COLORATION METRICS FOR HEADPHONE EQUALIZATION

Braxton Boren,1 Michelle Geronazzo,2 Fabian Brinkmann,3 Edgar Choueiri1

1) 3D Audio and Applied Acoustics Lab
Princeton University

2) Dept. of Information Engineering
University of Padova

3) Audio Communication Group
TU Berlin

1 Abstract
Headphone equalization is necessary for accurate binaural reproduction over headphones, but so far no metrics have been adopted for evaluating human perception of spectral coloration in post-equalization headphone transfer functions (HpTFs). A metric for peak error is proposed that represents the average HpTF error from narrow peaks per third octave band. In addition, a new metric for broadband error is defined by subtracting the average error from narrow peaks and notches from that of an auditory filter bank model. Used together, the peak error and broadband error terms are shown to represent the critical information necessary for transparent headphone reproduction.

2 Introduction
- The Headphone Transfer Function (HpTF)
- Variations based on
  - Different headphones
  - Different Listeners
  - Different positioning of same headphones!
- High frequency notches move around
- Naïve HpTF inversion makes things worse:
  - Ideal Inversion:
    \[ +6 \text{ dB} \]
    \[ -6 \text{ dB} \]
  - Realistic Inversion:
    \[ +6 \text{ dB} \]
    \[ -6 \text{ dB}, \text{ shifted} \]

3 Minimizing Coloration.
- Many HpTF equalizations have been proposed
  - Need to compare these over a large population
    - PHOnA dataset of HpTFs
  - This requires objective metrics of coloration
  - Existing method: ERB error (above)
  - Mainly captures broadband effects
  - Doesn’t account for ‘ringing’ from bad EQ
  - Peaks are more noticeable than notches

4 Peak Error
- Given an input power spectrum \( H \), we may calculate peak error \( E_{pk} \) as follows:
  - Input Power Spectrum \( H \)
  - Fractional Octave Smoothing
  - (Finely Smoothed) \( H_f \)
  - (Coarsely Smoothed) \( H_c \)
  - Difference Spectrum \( H_{pc} \)

- Taking the gradient of \( H_{pc} \) we can find changes in the sign to detect peaks, applying thresholding to keep those above 1 dB:
  - Mean Peak Error in dB
  - Mean ERB Error in dB

- Given peak locations \( p \), we define peak error as
  \[ E_{pk} = \sum_p H_d(p) \cdot \frac{1}{3[\log_2(f_{high}/f_{low})]} \]
  \( E_{pk} \) gives average peak height per 1/3 octave band
  \( E_{pk} \) increases when \( f_{high} \) increases
  \( E_{pk} \) is not perceptually relevant but is needed to calculate broadband error

5 Broadband Error
- In some cases \( E_{pk} \) and \( E_{ERB} \) are highly correlated, so we define broadband error by subtracting the mean of peak and notch error from ERB error:
  \[ E_{br} = E_{ERB} - \frac{E_{pk} + E_{n}}{2} \]

6 Future Work
- Using these metrics we can now compare peak and broadband performance of many different equalization algorithms over the entire PHOnA dataset
- This allows us to build new equalizations that will minimize error terms as well as variance over many different listeners and headphones

7 References


Contact: bboren@princeton.edu