PERCEPTUAL LOUDNESS COMPENSATION IN INTERACTIVE OBJECT-BASED AUDIO CODING SYSTEMS

Jouni Paulus
Fraunhofer IIS, Erlangen, Germany

MOTIVATION

- Object-based audio coding is gaining attention, e.g., ISO/MPEG-D SAOC and object part of ISO/MPEG-H 3D Audio
- Object-based delivery enables interactive rendering in the decoder
  - E.g., setting position or adjusting output gain of audio objects
- Recommendations in broadcast limit program average loudness for avoiding loudness jumps between programs
  - E.g., EBU R 128 requires program loudness to be -23.0 LUFS
- Interactivity affects the output signal loudness compared to the default mix
- **This paper studies loudness change in a dialogue enhancement application implemented with SAOC-DE**

DIALOGUE ENHANCEMENT

- Allows user to adjust the mixing balance between dialogue and background sound
  - E.g., increasing the dialogue level allows for better intelligibility and reduced listening effort for hearing-impaired and non-native listeners
  - Decreasing the dialogue level for increased feeling of "being there"
- The default mix is a sum of dialogue and background stems

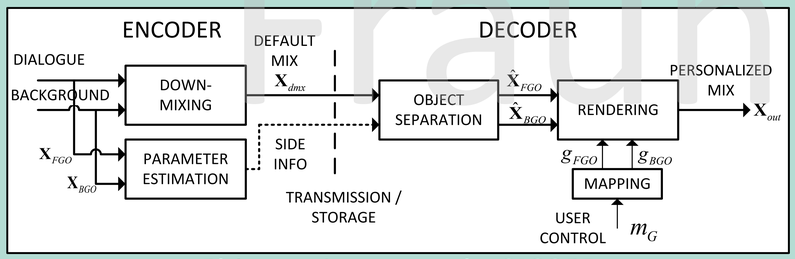
$$\mathbf{X}_{dmx} = \mathbf{X}_{FGO} + \mathbf{X}_{BGO}$$

- The output is a mixture with gains applied on the stems:

$$\mathbf{X}_{out} = g_{FGO}\hat{\mathbf{X}}_{FGO} + g_{BGO}\hat{\mathbf{X}}_{BGO}$$

- In SAOC-DE, a single dialogue modification gain (user control) is mapped into attenuating modification gains with

$$g_{FGO} = \min\left(1, m_G\right), g_{BGO} = \max\left(1, m_G^{-1}\right)$$



Dialogue enhancement using SAOC-DE

LOUDNESS ESTIMATION

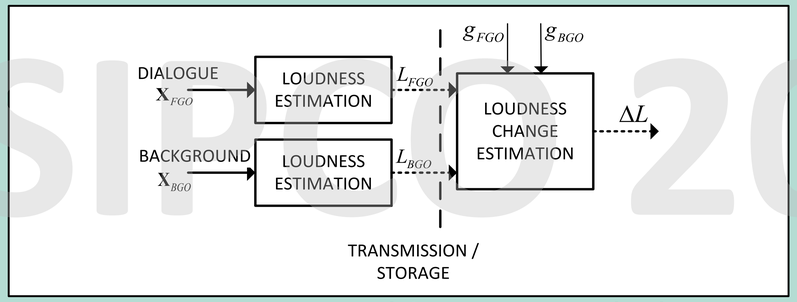
- Method from ITU-R BS.1770-3

$$L = c + 10\log_{10} E$$

where E is K-weighted energy and c is a constant offset

- Assuming independent dialogue and background with known loudness L_{FGO}, L_{BGO} the proposed method for estimating the loudness change from the default mix is

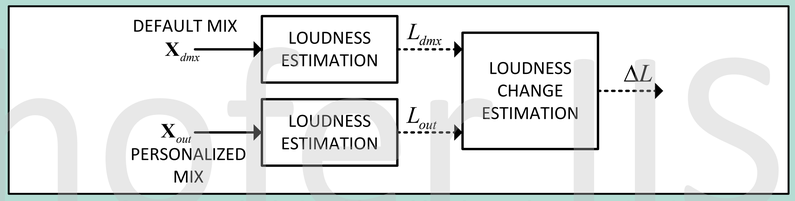
$$\Delta L = 10\log_{10}\frac{g_{FGO}^2 10^{L_{FGO}/10} + g_{BGO}^2 10^{L_{BGO}/10}}{10^{L_{FGO}/10} + 10^{L_{BGO}/10}}$$



Proposed method: parametric loudness change estimation

EVALUTION SETUP

- Material consisting of three stereo items similar to sports broadcast content with approximate length of 10 s
- Using an SAOC-DE system, varying the dialogue modification gain in 1 dB steps in the range -20..+20 dB
- Comparing parametric estimates with signal-based estimation



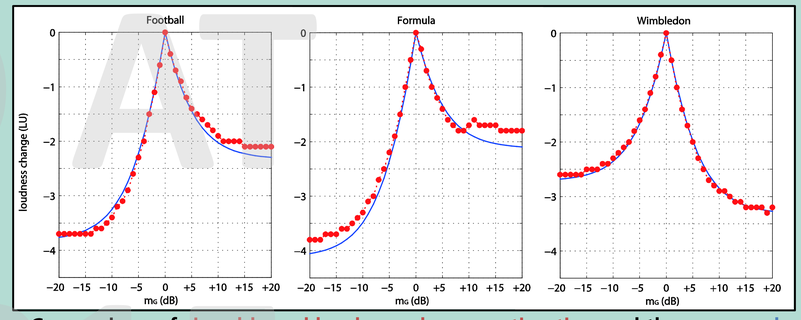
Reference method: signal-based loudness change estimation

SUBJECTIVE ADJUSTMENT TEST

- Test participants presented with the default mix and the dialogue enhancement system output
- Task is to adjust the gain of the modified output to maximize the pleasantness of switching between the two stimuli

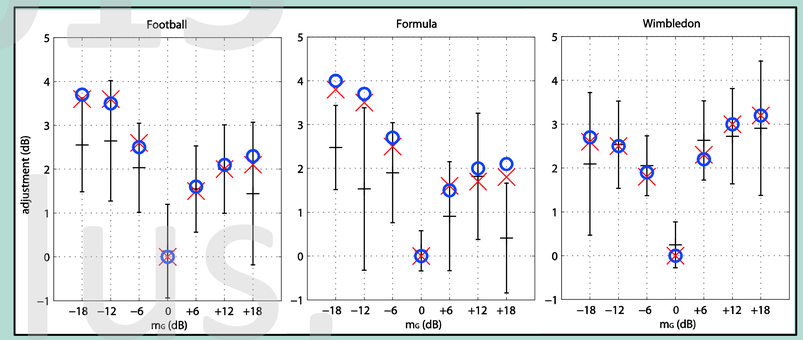
RESULTS

- Proposed vs reference system: mean absolute difference is 0.11 LU, and RMS difference is 0.14 LU



Comparison of signal-based loudness change estimation and the proposed parametric method

- Subjective results from 11 test participants



Mean adjustment, 95% confidence intervals, and values derived from signal-based loudness change estimation and the proposed method

CONCLUSIONS

- Interactivity in object-based audio coding systems may lead into changes in the overall loudness compared to default mix
- Change in the perceptual loudness in a dialogue enhancement application scenario is studied
- A parametric method for estimating the loudness change is proposed, and the results agree with signal-based estimation
- Overall gain adjustment in a DE scenario is studied with a subjective listening test, and the mean adjustments reflect the estimated loudness changes

Fraunhofer IIS