A Ground Truth on Room Acoustical Analysis and Perception (GRAP)

—

David Ackermann, Maximilian Ilse, Dmitry Grigoriev, Steffen Lepa, Stefan Weinzierl
Audio Communication Group, TU Berlin
Einsteinufer 17c, 10587 Berlin, Germany

Sönke Pelzer, Michael Vorländer
Institute of Technical Acoustics, RWTH Aachen University
Kopernikusstraße 5, 52074 Aachen, Germany

david.ackermann@tu-berlin.de
stefan.weinzierl@tu-berlin.de

September 10, 2020
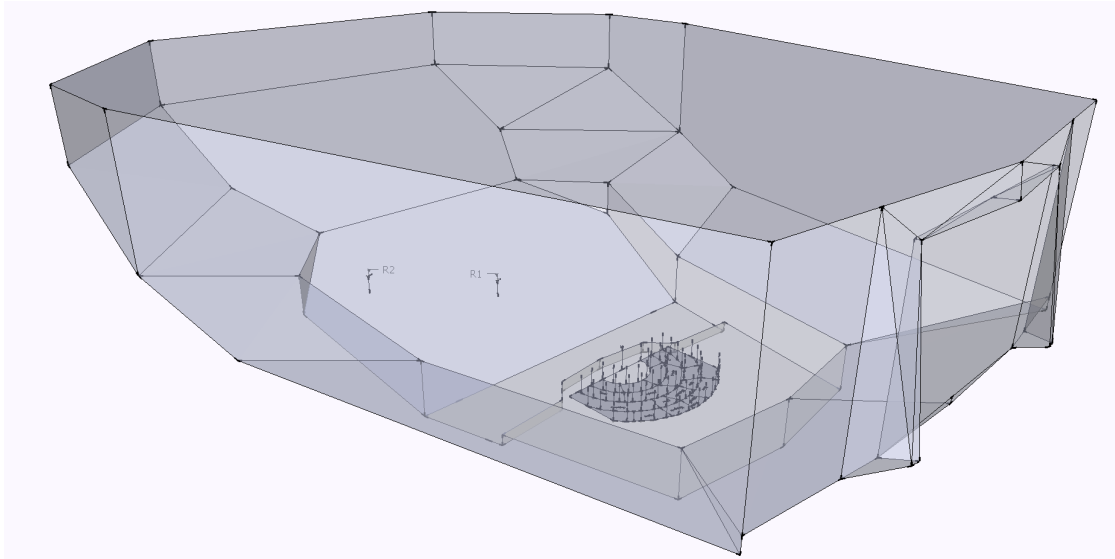
General Information

The following document describes the *Ground Truth on Room Acoustical Analysis and Perception* (GRAP) data base. It is provided under a Creative Commons BY-NC-SA licence, giving you the freedom to redistribute and edit the database for non-commercial purposes.

For more information visit http://creativecommons.org/licenses/by-nc-sa/4.0/ and contact the authors for commercial use.

The documentation is focused on the structure and content of the data base. More information on the acquisition, processing and validation can be found in the accompanying Journal of the Acoustical Society of America publication. If you use this data base please cite the authors, title and DOI number of the current electronic publication and

Weinzierl, S. and Lepa, S. and Ackermann, D. (2018): "A measuring instrument for the auditory perception of rooms. The Room Acoustical Quality Inventory (RAQI)." *Journal of the Acoustical Society of America*, **144**(3), 1245–1257. DOI: doi.org/10.1121/1.5051453

Database description

A database of 35 virtual room models was created that can serve as a ground truth for the future development of room acoustical parameters beyond ISO 3382-1 [1]. Some of the room models are based on existing performance venues, however without yielding a perfect match of measured and simulated acoustical parameters. Others are artificial designs which were made to systematically cover a wide variety of room acoustical properties.

Each of the 35 acoustical environments included in the GRAP database consists of three components: (1) the room model which specifies source and receiver positions and the acoustic properties of the surfaces, (2) the simulated monaural and binaural impulse responses, as well as (3) the item and factor scores, based on a listening test using the Room Acoustic Quality Inventory (RAQI). The data structure is outlined in Figure 1.

The CAD Software SketchUp was used to create 3D room models. Acoustical simulations were executed using the room simulation software RAVEN [2], developed at the Institute of Technical Acoustics, RWTH Aachen.

All simulations were done in third octave resolution, with the hybrid simulation algorithm using image sources up to third order and ray tracing using 200000 rays.

Figure 1: Data structure of the GRAP database.

Every room model features one source and two receiver positions. For each receiver position the database contains a monaural room impulse response, a set of binaural room impulse responses for dynamic binaural synthesis, the plane wave list and energy histogram, and room acoustical parameters according to ISO 3822-1.
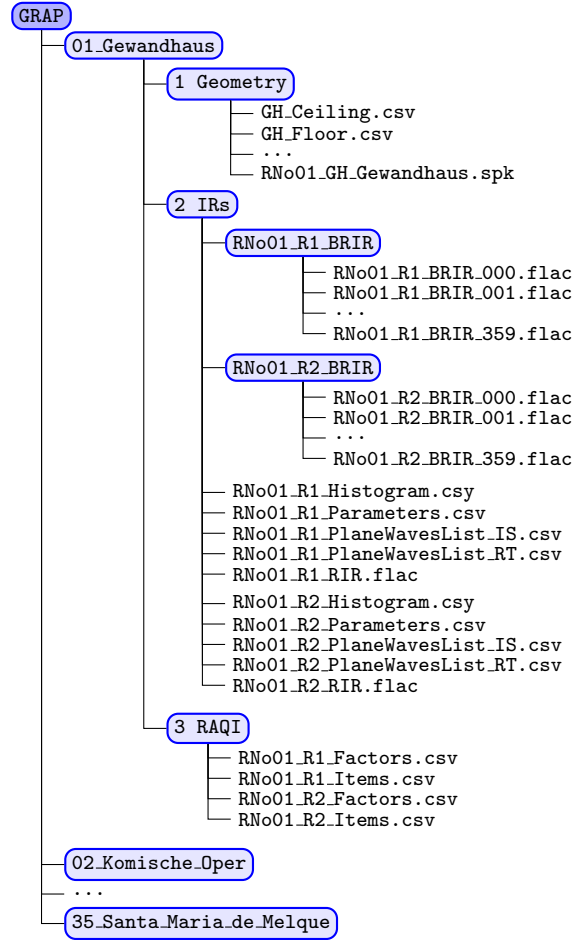
1 Scene geometry

The scene geometries and the surface properties can be found in the folder `1 Geometry` (Figure 2) and are provided as SketchUp Make models (free for educational purposes) and comma-separated values files. The models are named according to the scheme `roomNo_acronym_vanueName.spk`, the file `RNo01_GH_Gewandhaus.spk` defines the geometry of the *Neue Gewandhaus in Leipzig*.
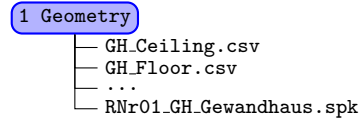

Figure 2: Data structure of the Geometry folder.

Each SketchUp files contains a 3D model of the room, the positions and orientation of the source and the two receivers (Figure 4). The colors assigned to the surfaces of the 3D models specify their material. For example, the color named `GH_Ceiling` links to a material whose surface properties can be found in the `GH_Ceiling.csv` file (Figure 3) in the same folder. Absorption (column 2) and scattering (column 3) coefficients are given in third octave bands for a frequency range from 20 Hz to 20 kHz (column 1).

```
 1   f in Hz,    absorption,   scattering,
 2   20,         0.2769,       0.01
 3   25,         0.2833,       0.01
 .
 .
 .
31   16000,      0.5566,       0.9
32   20000,      0.5954,       0.9
```

Figure 3: `GH_Ceiling.csv` data format.

To view the color of a surface in SketchUp, use the *Sample Point* option of the *Paint Bucket Tool*. If the object belongs to a group or a component, it is necessary to first go in the *edit mode* of the group by double clicking the object before being able to show the material. The degree of detail in the scene geometry has been adapted to common standards [3, p. 176].
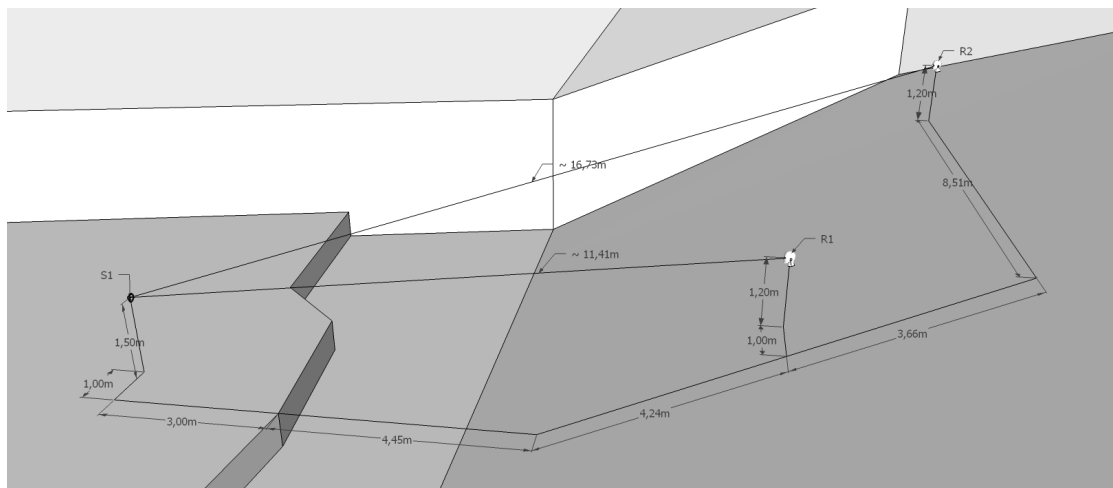

Figure 4: The positions and orientation of source (S1) and receivers (R1, R2) in the RNo01_GH_Gewandhaus.spk *SketchUp Make* model.

2 IRs

For each receiver the simulated monaural (`RIR.flac`) and binaural room impulse responses (`BRIR_xxx.flac`), along with energy histogram (`Histogram.csv`) and plane wave lists, separated for image source method (`PlaneWaveList_IS.csv`) and ray tracing (`PlaneWaveList_RT.csv`), as well as room acoustical parameters according to DIN ISO 3822-1 (`Parameters.csv`) are available in the folder `2 IRs` (Figure 5).

All files are named according to the scheme `roomNo_receiverNo_content` e.g. `RNo01_R1_BRIR_090.flac` defines the binaural impulse response (`BRIR`) for a head orientation of ninety degree *to the left* of receiver 1 (`R1`) in the performance venue *Neues Gewandhaus Leipzig* (`RNo01`). Whereas `RNo01_R2_BRIR_270.flac` defines the binaural impulse response for a head orientation of ninety degree *to the right* of receiver 2 (`R2`) in the same performance venue. An omni-directional characteristic was used for the source. Both receivers are frontally aligned to the sound source and the FABIAN head-related transfer functions [4] were used for simulation. The impulse response are provided in the *Free Lossless Audio Codec* (`.flac`) at a sample rate of 44.1 kHz. For each receiver the energy histograms (such as `RNo01_R1_Histogram.csv`) are stored as comma-separated values (Figure 6).
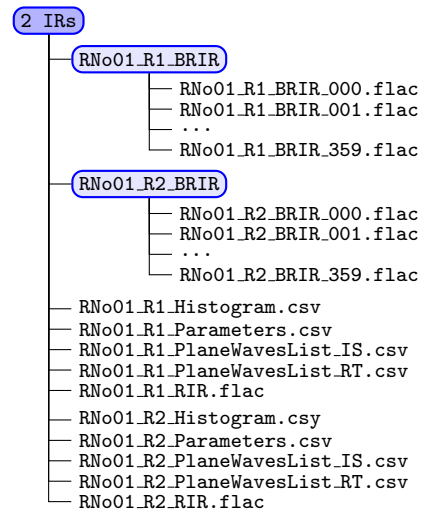

Figure 5: Data structure of the IRs folder.

```
    1   Direct sound delay in s, 0.0331,
    2
    3   f in Hz,           20,              25,              31.5,            40, ...
    4   energyLF,          0.00510,         0.00514,         0.00516,         0.00518, ...
    5   energyLFC,         0.00738,         0.00744,         0.00747,         0.00749, ...
    6
    7   Time in s,
    8   0.000000000000,    0.000000000000,  0.000000000000,  0.000000000000,  0.000000000000, ...
    9   0.001000000048,    0.000000000000,  0.000000000000,  0.000000000000,  0.000000000000, ...
    .
    .
 4658   4.650000095370,    0.000000000000,  0.000000000000,  0.000000000000,  0.000000000000, ...
 4659   4.651000022890,    0.000000000000,  0.000000000000,  0.000000000000,  0.000000000000, ...
```

Figure 6: `RNo01_R1_Histogram.csv` data format.

Line 1 indicates the direct sound delay in seconds between sound source and receiver. Line 4 shows the summed *energy* within the first 5 to 80 ms for a simulation with a figure-of-eight receiver characteristic, the null of the figure-of-eight pattern pointing towards the source position. Line 5 indicates also the energy arriving from lateral directions within the first 5 to 80 ms, but in this case weighted with the cosine of the angle [1, A.15]. The values can be used to calculate the fraction of energy arriving from lateral directions J_{LF} and J_{LFC} according to ISO 3822-1.

From line 8 down to the bottom the summed *energy* over the entire omni-directional detection sphere is provided for image source method and ray tracing in 1 ms time windows (column 1) in third octave band values (column 2 to 32) for the frequency range

from 20 Hz to 20 kHz (line 3, column 2 to 32). Diffuse showering [5] is taken into consideration in the energy histogram.

The separated plane waves lists, e.g., RNo01_R1_PlaneWavesList_IS.csv for image source method and RNo01_R1_PlaneWavesList_RT.csv for ray tracing, are shown in Figure 7 and Figure 8. From line 2 down to the bottom each line describes a single plane wave. Column 1 provides its time of arrival in seconds, column 2 the azimuth angle in [°] (front = 0°, left = 90° and right = 270°), column 3 the elevation angle in [°] (front = 0°, top = 90° and bottom = -90°) and the *linear absolute sound pressure* in third octave band values (column 4 to 34) for the frequency range from 20 Hz to 20 kHz (line 1, column 2 to 32). Pleas note that the diffuse showering method [5] is *not* included in both plane waves lists.

```
 1   Time in s,      Azimuth in deg,    Elevation in deg,  20,               25, ...
 2   0.0331486000,   359.979000000000,  0.011905100000,    0.087697200000,   0.087696400000, ...
 3   0.0355205000,   0.021224100000,    -44.0196000000,    0.075945000000,   0.075813100000, ...
 .
 .
 .
59   0.2301350000,   58.206400000000,   25.507800000000,   0.008017350000,   0.008045520000, ...
60   0.2307210000,   303.010000000000,  25.577400000000,   0.007996980000,   0.008025070000, ...
```

Figure 7: RNo01_R1_PlaneWavesList_IS.csv data format.

```
    1   Time in s,      Azimut in deg,     Elevation in deg,  20,               25, ...
    2   0.0426304000,   5.695060000000,    -6.2177500000,     0.001544710000,   0.001586170000, ...
    3   0.0562812000,   2.028260000000,    -0.5178170000,     0.000438636000,   0.000404255000, ...
    .
    .
    .
80321   4.6499800000,   90.000000000000,   90.000000000000,   0.000000000000,   0.000000000000, ...
80322   4.6500000000,   90.000000000000,   90.000000000000,   0.000000000000,   0.000000000000, ...
```

Figure 8: RNo01_R1_PlaneWavesList_RT.csv data format.

Figure 9 shows the comma-separated values file (RNo01_R1_Parameters.csv) with room acoustical parameters according to ISO 3822-1 (column 2 to 11). The parameters were calculated on the basis of the energy histogram and are provided in third octave band values (line 2 to 32) for the frequency range from 20 Hz to 20 kHz (column 1, line 2 to 32).

EDT (Early decay time), T_30 (Reverberation time) and Ts (Center time) values are given in seconds. G (Sound strength) and C_80 (Clarity) values are given in decibels. D_50 (Definition), LF and LFC (Early lateral energy) and IACC (Inter-aural cross correlation coefficient) values are without units.

```
 1   f in Hz,   EDT,       T_30,      C_80,      D_50,      G,         Ts,         LF,...
 2   20,        2.3551,    2.7103,    1.8272,    0.47384,   7.8536,    0.12253,    0.13869,...
 3   25,        2.3658,    2.7046,    1.9136,    0.47828,   7.8174,    0.12112,    0.13932,...
 .
 .
 .
31   16000,     0.38937,   0.39037,   17.252,    0.95289,   -2.3426,   0.0083259,  0.076545,...
32   20000,     0.22689,   0.29817,   22.909,    0.98098,   -4.4018,   0.0053348,  0.064826,...
```

Figure 9: RNo01_R1_Parameters.csv data format.

3 RAQI

The scores of 46 items of the *Room Acoustic Quality Inventory* (RAQI) and the scores for the 6-factor solution can be found for each receiver position separately in the folder 3 RAQI (Figure 10).
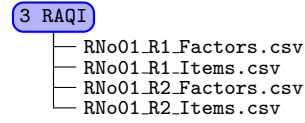

Figure 10: Data structure of the RAQI folder.

Figure 11 shows the comma-separated values RNo01_R1_Items.csv file with the item scores for the 46 attributes (column 1, from line 2 down to the bottom) for the three different content types speech (column 2), trumpet (column 3) and orchestra (column 4). Items not used for a particular content type are represented as NaN. Column 5 indicates the averaged item scores over all three content types.

```
1   Items,                    speech,   trumpet,   orchestra,     all,
2   Treble range characteristic,  -7.89,    16.67,       -6.08,    0.90,
3   Mid range characteristic,      2.63,    -2.80,        6.60,    2.14,
.
.
46  Blend,                         NaN,      NaN,        75.54,   75.54,
47  Richness of sound,           53.38,    66.71,        59.43,   59.84,
```

Figure 11: RNo01_R1_Items.csv data format.

The factors are stored in comma-separated value fiels e.g. RNo01_R1_Factors.csv (Figure 12), from line 2 down to the bottom the 6 factors are indicated in column 1, with the factor loadings for the three different content types speech (column 2), trumpet (column 3) and orchestra (column 4). Column 5 indicates the averaged factor load over all three content types.

```
1   Factors,        speech,   trumpet,   orchestra,     all,
2   Quality,         -0.46,     0.45,        0.01,   -0.00,
3   Strength,         0.07,     0.13,        0.45,    0.22,
4   Reverberation,    0.69,     0.08,        0.11,    0.29,
5   Brilliance,      -0.53,     0.43,        0.02,   -0.03,
6   IrregularDecay,   0.43,    -0.20,        0.04,    0.06,
7   Coloration,       0.15,    -0.41,        0.19,   -0.02,
```

Figure 12: RNo01_R1_Factors.csv data format.

References

[1] ISO 3382-1:2009 Acoustics - Measurement of room acoustic parameters - Part 1: Performance spaces, 2009.

[2] D. Schröder, M. Vorländer: RAVEN: A Real-Time Framework for the Auralization of Interactive Virtual Environments. Proc. of Forum Acusticum, Aalborg, Denmark, 2011, p. 1541–1546.

[3] M. Vorländer: Auralization. Fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality. Springer, Berlin, 2008.

[4] F. Brinkmann, A. Lindau, S. Weinzierl, G. Geissler, S. van de Par: A High Resolution and Full-Spherical Head-Related Transfer Function Database for Different Head-Above-Torso Orientations. J. Acoust. Soc. Am. **65(10)** (2017), p. 841–848.

[5] D. Schröder: Physically based real-time auralization of interactive virtual environments. Logos Verlag, Berlin, 2011.