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: $3, 7, 12, 8, 3, 4, 12, 1, 7, 21, 3, 9, 12, 3, 12, 11, \ldots$
- $n=2$
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

- **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%

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Human</th>
<th>Bird</th>
<th>Car</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>97.9</td>
<td>0</td>
<td>2.1</td>
</tr>
<tr>
<td>Bird</td>
<td>0</td>
<td>93.3</td>
<td>6.7</td>
</tr>
<tr>
<td>Car</td>
<td>5.2</td>
<td>0</td>
<td>94.8</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th></th>
<th>Human</th>
<th>Bird</th>
<th>Car</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>98.2</td>
<td>0</td>
<td>1.8</td>
</tr>
<tr>
<td>Bird</td>
<td>0</td>
<td>93.8</td>
<td>6.2</td>
</tr>
<tr>
<td>Car</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th>A/S matrix</th>
<th>Human</th>
<th>Bird</th>
<th>Car</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>87.6</td>
<td>2.2</td>
<td>10.2</td>
</tr>
<tr>
<td>Bird</td>
<td>0</td>
<td>94.4</td>
<td>5.6</td>
</tr>
<tr>
<td>Car</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 3: Rain

<table>
<thead>
<tr>
<th>A/S matrix</th>
<th>Human</th>
<th>Bird</th>
<th>Car</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>99.2</td>
<td>0</td>
<td>0.8</td>
</tr>
<tr>
<td>Bird</td>
<td>11.5</td>
<td>86.8</td>
<td>1.7</td>
</tr>
<tr>
<td>Car</td>
<td>2.1</td>
<td>9.9</td>
<td>88</td>
</tr>
</tbody>
</table>

Table 4: Wind

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

<table>
<thead>
<tr>
<th></th>
<th>Bird</th>
<th>Car</th>
<th>Animal</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bird</td>
<td>94</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Car</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Animal</td>
<td>8</td>
<td>0</td>
<td>92</td>
<td>0</td>
</tr>
<tr>
<td>Human</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 5: Confusion matrix
Results Study 3

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

<table>
<thead>
<tr>
<th></th>
<th>Bird</th>
<th>Car</th>
<th>Animal</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bird</td>
<td>94</td>
<td>2</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Car</td>
<td>3</td>
<td>97</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Animal</td>
<td>43</td>
<td>0</td>
<td>57</td>
<td>0</td>
</tr>
<tr>
<td>Human</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>98</td>
</tr>
</tbody>
</table>

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)


---

Oct 16, 2013

On the Design of an Acoustic Based Wildlife Intruder Detection System

Results Study 5

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

<table>
<thead>
<tr>
<th></th>
<th>Animal</th>
<th>Bird</th>
<th>Car</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animal</td>
<td>67</td>
<td>23</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Bird</td>
<td>0</td>
<td>95</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Car</td>
<td>4</td>
<td>2</td>
<td>94</td>
<td>0</td>
</tr>
<tr>
<td>Human</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>97</td>
</tr>
</tbody>
</table>

Table 7: Confusion matrix. Clean sounds

<table>
<thead>
<tr>
<th></th>
<th>Animal</th>
<th>Bird</th>
<th>Car</th>
<th>Human</th>
</tr>
</thead>
<tbody>
<tr>
<td>Animal</td>
<td>60</td>
<td>28</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Bird</td>
<td>1</td>
<td>93</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Car</td>
<td>4</td>
<td>2</td>
<td>94</td>
<td>0</td>
</tr>
<tr>
<td>Human</td>
<td>2</td>
<td>0</td>
<td>13</td>
<td>85</td>
</tr>
</tbody>
</table>

Table 8: Confusion matrix. Noise: Rain

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?