A Database of Anechoic Microphone Array Measurements of Musical Instruments

Recordings, Directivities, and Audio Features

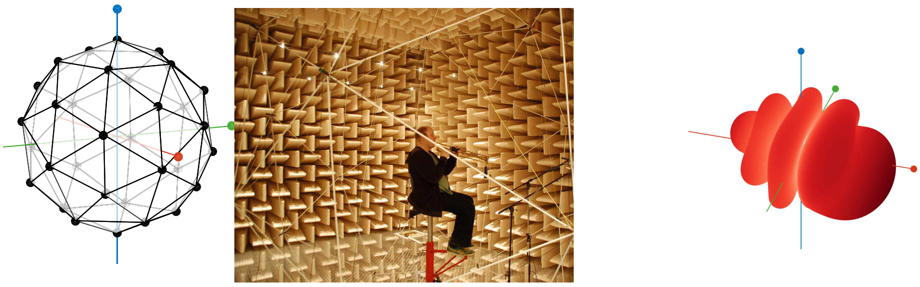
Stefan Weinzierl[1], **Michael Vorländer**[2]
Gottfried Behler[2], Fabian Brinkmann[1], Henrik von Coler[1],
Erik Detzner[1], Johannes Krämer[1], Alexander Lindau[1],
Martin Pollow[2], Frank Schulz[1], Noam R. Shabtai[2].

[1] *TU Berlin, Audio Communication Group*
Einsteinufer 17c, 10587 Berlin-Germany
stefan.weinzierl@tu-berlin.de

[2] *RWTH Aachen University, Institute of Technical Acoustics,*
Kopernikusstrae 5, 52074 Aachen-Germany
mvo@akustik.rwth-aachen.de

April 26, 2017

SEACEN ITA RWTH AACHEN UNIVERSITY audiokommunikation TU berlin



Spherical sampling grid, trumpet inside the microphone array, and corresponding third octave directivity at 4 kHz. Red, green, and blue axis point to positive x, y, and z direction.

General Information

A collection of 3305 single notes of 41 musical instruments of different historical periods was recorded and analyzed. The database includes the instrument recordings, radiation patterns (directivities), and audio features such as the sound power or spectral centroid along with information about the identity and the making of the instrument and its player. The database can be used in virtual reality applications such as room acoustic simulation and auralization, or for the study of musical instruments acoustics themselves. The recordings were made with 32-channel spherical microphone array. For details of the recording method see Table 1 and [1]. If the database is used for further analyses or applications, please cite the authors, title and DOI number of the current electronic publication and

> Noam R. Shabtai, Gottfried Behler, Michael Vorländer, and Stefan Weinzierl: *"Generation and Analysis of an Acoustic Radiation Pattern Database for forty-one Musical Instruments."* Journal of the Acoustic Society of America, vol. 141, no. 4, pp. 1246-1256, 2017.

Informations about the musical instruments and their players can be found in the accompanying document `0_Documentation_Musical_Instruments.pdf`.

wav-ch.	1	2	3	4	5	6	7	8
azimuth	36	36	72	0	72	0	72	36
elevation	37.4	79.2	100.8	100.8	142.6	142.6	63.4	116.6

wav-ch.	9	10	11	12	13	14	15	16
azimuth	108	108	144	144	0	144	108	0
elevation	37.4	79.2	100.8	142.6	0	63.4	116.6	180

wav-ch.	17	18	19	20	21	22	23	24
azimuth	180	252	180	252	216	216	216	180
elevation	37.4	37.4	79.2	79.2	100.8	142.6	63.4	116.6

wav-ch.	25	26	27	28	29	30	31	32
azimuth	324	324	288	288	0	288	252	324
elevation	37.4	79.2	100.8	142.6	63.4	63.4	116.6	116.6

Table 1: Microphone positions in degree corresponding to the channels in the wav-files.

Recordings

Instrument recordings were made in the anechoic chamber of Technical University Berlin using a surrounding spherical array of 32 microphones. During the recordings, musicians were looking at positive x direction (azimuth and elevation approx. 0°) with the main sound emitting part of their instrumented centered inside the array – if possible.

The recordings are located in the folder `1_Recordings`, with separate subfolders for each instrument. Within these folders, the single files are arranged in dedicated directories for the categories pianissimo (*pp*), fortissimo (*ff*), single tones (*et*), scales (*tl*), and special tones (*st*). Each of these directories contains one 32-channel wav-file for each played note, with the channels corresponding to the microphone position in Tab. 1. Wav-files are calibrated – i.e. a value of 1 corresponds to a pressure of 1 Pascal – and compensated for the frequency response of the microphone array. They are accompanied by identically named text files that list the sample indices for the tone onset and offset (#1, and #4), and the beginning and end of the steady part (#2, and #3). All indices were set manually. Markers #2 and #3 are not available for percussive instruments and pizzicato tones.

The single tone files are uniquely named, starting with the instrument name, followed by the tone category, the dynamic level and the pitch, separated by underscores, as shown by the following example:

```
Acoustic_guitar_modern_et_pp_a2.*
```

Scales were recorded in four different versions and labeled with the version number (e.g. `_v1`).

The recordings were conducted by Gottfried Behler, Erik Detzner, Johannes Krämer, Alexander Lindau, Martin Pollow and Frank Schulz.

Directivities

The directivities are represented in the 4th order spherical harmonics (SH) domain, using 25 coefficients. They are available in two dynamic levels (pianissimo and fortissimo) for 31 third-octave frequency bands, and separately for all played tones. Acoustic source centering is applied below 1 kHz, and used to align the acoustic center of the sound source to the physical center of the microphone array. Acoustic centering is applied using the center-of-mass approach [2] at frequencies below 0.5 kHz, and using the phase symmetry approach above 0.5 kHz [3]. In case of the single tone directivities, the uncentered data is also provided. The data is located in the folder `2_Directivities` with the subfolders `SingleTones`, and `ThirdOctaves`.

Directivity functions model the complex pressure in a spherical coordinate system following the coordinate and SH convention of Rafaely [4] (azimuth angle ϕ increasing counter clockwise from positive x in the xy-plain, and an elevation angle θ increasing from positive z to the xy-plain where ϕ is located), using the order-limited spherical Fourier series in the form of

$$p\left(k, r, \theta, \phi\right) = \sum_{n=0}^{N} \sum_{m=-n}^{n} p_{nm}\left(k, r\right) Y_n^m\left(\theta, \phi\right), \tag{1}$$

where $p_{nm}\left(k\right)$ are the radiation pattern spherical harmonics coefficients at wave number k and distance r. Every base function $Y_n^m\left(\cdot, \cdot\right)$ is referred to as the spherical harmonic of order n and degree m, given by

$$Y_n^m\left(\theta, \phi\right) \triangleq \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m\left(\cos\theta\right) e^{im\phi}, \tag{2}$$

where $P_n^m\left(\cdot\right)$ is the associated Legendre function [5] of order n and degree m. In practice, this database contains values for the order of $N = 4$ (25 coefficients) and $r = 2.1$ m, the radius at which the radiation pattern was measured. More detailed information can be found in Shabtai *et al.* (2017) [1].

The SH coefficients can for example be transformed to complex spectra and plotted using AKtools [6]:

```
sg      = AKgreatCircleGrid;
sg(:,2) = 90-sg(:,2);
```

generates a spherical sampling grid

```
h = AKisht(radiation.pnm, false, sg, 'complex');
```

computes the complex spectrum `h`, from the SH coefficients `radiation.pnm`, and the spatial sampling grid.

Plotting the log. magnitude spectrum can be done by calling:
`AKp(db(h(20,:)), 'x2', 'az', sg(:,1), 'el', sg(:,2), 'coord', 2)`,
which plots frequency bin number 20.

Acoustic source centering, and the generation of the directivities has been performed by Noam R. Shabtai.

Third-octave directivities

The MATLAB `*.mat` filenames have the following structure:

`<Instrument name>_<modern/historical>_et_<pp/ff>.mat` where

- `<Instrument name>` represents each one among the 41 musical instruments,
- `<modern/historical>` stands for modern musical instrument or an authentic one according to a historical manner of construction, and
- `<pp/ff>` stands for *pianissimo* or *fortissimo* dynamics.

Each `*.mat` file contains a variable `radiation` with the foloowing the fields:

- `radiation.bands.center_frequencies` contains the center frequencies of the 31 third-octave frequency bands,
- `radiation.bands.frequencies` contains the frequency limits of each third-octave frequency band, and
- `radiation.pnm` is a 25×31 matrix of the 25 spherical harmonics coefficients $\{p_{0,0}, p_{1,-1}, p_{1,0}, \ldots, p_{4,4}\}$ of the pressure function on a surface of a sphere at a radius of 2.1 m, arranged as a column vector at each third-octave frequency band.

Single tone directivities

The MATLAB `*.mat` filenames are identical to the third-octave name convention. Each `*.mat` file contains a variable `radiation` with the following fields:

- `radiation.pnm` is a $25 \times 10 \times M$ matrix of the 25 spherical harmonics coefficients $\{p_{0,0}, p_{1,-1}, p_{1,0}, \ldots, p_{4,4}\}$, for 1 fundamental frequency and 9 overtones/harmonics given for each of the M played note.
- `radiation.frequencies` is a $M \times 10$ matrix that contains the frequencies of the fundamental tone and the 9 overtones/harmonics for each of the M played notes.
- `radiation.midiNotes` contains integer midi note numbers corresponding to the fundamental frequency of each played note and the standard pitches as provided in the accompanying instruments table. An integer note of 69 corresponds to the standard pitch A.
- `radiation.noteNames` contains strings specifying the midi notes, where “A4” denotes the standard pitch A.

Audio Features

The sound power was calculated according to the enveloping surface method [7] according to which the sound pressure p is averaged for each microphone within the steady sound boundaries

$$L_p = 10 \log_{10} \left(\frac{\frac{1}{N} \sum_n p[n]^2}{p_0^2} \right) \text{ [dB]}, \tag{3}$$

with $p_0 = 2 \cdot 10^{-5}$ [Pa], averaged across microphones

$$\bar{L_p} = 10 \log_{10} \left(\frac{1}{M} \sum_m 10^{0.1 L_{p,m}} \right) \text{ [dB]}, \tag{4}$$

and referenced to a surface area of 1 m^2

$$L_w = \bar{L_p} + 10\log_{10}\left(\frac{S_1}{S_0}\right)\ [\mathrm{dB}], \tag{5}$$

with spherical surface areas $S_1 = 54.63\ \mathrm{m}^2$, and $S_0 = 1\ \mathrm{m}^2$. In case of transient sounds, sound pressures $p[n]^2$ in eq. (3) were subjected to time-weighted filtering (fast) [8] prior to averaging.

Timbre-describing audio features were calculated for for a subset of the audio files. Scales and special tones were excluded from this calculation.

Prior to the audio feature extraction, a main microphone was selected for each instrument, based on the highest RMS over all notes. The audio features were then calculated using the TimbreToolbox (TTB) [9]. The toolbox calculates various features describing the spectral distribution, respectively the harmonic content, as well as the temporal envelope.

The spectral distribution descriptors are first calculated as time-varying features, resulting from a frame-wise analysis of the audio data. Subsequently, the median and interquartile range are obtained from the trajectories as single values. Features related to the temporal envelope, are represented by single values for each recording. All features are stored in a Matlab cell array in the folder `3_Features`, together with the information for each file.

References

[1] N. R. Shabtai, G. Behler, M. Vorländer, and S. Weinzierl, "Generation and analysis of an acoustic radiation pattern database for forty-one musical instruments," *J. Acoust. Soc. Am.*, vol. 141, no. 2, pp. 1246–1256, 2017.

[2] I. Ben Hagai, M. Pollow, M. Vorländer, and B. Rafaely, "Acoustic centering of source measured by surrounding spherical microphone arrays," *J. Acoust. Soc. Am.*, vol. 130, no. 4, pp. 2003–2015, 2011.

[3] N. R. Shabtai and M. Vorländer, "Acoustic centering of sources with high-order radiation patterns," *J. Acoust. Soc. Am.*, vol. 137, no. 4, pp. 1947–1961, Apr. 2015.

[4] B. Rafaely, *Fundamentals of spherical array processing*, 1st ed. Berlin, Heidelberg, Germany: Springer, 2015.

[5] E. G. Williams, *Fourier Acoustics. Sound radiation and nearfield acoustical holography*, 1st ed. Academic Press, 1999.

[6] F. Brinkmann and S. Weinzierl. AKtools - an open toolbox for acoustic signal acquisition, processing, and inspection.

[7] DIN EN ISO 3745, *Determination of sound power levels and sound energy levels of noise sources using sound pressure.* Berlin, Germany: Beuth, Jul. 2012.

[8] DIN EN 61672-1, *Sound level meters – Part 1: Specifications.* Berlin, Germany: Beuth, Jul. 2014.

[9] G. Peeters, B. L. Giordano, P. Susini, N. Misdariis, and S. McAdams, "The timbre toolbox: Extracting audio descriptors from musical signals," *J. Acoust. Soc. Am.*, vol. 130, no. 5, pp. 2902–2916, Nov. 2011.