Sound Zones with a Cost Function based on Human Hearing
MSc Thesis Defence

Niels de Koeijer

Delft University of Technology, The Netherlands

Bang and Olufsen, Denmark

August 29, 2021

Preface:
About

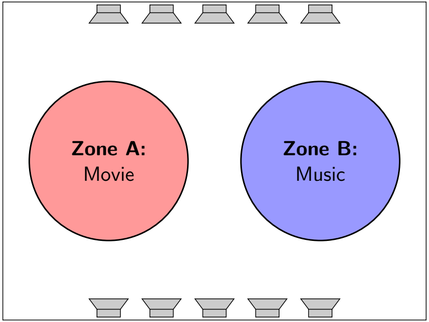
Niels de Koeijer
Master Student Thesis at the Research Department at Bang & Olufsen and the Delft University of Technology

This presentation will detail the work done during my MSc thesis.

Preface:
The Sound Zone Approach

The sound zone approach:

- **Given:**
  A room, an array of loudspeakers, and a number of zones.
- **Goal:**
  Reproduction distinct audio in the specified zones, with minimal interference.



Preface:
Introducing Perceptual Sound Zones

In order to improve sound zones, one recent approach is to include a model of the human auditory system which models how sound is perceived by humans, in the algorithms.

The motivation for this approach is as follows:

- Sound zone algorithms typically optimize over **physical measures** such as sound pressure, which do not always correspond well with how sound is actually perceived.
- By optimizing over **perceptual measure** instead, we are optimizing over what matters perceptually.

Preface:
Objectives & Research Questions

The goal of the project is to investigate the construction and benefits of including auditory perceptual information in sound zone algorithms.

To this end, two research questions are posed:

1. *"How can auditory perceptual models be included in sound zone algorithms?"*
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*

Structure:
Answering Research Questions

1. *"How can auditory perceptual models be included in sound zone algorithms?"*
    1. Determination of a suitable perceptual model for sound zone algorithms.
    2. Proposal of a perceptual sound zone framework using determined model.
    3. Proposal of perceptual sound zone algorithms through proposed framework.
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
    1. Simulation and analysis of proposed perceptual sound zone algorithms.

Structure:
Answering Research Questions

1. *"How can auditory perceptual models be included in sound zone algorithms?"*
    1. **Determination of a suitable perceptual model for sound zone algorithms.**
    2. Proposal of a perceptual sound zone framework using determined model.
    3. Proposal of perceptual sound zone algorithms through proposed framework.
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
    1. Simulation and analysis of proposed perceptual sound zone algorithms.

Determining a Suitable Perceptual Model:
Approach

In order to obtain a suitable perceptual model, various perceptual models from literature were considered.

- Mainly considered were algorithms that assign a perceptual “score” to audio.
- These could be used to propose sound zone algorithms that optimize over this score.

Determining a Suitable Perceptual Model:
Literature Review

To this end, two categories of perceptual models were considered:

- **Objective Audio Measures:** Perceptual models that seek to predict the outcomes of listening tests, e.g. PESQ, PEAQ, Distraction, and STOI.
- **Audio Coding Models:** Perceptual models that are used to make the quantization noise introduced by audio compression as minimally disturbing as possible, e.g. the MPEG perceptual models.

Determining a Suitable Perceptual Model:
Introduction to the Par Distortion Detectability

From this review, the “Par distortion detectability”[1] is selected as the most promising model because of its ease of integration into optimization problems.

The Par distortion detectability defines a mathematical function $D(x[n], \varepsilon[n])$ which models how easily a human listener can detect the disturbance signal $\varepsilon[n]$ in presence of the masking signal $x[n]$.

The detectability $D(x[n], \varepsilon[n])$ assigns a number between 0 and ∞, 0 representing that $\varepsilon[n]$ being completely undetectable in presence of $x[n]$.

[1] S. van de Par, A. Kohlrausch, R. Heusdens, J. Jensen, and S. H. Jensen (2005). “A perceptual model for sinusoidal audio coding based on spectral integration”. In: *EURASIP Journal on Advances in Signal Processing* 2005.9, pp. 1–13

Determining a Suitable Perceptual Model:
Perceptual Background for the Par Distortion Detectability

The detectability $D(x[n], \varepsilon[n])$ determines the detectability of $\varepsilon[n]$ by making use of two principals: the threshold of hearing and the masking threshold of the masking signal $x[n]$.

1. **Hearing Threshold:** The sound levels as a function of frequency below which humans cannot perceive sound.
   - E.g. if the sound is below this threshold, it is not detectable.

2. **Masking Threshold:** The sound levels required for another sound to be audible in presence of $x[n]$.
   - E.g. if $x[n]$ is loud tone, then it will overpower similar tones and they will not be audible.

Structure:
Answering Research Questions

1. *"How can auditory perceptual models be included in sound zone algorithms?"*
    1. **Determination of a suitable perceptual model for sound zone algorithms.**
    2. Proposal of a perceptual sound zone framework using determined model.
    3. Proposal of perceptual sound zone algorithms through proposed framework.
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
    1. Simulation and analysis of proposed perceptual sound zone algorithms.

Structure:
Answering Research Questions

1. *"How can auditory perceptual models be included in sound zone algorithms?"*
    1. Determination of a suitable perceptual model for sound zone algorithms.
    2. **Proposal of a perceptual sound zone framework using determined model.**
    3. Proposal of perceptual sound zone algorithms through proposed framework.
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
    1. Simulation and analysis of proposed perceptual sound zone algorithms.

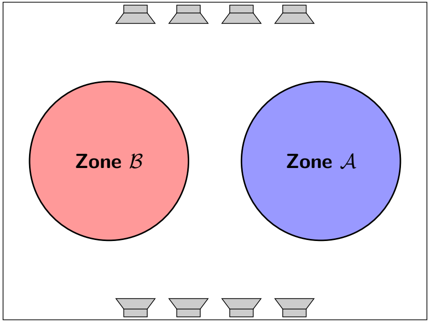
Proposal of Perceptual Sound Zone Framework:
Approach

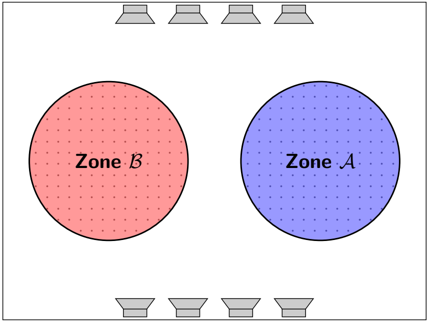
In the previous part we found that the Par distortion detectability is a suitable perceptual model. How can this perceptual model be used to pose perceptual sound zone algorithms?

In order to answer this question, a literature review was performed into various sound zone approaches. From this review it was found that a **Pressure Matching** sound zone approach can easily be adapted to use the Par distortion detectability.

Proposal of Perceptual Sound Zone Framework:
Review of Pressure Matching I

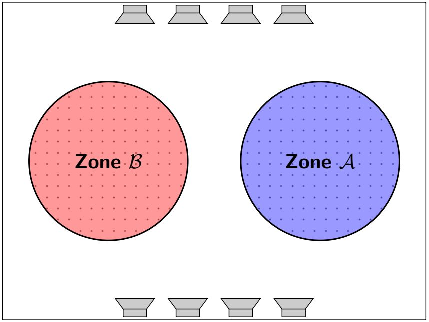
In order to solve the sound zone problem, a pressure matching approach samples the zones into a number of **control points**:





Proposal of Perceptual Sound Zone Framework:
Review of Pressure Matching II

- Per control point m, a **target sound pressure** is assigned.
- The goal is to use the loudspeakers to **attain the target sound pressure** in all points.
- At the same time, we want to have **minimal interference** in all control points.



Proposal of Perceptual Sound Zone Framework:
Review of Pressure Matching III

Mathematically, the optimization problem can be given as follows:

$$\arg\min \quad \sum_{m} \left(\mathrm{RE}^{(m)} + \mathrm{LE}^{(m)} \right) \tag{1}$$

Here, the following quantities are minimized:

- $\mathrm{RE}^{(m)}$ is the **reproduction error** for control point m. It quantifies the deviation from the target sound pressure.
- $\mathrm{LE}^{(m)}$ is the **leakage error** for control point m. It quantifies how much interference is present.

Proposal of Perceptual Sound Zone Framework:
Proposal of Pressure Error Detectability Framework

Rather than optimizing over the reproduction error and the leakage error per control point, the proposed approach is to instead optimize over the reproduction error detectability and the leakage error detectability.

Structure:
Answering Research Questions

1. *"How can auditory perceptual models be included in sound zone algorithms?"*
    1. Determination of a suitable perceptual model for sound zone algorithms.
    2. **Proposal of a perceptual sound zone framework using determined model.**
    3. Proposal of perceptual sound zone algorithms through proposed framework.
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
    1. Simulation and analysis of proposed perceptual sound zone algorithms.

Structure:
Answering Research Questions

1. *"How can auditory perceptual models be included in sound zone algorithms?"*
    1. Determination of a suitable perceptual model for sound zone algorithms.
    2. Proposal of a perceptual sound zone framework using determined model.
    3. **Proposal of perceptual sound zone algorithms through proposed framework.**
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
    1. Simulation and analysis of proposed perceptual sound zone algorithms.

Proposal of Perceptual Sound Zone Algorithms:
Approach

To propose perceptual sound zone algorithms, the previously introduced framework is used:

- $\mathrm{RED}^{(m)}$ is the reproduction error detectability which quantifies how noticeable the deviation from the target sound pressure is for point m.
- $\mathrm{LED}^{(m)}$ is the leakage error detectability which quantifies how noticeable the interference is for point m.

Proposal of Perceptual Sound Zone Algorithms:
Algorithm 1: Unconstrained Perceptual Pressure Matching

The first algorithm minimizes the total error detectability.

$$\arg\min \quad \sum_{m} \left(\mathrm{RED}^{(m)} + \mathrm{LED}^{(m)} \right) \tag{2}$$

Proposal of Perceptual Sound Zone Algorithms:

Algorithm 2: Constrained Perceptual Pressure Matching

The second algorithm minimizes the leakage error detectability, while constraining the reproduction error detectability.

$$\operatorname*{arg\,min} \quad \sum_m \mathrm{LED}^{(m)} \tag{3}$$

$$\text{subject to} \quad \mathrm{RED}^{(m)} \leq D_0 \quad \forall\, m \tag{4}$$

Structure:
Answering Research Questions

1. *"How can auditory perceptual models be included in sound zone algorithms?"*
   1. Determination of a suitable perceptual model for sound zone algorithms.
   2. Proposal of a perceptual sound zone framework using determined model.
   3. **Proposal of perceptual sound zone algorithms through proposed framework.**
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
   1. Simulation and analysis of proposed perceptual sound zone algorithms.

Structure:
Answering Research Questions

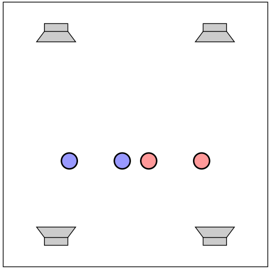
1. *"How can auditory perceptual models be included in sound zone algorithms?"*
    1. Determination of a suitable perceptual model for sound zone algorithms.
    2. Proposal of a perceptual sound zone framework using determined model.
    3. Proposal of perceptual sound zone algorithms through proposed framework.
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
    1. **Simulation and analysis of proposed perceptual sound zone algorithms.**

Evaluation of Perceptual Sound Zone Algorithms:
Approach

In order to determine the benefits of the perceptual sound zone approach, the previously proposed perceptual sound zone algorithms are simulated.

Evaluation of Perceptual Sound Zone Algorithms:
Simulation Setup

- A 5 by 5 meter square room with 4 loudspeakers is used for the evaluation.
- The zones, each consisting of two points, are assigned speech content for the simulations.



Evaluation of Perceptual Sound Zone Algorithms:
Evaluation Measures

In order to effectively compare the reference and the perceptual approach, perceptual measures are used.

This presentation will use the Perceptual Evaluation of Speech Quality (PESQ)[2] and Distraction[3] perceptual measures.

---

[2] A. W. Rix, J. G. Beerends, M. P. Hollier, and A. P. Hekstra (2001). "Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs". In: *2001 IEEE international conference on acoustics, speech, and signal processing. Proceedings (Cat. No. 01CH37221)*. Vol. 2. IEEE, pp. 749–752

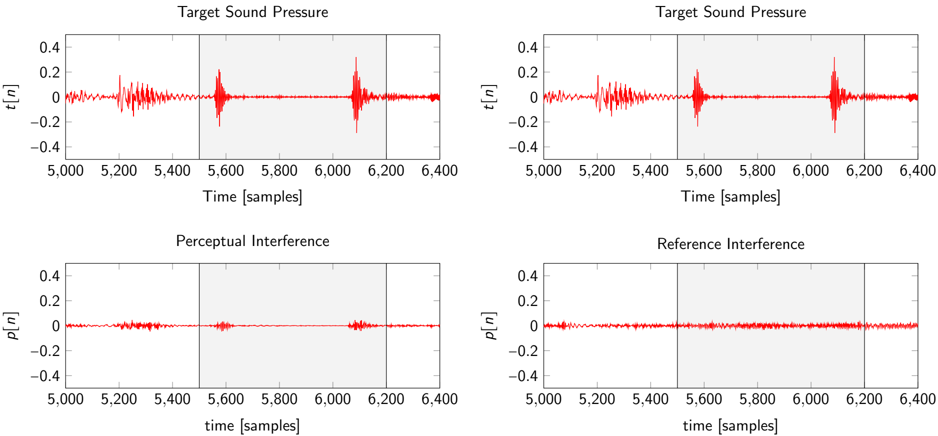
[3] J. Rämö, S. Bech, and S. H. Jensen (2017). "Real-time perceptual model for distraction in interfering audio-on-audio scenarios". In: *IEEE Signal Processing Letters* 24.10, pp. 1448–1452

Evaluation of Perceptual Sound Zone Algorithms:
Evaluation of Unconstrained Perceptual Pressure Matching

Measure	Unconstrained Perceptual PM Mean ($\pm$ 95% CI)	Reference PM Mean ($\pm$ 95% CI)
PESQ (No interference)	3.345 ± 0.087	4.107 ± 0.051
PESQ	3.154 ± 0.081	2.609 ± 0.084
Distraction	7.828 ± 1.868	12.693 ± 3.405

The unconstrained perceptual pressure matching approach outperforms the reference when interference is taken into account. It seems to make a better perceptual trade-off between minimizing the reproduction error and minimizing interference.

Evaluation of Perceptual Sound Zone Algorithms:
Evaluation of Unconstrained Perceptual Pressure Matching



Evaluation of Perceptual Sound Zone Algorithms:

Evaluation of Constrained Perceptual Pressure Matching

Structure:
Answering Research Questions

1. *"How can auditory perceptual models be included in sound zone algorithms?"*
    1. Determination of a suitable perceptual model for sound zone algorithms.
    2. Proposal of a perceptual sound zone framework using determined model.
    3. Proposal of perceptual sound zone algorithms through proposed framework.
2. *"What are benefits of including auditory perceptual models in sound zone algorithms?"*
    1. **Simulation and analysis of proposed perceptual sound zone algorithms.**

Conclusions:

Future Work:

Sources:

Par, S. van de, A. Kohlrausch, R. Heusdens, J. Jensen, and S. H. Jensen (2005). "A perceptual model for sinusoidal audio coding based on spectral integration". In: *EURASIP Journal on Advances in Signal Processing* 2005.9, pp. 1–13.

Rämö, J., S. Bech, and S. H. Jensen (2017). "Real-time perceptual model for distraction in interfering audio-on-audio scenarios". In: *IEEE Signal Processing Letters* 24.10, pp. 1448–1452.

Rix, A. W., J. G. Beerends, M. P. Hollier, and A. P. Hekstra (2001). "Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs". In: *2001 IEEE international conference on acoustics, speech, and signal processing. Proceedings (Cat. No. 01CH37221)*. Vol. 2. IEEE, pp. 749–752.