Acoustic Vector Sensors

DOA Estimation, Beamforming and Applications

Sharon Gannot

joint work with Dovid Levin and Emanuël Habets

Faculty of Engineering and Data Science Institute, Bar-Ilan University, Israel

China Computer Federation, January 6, 2021

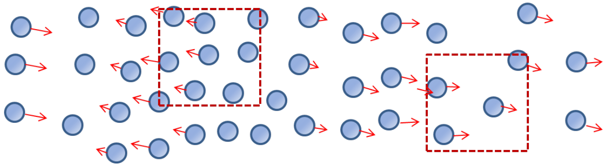
Motivation

- Conventional methods for DOA estimation and beamforming involve multiple microphones and entail spacing requirements.
- Compact superdirective devices have been utilized for these tasks.
- Vector sensors are inherently superdirective devices.
- We describe new algorithms and analysis using a single acoustic vector sensors (AVS).
- We also discuss a smart-glasses application using two AVS's.

Outline

Outline

Pressure and particle velocity



- **Pressure** is determined by the *density* of air particles.
- **Particle velocity** describes the *motion* of air particles.

A conventional sensor does not measure velocity. $\ddot{\frown}$

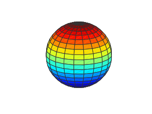
What is a vector-sensor?

A vector sensor measures both pressure and particle velocity.

Conventional microphone:

Channels: 1

Directivity: monopole (*typically*)

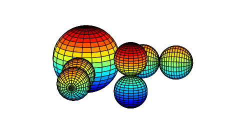


Schematic:

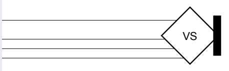
Vector-sensor:

Channels: 4

Directivity: monopole ($\times 1$)
dipole ($\times 3$)



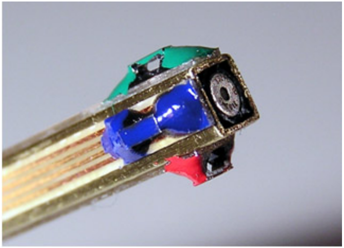
Schematic:



Construction of vector-sensor

Dipole elements may be obtained from:

1. Particle velocity sensors
2. Differential-microphone arrays

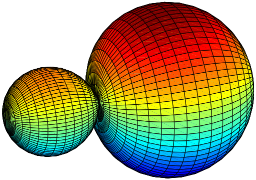


Notation for measurements

The measurements of the vector-sensor are denoted:

$$\mathrm{y}[n] = \begin{bmatrix} p[n] \\ v_x[n] \\ v_y[n] \\ v_z[n] \end{bmatrix} = \begin{bmatrix} p[n] \\ \mathrm{v}[n] \end{bmatrix}$$

A linear combination of the sensor signals produces a limaçon (to be explained later).



Outline

DOA Estimation using a single AVS [Levin et al., 2011]

Problem formulation

Goal

Estimate the direction of arrival (DOA) of a single acoustic source.

- The source is located in the far-field.
- The signal produced is denoted $s[n]$.
- The DOA is described by a unit vector u.
- Noise components are assumed to be uncorrelated which applies to sensor-noise and isotropic fields.

Signal and noise notation

The measurements consist of signal and noise components:

$$\begin{bmatrix} p[n] \\ v_x[n] \\ v_y[n] \\ v_z[n] \end{bmatrix} = \begin{bmatrix} 1 \\ \mathrm{u} \end{bmatrix} s[n] + \begin{bmatrix} e_p[n] \\ e_{v_x}[n] \\ e_{v_y}[n] \\ e_{v_z}[n] \end{bmatrix}$$

or more briefly:

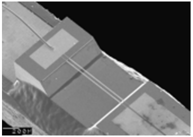
$$\begin{bmatrix} p[n] \\ \mathrm{v}[n] \end{bmatrix} = \begin{bmatrix} 1 \\ \mathrm{u} \end{bmatrix} s[n] + \begin{bmatrix} e_p[n] \\ \mathrm{e}_v[n] \end{bmatrix}$$

Note: The particle-velocity has been scaled to produce normalized dipoles.

Noise models

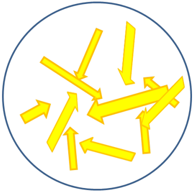
Sensor noise:

- Noise of all sensors are uncorrelated.
- Monopoles and dipoles may have *different* variances.



Diffuse noise:

- Noise of all sensors are *uncorrelated*, since all sensors are colocated
- $\sigma_{e_v}^2 = \frac{1}{3}\sigma_{e_p}^2$.



Statistical characterization

- The signal and noise are statistically independent.
- The signal is a white zero-mean WSS process with variance σ_s^2.

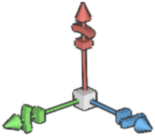
$$E\{s[n]s[m]\} = \sigma_s^2 \delta[n-m]$$

- The noise is a zero-mean WSS process with variances $\sigma_{e_p}^2$ and $\sigma_{e_v}^2$.

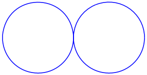
$$E\left\{\begin{bmatrix} e_p[n] \\ \mathrm{e}_v[n] \end{bmatrix}\begin{bmatrix} e_p[m] \\ \mathrm{e}_v[m] \end{bmatrix}^T\right\} = \begin{bmatrix} \sigma_{e_p}^2 & 0_{1\times 3} \\ 0_{3\times 1} & \sigma_{e_v}^2 \cdot \mathrm{I}_{3\times 3} \end{bmatrix} \delta[n-m]$$

Steering

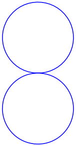
- The orientation of the dipole sensors are fixed.
- A linear combination of these sensors allows for steering in any direction q.
- DOA estimation can be obtained by steering the AVS to all possible directions.
- $v_{\mathrm{q}}[n] = \mathrm{q}^T \mathrm{v}[n]$



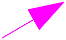
$\mathbf{v}_x$



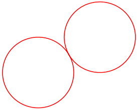
$\mathbf{v}_y$



$$\mathbf{q} = \begin{bmatrix}\sqrt{3}/2 & 1/2\end{bmatrix}^T$$



$$\mathbf{v}_{\mathbf{q}} = \frac{\sqrt{3}}{2}\mathbf{v}_x + 0.5\mathbf{v}_y$$

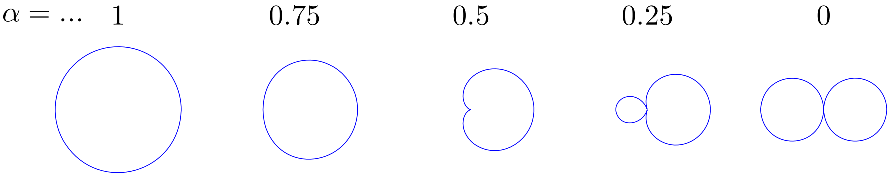


Beam shaping

- A weighted combination of monopole and steered dipole produces a limaçon pattern.
- The combination

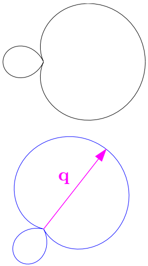
$$y_{\mathrm{q}}[n] = \alpha p[n] + (1-\alpha) v_{\mathrm{q}}[n] .$$

ensures unity boresight response.



Steered response power (SRP)

1. We choose a beampattern paramater α.
2. We steer the beam in the direction q.
3. We measure the response power over N samples.



The SRP is defined as the average power:

$$\mathrm{SRP}(\alpha, \mathsf{q}) = \frac{1}{N} \sum_{n=0}^{N-1} \left(\alpha p[n] + (1 - \alpha) \mathsf{q}^T \mathsf{v}[n] \right)^2 .$$

DOA estimation

For a given α we search for the direction q corresponding to maximum SRP.

$$\hat{\mathsf{u}} = \underset{\mathsf{q}}{\operatorname{argmax}} \{\mathrm{SRP}(\alpha, \mathsf{q})\}$$
$$\text{subject to} \quad \mathsf{q}^T \mathsf{q} = 1 .$$

Simplified form

$$\hat{\mathrm{u}} = \underset{\mathrm{q}}{\operatorname{argmax}} \left\{ \alpha\, \mathrm{q}^T \hat{\mathrm{r}}_{pv} + \frac{1-\alpha}{2} \mathrm{q}^T \hat{\mathrm{R}}_{vv} \mathrm{q} \right\}$$

$$\text{subject to} \quad \mathrm{q}^T \mathrm{q} = 1\,,$$

where:

$$\hat{\mathrm{R}}_{vv} = \frac{1}{N} \sum_{n=0}^{N-1} \mathrm{v}[n] \mathrm{v}^T[n]$$

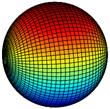
$$\hat{\mathrm{r}}_{pv} = \frac{1}{N} \sum_{n=0}^{N-1} \mathrm{v}[n] p[n]\,.$$

- $\hat{\mathrm{r}}_{pv}$ is also known as *average intensity vector*.
- Let's inspect two extreme cases . . .

Case I: Near-monopole

- For a near-monopole beampattern $(\alpha \to 1^-)$, the problem becomes:

$$\hat{\mathrm{u}} = \underset{\mathrm{q}}{\operatorname{argmax}}\ \mathrm{q}^T \hat{\mathrm{r}}_{pv}$$
$$\text{s.t.} \quad \mathrm{q}^T \mathrm{q} = 1 \,.$$



- The solution is:

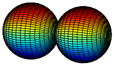
$$\hat{\mathrm{u}} = \frac{\hat{\mathrm{r}}_{pv}}{\left\|\hat{\mathrm{r}}_{pv}\right\|} \,.$$

Note: This estimator was proposed by [Davies, 1987] and later by [Nehorai and Paldi, 1994] under the name Intensity-based algorithm.

Case II: Dipole

- For a dipole beampattern ($\alpha = 0$), the problem becomes:

$$\hat{\mathrm{u}} = \underset{\mathrm{q}}{\operatorname{argmax}}\ \mathrm{q}^T \hat{\mathrm{R}}_{vv}\mathrm{q}$$
$$\text{s.t.} \quad \mathrm{q}^T\mathrm{q} = 1\,.$$



- The solution is:

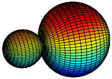
$$\hat{\mathrm{u}} = \text{eigenvector corresponding to largest eigenvaue of } \hat{\mathrm{R}}_{vv}\,.$$

Note: This estimator was proposed by [Nehorai and Paldi, 1994] under the name Velocity-Covariance-based algorithm.

General case: Limaçon beampattern

- In the general case ($0 \leq \alpha < 1$), the problem involves both linear and quadratic terms:

$$\hat{u} = \underset{q}{\operatorname{argmax}} \left\{ \alpha\, q^T \hat{r}_{pv} + \frac{1-\alpha}{2} q^T \hat{R}_{vv} q \right\}$$
$$\text{subject to} \quad q^T q = 1\,,$$



- We propose an iterative gradient based solution. The gradient of the target function is:

$$\nabla_q T(q) = \alpha\, \hat{r}_{pv} + (1-\alpha)\hat{R}_{vv} q\,,$$

Algorithm

Constrained gradient ascent algorithm for SRP maximization

Input: $\hat{\mathrm{R}}_{vv}$, $\hat{\mathrm{r}}_{pv}$, α

$\mathrm{q}_0 := \hat{\mathrm{u}} = \hat{\mathrm{r}}_{pv}/\|\hat{\mathrm{r}}_{pv}\|$

$k := 0$

$K :=$ maximum number of iterations

$\epsilon :=$ tolerance parameter

$\mu :=$ step size parameter

repeat

$\quad \mathrm{q}_{k+1} := \mathrm{q}_k + \mu\big(\alpha\,\hat{\mathrm{r}}_{pv} + (1-\alpha)\hat{\mathrm{R}}_{vv}\mathrm{q}_k\big)$

$\quad \mathrm{q}_{k+1} := \mathrm{q}_{k+1}/\|\mathrm{q}_{k+1}\|$

$\quad k := k+1$

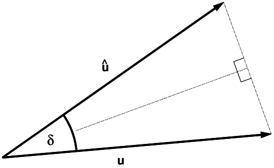
until $(k = K)$ *or alternatively* $(\|\mathrm{q}_k - \mathrm{q}_{k-1}\|^2 < \epsilon^2)$

Output: $\hat{\mathrm{u}} := \mathrm{q}_k$

Mean square angular error

- The *angular error* is the angle δ by which $\hat{\mathrm{u}}$ deviates from u, defined formally as:

$$\mathrm{AE} \equiv 2\sin^{-1}\left(\frac{\|\hat{\mathrm{u}}-\mathrm{u}\|}{2}\right)$$



- The *mean square angular error* (MSAE) describes convergence rate of $\hat{\mathrm{u}}$ towards the true DOA:

$$\mathrm{MSAE} \equiv \lim_{N\to\infty}\left(N\cdot E\{\mathrm{AE}^2\}\right).$$

Monte Carlo simulation test

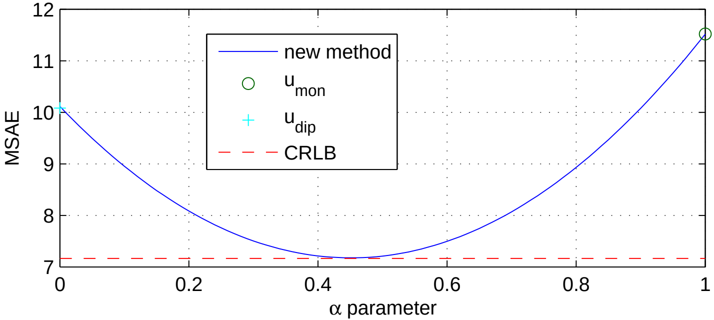
Steps for single Monte Carlo trial

1. A DOA u is selected randomly.
2. The signals and noise components are generated ($N = 8000$ time instants), and are used to produce sensor measurements $p[n]$, $\mathrm{v}[n]$.
3. The estimators $\hat{\mathrm{u}}$ are calculated for values of α ranging from 0 to 1; the corresponding angular errors are recorded.

A total of $MC = 100,000$ independently conducted trials creates records from which the sample MSAEs are obtained.

Monte Carlo results

Signal and noise power: $\sigma_s^2 = 0.5$, $\sigma_{e_p}^2 = 1.1$, $\sigma_{e_v}^2 = 0.9$:



Note: the estimator $\hat{u}$ approaches the Cramér-Rao lower bound.

Maximum likelihood [Levin et al., 2012] I

- Consider N sensor measurements:

$$\mathrm{y}[n] = \begin{bmatrix} p[n] \\ \mathrm{v}[n] \end{bmatrix};\ 0 \le n \le N-1$$

- The corresponding covariance matrix is:

$$\mathrm{C} = \mathrm{Cov}\{\mathrm{y}[n]\} = \begin{bmatrix} \sigma_{e_p}^2 & & & \\ & \sigma_{e_v}^2 & & \\ & & \sigma_{e_v}^2 & \\ & & & \sigma_{e_v}^2 \end{bmatrix} + \sigma_s^2 \begin{bmatrix} 1 \\ \mathrm{u} \end{bmatrix} \begin{bmatrix} 1 \\ \mathrm{u} \end{bmatrix}^T,$$

Maximum likelihood [Levin et al., 2012] II

- $\mathrm{y}[n]$ and $\mathrm{y}[m]$ are statistically independent for $n \neq m$:

$$f_{\mathrm{y}[0],\ldots,\mathrm{y}[N-1]}(\mathrm{y}[0],\ldots,\mathrm{y}[N-1]) = \frac{1}{(2\pi)^{N/2}|\mathrm{C}|^{\frac{1}{2}}}\prod_{n=0}^{N-1}\exp\left\{-\frac{1}{2}\mathrm{y}^T[n]\mathrm{C}^{-1}\mathrm{y}[n]\right\}$$

- Maximum likelihood DOA estimation:

$$\hat{\mathrm{u}}_{\mathrm{ML}} = \underset{\mathrm{u}}{\operatorname{argmax}} f_{\mathrm{y}[0],\ldots,\mathrm{y}[N-1]}(\mathrm{y}[0],\ldots,\mathrm{y}[N-1]) \quad \text{subject to} \quad \mathrm{u}^T\mathrm{u} = 1\,.$$

Maximum likelihood [Levin et al., 2012] III

- The likelihood can be recast as:

$$\hat{u}_{\mathrm{ML}} = \underset{u}{\operatorname{argmax}} \left\{ \alpha_0 \, u^T \hat{r}_{pv} + (1 - \alpha_0) \frac{1}{2} u^T \hat{R}_{vv} u \right\} \quad \text{s.t.} \quad u^T u = 1 \, ,$$

with $\alpha_0 = \frac{\sigma_{e_v}^2}{\sigma_{e_p}^2 + \sigma_{e_v}^2}$.

The ML estimator identifies with the SRP maximizer for $\alpha = \frac{\sigma_{e_v}^2}{\sigma_{e_p}^2 + \sigma_{e_v}^2}$.

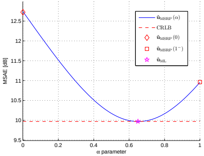
Comparison of the MSAE of presented estimators

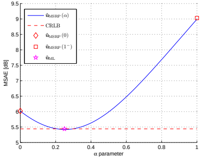
MSAEs of various estimators: [Nehorai and Paldi, 1994]

$$\mathrm{MSAE}_{\hat{\mathbf{u}}_{\mathrm{MSRP}}(1^-)} = 2\frac{\sigma_{e_v}^2}{\sigma_s^2}\left(1 + \frac{\sigma_{e_p}^2}{\sigma_s^2}\right)$$

$$\mathrm{MSAE}_{\hat{\mathbf{u}}_{\mathrm{MSRP}}(0)} = 2\frac{\sigma_{e_v}^2}{\sigma_s^2}\left(1 + \frac{\sigma_{e_v}^2}{\sigma_s^2}\right)$$

$$\mathrm{MSAE}_{\mathrm{CRLB}} = 2\frac{\sigma_{e_v}^2}{\sigma_s^2}\left(1 + \frac{\left(\sigma_{e_p}^{-2} + \sigma_{e_v}^{-2}\right)^{-1}}{\sigma_s^2}\right)$$





- When $\sigma_{e_v}^2 > \sigma_{e_p}^2$, intensity-based estimator outperforms velocity-covariance-based estimator (and vice-versa).
- When $\sigma_{e_v}^2 \gg \sigma_{e_p}^2$, intensity based estimator approaches the CRLB.

Minimum power distortionless response (MPDR) I

Goal:

Estimate $s[n]$ as a weighted combination of the available measurements:

$$\hat{s}[n] = \mathrm{w}^T \mathrm{y}[n],$$

Measurements:

$$\mathrm{y}[n] = \mathrm{h}s[n] + \mathrm{e}[n],$$

where $\mathrm{h} = [1\ \mathrm{u}^T]^T$ is the array-manifold vector corresponding to the correct DOA , and $\mathrm{e}[n] = [e_p[n]\ \mathrm{e}_v[n]^T]^T$.

MPDR criterion:

$$\mathrm{w}_{\mathrm{MPDR}} = \underset{\mathrm{w}}{\operatorname{argmin}} \left\{ \mathrm{w}^T \mathrm{C} \mathrm{w} \quad \mathrm{s.t.} \quad \mathrm{w}^T \mathrm{h} = 1 \right\}$$

Minimum power distortionless response (MPDR) II

Using

$$\mathbf{C} = \mathrm{Cov}\{\mathbf{y}[n]\} = \begin{bmatrix} \sigma_{e_p}^2 & & & \\ & \sigma_{e_v}^2 & & \\ & & \sigma_{e_v}^2 & \\ & & & \sigma_{e_v}^2 \end{bmatrix} + \sigma_s^2 \mathbf{h}\mathbf{h}^T$$

Results in:

$$\mathbf{w}_{\mathrm{MPDR}} = \begin{bmatrix} \alpha_0 \\ (1-\alpha_0)\mathbf{u} \end{bmatrix}, \text{with } \alpha_0 = \frac{\sigma_{e_v}^2}{\sigma_{e_p}^2 + \sigma_{e_v}^2}.$$

Minimum power distortionless response (MPDR) III

MPDR beamformer

$$\hat{s}[n] = \alpha_0 p[n] + (1 - \alpha_0)\mathrm{u}^T \mathrm{v}[n]$$

The MPDR beamformer conforms to the mold of the SRP maximizer with:

- The look direction specified as the true DOA.
- The shape parameter specified as α_0 (the optimal ML parameter).

Conclusions

- A single vector-sensor provides direction-sensitive information.
- We derive a method for DOA estimation in the presence of noise.
- The method generalizes two previously suggested methods.
- The proposed method attains lower MSAE than previously suggested methods.
- Optimal choice of α (in the ML sense) depends on $\sigma_{e_v}^2/\sigma_{e_p}^2$ and asymptotically achieves CRLB.
- The estimator can be recast as an MPDR beamformer.

Outline

Effects of reverberation on the intensity vector [Levin et al., 2010b]

The Intensity vector was shown to provide a good estimate of the source direction. Specifically, when $\sigma^2_{e_v} \gg \sigma^2_{e_p}$, the intensity based estimator approaches the CRLB.

Sources of error:

- The DOA estimate is prone to errors due to sensor noise, ambient noise, and reverberation.
- A scenario of sensor noise has been analyzed [Nehorai and Paldi, 1994]. The estimator $\hat{u}$ was shown to be *unbiased* and *consistent*.
- The effects of reverberation on estimation accuracy have been hitherto unanalyzed.
- We provide a statistical analysis and an experimental evaluation for the DOA bias in reverberant environments.

DOA estimation based on the intensity vector I

Definitions:

- The intensity vector is the product of pressure and particle velocity:

$$\mathrm{i}[n] = p[n]\mathrm{v}[n]$$

  and represents the transport of acoustical energy (the acoustical equivalent of the Poynting vector).

- Utilization of the time-averaged intensity vector $\bar{\mathrm{i}}$ to estimate DOA was proposed by [Nehorai and Paldi, 1994]:

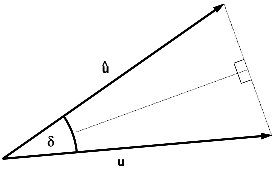
$$\bar{\mathrm{i}} = \frac{1}{N}\sum_{n=1}^{N}\mathrm{i}[n] = \frac{1}{N}\sum_{n=1}^{N}p[n]\mathrm{v}[n]$$

  The estimated DOA is given by $\hat{\mathrm{u}} = \frac{\bar{\mathrm{i}}}{\|\bar{\mathrm{i}}\|}$.

DOA estimation based on the intensity vector II

- The accuracy of a DOA estimation can be evaluated by the *AE* which is defined as the angle δ by which $\hat{\mathrm{u}}$ deviates from the true DOA u:

$$\mathrm{AE} \equiv 2\sin^{-1}\left(\frac{\|\hat{\mathrm{u}} - \mathrm{u}\|}{2}\right)$$



Reverberant Scenario

- As sound obeys the wave equation which is an LTI system, $p[n]$ and $\mathrm{v}[n]$ can be expressed as convolutions of a source signal $s[n]$ with RIRs:

$$\begin{bmatrix} p[n] \\ v_x[n] \\ v_y[n] \\ v_z[n] \end{bmatrix} = \begin{bmatrix} (s * h_p)[n] \\ (s * h_{v_x})[n] \\ (s * h_{v_y})[n] \\ (s * h_{v_z})[n] \end{bmatrix}.$$

- The source signal is assumed to be WSS process described by:

$$E\{s[n]\} = 0; \quad E\{s[n_1]s[n_2]\} = R_{ss}[n_1 - n_2].$$

- The expectation of the intensity vector w.r.t. the signal, assumed here to be white is shown to have the form:

$$\Psi(\mathrm{h}) = E\{\bar{\mathrm{i}}[n] | \mathrm{h}\} = \sum_{\ell} R_{ss}[\ell] \sum_{m} h_p[m]\mathrm{h}_v[m+\ell] = \sigma_s^2 \sum_{m} h_p[m]\mathrm{h}_v[m]$$

Note, that the RIRs h are fixed here, and assumed to be an instance from a random vector, as will be explained in the sequel.

The intensity accumulation vector (IAV) I

- The instantaneous product of the pressure and velocity RIRs:

$$\mathrm{h}_i[n] \equiv h_p[n]\mathrm{h}_v[n].$$

- The expected intensity vector can be represented as a contribution of the direct-path (arbitrarily defined at $-n_{\mathrm{d}}$) and the reverberant tail:

$$\Psi(\mathrm{h}) = E\{\bar{\mathrm{i}}|\,\mathrm{h}\} = \sigma_s^2\left(\mathrm{h}_i[-n_{\mathrm{d}}] + \sum_{n=0}^{\infty}\mathrm{h}_i[n]\right).$$

- The well-known energy-decay curve (EDC) [Schroeder, 1965] is defined as:

$$\mathrm{EDC}[n] = \sum_{m=n}^{\infty} h_p^2[m].$$

The intensity accumulation vector (IAV) II

- Similarly, we define the intensity accumulation vector (IAV):

$$\mathrm{IAV}[n] = \sum_{m=n}^{\infty} \mathrm{h}_i[m]$$ [there was a typo in the presented version]

- Hence, reverberation, as summarized by $\mathrm{IAV}[0]$, distracts the intensity vector from pointing towards the direct-arrival direction:

$$\Psi(\mathrm{h}) = E\{\bar{\mathrm{i}} | \mathrm{h}\} = \sigma_s^2 \left(\mathrm{h}_i[-n_\mathrm{d}] + \mathrm{IAV}[0] \right).$$

Statistical RIR model I

- Although the expected intensity is a function of h, the values of these RIRs are generally unknown.
- General properties of these functions (e.g. decay rate) can be utilized to provide a statistical model of the RIRs.
- A new model has been suggested [Levin et al., 2010] which extends Polack and Moorer model to incorporate particle velocity.
- The RIRs are presented as:

$$\mathrm{h}[n] \equiv \begin{bmatrix} h_{\mathrm{p}}[n] \\ h_{v_x}[n] \\ h_{v_y}[n] \\ h_{v_z}[n] \end{bmatrix} = A_d \begin{bmatrix} 1 \\ \mathrm{u} \end{bmatrix} \delta[n+n_d] + \sigma_0 \begin{bmatrix} & w_1[n] \\ 1/\sqrt{3} & w_2[n] \\ 1/\sqrt{3} & w_3[n] \\ 1/\sqrt{3} & w_4[n] \end{bmatrix} e^{-\alpha n} u[n]$$

Statistical RIR model II

- A_d – the amplitude of the direct arrival coefficients,
- σ_0 – reverberation amplitude,
- $u[n]$ – discrete-time unit step function,
- n_d – direct arrival $\mathrm{h}[-n_d]$,
- $\mathrm{w}[n] = [w_1[n] \; w_2[n] \; w_3[n] \; w_4[n]]^T$ – i.i.d. Gaussian process:

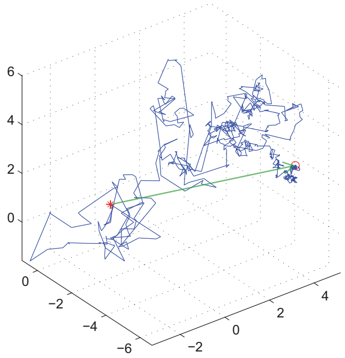
$$E\{\mathrm{w}[n]\} = 0$$
$$E\{\mathrm{w}[n_1]\mathrm{w}^T[n_2]\} = \mathrm{I}_4 \cdot \delta[n_1 - n_2] \, .$$

- $\Psi(\mathrm{h}) = E\{\bar{\mathrm{i}}[n] | \, \mathrm{h}\}$ (with h drawn from a random distribution), can be viewed as a direct arrival, followed by a 3-dimensional random walk consisting of independent random steps with exponentially decaying magnitude. In the sequel, we provide a statistical analysis of the distraction from the direct arrival.

Statistical RIR model III

- The norm of these reflections amounts to a Maxwell variable:

$$f_Q(q) = \epsilon(q)\,\frac{1}{\sigma_G^3}\sqrt{\frac{2}{\pi}}\,q^2 e^{\frac{-q^2}{2\sigma_G^2}}\ .$$

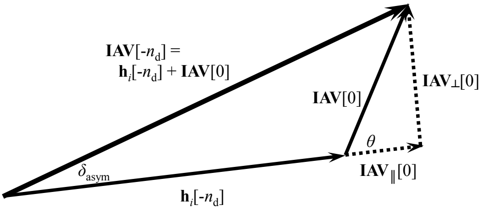


This random walk in three dimensions corresponds to the summation of 8000 terms from $h_p[n]\mathrm{h}_v[n]$. The walk commences at the origin (*) eventually arriving at ($\circ$). It corresponds to the reverberant part of $h_p[n]\mathrm{h}_v[n]$ and does not include the direct arrival.

Statistical RIR model IV

- The angular error can be approximated by:

$$\delta_{\mathrm{asym}} \approx \frac{\|\mathsf{IAV}[0]\| \sin(\theta)}{\|\mathsf{h}_i[-n_{\mathrm{d}}]\|}$$



- The projection of the Maxwell variable onto a 2-dimensional plane, produces a Rayleigh variable:

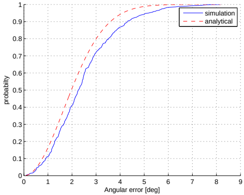
$$f_R(r) = \epsilon(r)\,\frac{1}{\sigma_G^2}\, r e^{\frac{-r^2}{2\sigma_G^2}}\,.$$

Statistical RIR model V

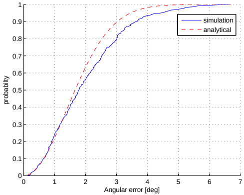
- The asymptotic bias δ_{asym} can thus be approximated, in high direct to reverberant ratio (DRR), by a scaled version of the Rayleigh variable, with standard deviation dependent on the DRR and the decaying factor of the RIRs α.

Evaluation I

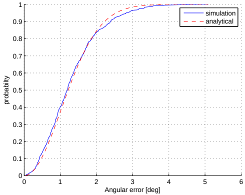
- We have used simulated RIRs from [Habets, 2006] to evaluate our statistical model.
- Evaluation carried out by comparing theoretical and the simulated c.d.f. of the angular error and its respective mean.
- Since the early reflections are not modelled by our statistical analysis, we examined removing them for better match.



(a) Full RIR



(b) 10 msec removed



(c) 20 msec removed

Evaluation II

Table: Statistical results of AE (in degrees) for original and modified data.

Data set	Theoretical AE mean	Sample AE mean	Relative estimation error
Unmodified data	2.0995	2.4748	-15.17%
10 msec removed	1.7657	2.0088	-12.10%
20 msec removed	1.2979	1.3112	-1.02%

A very good correspondence between theory and simulated RIRs, provided early are reflections removed!

Conclusions

1. A single vector sensor is capable of DOA estimation.
2. Reverberation causes estimation bias.
3. Bias distribution may be obtained through statistical room acoustics and is shown to obey Rayleigh distribition.
4. Increased signal coloration tends to increase bias [Levin et al., 2010a].

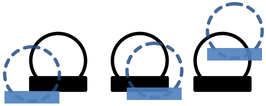
Outline

Robust beamforming [Levin et al., 2013]

- The actual performance of a beamformer can differ greatly from its ideal performance.
- This degradation is due to sensitivity towards deviation from the assumed scenario.
- Diagonal loading is a known cure for robustifying beamformers.
- We show that a different loading is more successful for arrays of nonidentical sensors.

Examples of deviations:

- weight values
- sensor locations
- internal sensor noise
- steering vector



Directivity Patterns of sensors

- We discuss arrays containing elements with nonidentical directivity patterns.

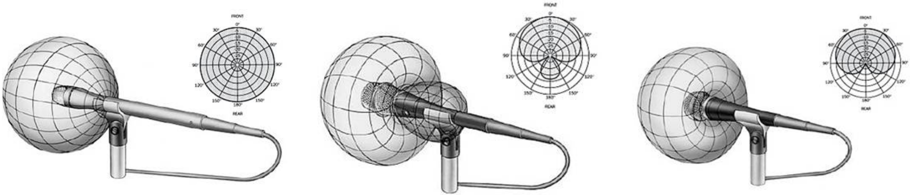
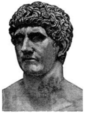


Image credits: Shure Inc. and ProSoundWeb (not for commercial use).

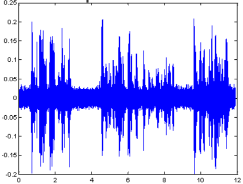
- Different orientation means nonidentical.

Example: robust vs. nonrobust

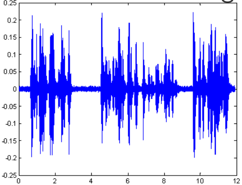
- Simulation of a linear array of 4 vector-sensors (16 subsensors).
- Signal in presence of diffuse noise.



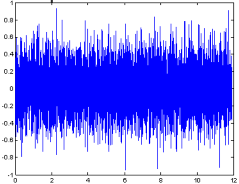
unprocessed



robust beamforming



"optimal" MVDR



Sound credits: original recording from HarperAudio (not for commercial use).

Diagonal loading as modification of MVDR

MVDR beamformer

Objective: minimize noise.

Constraint: unity response in look direction.

Formulation:

$$\text{minimize:} \quad w^H \Phi w$$
$$\text{s.t.} \quad w^H v = 1$$

$$w_{MVDR} = \frac{\Phi^{-1} v}{v^H \Phi^{-1} v}$$

Robust DL beamformer

Objective: minimize noise.

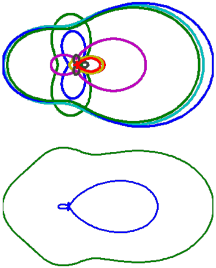
Constraint: 1. unity response in look direction.
2. constrain $\|w\|^2$.

Formulation:

$$\text{minimize:} \quad w^H \Phi w$$
$$\text{s.t.} \quad w^H v = 1$$
$$\|w\|^2 \leq C$$

$$w_{DL} = \frac{(\Phi + \mu I)^{-1} v}{v^H (\Phi + \mu I)^{-1} v}$$

Impact of weight and placement errors



THE BELL SYSTEM TECHNICAL JOURNAL, MAY 1955

Optimum Design of Directive Antenna Arrays Subject to Random Variations

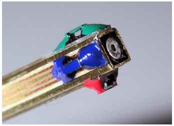
By E. N. GILBERT and S. P. MORGAN

- Each realization of errors creates a different beam-power pattern.
- The **mean beam-power** is used to evaluate extent of deviation from design.

> Sensitivity is proportional to $\|\mathrm{w}\|^2$;
> diagonal loading increases robustness.

Applicability to nonidentical sensors?

- Conventional assumption of sensors with identical directivities does not always hold.



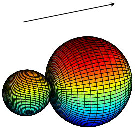
microflown

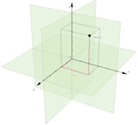
soundfield microphone

- Should the diagonal loading method be applied to an array with nonidentical beampatterns? Are modifications in place?
- We extend the classical work of [Gilbert and Morgan, 1955]

Notation

- N sensors
- Direction of arrival (DOA) unit vector: u
- Directivity pattern of the sensors ($N \times 1$ vector): $\mathrm{b}(\mathrm{u}, \omega)$
- Wave number: $k = \frac{\omega}{c}$
- Position of sensors ($N \times 3$ matrix): P





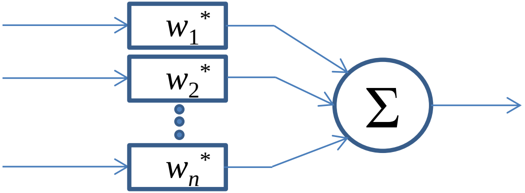
Array manifold vector

$$\mathrm{v}(\mathrm{u}, \omega) = \mathrm{b}(\mathrm{u}) \odot \exp\{j\, k \cdot \mathrm{P}^T \mathrm{u}\}$$

ω will henceforth be omitted for brevity

Weights and beampattern / beam-power

- Vector of weights for each channel: $\mathbf{w} = [w_1 \; w_2 \ldots w_n]^T$



Beampattern

$$\mathrm{BP}(\mathbf{u}) = \mathbf{w}^H \mathbf{v}(\mathbf{u})$$

Beam-power

$$\Psi(\mathbf{u}) = |\mathbf{w}^H \mathbf{v}(\mathbf{u})|^2$$

Random perturbation of weights

- Weights are modeled as: nominal value + error

$$\mathrm{w} = \mathrm{w}_0 + \mathrm{w}_e$$

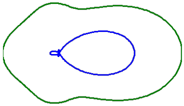
- Errors are random with zero-mean.
- Mean beampattern is unaffected by perturbations:

$$E_{\mathrm{w}}\{\mathrm{BP}\} = E\left\{(\mathrm{w}_0 + \mathrm{w}_e)^H \mathrm{v}(\mathrm{u})\right\} = \mathrm{w}_0^H \mathrm{v}(\mathrm{u})$$

- Mean beam-power is affected by perturbations:

$$\begin{aligned} E\{\Psi\} &= E_{\mathrm{w}}\{|(\mathrm{w}_0 + \mathrm{w}_e)^H \mathrm{v}(\mathrm{u})|^2\} \\ &= |\mathrm{w}_0^H \mathrm{v}(\mathrm{u})|^2 + E_{\mathrm{w}}\{|\mathrm{w}_e^H \mathrm{v}(\mathrm{u})|^2\} \end{aligned}$$

nominal beam-power | *excess beam-power*



Excess beam-power due to weight errors

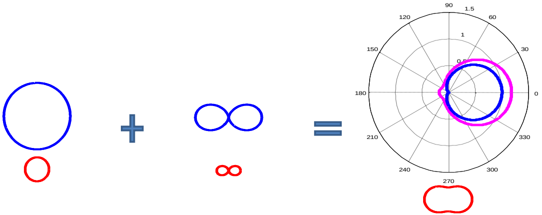
Assuming that weight errors are uncorrelated:

$$\Psi_{\text{ex}} = \sum_{n=1}^{N} E_{\text{w}}\{|w_{\text{e}_n}|^2\} \cdot |b_n(\text{u})|^2$$

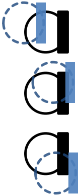
Interpretation:
Weight errors cause excess mean beam-power proportional to directivity power of sensors.

Example:

nominal,
excess,
total



Random perturbation of sensors location



- Location is modeled as: nominal value + error
  $$\mathrm{P} = \mathrm{P}_0 + \mathrm{P}_e$$
- Errors are random with zero-mean.
- Mean beampattern is affected by perturbations:

$$E_{\mathrm{P}}\{\mathrm{BP}(\mathrm{u})\} = \mathrm{w}^H(\mathrm{v}_0(\mathrm{u}) \odot E_{\mathrm{P}}\{\exp\{j\, k \cdot \mathrm{P}_e^T \mathrm{u}\}\}$$

- Assuming that errors have identical spherical distributions,

$$E_{\mathrm{P}}\{\mathrm{BP}(\mathrm{u})\} = \kappa \boxed{\mathrm{w}^H \mathrm{v}_0(\mathrm{u})} \quad \textit{nominal beampattern}$$

Mean beampattern attenuated via constant coefficient but retains shape, since distributions are spherical:

$$\kappa = E_{\mathrm{P}}\left\{\exp\{j\, k \cdot \mathrm{p}_{e_m}^T \mathrm{u}\}\right\}$$

Excess beam-power due to location errors I

- The mean beam-power $E_{\mathrm{P}}\{\Psi\}$ is:

$$E_{\mathrm{P}}\left\{|\mathrm{w}^H \mathrm{v}(\mathrm{u})|^2\right\} = \mathrm{w}^H E_{\mathrm{P}}\left\{\mathrm{v}(\mathrm{u})\mathrm{v}^H(\mathrm{u})\right\}\mathrm{w}\,.$$

- The term $E_{\mathrm{P}}\left\{\mathrm{v}(\mathrm{u})\mathrm{v}^H(\mathrm{u})\right\}$ can be simplified:

$$E_{\mathrm{P}}\left\{\mathrm{v}(\mathrm{u})\mathrm{v}^H(\mathrm{u})\right\} = \mathrm{v}_0(\mathrm{u})\mathrm{v}_0^H(\mathrm{u}) \odot E_{\mathrm{P}}\left\{\exp\{j\,k\cdot \mathrm{P}_{\mathrm{e}}^T\mathrm{u}\}(\exp\{j\,k\cdot \mathrm{P}_{\mathrm{e}}^T\mathrm{u}\})^H\right\}.$$

- We assume a scenario with shared packaging, namely that some of the elements are displaced together; otherwise displacement errors are independent.
- Packaging denoted by matrix Ξ, then

$$E_{\mathrm{P}}\left\{\exp\{j\,k\cdot \mathrm{P}_{\mathrm{e}}^T\mathrm{u}\}(\exp\{j\,k\cdot \mathrm{P}_{\mathrm{e}}^T\mathrm{u}\})^H\right\} = \kappa^2\cdot 1_{N\times N} + (1-\kappa^2)\cdot\Xi$$

Excess beam-power due to location errors II

- An example of shared packing and the associated matrix Ξ:

$$\Xi = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

- Mean beam-power is then given by:

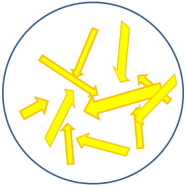
$$E_{\mathrm{P}}\{\Psi\} = \kappa^2 \cdot \Psi_0(\mathrm{u}) \qquad \textit{nominal beam-power}$$
$$+ (1-\kappa^2) \cdot \mathrm{w}^H [\Xi \odot \mathrm{v}_0(\mathrm{u}) \mathrm{v}_0^H(\mathrm{u})] \mathrm{w} \qquad \textit{excess beam-power}$$

- If sensors do not share packaging then Ξ becomes the identity matrix, resulting in:

$$\Psi_{\mathrm{ex}} = (1-\kappa^2) \sum_{n=1}^{N} w_n^2 \cdot |b_n(\mathrm{u})|^2 .$$

Reducing mean noise power

Goal: Reduce mean noise level.



- The designer controls nominal weights (w_0).
- This implicity affects the weight errors (w_e)
  under the assumption:

$$E_\mathrm{w}\{|w_{\mathrm{e}_n}|^2\} = \beta^2 |w_{0_n}|^2$$

- Integrating mean beam-power over all relevant directions of noise,

$$\Phi_0 = \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} A(\mathrm{u})\mathrm{h}(\mathrm{u})\mathrm{h}^H(\mathrm{u}) \sin(\theta) d\theta \, d\phi$$

yields:

$$\kappa^2 \mathrm{w}_0^H \Phi_0 \mathrm{w}_0 + \beta^2 \mathrm{w}_0^H (\Phi_0 \odot \mathrm{I}) \mathrm{w}_0 + (1-\kappa^2) \cdot \mathrm{w}_0^H (\Xi \odot \Phi_0) \mathrm{w}_0 \,.$$

excess noise

nominal noise

Sensitivity to errors

- Typical MVDR design seeks to minimize the term $\mathbf{w}_0^H \Phi_0 \mathbf{w}_0$, while maintaining unity gain towards the desired look-direction.
- A robust design will seek to minimize the excess noise terms as well:

$$\Phi_{\mathrm{rob}} = \Phi_0 \odot \left[\kappa^2 \cdot \mathbf{1}_{N \times N} + \beta^2 \cdot \mathbf{I} + (1 - \kappa^2) \cdot \Xi\right)$$

- Excess noise can be written as:

$$\|\mathbf{w}_0\|_{\mathbf{L}}^2 = \mathbf{w}_0^H \mathbf{L} \mathbf{w}_0$$

- Reducing sensitivity to excess noise can be obtained by constraining $\|\mathbf{w}_0\|_{\mathbf{L}}^2$. However, in practice β and κ are unknown.
- We propose using $\mathbf{L}$ as a (not necessarily) diagonal loading matrix:

$$\mathbf{L} = \alpha \cdot \mathbf{w}_0^H (\Phi_0 \odot \mathbf{I}) \mathbf{w}_0 + (1 - \alpha) \cdot \mathbf{w}_0^H (\Xi \odot \Phi_0) \mathbf{w}_0$$

- If $\Xi = \mathbf{I}$ and $\mathbf{L}$ is not dependent on α.

Loading matrix

Robust beamformer

Objective: minimize noise.

Constraint: 1. unity response in look direction.

2. constrain $\|\mathrm{w}\|_{\mathrm{L}}^{2}$.

Formulation:

$$\text{minimize:} \quad \mathrm{w}^{H}\Phi\mathrm{w}$$

$$\text{s.t.} \quad \mathrm{w}^{H}\mathrm{v}=1$$

$$\|\mathrm{w}\|_{\mathrm{L}}^{2} \leq C$$

$$\mathrm{w}_{\text{robust}} = \frac{(\Phi+\mu\mathrm{L})^{-1}\mathrm{v}}{\mathrm{v}^{H}(\Phi+\mu\mathrm{L})^{-1}\mathrm{v}}$$

Notes:

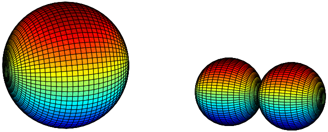
1. Loading matrix L formed from elements of nominal noise matrix (Φ_0).
2. L is not necessarily diagonal (shared packaging).
3. Main diagonal may be nonuniform.
4. Conventional case reverts to traditional DL.

Evaluation – 1

- Inverse of sensitivity to errors used as robustness metric.

$$\mathrm{RM} = 1/\|\mathrm{w}_0\|_{\mathrm{L}}^2$$

- **Scenario 1**: Collocated monopole and dipole in diffuse noise field.

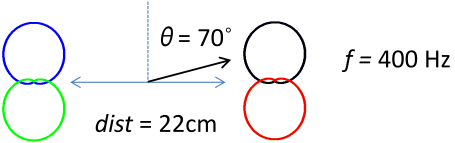


- Attempting to attain maximum robustness ($\mu \to \infty$):

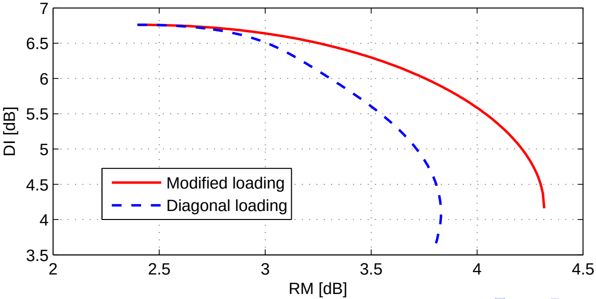
	Loading matrix	DI	RM	Beampattern
Diagonal loading	$\mathbf{I}$	4.8 dB	3.0 dB	cardioid
Modified loading	$\mathbf{L} = \begin{bmatrix} 1 & 0 \\ 0 & 1/3 \end{bmatrix}$	6.0 dB	3.6 dB	hypercardioid

Evaluation – 2

- **Scenario 2**: Two pairs of collocated cardioids in diffuse noise:



- DI-RM tradeoffs for both methods (for position errors, $\alpha = 0$):



Conclusions

- **Topic**: arrays with nonidentical sensors.
- ☹ Weight and location errors increase mean beam-power.
- ☺ Robust design with modified loading matrix.
- Loading matrix produced from noise covariance matrix.
- ☺ Method improves directivity/robustness tradeoff.
- Method does not involve extra computational burden.

Outline

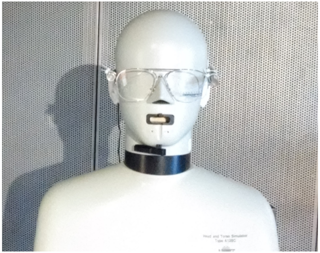
Smartglasses: own voice enhancement [Levin et al., 2016] I

Motivation

- Recent years have witnessed an increased interest in wearable devices, consisting of miniature computers worn by users which can perform certain sensing, networking and computing tasks.
- One specific type of wearable computer is the smartglasses: a device which displays computer generated information supplementing the user's visual field (augmented reality), e.g., Google Glass and Microsoft HoloLens.
- In addition to their visual-output capabilities, smartglasses may incorporate acoustic sensors.
- These sensors are used for hands-free mobile telephony applications, and for applications using a voice-control interface to convey commands and information to the device.

Smartglasses: own voice enhancement [Levin et al., 2016] II

Motivation



- We propose a system for the acquisition of the desired near-field speech in a noisy environment, based on an acoustic array embedded in eyeglasses frames worn by the desired speaker.
- Glasses frames constitute a spatially compact platform, with little room to spread the sensors out.
- We choose to use AVS due to their compact dimensions and inherit super-directivity, especially in near-field scenarios.

Problem Formulation I

Microphone signals:

$$x_m[n] = (s * h_m)[n] + \left(\sum_{p=1}^{P} (z_p * g_{m,p})[n] \right) + \epsilon_m[n] \, .$$

with $M = 8$ sensors (2 AVS's), P interference sources, h_m RIR relating the desired source and sensor m, and $g_{m,p}$ RIRs relating the pth source and sensor m.

Vector form:

$$\mathrm{x}[n] = (s * \mathrm{h})[n] + \left(\sum_{p=1}^{P} (z_p * \mathrm{g}_p)[n] \right) + \boldsymbol{\epsilon}[n] \, .$$

Problem Formulation II

Direct-path and interference:

$$\mathrm{x}[n] = (s * \mathrm{h}_{\mathrm{d}})[n] + \mathrm{e}[n]\,,$$

with

- $\mathrm{h}[n] = \mathrm{h}_{\mathrm{d}}[n] + \mathrm{h}_{\mathrm{r}}[n]$,
- $\mathrm{e}[n]$ - undesired sound sources, ambient and sensor noise, and reverberation.

STFT domain:

$$\mathrm{x}[n] \mapsto \mathrm{x}(\ell, k) = \begin{bmatrix} X_1(\ell, k) \\ X_2(\ell, k) \\ \vdots \\ X_M(\ell, k) \end{bmatrix} = \mathrm{h}_{\mathrm{d}}(k) s(\ell, k) + \mathrm{e}(\ell, k)\,.$$

Problem Formulation III

Normalized form using relative transfer function (RTF):

$$\mathrm{x}(\ell, k) = \widetilde{\mathrm{h}}_{\mathrm{d}}(k)\widetilde{s}(\ell, k) + \mathrm{e}(\ell, k)\,,$$

with

$$\widetilde{s}(\ell, k) = \mathrm{c}^H(k)\mathrm{h}_{\mathrm{d}}(k)s(\ell, k)$$

and

$$\widetilde{\mathrm{h}}_{\mathrm{d}}(k) = \frac{\mathrm{h}_{\mathrm{d}}(k)}{\mathrm{c}^H(k)\mathrm{h}_{\mathrm{d}}(k)}\,.$$

where the vector $\mathrm{c}(k)$ determines the linear combination/microphone selection, e.g., $\mathrm{c}(k) = [1\,0 \cdots 0]^T$.

Proposed beamforming framework I

Beamforming:

$$y(\ell, k) = \mathrm{w}^H(\ell, k)\mathrm{x}(\ell, k)\,,$$

MVDR beamformer:

$$\mathrm{w}_{\mathrm{MVDR}}(\ell, k) = \frac{\Phi_{\mathrm{ee}}^{-1}(\ell, k)\widetilde{\mathrm{h}}_{\mathrm{d}}(k)}{\widetilde{\mathrm{h}}_{\mathrm{d}}^{H}(k)\Phi_{\mathrm{ee}}^{-1}(\ell, k)\widetilde{\mathrm{h}}_{\mathrm{d}}(k)}\,.$$

MWF beamformer:

$$\mathrm{w}_{\mathrm{MWF}}(\ell, k) = \mathrm{w}_{\mathrm{MVDR}}(\ell, k) \cdot W(\ell, k)\,.$$

with $W(\ell, k)$, a single-channel Wiener postfilter:

$$W(\ell, k) = \frac{1}{1 + \mathrm{SNR}^{-1}(\ell, k)}\,.$$

Proposed beamforming framework II

Improved robustness:

$$\mathbf{w}_{\mathrm{reg}}(\ell, k) = \begin{cases} \mathbf{w}_{\mathrm{MVDR}}(\ell, k), & \text{if } \|\mathbf{w}_{\mathrm{MVDR}}(\ell, k)\|_2 \leq \rho \\ \rho \frac{\mathbf{w}_{\mathrm{MVDR}(\ell,k)}}{\|\mathbf{w}_{\mathrm{MVDR}(\ell,k)}\|}, & \text{otherwise.} \end{cases}$$

Noise covariance estimation:
Covariance matrix is subject to rapid changes due to moving interfering sources and/or abrupt head movement:

$$\widehat{\mathrm{R}}_{\mathrm{ee}}(\ell, k) = \alpha(\ell, k)\widehat{\mathrm{R}}_{\mathrm{ee}}(\ell - 1, k) + \big(1 - \alpha(\ell, k)\big)\mathbf{x}(\ell, k)\mathbf{x}^H(\ell, k)$$

with

$$\alpha(\ell, k) = \begin{cases} 1, & \text{if desired speech is detected} \\ \alpha_0, & \text{otherwise.} \end{cases}$$

Proposed beamforming framework III

RTF estimation:
Since the mouth-sensors constellation is fixed, the RTF vector is estimated during noise-free training stage:

$$\widetilde{\mathrm{h}}_{\mathrm{d}}(k) = \frac{\sum_{\ell=1}^{L_{\mathrm{b}}} \mathrm{x}(\ell,k)\mathrm{x}^H(\ell,k)\mathrm{c}(\ell,k)}{\sum_{\ell=1}^{L_{\mathrm{b}}} |\mathrm{c}^H(\ell,k)\mathrm{x}(\ell,k)|^2} .$$

Near-field speech detection:

$$T(\ell,k) = \frac{|\mathrm{x}^H(\ell,k)\widetilde{\mathrm{h}}_{\mathrm{d}}(k)|^2}{|\mathrm{x}(\ell,k)|^2 |\widetilde{\mathrm{h}}_{\mathrm{d}}(k)|^2} ,$$

Geometrically, $T(\ell,k)$ corresponds to the square of the cosine of the angle between the two vectors x and $\widetilde{\mathrm{h}}_{\mathrm{d}}$.

Proposed beamforming framework IV

Post-filtering:
SNR is estimated using a variant of [Ephraim and Malah, 1984] "decision-directed" approach:

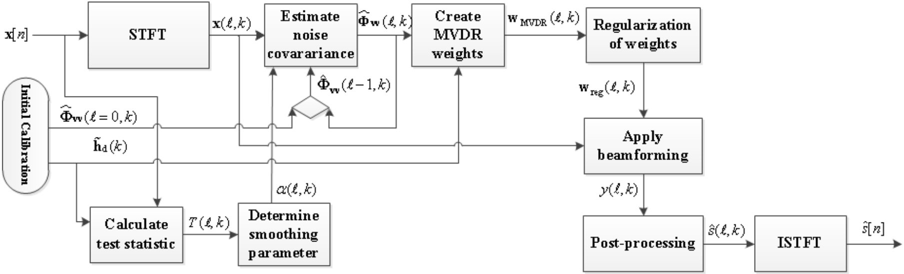
$$\gamma(\ell,k) = \frac{|\mathbf{w}_{\mathrm{reg}}^H(\ell,k)\mathbf{x}(\ell,k)|^2}{\mathbf{w}_{\mathrm{reg}}^H(\ell,k)\widehat{\mathbf{R}}_{\mathbf{ee}}(\ell,k)\mathbf{w}_{\mathrm{reg}}(\ell,k)}$$

$$\widehat{\mathrm{SNR}}(\ell,k) = \beta|\widehat{W}(\ell-1,k)|^2\gamma(\ell-1,k) + (1-\beta)\max\{\gamma(\ell,k)-1, \mathrm{SNR}_{\min}\}$$

$$\widehat{W}(\ell,k) = \max\left\{\frac{\widehat{\mathrm{SNR}}(\ell,k)}{1+\widehat{\mathrm{SNR}}(\ell,k)}, W_{\min}\right\}.$$

The parameters $\mathrm{SNR}_{\min}$ and $W_{\min}$ are used to eliminate the musical noise phenomenon.

Block diagram



Sound examples

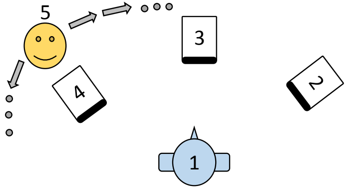


Figure: #1:HATS; #2–#4: static loudspeakers; #5: moving human speaker. HATS-loudspeakers distance 1 m. Medium reverb.

Scenario A: HATS + 3 static interferes #2,#3,#4:

Microphone MVDR MVDR+Post1 MVDR+Post2

Scenario B: HATS + moving speaker #5:

Microphone MVDR

Conclusions

- We proposed an array which consists of two AVS's mounted on an eyeglasses frame.
- The specific configuration circumvents the need to re-estimating the steering vector.
- The proposed MVDR algorithm adapts to changes of the noise characteristics by continuously estimating the noise covariance matrix.
- A speech detection scheme is used to identify the presence of time-frequency bins containing desired speech and preventing them from corrupting the estimation of the noise-covariance matrix.
- Experiments confirm that the proposed system performs well in both static and changing scenarios.
- The proposed system may be used to improve the quality of speech acquisition in smartglasses.

References and Further Reading I

Davies, S. (1987).
Bearing accuracies for arctan processing of crossed dipole arrays.
In *OCEANS'87*, pages 351–356. IEEE.

Ephraim, Y. and Malah, D. (1984).
Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator.
IEEE Transactions on acoustics, speech, and signal processing, 32(6):1109–1121.

Gilbert, E. and Morgan, S. (1955).
Optimum design of directive antenna arrays subject to random variations.
Bell System Technical Journal, 34(3):637–663.

Habets, E. A. (2006).
Room impulse response generator.
Technische Universiteit Eindhoven, Tech. Rep, 2(2.4):1.

Levin, D., Gannot, S., and Habets, E. (2010a).
Impact of source signal coloration on intensity vector based DOA estimation.
In *The International Workshop on Acoustic Echo and Noise Control (IWAENC)*, Tel-Aviv, Israel.

Levin, D., Gannot, S., and Habets, E. (2011).
Direction-of-arrival estimation using acoustic vector sensors in the presence of noise.
In *The IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 105–108, Prague, Czech Republic.

Levin, D., Habets, E., and Gannot, S. (2010b).
On the angular error of intensity vector based direction of arrival estimation in reverberant sound fields.
The Journal of the Acoustical Society of America, 128:1800–1811.

References and Further Reading II

Levin, D., Habets, E., and Gannot, S. (2012).
Maximum likelihood estimation of direction of arrival using an acoustic vector-sensor.
The Journal of the Acoustical Society of America, 131(2):1240–1248.

Levin, D., Habets, E., and Gannot, S. (2013).
Robust beamforming using sensors with nonidentical directivity patterns.
In *The IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vancouver, Canada.

Levin, D. Y., Habets, E. A. P., and Gannot, S. (2016).
Near-field signal acquisition for smartglasses using two acoustic vector-sensors.
Speech Communication, 83:42–53.

Nehorai, A. and Paldi, E. (1994).
Acoustic vector-sensor array processing.
IEEE Transactions on signal processing, 42(9):2481–2491.

Schroeder, M. R. (1965).
New method of measuring reverberation time.
The Journal of the Acoustical Society of America, 37(6):1187–1188.