University of Huddersfield Repository

Lee, Hyunkook

Investigation on the Phantom Image Elevation Effect

Original Citation


This version is available at http://eprints.hud.ac.uk/26558/

The University Repository is a digital collection of the research output of the University, available on Open Access. Copyright and Moral Rights for the items on this site are retained by the individual author and/or other copyright owners. Users may access full items free of charge; copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational or not-for-profit purposes without prior permission or charge, provided:

- The authors, title and full bibliographic details is credited in any copy;
- A hyperlink and/or URL is included for the original metadata page; and
- The content is not changed in any way.

For more information, including our policy and submission procedure, please contact the Repository Team at: E.mailbox@hud.ac.uk.

http://eprints.hud.ac.uk/
Investigation on Phantom Image Elevation

Hyunkook Lee
h.lee@hud.ac.uk

Applied Psychoacoustics Lab (APL)
University of Huddersfield, UK
Contents

- Research background
- Experiment
- Summary
Pitch-Height Effect for “Real” Source

• The higher the frequency of a pure tone is, the higher the perceived image position is, regardless of the physical height of the loudspeaker. (Pratt 1930).

• Confirmed by Trimble (1934), Roffler and Butler (1968), etc.
Pitch-Height Effect for “Real” Source

- For **band-passed noise** signals, high frequency components (above 7kHz) are essential for accurate vertical localisation. (Roffler and Butler 1968b)

From Roffler and Butler (1968b)
- Low passed & high passed noise
  - 4.8kHz tone
  - 600Hz tone
  - LPF noise < 2kHz

HPF noise > 8k
HPF noise > 2k
Broadband
Pitch-Height Effect for “Real” Source

- Pitch height effect for **octave band pink noise**
  - Simplified from Cabrera and Tiley (2003); median plane results
Pitch-Height Effect for “Phantom” Source

- Pitch-height effect for horizontal \textit{phantom} images from main and height layers (Lee 2015)

(a) Front view

(b) Side view

Acoustic curtain

Scale (0 – 300)

1.2m

1.78m

60°

30°

Floor level

0°

30°
Pitch-Height Effect for Phantom Source

- Pitch-height effect for horizontal **phantom** image (Lee 2015)

- Overall, the pitch-height effect operates in two separate regions.
- Reset at 1kHz → Back localisation (Blauert’s Directional bands)
Directional bands

- Blauert (1968): physical mapping between frequency bands and their perceived positions in the median plane.
• **Horizontal plane phantom images are elevated**, not only for high frequencies but also for low frequencies (125Hz, 250Hz, 500Hz) → different from “real” source situations.
Pitch-Height Effect for Real Source

• Pitch height effect for octave bands
  – Simplified from Cabrera and Tiley (2003); median plane results

Vertical Image Position (m)

-0.6
-0.3
0
0.3
0.6

125 Hz  500 Hz  2 kHz  8 kHz  Full band

Loudspeaker position

Image position
Horizontal Phantom Image Elevation

Vs.

Loudspeaker Base Angle & Sound Source
Previous studies

• de Boer (1947): Phantom centre image is perceived to be elevated, and the elevation angle increases as the loudspeaker base angle increases. (180° → overhead region)
Previous studies reporting the elevation effect are limited in terms of sound sources or loudspeaker base angles tested.

<table>
<thead>
<tr>
<th>Source</th>
<th>Base angles</th>
</tr>
</thead>
<tbody>
<tr>
<td>de Boer (1947)</td>
<td>0° to 180°</td>
</tr>
<tr>
<td>Damaske and Mellert (1969/1970)</td>
<td>0° to 360°</td>
</tr>
<tr>
<td>Jo et al. (2010)</td>
<td>60°, 220°</td>
</tr>
<tr>
<td>Frank (2014)</td>
<td>40°</td>
</tr>
<tr>
<td>Lee (2015)</td>
<td>60°</td>
</tr>
</tbody>
</table>
Aim of the Current Experiment

- To investigate the phantom image elevation effect for a wide range of sound sources, with base angles covering from 0° to 360°.

- Sound sources
  - Speech, Helicopter, Aeroplane, Thunder, Rain, Bird, Church Bell
  - Broadband pink noises (continuous and transient)
  - Broadband white noises (continuous and transient)
Test Method

• **Loudspeaker arrangement**
  – At the ear height in the horizontal plane, 0° to 360° at 30° interval.
Critical listening room at the University of Huddersfield (ITU-R BS.1116-Compliant)
Test Method

- GUI written in Max
  - Response method similar to Blauert (1968) but in a finer resolution
Test Method

- Subjects
  - 10 people comprising researchers and post-graduate students from the University of Huddersfield’s music technology courses.
  - All were much experienced in spatial quality evaluation but not trained for the particular task of the experiment.
Results

- Responses for all sources
  - The general trend agrees with the suggestions from the past research.
Phantom Image Elevation

- Sound source dependency
  - Responses are most linear and consistent for source with a broad and flat spectrum.

![White noise graph]

![Rain graph]
Phantom Image Elevation

- Sound source dependency
  - Responses are most linear and consistent for source with a broad and flat spectrum.

\[
\text{Perceived elevation angle} = \frac{\text{Loudspeaker base angle}}{2}
\]
Phantom Image Elevation

- Sound source dependency
  - The elevation effect is weaker for sources with more low frequency energy. (no strong “aboveness”)

![Graphs of Pink noise and Speech signals](image-url)
Phantom Image Elevation

- Sound source dependency
  - The elevation effect is weaker for sources with more low frequency energy. (no strong “aboveness”)

![Graph showing perceived image region vs. loudspeaker base angle for Continuous Pink noise and Speech. The graph indicates a trend where the perceived image region changes with the loudspeaker base angle.](image-url)
Phantom Image Elevation

- Sound source dependency
  - The elevation effect is weaker for sources with more low frequency energy. (no strong “aboveness”)

![Airplane Graph](image1)

![Thunder Graph](image2)
Phantom Image Elevation

- Sound source dependency
  - The elevation effect is weaker for sources with more low frequency energy. (no strong “aboveness”)

Airplane

Thunder
• Sound source dependency
  – Responses are most inconsistent for sources with narrow spectrum or steady-state nature.
• Sound source dependency
  – Responses are most inconsistent for sources with narrow spectrum or steady-state nature.
• Expectancy bias
  – Subjective responses affected by the likely auditory or visual positions of the sound sources in real life.

No “directly above” for speech
Perceived lower than rain
Theoretical explanations

• Spectral energy distribution of ear signal

(a) 60 degrees

Delta HRTF (60°– 0°)

(b) 120 degrees

(120°– 0°)

8k

4k

(c) 180 degrees

(180°– 0°)

8k

4k

• As the base angle increases up to 240°, 8kHz energy increases while 4kHz energy decreases. → Increasing “aboveness” & decreasing “frontness”.

(d) 240 degrees

(240°– 0°)

8k

4k

(e) 300 degrees

(300°– 0°)

8k

4k

(f) 360 degrees

(360°– 0°)

4k

8k
Theoretical explanations

- However, spectral energy distribution does not explain the phantom image elevation for **low frequencies**.
  - Phantom image elevation is also perceived for low-frequency dominant sources and for octave-bands such 250Hz and 500Hz bands.
Theoretical explanations

- A new hypothesis from a **cognitive** perspective
  - The brain interprets the acoustic crosstalk delay as a shoulder reflection delay for a real elevated source.
  - Shoulder reflection delay is the main cue for elevation perception for low frequencies in the median plane (Algazi et al. 2001)
Theoretical explanations

• A new hypothesis from a **cognitive** perspective
  – As the loudspeaker base angle increases, acoustic crosstalk delay increases (max. around 0.7ms for 180°)
  – As the real source elevation angle increases, should reflection delay increases (max. around 0.7ms for right above).
Theoretical explanations

- A new hypothesis from a **cognitive** perspective
  - As the loudspeaker base angle increases, acoustic crosstalk delay increases (max. around 0.7ms for 180°)
  - As the real source elevation angle increases, should reflection delay increases (max. around 0.7ms for a source right above).
Theoretical explanations

- A new hypothesis from a cognitive perspective
  - Low frequencies: Cognitive effect (crosstalk – shoulder delay)
  - High frequencies: Hard-wired effect (HRTF, directional bands, etc.)
Applications for 3D music production

- Simply routing overhead sources to the side or rear speaker pair in the conventional 5.1 or 7.1 format can create a virtual overhead image.
  - 3D mix without overhead speakers
  - 3D to 2D downmixing
  - 2D to 3D upmixing
  - Etc.
Conclusions

• Phantom image elevation effect depends on the loudspeaker base angle and sound source characteristics.

• Base angles around $180^\circ$ produces a virtual overhead image.
  → This is most effective for sound sources with a broad and flat frequency spectrum. (e.g. rain, white noise like sources)
  → Phantom image elevation is weaker for sources with low frequency dominance, narrow bandwidth or steady-state characteristics.

• Phantom image elevation can be explained by spectral energy distribution at HF, whereas it is more of a cognitive effect at LF.
Ongoing work

- Relative weighting between different frequency bands in terms of phantom image elevation
- Verification of the cognitive hypothesis
- Virtual overhead panning method


Free download links for useful tools

- **HULTI-GEN**  http://eprints.hud.ac.uk/24809
- **HAART**     http://eprints.hud.ac.uk/24579
- **IAR**       http://eprints.hud.ac.uk/25547

Please contact us for more information:

Hyunkook Lee
h.lee@hud.ac.uk