Planarity Panning for Listener-Centered Spatial Audio

Philip Coleman, Philip J. B. Jackson

Centre for Vision, Speech and Signal Processing,
University of Surrey, Guildford, UK

p.d.coleman@surrey.ac.uk

28th August 2014
AES 55th International Conference on Spatial Audio, Helsinki

Future Spatial Audio
for an Immersive Listener Experience
Introduction

- **Methods to generate directional sound [1]**
  - Panning (Amplitude panning, VBAP)
  - Sound field synthesis (WFS, HOA)
  - Multi-point optimization of sound pressures
  - *Wavenumber domain point focusing [2]*


Introduction

• Desired properties of a spatial audio method
  – Listener perceives sound at the intended location
  – Arbitrary loudspeaker arrangement
  – Efficient
  – Applications: home, car, mobile
  – Listener-centered
Introduction

• “Listener-centered” approach
  – Listener occupies a zone (and not elsewhere)
  – Target directions specified at listener position
  – Sound field measurements at control microphones in zone
  – Directional information obtained from the listening position
  – Precision of sound image placement is relaxed
    • Limited resolution of human localization
    • Reduce aliasing effect with fewer loudspeakers
Introduction

• Efficient sound focusing
  – Brightness control [3]
  – Offers no control over the spatial image

• Multi-point optimization
  – Requires strict sound field definition
  – All deviations treated as equal

• Planarity concept
  – Useful for personal audio [4]

Background

- Notation

\[
p(x, \omega) = - \int_{\partial V} G(x|x_0, \omega) q(x_0, \omega) dA(x_0)
\]
Background

- **Notation (discrete)**
  - Vector notation for a single frequency
  - Source weights
    \[ q = [q(x^1_c), \ldots, q(x^L_c)]^T \]
  - Microphone pressures
    \[ p = [p(x^1_s), \ldots, p(x^N_s)]^T \]
  - Transfer function matrix
    \[
    G = \begin{pmatrix}
    G(x^1_s|x^1_c) & \cdots & G(x^1_s|x^L_c) \\
    \vdots & \ddots & \vdots \\
    G(x^N_s|x^1_c) & \cdots & G(x^N_s|x^L_c)
    \end{pmatrix}
    \]
  - Sound pressures
    \[ p = Gq \]
Planarity panning

- **Angular spectrum**
  - Observed energy in zone

\[
\Gamma = \text{diag}[\gamma_1, \gamma_2, \ldots, \gamma_I]
\]
Planarity panning

- Superdirective beamforming
  - Steering matrix relates sound pressures to azimuth

\[
Y = \begin{bmatrix}
\text{angles} \\
\text{microphones}
\end{bmatrix}
\]

- Cost Function

\[
J = p^H Y^H \Gamma Y p - \lambda (q^H q - Q)
\]

constraint on source weights

listening zone energy projected in to angular domain [5]

Method

• Simulation geometry
  – Up to 60 channel circular array (2-D reproduction)
  – Centre/Off-centre listening positions
  – Virtual 5.0 system

• Pass range
  – Single target angle with ±5° raised-cosine roll-off
Method

• **Evaluation metrics:**
  
  – RMSE of energy direction
    
    \[
    \epsilon = \sqrt{\frac{1}{F} \sum_{f=1}^{F} |\alpha(f) - \varphi|^2}
    \]
  
  – Planarity
    
    \[
    \eta = \frac{\sum_i w_i \mathbf{u}_i \cdot \mathbf{u}_\alpha}{\sum_i w_i}
    \]
  
  – Control effort
    
    \[
    E = 10 \log_{10} \left( \frac{q^H q}{|q_r|^2} \right)
    \]
Method

- **Methods for comparison**
  - Amplitude panning (AP)
  - Wave field synthesis (WFS)* [6]
  - Higher-order ambisonics (HOA)* [7]
  - Pressure matching (PM) [8]

*Source weights generated using SFS toolbox [9]*

Simulations

- Planarity panning performance
  - 60 loudspeakers, 1 kHz, 0 deg. virtual source

![Central position](image1.png) ![Side position](image2.png)
Simulations

• **Planarity panning performance**
  – 60 loudspeakers, 1 kHz, 0 deg. virtual source
  – Sound energy arrives from correct direction
  – Energy focused on zone
Simulations

- **Planarity panning performance**
  - 60 loudspeakers, central position, 0 deg. virtual source

- Planar sound field maintained over a range of frequencies

![Simulations](image)
Virtual 5.0 surround

- Virtual source placed at 5.0 locations
  - 60 loudspeakers, 1 kHz
  - Off-centre zone (approx. listener position!)

---

LS  
Error: 1°

L  
0°

C  
0°

R  
0°

RS  
1°
Method comparison

- **Central position**
  - 20 loudspeakers, 1 kHz, 0 deg. virtual source
  - 325 Hz spatial aliasing limit
  - WFS worst spatial aliasing artifacts
  - PP most efficient
  - PM and HOA have similar sound fields

![Graphs comparing different methods](image-url)
Method comparison

- Decreasing loudspeakers
  - Averaged over direction and frequency
  - RMSE increases for all methods as loudspeakers are removed
  - WFS aliasing artifacts affect whole listening region
  - HOA planarity reduces more than other methods
  - PP always least-effort
Method comparison

- **Side position**
  - 20 loudspeakers, 1 kHz, 0 deg. virtual source
  - 325 Hz spatial aliasing limit
  - HOA, PP and PM have similarities
  - HOA slightly lower planarity
  - PP most efficient
Method comparison

• **Zone position**
  – Averaged over direction and frequency
  – RMSE increases for off-centre position (all methods)
  – Effort decreases for off-centre position (all methods)
Summary and future work

• Planarity panning (PP) proposed
• Efficient directional sound over a listener-sized zone
  – Best $1.7^\circ$ RMSE (PM $1.2^\circ$); -5 dB effort (AP -0.8 dB)
• PP behaves similarly to HOA
  – Lower effort, higher HF planarity
• Future work
  – Integration with listener tracking
  – Use measured RIR for room compensation in filters
  – Equalization filters to improve sound quality
  – Formal listening test scores
News!

• New resources freely available via
  www.cvssp.org/soundzone/resource

• Surrey Studio 2 RIR Dataset (SOFA format)
  – 60 channel circular array in recording studio
    (used for AES 52\textsuperscript{nd} conference demonstrations)
  – RIRs to each of 864 microphone positions (three sampled zones)

• Planarity metric
  – Matlab function to evaluate the planarity of a sound field
  – Useful for evaluation when there is no target field
Thanks for your attention!

✉️ p.d.coleman@surrey.ac.uk

🔗 www.linkedin.com/in/philipcolemanaudio
Listening test

• Compared PP, WFS in Surrey Sound Sphere
  – Up to 60 loudspeakers in circular array
Listening test

- Compared PP, WFS in Surrey Sound Sphere
  - Up to 60 loudspeakers in circular array
  - PP implemented in frequency domain based on free-field transfer functions
  - WFS reproduced using Soundscape Renderer
    http://spatialaudio.net/ssr/
  - Loudspeakers obscured from listening positions
• **Localization**
  - Subjects asked to mark the direction of three sound sources at 0, -29 and -63 degrees azimuth
  - Tested for centre and off-centre listening zones

**Localization error**

<table>
<thead>
<tr>
<th>Target region</th>
<th>RMSE (deg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zone A</td>
<td></td>
</tr>
<tr>
<td>Zone B</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td></td>
</tr>
</tbody>
</table>

- WFS
- PP

Image: [Diagram of Localization Error](image-url)