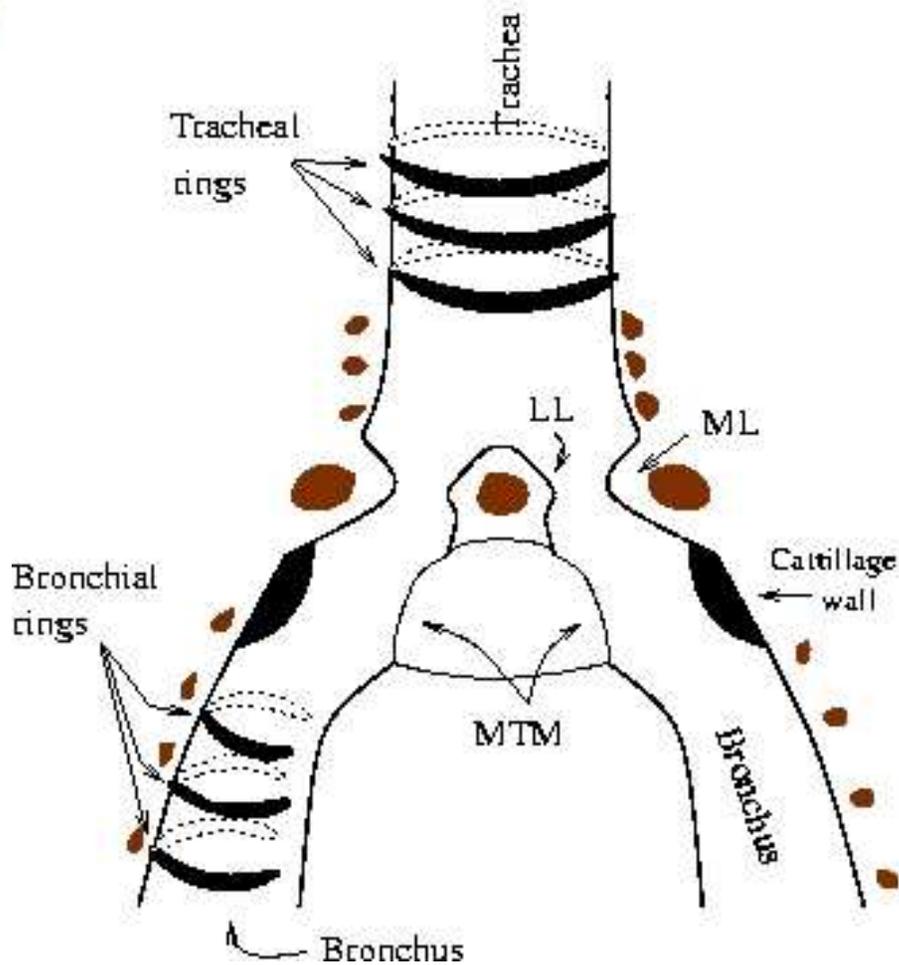
- Approx 9000 bird species in the world, in Finland 443 natural species
 - Spectrum of different sounds birds are able to produce is large
 - Bird sounds are divided by the function and by hierarchical level
 - Recognition of bird species is typical pattern recognition problem
 - In this work focus is in inharmonic bird sounds
- 

Bird sounds – production mechanism

- Avian vocal tract
- Main source of sound is syrinx
- Syrinx is unique organ to birds
- Diversity between species is large
- Postsyringeal part of tract as resonator and transfer channel



Bird sounds – production mechanism - the syrinx

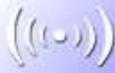


- Location of main sound source

- Bronchi
- Trachea
- Junction of bronchi and trachea

- Sound is produced by:

- Tympaniform membrane (MTM)
- Two soft tissues similar to human vocal folds (LL and ML)



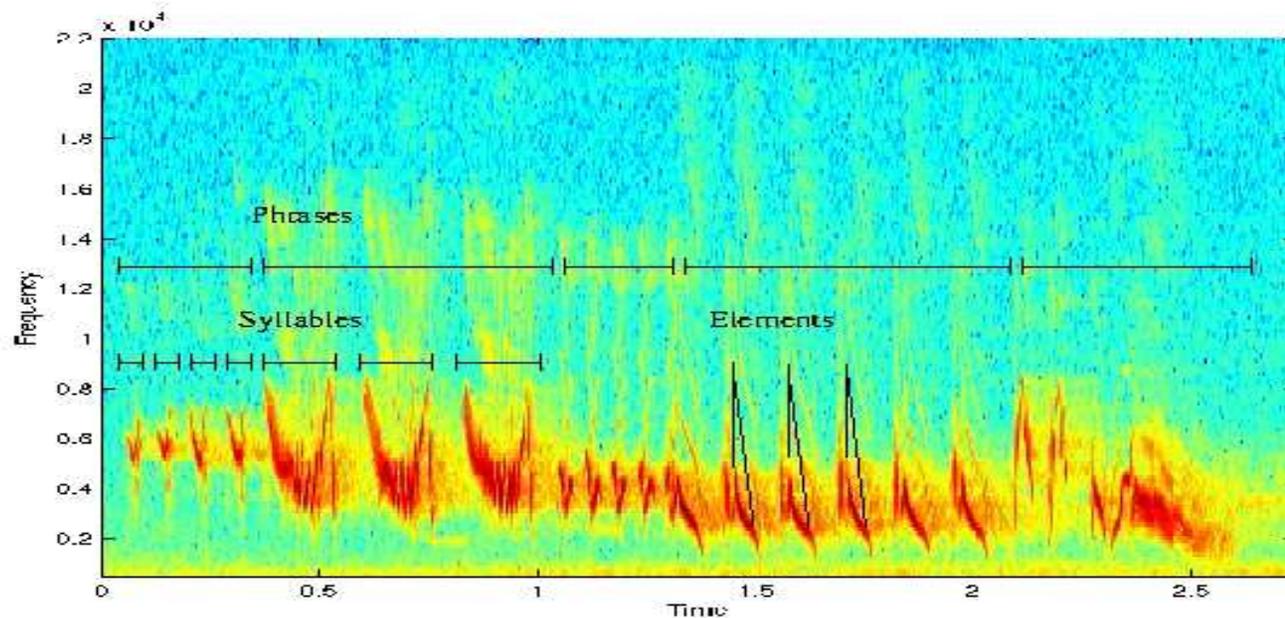
Bird sounds – organization

- Bird sounds are divided into songs and calls
- They are divided into hierarchical levels of phrases, syllables and elements
- Automatic detection of syllables more accurate than detection of elements
- Phrases and songs (and calls) include more regional and individual variability
- Thus syllable is suitable unit for recognition of species



Bird sounds – Organization

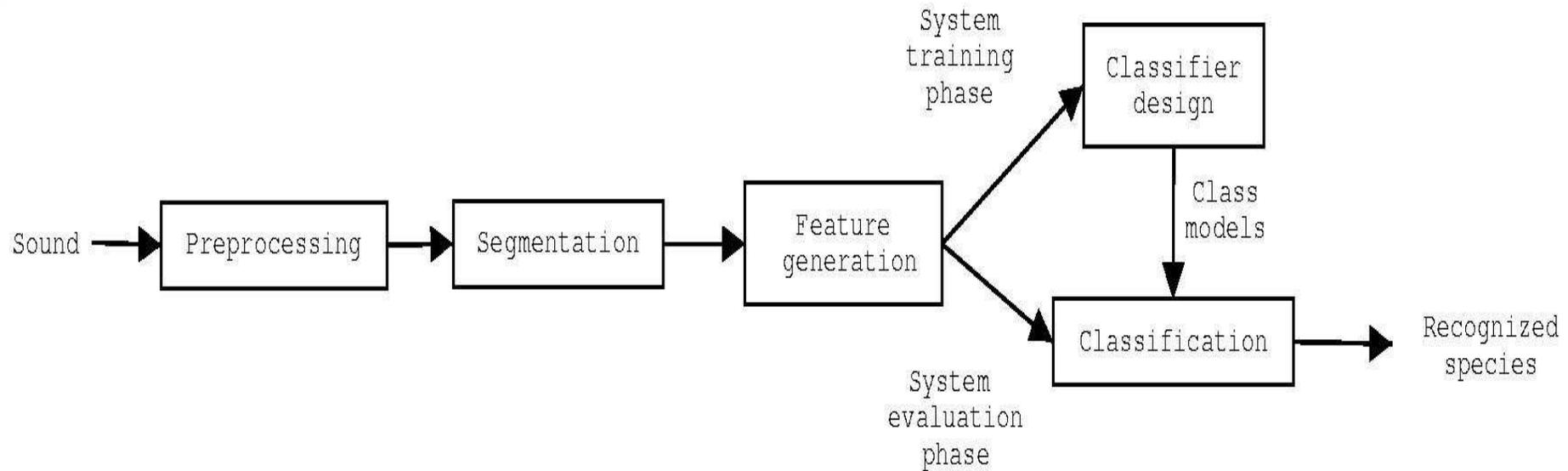
- Hierarchical levels of bird sounds: song/call, phrase, syllable and element or note
 - element is smallest separable unit in spectrogram
 - syllables are produced by one or more elements
 - series of syllables in particular pattern is a phrase



Song of Common
Chaffinch (Peippo)

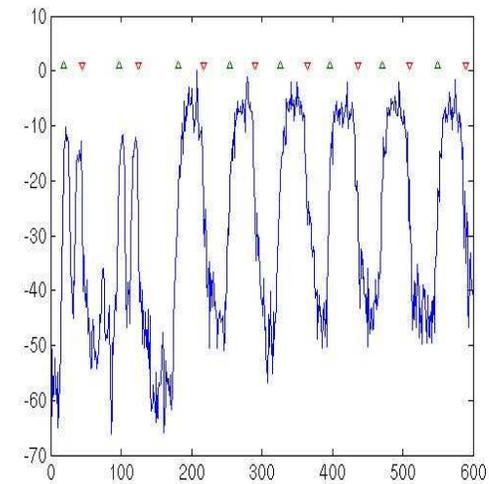
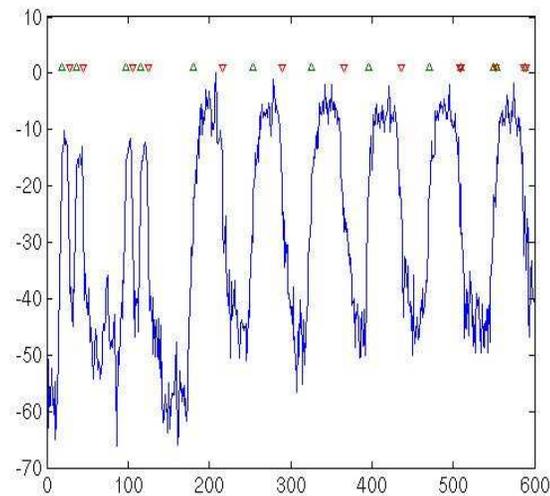
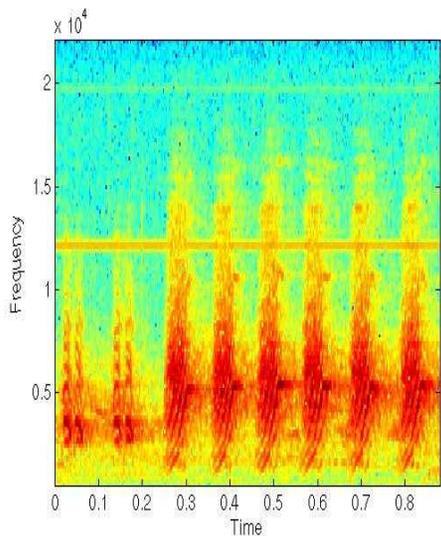
Recognition of bird species – overview of classification system

- Typical pattern recognition problem



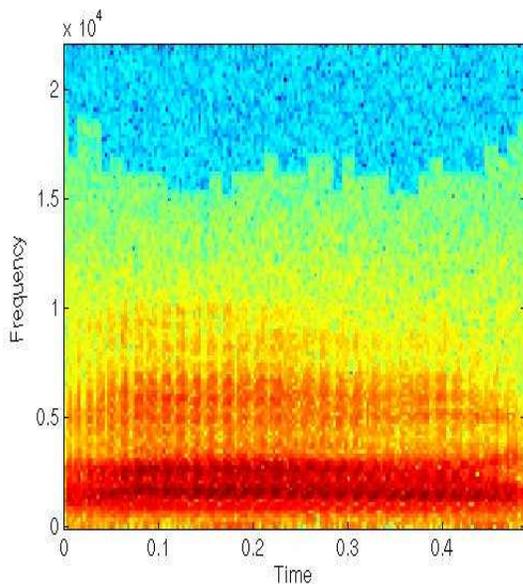
Recognition of bird species – segmentation

- Based on short-time signal energy
 - Overlapping frames of size 128 samples
 - Energy on frames -> Energy envelope of syllable
 - Iterative noise estimate to set threshold for syllables
 - Merging and omitting syllable candidates



Recognition of bird species – feature extraction (data reduction)

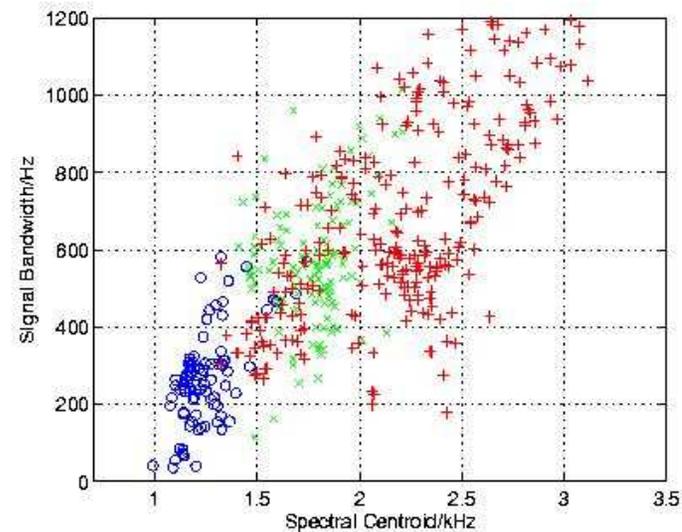
- Syllables are represented with features
- 19 low-level acoustical signal parameters
- Classes map to different position in feature space



Feature vector:

$$x = \begin{bmatrix} 0.1820 \\ 0.0312 \\ 0.2619 \\ 0.0006 \\ -0.0005 \\ \vdots \end{bmatrix}$$

Feature space
(3 species; 2-dim):



Recognition of bird species – classification

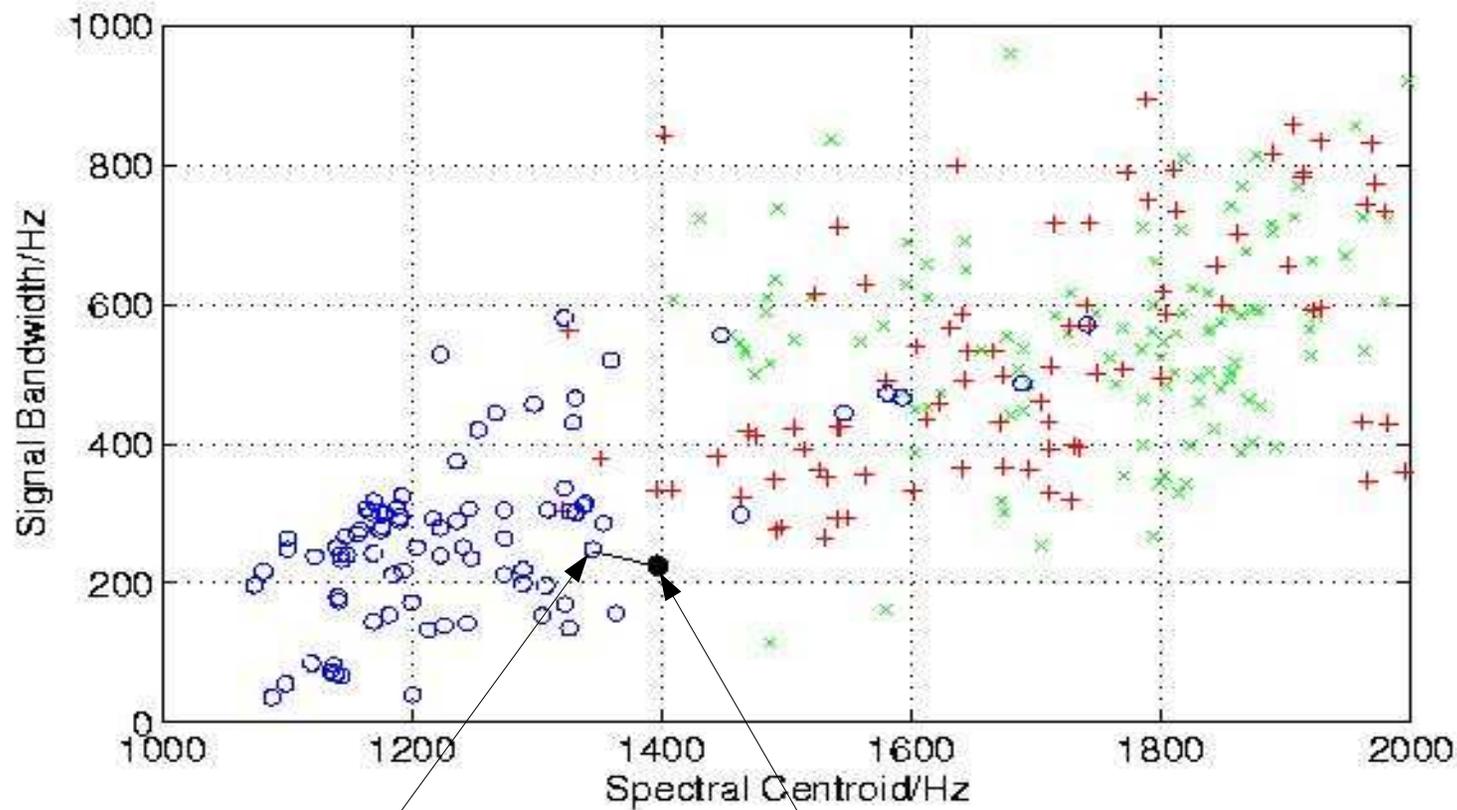
- Patterns (syllables) are assigned to classes (species)
- Training and testing data sets by leave-k-out method
- K-Nearest-Neighbour (k-NN) classifier
 - Simple to implement, flexible, does not need clustering
 - Computationally demanding, favour classes with high distribution density
- Neighbourhood based on distance measure:
 - Euclidean and Mahalanobis distance measures

$$d_E(x, y) = \sqrt{(x - y)^T (x - y)}$$

$$d_M(x, y) = \sqrt{(x - y)^T \Sigma (x - y)}$$

Σ =covariance matrix of
training data

Recognition of bird species – classification-1-NN



Nearest neighbour

Test sample

Recognition of bird species – simulation results

- Species:

Common name	Latin abbreviation	Finnish name
Common Raven	CORRAX	Korppi
Hooded Crow	CORNIX	Varis
Magpie	PICPIC	Harakka
Eurasian Jay	GARGLA	Närhi
Sedge Warbler	ACRSCH	Ruokokerttunen
March Warbler	ACRRIS	Luhtakerttunen

- Recognition results (1-NN):

Euclidean distance (49%)

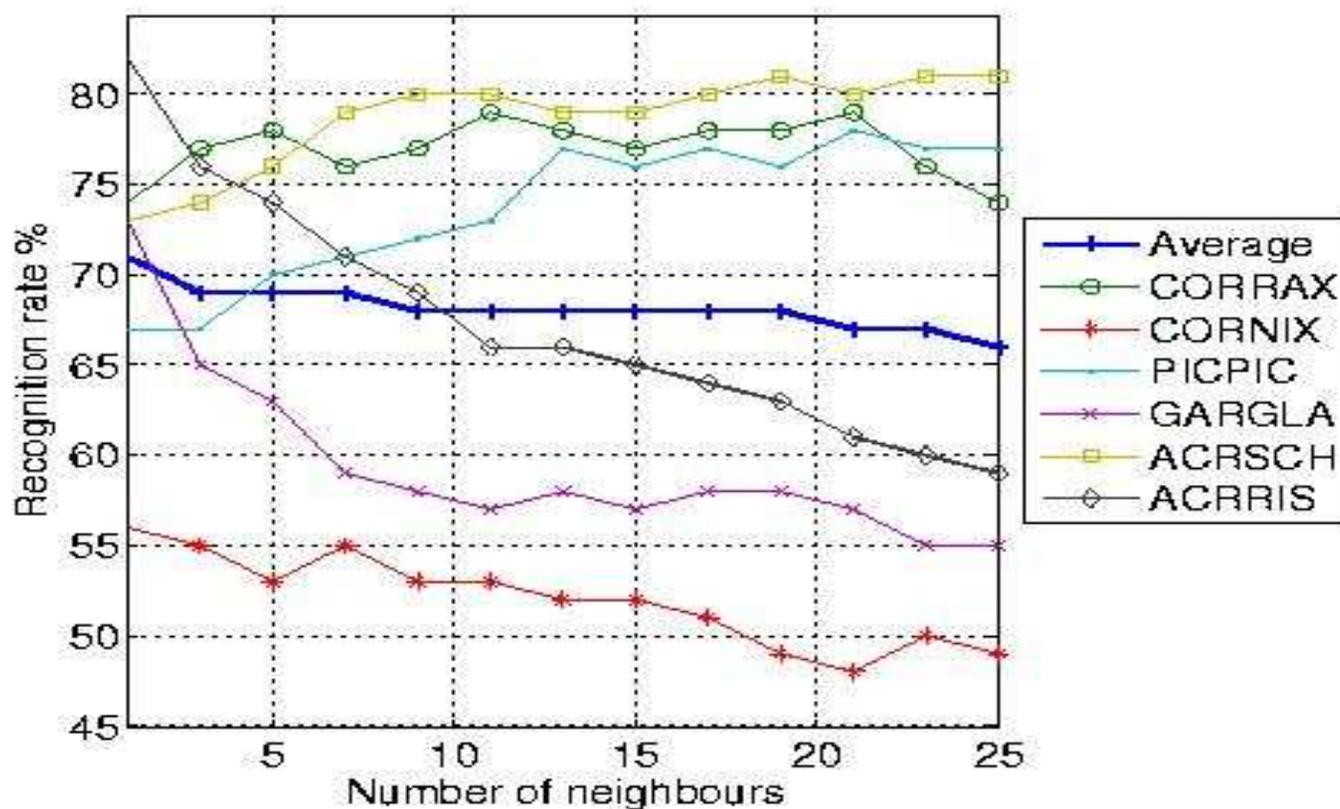
	CORRAX	CORNIX	PICPIC	GARGLA	ACRSCH	ACRRIS
CORRAX	69	14	4	3	0	0
CORNIX	19	36	24	7	0	3
PICPIC	10	36	41	41	7	12
GARGLA	2	7	16	36	5	3
ACRSCH	0	1	6	4	56	29
ACRRIS	0	5	10	8	32	53

Mahalanobis distance (71%)

	CORRAX	CORNIX	PICPIC	GARGLA	ACRSCH	ACRRIS
CORRAX	74	5	5	0	0	0
CORNIX	10	56	12	21	2	1
PICPIC	14	28	67	5	4	5
GARGLA	0	9	7	73	0	1
ACRSCH	0	1	2	0	73	10
ACRRIS	2	2	6	1	23	82

Recognition of bird species – simulation results

- Results as function of number of neighbours





Conclusion and future work

- Recognition tested for one type (inharmonic) of bird sounds
 - Recognition results relatively good with low-level features and simple classifier
 - Biggest challenge in future is in large number of classes and different sounds
 - Lots of work in all stages of the recognition system is needed
- 



Questions?

