



DAFxTra Evaluation

Graham Coleman, Jordi Bonada,
Xavier Serra

MTG, Universitat Pompeu Fabra

Introduction

- How can audio effects be evaluated?
 - multiple attributes
 - prior listening tests "quality"-based
 - lack of objective criteria
 - needed for evaluation of academic work

Introduction

- Lots of work, focuses on transmission quality
 - telecommunications, coding systems
 - subjective tests (listening)
 - objective tests
 - predictive
 - invasive
 - non-invasive
- PEAQ (Perceptual Evaluation of Audio Quality)
- PESQ (Perceptual Evaluation of Speech Quality)
- MUSHRA (MULTi Stimulus test with Hidden Reference and Anchor) from the International Telecommunications Union (ITU) - [ITU BS.1534](#)
 - "the comparison of high quality reference sounds with several lower quality test sounds"

We propose the DAFxTra initiative (DAFx Transformation Rating)

aims:

- evaluate and compare audio effects algorithms
- define specific and general evaluation strategies
- involve audio effects research community
- bring together users, researchers, and producers
- learn from the process and improve our algorithms

We believe that the DAFx conference provides the perfect context for DAFxTra!

Similar Initiatives

in other research communities:

- MIREX - Music Information Retrieval
- SiSec - Signal Source Separation
- NIPS Feature Selection Challenge - Machine Learning

The Plan

- Propose several effects categories:
 - time-scaling, pitch-shifting, upmixing, voice conversion,
 - ...
- For each category, find people to be responsible for:
 - choosing attributes
 - writing objective tests in MATLAB, octave
 - design overall evaluation procedure
- For each category, find participants
- Tentative Timeline:
 - Dec 2008 - establish evaluation team
 - June 2009 - finalize evaluations for all categories
 - Sept 2009 - run evaluations (DAFx09?)

Plan Details

- In each task, evaluators propose attributes
- For each task we test samples with both objective and subjective tests.
- Statistical analysis to evaluate relevance of objective attributes to perception
- For the objective tests, we can use synthetic signals generated from Matlab code.
- For the subjective tests, we can use sounds from freesound.

Progress

We currently have a wiki page and a email list for discussion and organization:

<http://smcnetwork.org/wiki/Dafxtra2008>

We already have contributions from several researchers, but we need your feedback and ideas!

So far, we have the following people involved in designing the evaluations for effect categories:

- Time-scale modification - Jordi Bonada, Axel Roebel, Yizhar Lavner
- Voice Conversion - Fernando Villavicencio
- Upmixing - Peter Driessen

References

- PEAQ, ICU, [BS.1387-1](#)
- PESQ, ITU-T [P.862](#)
- Rix, A.W., "Perceptual speech quality assessment - a review," *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on* , vol.3, no., pp. iii-1056-9 vol.3, 17-21 May 2004
- MUSHRA, [ITU BS.1534](#)

thank you!

but...

who wants to participate?