TECHNICAL UNIVERSITY
OF CLUJ-NAPOCA

Signal Processing Group

On the Design of an Acoustic Based Wildlife Intruder Detection System

Corneliu Rusu, Marius Ghiurcau, Radu Bilcu

Marius Sirbu, Jaakko Astola

-SpeD 2013-

Contents

- Introduction
- On the Design of System
- Theoretical Background
- Practical Work
- Results
- Conclusions
- Final Remarks

Introduction

- detect intruders in wildlife regions
- motivation: ensure integrity of wildlife
- monitoring system -> acoustic eye
- standalone systems must be utilized
- the intrusion of people for surveillance should be limited
- low complexity algorithms and hardware with low power consumption
- fast and consistent development of wireless communication systems
- sensor networks

Introduction

- performance of the sensor networks depends on several factors:
  a) sensors accuracy
  b) network topology
  c) communication system planning
  d) power source autonomy period
- A combination of feasible and solid communication system planning and design as well as usage of advanced signal processing techniques at the receiver side will give a strong component on the effective implementation of a sensor network

On the Design of System

- proposed sensor network consists of several remote sensors
- collect the information and transmit it to a central collection knot of the network via low data rate - low power radio protocols
- fast Internet connection enables the sending of collected data to a central database.
- signal processing algorithms in order to extract valuable information
- the smart sensors need to satisfy at least two conditions:

  a) low power consumption

  b) redundant communication mode

On the Design of System

- appropriate solution might be the ZigBee protocol
- Zig-Bee transmitter is sending data only when programmed to do it or when the data is available
- it sends the collected data on short data bursts and in rest it is functioning on a sleeping mode with very low power consumption
- we are collecting data from processes that change very slight over the time: presence of human voice, artificial noise, temperature, humidity etc.
- a ZigBee protocol seems the most convenient to be used as a communication link

On the Design of System

- the particular network topology implementation
- each sensor communicates with the neighbor sensors
- if sensor is not working properly, the communication ways that are routed through that sensor are diverted to another one and therefore only the defected sensor data will not be collected
- this topology is giving an automatic diagnosis of the sensors health
- if one sensor is unable to communicate with its neighbors, then these neighbor sensors will signal to the collecting knot this situation

On the Design of System

- information that can be used in order to:
  a) detect human presence in a wild environment
  b) detect artificial noises (chainsaw or boat noise)
  c) monitoring of environmental condition (temperature, humidity, rain volume, wind strength)
  d) detection of climatic conditions for natural disasters - flooding, hail storm
- it is not feasible to use advanced signal processing at the sensors level
- a standalone system with very low complexity and low power consumption

Theoretical Background

Features of sound:

- Mel-frequency cepstral coefficients (MFCC)
- DELTA coefficients

Sound classification:

- Time Encoded Signal Processing and Recognition (TESPAR)
- Gaussian Mixture Models (GMM)
- Support Vector Machines (SVM)

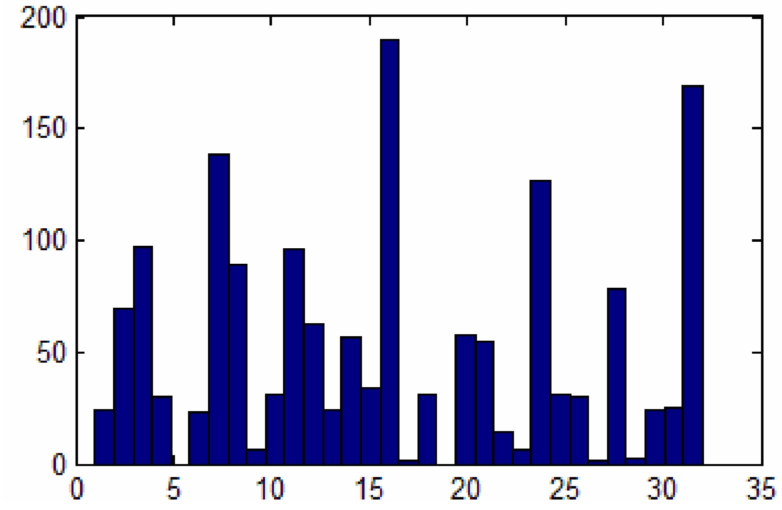
Theoretical Background

- TESPAR (Time Encoded Signal Processing and Recognition)
- uses the zero-crossings -> the infinite clipping theory
- simplest implementation of a TESPAR coder -> use of 2 descriptors for each segment:
  - *duration* (epoch) between successive real zeros (D)
  - *shape* between successive real zeros (S)
- for each D/S pair -> assign a symbol (using an alphabet)
- array of symbols

Theoretical Background

S Matrix

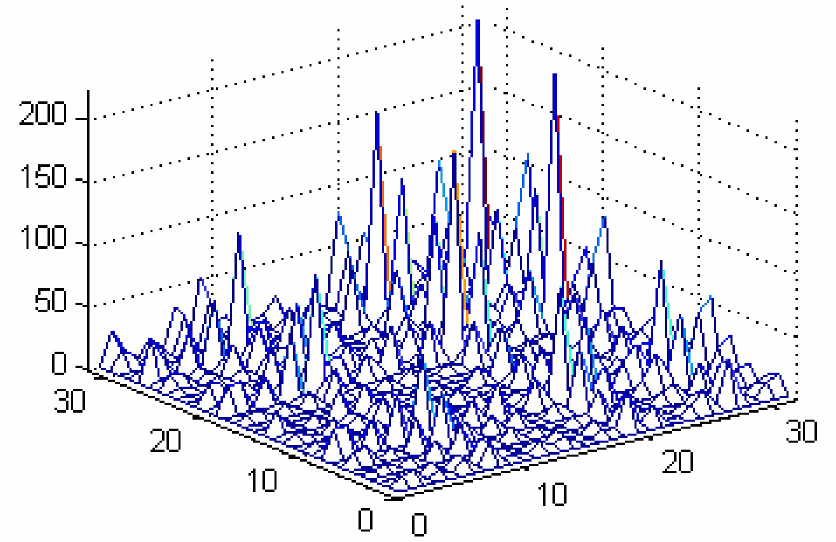
- symbols: 4,6,1,24,2,4,1,4,18,9,4,…



Theoretical Background

A Matrix

- symbols: 3,7,12,8,3,4,12,1,7,21,3,9,12, 3,12,11…
- n=2

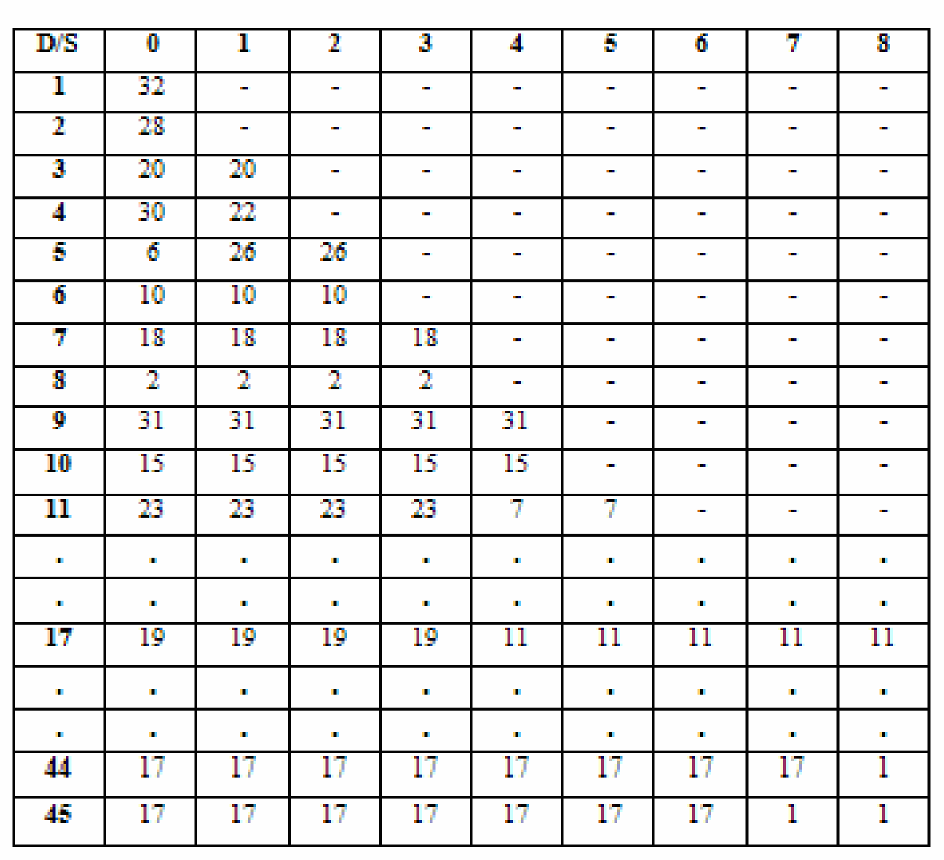


Theoretical Background

- **Archetypes**
  - normalize matrixes
  - adding together and then averaging matrixes (mean)
  - cross-validation
  - training set and validation set
- **Classification process**
  - new sound sample -> TESPAR coding process
  - generate S and/or A matrix
  - compare with archetypes (city block distance)
  - closest <=> winner (closed set)

Practical Work

- **Alphabet**
  - Linde-Buzo-Gray VQ
  - 100.000 D/S pairs
  - 32 symbols
  - coding table



D/S	0	1	2	3	4	5	6	7	8
1	32	-	-	-	-	-	-	-	-
2	28	-	-	-	-	-	-	-	-
3	20	20	-	-	-	-	-	-	-
4	30	22	-	-	-	-	-	-	-
5	6	26	26	-	-	-	-	-	-
6	10	10	10	-	-	-	-	-	-
7	18	18	18	18	-	-	-	-	-
8	2	2	2	2	-	-	-	-	-
9	31	31	31	31	31	-	-	-	-
10	15	15	15	15	15	-	-	-	-
11	23	23	23	23	7	7	-	-	-
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
17	19	19	19	19	11	11	11	11	11
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
44	17	17	17	17	17	17	17	17	1
45	17	17	17	17	17	17	17	1	1

- **Database**
  - 4x100 recordings

  (birds, human, car, animal)

Practical Work

- **Band-pass filters**
  - 10th order band pass Butterworth filters
  - low -> 70 Hz and high -> 3 kHz
  - new alphabet and archetypes
- **Downsample**
  - from 8 kHz to 6 kHz / 4 kHz
- **Noise add**
  - rain
  - wind
  - AWGN
  - training -> clean sounds

Results Study 1

- **Classification rates**
  - 95.33% (S) and 97.33% (A)
- **Classification rates for BPF**
  - less than 1% decrease for S
  - 1.33% increase for A
- **Noisy environments:**
  - AWGN - 74% (S) and 84% (A)
  - rain - 74.66% (S) and 89.33% (A)
  - wind - 71.33% (S) and 88.66% (A)
  - 6 kHz: decrease with 2-3%

S matrix	Human	Bird	Car
Human	97.9	0	2.1
Bird	0	93.3	6.7
Car	5.2	0	94.8

Table 1

A matrix	Human	Bird	Car
Human	98.2	0	1.8
Bird	0	93.8	6.2
Car	0	0	100

Table 2

M. V. Ghiurcau, C. Rusu and R. Bilcu, *Wildlife intruder detection using sounds captured by acoustic sensors*, in Proceedings of ICASSP 2010, Dallas, USA, March 2010.

Results Study 2

- clean sounds - improvement less than 2% for S, less than 1% for A
- for rain: 94% both A and S (Table 1)
- wind: 91.33% (A and S - Table 2)
- AWGN - increase, less than 1% for A, greater for S

A/S matrix	Human	Bird	Car
Human	**87.6**	**2.2**	**10.2**
Bird	**0**	**94.4**	**5.6**
Car	**0**	**0**	**100**

Table 3: Rain

A/S matrix	Human	Bird	Car
Human	**99.2**	**0**	**0.8**
Bird	**11.5**	**86.8**	**1.7**
Car	**2.1**	**9.9**	**88**

Table 4: Wind

M. V. Ghiurcau, C. Rusu and R. Bilcu, *A modifed TESPAR algorithm for wildlife sound classification*, in Proceedings of ISCAS 2010, Paris, France, June 2010.

Results Study 3

- **Gaussian Mixture Models**
  - clean sounds: 99.66% (97.33% previously)
  - rain: 97.33% (94% TESPAR)
  - wind: 93.33% (89.33% TESPAR)

- **4 databases**
  - clean sounds: 96.50%
  - DR=1
  - FAR=0

	Bird	Car	Animal	Human
Bird	94	0	6	0
Car	0	100	0	0
Animal	8	0	92	0
Human	0	0	0	100

Table 5: Confusion matrix

M. V. Ghiurcau, C. Rusu, J. Astola and R. Bilcu, *Towards an application for detecting intruders in wildlife regions*, in Proceedings of EUSIPCO 2010, Aalborg, Denmark, August 2010.

Results Study 3

- **Gaussian Mixture Models**
  - rain: 86.50%
  - wind: 82.25%
  - DR=0.975
  - FAR=0.01

	Bird	Car	Animal	Human
Bird	94	2	4	0
Car	3	97	0	0
Animal	43	0	57	0
Human	0	0	2	98

Table 6: Confusion matrix. Noise: Rain

Results Study 4

- **Support Vector Machines**
  - clean sounds 98.66% (99.66% GMM)
  - rain: 95% (97.33% GMM)
  - wind: 90.66% (93.33% GMM)
- **4 databases**
  - clean: 94.50% (96.50% GMM)
  - rain: 91.50 (86.50% GMM)
  - wind: 86.50 (82.25% GMM)

M. V. Ghiurcau and C. Rusu, *About classifying sounds in protected environments*, in Proceedings of ISEEE 2010, Galati, Romania, September 2010.

Results Study 5

- **TESPAR revisited**
  - clean sounds: 88.25%
  - rain: 83%
  - wind: 81.25%
  - DR = 0.96
  - FAR = 0.075

	Animal	**Bird**	**Car**	**Human**
Animal	**67**	**23**	**6**	**4**
Bird	**0**	**95**	**5**	**0**
Car	**4**	**2**	**94**	**0**
Human	**2**	**0**	**1**	**97**

Table 7: Confusion matrix. Clean sounds

	Animal	**Bird**	**Car**	**Human**
Animal	**60**	**28**	**10**	**2**
Bird	**1**	**93**	**6**	**0**
Car	**4**	**2**	**94**	**0**
Human	**2**	**0**	**13**	**85**

Table 8: Confusion matrix. Noise: Rain

M. V. Ghiurcau, C. Rusu, J. Astola and R. Bilcu, *Audio based solutions for detecting intruders in wild areas*, Signal Processing, 92(3), 2012, 829-840.

Conclusions and Future Work

- **Wildlife intruder detection**
    - TESPAR -> fairly robust for noisy environments
    - GMM - SVM -> more robust, increased complexity
    - solutions -> first effort
- **Future**
    - intruder verification
    - sounds database
    - noise reduction
    - mixed sounds

Thank you for your attention!

Questions?