The Ambisonics Recordings of Typical Environments (ARTE) Database

Adam Weisser and Jörg M. Buchholz

Overview

The ARTE database is described in detail in Weisser et al. (2019) and was designed primarily to:

1. Provide to the research community accessible multichannel recordings of a range of realistic acoustic everyday scenes that can be used in a large variety of auditory perception tests with improved ecological validity and played back in loudspeaker arrays of different geometries as well as on headphones.
2. Enable standardization and replication of auditory perception tests that utilize realistic noisy environments.
3. Complement the multichannel recordings with measured multichannel Room Impulse Responses (RIRs) as well as basic derived acoustic data.

The ARTE database, so far, contains 13 acoustic environments that were recorded with a purpose-built 62-channel microphone array in various locations around Sydney (Australia), and was decoded into the higher-order Ambisonics (HOA) format.

For each acoustic environment the following files are provided:

HOA environment files: The recorded environments were decoded into 31 HOA channels and saved as WAV-files with a sampling frequency of 44.1 kHz and 32 bits per sample. Thereby, channels 1-25 refer to the 3D HOA components up to the order of $M = 4$, and channels 26-31 refer to additional 2D components (i.e., $m = n$) up to the order of $M = 7$.

HOA RIR files: In each environment, Room Impulse Responses (RIRs) were measured with a Tannoy V8 dual-concentric loudspeaker at a number of positions relative to the microphone array. Currently, only a single RIR is provided in each environment which was measured with a loudspeaker in front of the microphone array (0 degree azimuth) at a distance of 1.3 m. Similar to the noise files, the RIRs are provided as 31-channel WAV-files with a sampling frequency of 44.1 kHz and 32 bits per sample. In addition to the "standard" RIR, a second version is provided in which the RIR was split into a direct sound (DS) component as well as a reverberation component (REV). The separated version of the RIR can be useful for enhancing the directionality (and frequency response) of the direct sound by decoding it into a single loudspeaker channel (i.e., a loudspeaker at an azimuth angle of 0 degrees) and then adding it back to the reverberant component, which is decoded normally. This process has been shown to be particularly useful when evaluating the benefit provided by directional signal enhancement methods (e.g., beamformers) in hearing aids.

Binaural environment files: The HOA noise files were transformed into binaural headphone signals by simulating their playback via a 41-channel loudspeaker array to the in-ear microphones of a calibrated Bruel & Kjaer Head and Torso Simulator (HATS type 4128C). These binaural signals are provided in two versions: (a) an unprocessed version that needs to be presented via headphones that are equalized using an artificial ear and (b) a version that can be directly played back via any diffuse-field equalized headphones.

Binaural RIRs: The HOA RIRs were transformed into binaural RIRs in the same way as the HOA noise files (see above) and were saved both unequalized and diffuse-field equalized.

Basic acoustic measures: A number of basic acoustic measures are provided by a separate PDF-file for each environment, including: (a) unweighted sound pressure levels (dB SPL), (b) A-weighted sound pressure levels (dBA), (c) reverberation time (RT60), (d) third-octave power spectra in dB SPL, (e) temporal envelopes, (f) amplitude modulation spectra, and (g) directional characteristics in the horizontal plane. The acoustic measures were derived by simulating the playback of the MOA noise files (and RIRs) via a 41-channel loudspeaker array to a calibrated omni-directional 1/4" GRAS microphone (Type 46BL).

Apart from the acoustic environment specific files, the ARTE database includes a number of Matlab™ functions that help decoding the provided HOA files into a format that can be played back via a given loudspeaker array, and includes a number of examples.

Further technical details are described in Weisser, et al. (2019).

Supporting material

The provided Matlab™ scripts and examples assume that the downloaded files are organized in the below directory structure. This structure is generated automatically when downloading (and unzipping) the below zip-file. Note that this zip-file contains all required functions except the MOA and binaural sound files and RIRs. Due to their file size (about 10 GBytes in total), these sound files can be downloaded, one by one, from the individual links provided below.

Matlab™ functions with examples (see: ARTE database download.7z)

- ARTE database
  - acousticMeasures — pdf-files for acoustic measures
  - baseFunctions — Matlab functions including examples
  - binaural
    - noiseFiles — Binaural noise files
      - diffusefieldEQ
      - noEQ
    - RIRs — Binaural Room Impulse Responses
      - diffusefieldEQ
      - noEQ
  - decoded — Directory for saving loudspeaker sound files
  - MOA31ch
    - noiseFiles — 31 channel MOA noise files
    - RIRs — 31 channel MOA Room Impulse Responses
      - DS — Direct sound component only
      - REV — Reverberant component only
      - RIR — Entire Room Impulse Response

Some notes on calibration

All HOA noise files were normalized in the same way such that they correctly maintain their original differences in sound pressure level. Hence, once the sensitivity of the loudspeaker playback system is known, the same playback gain must be applied to all noise files. Even though this playback gain can be derived using any of the provided noise files, the easiest noise file for calibrating the loudspeaker playback system is the provided diffuse noise due to its steady-state behavior. Given that most playback environments contain significant low-frequency background noise, and loudspeakers have different low-frequency roll-offs, the provided A-weighted sound pressure levels should be best used for calibration. Also, it is assumed here that all loudspeakers in the playback array have the same distance to the listener, identical sensitivity, and a flat frequency response. If this is not the case the loudspeakers need to be equalized individually. Also, reverberation of the playback room should be as low as possible.

Acknowledgement

The research related to the ARTE database was financially supported by the Oticon foundation as well as the HEARing CRC, established and supported under the Cooperative Research Centres Program – an initiative of the Australian Government.

References

Weisser, A., Buchholz, J. M., Oreinos, C., Badajoz-Davila, J., Galloway, J., Beechey, T., Freeston, K., Keidser, G. (2019). The Ambisonics Recordings of Typical Environments (ARTE) database. Acta Acustica united with Acustica. (see: Weisser_et_al_AAuA.pdf)

Acoustic environments (downloads)

ID	Scene	Brief description	SPL		RT60
			dB SPL	dB A	(sec)
1	Library	University study area in the main library, off-peak hours, quiet	53.0	46.1	0.6
2	Office	Open plan office, people typing, chatting and talking on the phone.	56.7	51.4	0.2
3	Church (1)	Small church space, people entering and chatting quietly before service	60.5	54.7	1.2
4	Living Room	Living room with access to kitchen in the back, loud television and sounds from the kitchen	63.3	58.7	0.2
5	Church (2)	Same as 3, but busier and louder conversations	65.9	60.9	1.2
6	Diffuse noise	Speech-weighted broadband diffuse noise	70.0	65.9	N/A
7	Café (1)	Indoor café at medium occupancy	71.0	67.3	1.1
8	Café (2)	Indoor (company) café at medium occupancy before lunch, recorded next to a wall	71.7	66.2	1.2
9	Dinner Party	Small living room with 8 people chatting over the table with background music	72.8	68.7	0.4
10	Street / Balcony	Apartment balcony on 1st floor over a busy arterial road; Mainly traffic noise with some noise from within the apartment	74.5	71.1	N/A
11	Train Station	Central Station, main concourse – large space, open to the platforms with people walking and talking at peak hour; loud announcements and train sounds	77.1	73.6	1.0*
12	Food Court (1)	Busy University food court during lunch	78.2	74.9	0.9
13	Food Court (2)	Very noisy food court in a shopping mall during lunch	79.6	76.7	1.0