Sound Source Localization in a Security System using a Microphone Array

V. Behar\textsuperscript{1}, Chr. Kabakchiev\textsuperscript{2}, I. Garvanov\textsuperscript{3}

\textsuperscript{1}Institute of Information & Communication Technologies, BAS, 1113 Sofia, Bulgaria
e-mail: behar@bas.bg

\textsuperscript{2}Sofia University “St. Kliment Ohridski”, 1164 Sofia, Bulgaria
e-mail: ckabakchiev@fmi.uni-sofia.bg

\textsuperscript{3}University of Library Studies & Information Technologies, 1784 Sofia, Bulgaria
e-mail: igarvanov@yahoo.com
Outline

- Introduction
- Microphone Arrays
- Signal Model
- Signal Processing in a Security System
- Sound Source Localization
- Parallel Algorithm
- Simulation Results
- Conclusion
Security and video surveillance systems are a standard part of protection of public buildings and becoming more relevant in the protection of private property in Bulgaria.

Depending on user requirements, the security surveillance systems are different, as a structure and algorithm, and depend on the plan of security, protected areas etc.

There is no universal method for the design of a security surveillance system
A typical security and surveillance system consists of:

- **Control Panel (Display)** – user defined. The secured object is divided in zones. Each zone can be controlled differently depending on the type of sensors and user requirements.

- **Event Sensors** – detect undesired event (door opening, glass breaking, fire, flame within the secure zone).

- **Alarm devices** – alarm about undesired events. Typical alarm devices are sirens, blinking devices etc. The alarm device communicates with the sensor through wire or wireless connection.

- **Register devices** – register the undesired events (video cameras).

- **Program support** – for processing of the information from the alarm devices, decision making and control of the security system.
In a security and surveillance system,

- Many sensors for fire detection or building surveillance are equipped with sound alarm devices. In case of alarm event (smoke, flame, intrusion, glass breaking, unauthorized car opening) the alarm device emits powerful sound signal with various parameters (duration, modulation, frequency, power).

- The localization of sound signal direction could be used for pointing the additional video surveillance devices, which record the additional information and send it to control center for further analysis.

- Sound direction localization could be performed through:
  - parabolic microphones;
  - microphone arrays (linear, rectangular, circular);
The **parabolic** microphone uses a parabolic reflector to collect and focus sound waves onto a receiver, and acts in the same way as a parabolic antenna. The parabolic reflector is made of optical material.

![Diagram of a parabolic microphone system]

However, the usage of parabolic microphones for sound source localization has some **disadvantages**:  
- the microphone diagram depends on the microphone physical properties;  
- the microphone needs to be pointed exactly at the direction of control. It is complicated: each sound alarm device needs an own microphone or a mechanical scanning algorithm in case of a single parabolic microphone device.
Microphone arrays are a set of microphones (A1, A2, A3...) arranged in some geometric configuration. At the output of a microphone array signals from microphones are summed according to the certain algorithm.
Introduction (6)

Microphone arrays have some **important advantages**:

- The beamforming can be done digitally to control all dangerous directions using only a single microphone array.
- All noise signals coming from undesired directions are adaptively rejected (speaking of people, banging on the walls, etc.), which increases the detectability of signals from sensors.
- A microphone array can simultaneously generate a set of independent beam patterns and collect the information from multiple sound sources.
- 3D area of observation is possible (by 2D arrays).
- Microphone arrays can be easy adapted to detect sound signals with different carrier frequency only by change of the inter-microphone distance in the array.
The aim of this paper is:

- to develop, using commercial microphone arrays, the computational algorithm according to the adaptive Minimum Variance Distortionless Response (MVDR) beamforming method in order to estimate DOA of sound signals coming from sensors or other sound sources.

- to develop the parallel version of the computational algorithm in the computational environment with MPI interface (SoundDetect).

- to test the parallel algorithm (SoundDetect) using simulation of sound signals generated by commercial sensors produced by three well-known companies (SONITRON, E2S and SYSTEM SENSORS)
Microphone Arrays (1)

- For the sake of simplicity, we consider three types of arrays – Uniform Linear array (ULA), Uniform Rectangular Array (URA) and Uniform Circular Array (UCA).
- The ULA beam pattern can be controlled only in one dimension (azimuth or elevation), however, URA and UCA with microphones located in two dimensions can control the beam pattern in both azimuth and elevation.
Array response vector $a_c$

$$a_c(\varphi, \theta) = [1, \exp(j \frac{2\pi}{\lambda} d_1), \exp(j \frac{2\pi}{\lambda} d_2), \ldots, \exp(j \frac{2\pi}{\lambda} d_{m(i,k)}), \ldots, \exp(j \frac{2\pi}{\lambda} d_M)]$$

$$d_{m(i,k)} = \cos \theta \cdot d \cdot [\sin \varphi(i-1) + \cos \varphi(k-1)]$$

$i$ and $k$ are element positions along the $y$- and the $x$-axis, respectively

$$d_m = d \cos \theta \cos(\varphi - \varphi_m)$$

$$\varphi_m = 2\pi(m-1)/M$$, where $m=1, \ldots, M$

$$a_c(\varphi, \theta) = [\exp(j \frac{2\pi}{\lambda} d_1), \exp(j \frac{2\pi}{\lambda} d_2), \ldots, \exp(j \frac{2\pi}{\lambda} d_m), \ldots, \exp(j \frac{2\pi}{\lambda} d_M)]$$
Signal Model (1)

We consider the scenario, in which \( L \) sensor signals combined with some sound noise arrive at the microphone array with \( M \) microphones. The output signal of each microphone is a sum of sound-source-generated signals and thermal noise. Before beamforming, the vector of complex samples of the signal at the microphone array output at time instant \( k \) can be mathematically described as:

\[
x(k) = \sum_{l=1}^{L} b_l s_l(k) + n(k)
\]

- \( x \) - the \( M \)-element complex data vector
- \( s \) - the complex signal from a sensor
- \( b \) - the \( M \)-element microphone array response vector in direction of a sensor
- \( n \) - the \( M \)-element complex noise vector
  (noise occupies the entire frequency bandwidth of a microphone and can be represented mathematically as band-limited white additive Gaussian noise (AWGN)).
Signal Model (2)

The signal incoming from the \( l \)-th sound source is given by:

\[
s_l(k) = \sqrt{P_l} A_l(k) \exp[j(2\pi f_0 t + \Phi_l)]
\]

where:

- \( P_l \): the received signal power from a sensor
- \( A \): the modulating function of a sensor signal
- \( f_0 \): the carrier frequency of a sensor signal
- \( \Phi \): the initial phase of a sensor signal

For the sake of simplicity, we assume that all sensors generate signals at the same carrier frequency \( f_0 \).
The DOA estimation is used for pointing the video surveillance device (video camera) in the needed direction. The estimated signal parameters (duration, frequency, modulation, type and power) are used for signal identification – warning, alarm or emergency. The priority direction for pointing of a video camera is determined in result of analysis of the identified signals received from all detected sound sources.
The idea is to digitally scan the microphone array in the surveillance area. The microphone array simultaneously creates the main receiver lobes in a given set of directions. The power of signals received from each direction is estimated and compared with a fixed threshold.

Example:
If the video camera is mounted above the microphone array and firstly is pointed at zero azimuth, then after the digital scan it is pointed at the direction, from where the incoming sound power exceeds a certain threshold and the priority of the signal is highest.
Sound Source Localization (2)

The output signals of all microphones \((x_1, x_2, \ldots, x_M)\) are converted in the complex form by the Hilbert filter. For each direction \((\beta, \theta)\), the vector of weights \(W(\beta, \theta)\) is adaptively calculated, and the output signal \(Y(\beta, \theta)\) is formed as a weighted sum of the microphone signals \(x_1, x_2, \ldots, x_M\). The power of the output signal \(P(\beta, \theta)\) is compared with the threshold \(H\) in order to detect DOA \((\beta^*, \theta^*)\) if the signal power \(P(\beta, \theta)\) exceeds the threshold.
The purpose of the beamforming is to maximally amplify the signal incoming from the certain direction and at the same time to maximally suppress signals from other directions. The output signal of the $M$ – element array ($Y$) is formed as a weighted sum of $M$ microphone signals.

$$Y = W^T X$$
According to the Conventional Beamforming method (BF), the complex vector of weights of the microphone array \( \mathbf{W} \) is equal to the array response vector \( \mathbf{a}_c \), which is determined by the geometric array configuration (ULA, URA, UCA) and the number of elements:

\[
\mathbf{W}_{BF} = \mathbf{a}_c
\]

Disadvantages:
- Non-adaptive
- does not reject signals from undesired directions
- forms a beam pattern with high side lobe level
In the adaptive beamforming, the optimal weight vector \( W \) is chosen to maximize the signal-to-interference-plus-noise ratio (SINR) in the certain direction.

\[
SINR = \frac{\sigma_s^2 |W^H a_c|^2}{W^H K_{j+n} W}
\]

\( K_{j+n} \) is the “interference + noise” covariance matrix of size \((M \times M)\).

According to the Minimum Variance Distortionless Response (MVDR) beamforming method, the optimal weight vector \( W \) is determined through linear constrained optimization. The criterion of optimization is:

\[
\min_W W^H K_{j+n} W \quad \text{Subject to} \quad W^H a_c = 1
\]
Sound Source Localization (6)

In result of solving the optimization problem, the optimal weight vector $W$ is calculated as:

$$W_{MVDR} = \frac{K^{-1} a_c}{a_c^H K^{-1} a_c}$$

- $\alpha_c$ - array response vector for the desired direction
- $K$ – covariance matrix

**Advantages:**
- Adaptive
- Rejects signals coming from undesired directions
- Reduces side lobe level
- Improves angular resolution
Adaptive MVDR through QR decomposition

In many practical applications, the calculation of the weights $W_{MVDR}$ using estimation and inversion of the covariance matrix $K$ is very time-consuming and unstable, if the sample covariance matrix $K$ is ill-conditioned.

A numerical stable and efficient algorithm can be obtained by using QR decomposition of the input signal matrix $X$.

The signal matrix is decomposed as a product ($X=QR$) of the unitary matrix ($Q$) and the upper triangular matrix ($R$). Through the QR-decomposition of $X$, we obtain:

$$K^{-1}a_c = (X^H X)^{-1} a_c = (R^H Q^H QR)^{-1} a_c = R^{-1} (R^H)^{-1} a_c$$
The QR-decomposition of the signal matrix $X$ gives the other expression for calculation of the weight vector $W$:

$$W_{MVDR} = \frac{K_{j+n}^{-1} a_c}{a_c^H K_{j+n}^{-1} a_c}$$

$$W_{MVDR,QR} = \frac{R^{-1}(R^H)^{-1} a_c}{a_c^H R^{-1}(R^H)^{-1} a_c}$$

**Stage 1:** QR decomposition of $X$

$$X = QR$$

**Stage 2:** Solving of the system of equations

$$R^H z_1 = a_c$$

$$z_1 = (R^H)^{-1} a_c$$

**Stage 3:** Solving of the system of equations

$$R z_2 = z_1$$

$$z_2 = R^{-1} z_1$$

**Stage 4:** The optimal vector is:

$$\hat{W} = z_2^* / (a_c^H z_2^*)$$
Sound Source Localization (9)

Finally, the computational algorithm for sound source localization in a security system has the following stages:

• **Stage 1:** The microphone array is digitally steered in each angular direction \((\beta, \theta)\).
  As a result, the output signal of a microphone array is formed as:

\[
Y(\beta, \theta) = W_{MVDR,QR}^TX(\beta, \theta)
\]

**Stage 2:** The signal power is estimated at the output of a microphone array in each angular direction \((\beta, \theta)\):

\[
P(\beta, \theta) = |Y(\beta, \theta)|^2
\]

**Stage 3:** The estimated signal power is compared with a fixed predetermined threshold \(H\). If the signal power, corresponding to some angular direction \((\beta^*, \theta^*)\) exceeds the threshold \(H\), then this angular direction \((\beta^*, \theta^*)\) is the DOA estimate.

\[
(\beta^*, \theta^*) = (\beta, \theta), \text{ if } P(\beta, \theta) > H
\]
Parallel Algorithm (SoundDetect)

- The parallel version of the algorithm is implemented as a program in IBM Blue Gene/P environment using the interface MPI. The parallel program calculates the signal power at the output of a microphone array simultaneously from all directions of observation.

- The azimuth diapason [-90°,90°] is divided into \( N \) fixed directions, \( \beta_1, \beta_2, \ldots \). Each processor \( (n) \) performs the beamforming for the certain direction \( (\beta_n) \).

- The server loads the same copy of the program on all processors from 0 to \((N-1)\) where \( N \) is the number of processors allocated to the program.

- The master processor \( (0) \) performs initialization of parameters, prepares the signal data for all processors and sends all the information to each slave processor.

- All directions \( (\beta_1, \beta_2, \ldots) \), in which the microphone array is steered, are sent from the master processor to each slave processor.

- Each slave processor calculates the signal power after beamforming in the given direction \( (\gamma_n) \) and sends the result to the master processor.

- The master processor compares the signal power \( (\gamma_1, \gamma_2, \ldots) \) received from each slave processor with a predefined threshold \( H \) and estimates the DOA.
The scenario includes:
- sensors (A, B and C) located respectively at 50m, 60m and 70m away from the microphone array.
- a car as a source of natural noise located at 90 m away from the microphone array. The horn of a car generates sound with the power of 110dB.

Restrictions
- wave front is flat: $r > 2D^2 / \lambda$
- $r \gg D$
- $r \gg \lambda$

All sensors (A, B and C) are identical.
(the same carrier frequency of sound)
Considered sound alarm devices are characterized with parameters:

- Sound power (LW), dB;
- Carrier frequency (Hz);
- Signal waveform (continuous harmonic, intermittent harmonic, increasing and decreasing chirp signal, constant);
- Number of different signals;

### 1. Company – SONITRON (Belgium)

<table>
<thead>
<tr>
<th>Model</th>
<th>Signal waveform</th>
<th>Working Voltage</th>
<th>Frequency</th>
<th>Pulse frequency</th>
<th>Consumption</th>
<th>Sound power</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Min, V</td>
<td>Max, V</td>
<td>Hz</td>
<td>Min, mA</td>
<td>Max, mA</td>
</tr>
<tr>
<td>SCI 535</td>
<td>Multimode</td>
<td>5</td>
<td>35</td>
<td>2500</td>
<td>1</td>
<td>1.4</td>
</tr>
<tr>
<td>A1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCI 535</td>
<td>Multimode</td>
<td>5</td>
<td>35</td>
<td>3500</td>
<td>1</td>
<td>1.4</td>
</tr>
<tr>
<td>B1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCI 535</td>
<td>Multimode</td>
<td>5</td>
<td>35</td>
<td>2500</td>
<td>5</td>
<td>1.4</td>
</tr>
<tr>
<td>A5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCI 535</td>
<td>Multimode</td>
<td>5</td>
<td>35</td>
<td>3500</td>
<td>5</td>
<td>1.4</td>
</tr>
<tr>
<td>B5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Working temperature: -35°C...+75°C
## Simulation Results (3)

2. Company – SYSTEM SENSOR (USA)

- Combined multi alert home and strobes
  - EMA24FRSSR

- LW = 103 dB
- 32 different sound signals

### Type 1

<table>
<thead>
<tr>
<th>Hz</th>
<th>Screen</th>
<th>Hz</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td></td>
<td>1200</td>
<td>0.015</td>
</tr>
<tr>
<td>2400</td>
<td></td>
<td>2400</td>
<td>-</td>
</tr>
<tr>
<td>1200</td>
<td>0</td>
<td>1200</td>
<td>0.02</td>
</tr>
<tr>
<td>1200</td>
<td>500</td>
<td>1200</td>
<td>0.10</td>
</tr>
<tr>
<td>800</td>
<td>800</td>
<td>800</td>
<td>-</td>
</tr>
<tr>
<td>500</td>
<td></td>
<td>1200</td>
<td>0.50</td>
</tr>
<tr>
<td>800</td>
<td></td>
<td>1000</td>
<td>0.05</td>
</tr>
<tr>
<td>2400</td>
<td>0</td>
<td>1200</td>
<td>0.05</td>
</tr>
<tr>
<td>500</td>
<td></td>
<td>1200</td>
<td>0.12</td>
</tr>
<tr>
<td>2400</td>
<td>2400</td>
<td>800</td>
<td>0.50</td>
</tr>
<tr>
<td>1200</td>
<td>0</td>
<td>1200</td>
<td>0.50</td>
</tr>
<tr>
<td>1200</td>
<td>500</td>
<td>1200</td>
<td>1.00</td>
</tr>
<tr>
<td>800</td>
<td>800</td>
<td>800</td>
<td>-</td>
</tr>
<tr>
<td>500</td>
<td></td>
<td>1200</td>
<td>1.00</td>
</tr>
<tr>
<td>800</td>
<td></td>
<td>1000</td>
<td>0.50</td>
</tr>
<tr>
<td>2400</td>
<td>0</td>
<td>1200</td>
<td>0.50</td>
</tr>
</tbody>
</table>

### Type 2

<table>
<thead>
<tr>
<th>Hz</th>
<th>Screen</th>
<th>Hz</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>1200</td>
<td></td>
<td>500</td>
<td>0.10</td>
</tr>
<tr>
<td>800</td>
<td></td>
<td>1000</td>
<td>0.05</td>
</tr>
<tr>
<td>500</td>
<td></td>
<td>1200</td>
<td>0.50</td>
</tr>
<tr>
<td>2400</td>
<td></td>
<td>1000</td>
<td>0.50</td>
</tr>
<tr>
<td>1200</td>
<td></td>
<td>500</td>
<td>0.10</td>
</tr>
<tr>
<td>1200</td>
<td>500</td>
<td>500</td>
<td>1.00</td>
</tr>
<tr>
<td>800</td>
<td>800</td>
<td>1000</td>
<td>0.50</td>
</tr>
<tr>
<td>500</td>
<td>1200</td>
<td>800</td>
<td>0.50</td>
</tr>
<tr>
<td>800</td>
<td>1000</td>
<td>1000</td>
<td>0.50</td>
</tr>
<tr>
<td>2400</td>
<td>0</td>
<td>500</td>
<td>1.00</td>
</tr>
</tbody>
</table>

http://www.iict.bas.bg

AComIn: Advanced Computing for Innovation

http://www.iict.bas.bg
### Simulation Results (4)

#### 3. Company – E2S (UK)

<table>
<thead>
<tr>
<th>Tone</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tone 1</td>
<td>340 Hz Continuous</td>
</tr>
<tr>
<td>Tone 2</td>
<td>800/1000Hz @ 0.25 sec Alternating</td>
</tr>
<tr>
<td>Tone 3</td>
<td>500/1200Hz @ 0.3Hz 0.5 sec Slow Whoop</td>
</tr>
<tr>
<td>Tone 4</td>
<td>800/1000Hz @ 1Hz Sweeping</td>
</tr>
<tr>
<td>Tone 5</td>
<td>2400Hz Continuous</td>
</tr>
<tr>
<td>Tone 6</td>
<td>2400/2900Hz @ 7Hz Sweeping</td>
</tr>
<tr>
<td>Tone 7</td>
<td>2400/2900Hz @ 1Hz Sweeping</td>
</tr>
<tr>
<td>Tone 8</td>
<td>500/1200/500Hz @ 0.3Hz Sweeping</td>
</tr>
<tr>
<td>Tone 9</td>
<td>1200/500Hz @ 1Hz - DIN / PFEER P.T.A.P.</td>
</tr>
<tr>
<td>Tone 10</td>
<td>2400/2900Hz @ 2Hz Alternating</td>
</tr>
<tr>
<td>Tone 11</td>
<td>1000Hz @ 1Hz Intermittent</td>
</tr>
<tr>
<td>Tone 12</td>
<td>800/1000Hz @ 0.875Hz Alternating</td>
</tr>
<tr>
<td>Tone 13</td>
<td>2400Hz @ 1Hz Intermittent</td>
</tr>
<tr>
<td>Tone 14</td>
<td>800Hz 0.25sec on, 1 sec off Intermittant</td>
</tr>
<tr>
<td>Tone 15</td>
<td>800Hz Continuous</td>
</tr>
<tr>
<td>Tone 16</td>
<td>660Hz 150mS on, 150mS off Intermittant</td>
</tr>
<tr>
<td>Tone 17</td>
<td>544Hz (100mS)/440Hz (400mS) - NF S 32-001</td>
</tr>
<tr>
<td>Tone 18</td>
<td>660Hz 1.8sec on, 1.8sec off Intermittant</td>
</tr>
<tr>
<td>Tone 19</td>
<td>1.4KHz-1.6KHz 1s, 1.6KHz-1.4KHz 0.5s - NFC48-266</td>
</tr>
<tr>
<td>Tone 20</td>
<td>660Hz Continuous</td>
</tr>
<tr>
<td>Tone 21</td>
<td>554Hz/440Hz @ 1Hz Alternating</td>
</tr>
<tr>
<td>Tone 22</td>
<td>544Hz @ 0.875 sec. Intermittant</td>
</tr>
<tr>
<td>Tone 23</td>
<td>800Hz @ 2Hz Intermittant</td>
</tr>
<tr>
<td>Tone 24</td>
<td>800/1000Hz @ 50Hz Sweeping</td>
</tr>
<tr>
<td>Tone 25</td>
<td>2400/2900Hz @ 50Hz Sweeping</td>
</tr>
<tr>
<td>Tone 26</td>
<td>Bell</td>
</tr>
<tr>
<td>Tone 27</td>
<td>554Hz Continuous</td>
</tr>
<tr>
<td>Tone 28</td>
<td>440Hz Continuous</td>
</tr>
<tr>
<td>Tone 29</td>
<td>800/1000Hz @ 7Hz Sweeping</td>
</tr>
<tr>
<td>Tone 30</td>
<td>420Hz @ 0.625 sec Australian Alert</td>
</tr>
<tr>
<td>Tone 31</td>
<td>660/1200Hz @ 1Hz Sweeping</td>
</tr>
<tr>
<td>Tone 32</td>
<td>500-1200Hz 3.75sec /0.25sec. Australian Evac.</td>
</tr>
</tbody>
</table>

LW=100 dB

32 different sound signals

7/15/2013
Considered microphone arrays are characterized with parameters:

- configuration (linear, rectangular, square);
- number of microphones;
- frequency band $[100 - 5000]$ Hz;
- microphone noise $-35$dB

Company – Brüel&Kjær (Denmark)
**Simulation Results (6)**

**Microphone arrays**
- type WA 0807 (Brüel&Kjær);
- microphone 4935 (Brüel&Kjær);
- frequency band [100 – 5000] Hz;
- geometrical configuration:
  - ULA -11 (SONITRON- sound generators);
  - URA -11x4 (SONITRON- sound generators);
  - ULA -4 (E2S- sound generators);
  - URA -4x4 (E2S- sound generators);
- ULA -8 (SYSTEM SENSOR- sound generators);
- URA -8x4 (SYSTEM SENSOR- sound generators);
- microphone noise – 35dB
- approximate array length- 50cm,
Simulation Results (7)

Simulated signal parameters:

<table>
<thead>
<tr>
<th>Producer</th>
<th>Sound power LW [dB]</th>
<th>Sound frequency [Hz]</th>
</tr>
</thead>
<tbody>
<tr>
<td>SONITRON</td>
<td>77</td>
<td>2500</td>
</tr>
<tr>
<td>E2S</td>
<td>100</td>
<td>1000</td>
</tr>
<tr>
<td>SYSTEM SENSOR</td>
<td>103</td>
<td>2400</td>
</tr>
</tbody>
</table>

\[ A = 2 \cdot 10^{-5} \cdot 10^{L_p/20} \]

\[ L_p = LW - 11 - 20 \log(R) \]

Modulation

<table>
<thead>
<tr>
<th>Continuous (warning), A</th>
<th>Intermittent -1 (alarm), B</th>
<th>Intermittent -2 (emergency), C</th>
</tr>
</thead>
<tbody>
<tr>
<td>f_int=0 Hz</td>
<td>f_int=5 Hz</td>
<td>f_int=1 Hz</td>
</tr>
<tr>
<td>T_sig=10s</td>
<td>T_sig=30s</td>
<td>T_sig=60s</td>
</tr>
</tbody>
</table>

Signal processing

- Sampling frequency – 5.5kHz;
- Scan period – 2сек
- Number of directions -91 in [-90°, 90°] by step of 2°
Simulation Results (8)

Signals, interference, microphone noise

**Signals:**
- a)-SONITRON
- b) -SYSTEM SENSOR
- c)-E2S

**Interference**
(sound of a car)

**Microphone noise**
**Simulation Results (9)**

**Numerical results**

<table>
<thead>
<tr>
<th>Sound generators</th>
<th>Microphone array</th>
<th>Source sound azimuth (real) [°]</th>
<th>Source sound azimuth (estimated) [°]</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SONITRON</strong></td>
<td>ULA (11x1)</td>
<td>-14; 0; 14; 28</td>
<td>-14; 0; 14; 28</td>
</tr>
<tr>
<td></td>
<td>URA (11x4)</td>
<td>-14; 0; 14; 28</td>
<td>-14; 0; 14; 28</td>
</tr>
<tr>
<td><strong>E2S</strong></td>
<td>ULA (4x1)</td>
<td>-26; 0; 26</td>
<td>-26; 0; 26</td>
</tr>
<tr>
<td></td>
<td>URA (4x4)</td>
<td>-26; 0; 26; 52</td>
<td>-26; 0; 26; 52</td>
</tr>
<tr>
<td><strong>SYSTEM SENSOR</strong></td>
<td>ULA (8x1)</td>
<td>-14; 0; 14; 28</td>
<td>-14; 0; 14; 28</td>
</tr>
<tr>
<td></td>
<td>URA (8x4)</td>
<td>-14; 0; 14; 28</td>
<td>-14; 0; 14; 28</td>
</tr>
</tbody>
</table>
Simulation Results (10)

Microphone array beam pattern calculation
SONITRON

Microphone array: ULA -11
Microphone array: URA -11 x 4

Linear microphone arrays (ULA) should be used in cases where it is important to control the movement of the video camera only in azimuthal direction.
Simulation Results (11)

Microphone array beam pattern calculation
SYSTEM SENSOR

Microphone array: ULA -8

Microphone array: URA -8 x 4
Microphone array beam pattern calculation

E2S

ULA-4 can separate maximum 3 sound signals received from different directions, i.e. (M-1). For example, if we want control 10 directions from sensors, the microphone array must be with at least 11 microphones.
Conclusions

• The computational algorithm for sound source localization in a security system with commercial devices (sensors, microphone arrays) is proposed and tested.

• It is shown that the accurate sound source localization is possible using the adaptive MVDR-beamforming algorithm.

• The maximal number of controlled directions from sensors depends on the number of array microphones.

• All desired directions can be controlled simultaneously using a single microphone array.

The future studies will be conducted with the help of equipment for noise source identification purchased under the project AComIn "Advanced Computing for Innovation", grant 316087, funded by the FP7 Capacity Programme.
This work is financially supported:

• by the Bulgarian Science Fund (projects DTK 02/28.2009, DDVU 02/50/2010)

• partly by the project AComIn "Advanced Computing for Innovation", grant 316087, funded by the FP7 Capacity Programme (Research Potential of Convergence Regions).
THANK YOU FOR YOUR ATTENTION!