High Resolution Sound Field Analysis with Directional Room Impulse Responses

Stefan Ziesemer, Eric Kurz

Signal Processing and Speech Communication Laboratory (SPSC), Graz University of Technology, 8010 Graz, Austria

stefan.ziesemer@student.tugraz.at, eric.kurz@tugraz.at

Introduction

Until today, room acoustic measurements are mostly performed with a few, sparsely distributed, omnidirectional microphones. Furthermore, these measurement techniques are based on the assumption of a diffuse and homogeneous sound field. As shown by e.g. De Vries [1], this assumption often does not correspond with reality. Moreover, since the diffuseness properties of the sound field, namely isotropy and incoherence, are directional-dependent, conventional measurement techniques can not obtain them sufficiently even though they have influence on the measurement results. The recent discussion about the revision of ISO-354 [2] and works by Del Galdo et al. [3] and others [4, 5] show that this is an highly topical subject.

In order to examine homogeneity and diffuseness more precisely, more sophisticated, i.e. spatially dense and directional measurement methods are necessary. Fortunately, measures already exist for estimating isotropy and incoherence using directional room impulse responses (RIRs) obtained with such measurements. However, since the purpose of these measures is either mainly targeting sound field reproduction [6, 7, 8] or is more of a theoretical consideration [3], they have not made it to room acoustic measurement techniques yet.

In the first part of this paper, an attempt is made to outline the theoretical concept of homogeneity and diffuseness more clearly. Next, the covariance matrix eigenvalue diffuseness estimate (COMEDIE) [9] as an advanced representative of diffuseness measures is introduced. In addition to a measurement series, which is presented in the third part, a simulation of the same room was also carried out. The comparison between measurement and simulation is the last part of the paper.

Theoretical Background

The term "diffuseness" is probably one of the most used terms in room acoustics. However, there is no uniform definition to it. Depending on literature and context, "diffuseness" is sometimes used synonymously with "homogeneity". In this paper a distinction is made, where homogeneity concerns the space as a whole, whereas diffuseness considers a single point in space.

Homogeneity

$$\langle p^2(\boldsymbol{x})\rangle = const. \; \forall \, \boldsymbol{x} \tag{1}$$

A scalar field is homogeneous if its measurable quantities (e.g. mean squared sound pressure $\langle p^2(\boldsymbol{x})\rangle$, eq. (1)) do not depend on location, i.e., they are invariant against a shift of the spatial coordinates $\boldsymbol{x}$. Therefore, the energy of the sound field is distributed evenly over space.

Diffuseness

1. Isotropy

$$\langle \boldsymbol{v}'(\boldsymbol{x})\rangle = \langle \boldsymbol{R}\boldsymbol{v}(\boldsymbol{x})\rangle = \langle \boldsymbol{v}(\boldsymbol{x})\rangle = const. \; \forall \, \boldsymbol{R} \tag{2}$$

A vector field is isotropic if its measurable quantities (e.g. mean sound particle velocity $\langle \boldsymbol{v}(\boldsymbol{x})\rangle$, eq. (2)) do not depend on direction, i.e., they are invariant against a rotation of the directional components $\langle \boldsymbol{v}(\boldsymbol{x})\rangle$ by an arbitrary rotation matrix $\boldsymbol{R}$ around the constant reference point $\boldsymbol{x}$. Therefore, in a plane wave model, the incident plane waves have equal power and an evenly distributed direction of arrival for a single point in space.

2. Incoherence

A field is coherent if, for a given frequency f, the phase shift among its (directional) wave components do not depend on location or time. It is (totally) incoherent if these phase shifts have a probability distribution that is invariant against a cyclic shift. When looking at two different, arbitrary plane waves with pressure signal $p_n(f)$ and $p_m(f)$, their cross power spectral density $G_{p,nm}(f)$ (eq. (3)) is always zero:

$$G_{p,nm}(f) = 0 \; \forall \, n, m, n \neq m \tag{3}$$

Therefore, in addition to the evenly powerful plane waves arriving in terms of isotropy, the incident plane waves need to have random phase for diffuseness in a single point in space.

Diffuseness estimator d

In order to examine the diffuseness of the sound field at discrete receiver positions the diffuseness estimator d by Epain and Jin [9] is the measure of choice. It quantifies diffuseness by analysing the correlations between microphone signals of a spaced microphone array. Since those microphone signals are translated into the spherical harmonics (SH) domain, it takes also the spatial distribution of the incident sound into account. Therefore it analyses isotropy and incoherence at the same time. This is desirable as there are situations, where isotropy does not imply incoherence. If, hypothetically, six plane waves arrive at the receiver position with the same energy and the same phase, each from the positive and negative spatial direction, the criterion for isotropy is fulfilled, but that for incoherence is not. Accordingly, diffuseness is no longer met.

Like other diffuseness estimates (e.g. ϕ [10], ψ [6, 8]), it is in the range of $d(t) \in [0, 1]$. If $d(t) = 0$, either the incident plane waves are all same phase arriving from evenly distributed directions equally powerful or there

is just one single plane wave arriving. If $d(t) = 1$, the sound field at the receiver position is totally isotropic and incoherent, therefore theoretically maximum diffuse. $d(t)$ can be easily determined from any SH signal of an arbitrary order L:

$$d(t) = 1 - \frac{\gamma(t)}{\gamma_0} \tag{4}$$

Where

- γ_0 ... most non-diffuse case: $\gamma_0 = 2[(L+1)^2 - 1]$
- $\gamma(t)$ ... mean of the deviation of the eigenvalues of the SH covariance matrix from their mean:

$$\gamma(t) = \frac{1}{\overline{\lambda}(t)} \sum_{i=1}^{(L+1)^2} |\lambda_i(t) - \overline{\lambda}(t)|$$

  - $\lambda_i(t)$ ... eigenvalues of the covariance matrix of the SH signals
  - $\overline{\lambda}(t)$ ... mean of the eigenvalues

Case Study: Lecture Hall at TU Graz

To see this measure in a practical, real world application at work, a case study at the lecture hall HSi2 at Graz University of Technology was carried out. It hast a fixed seating capacity for up to 91 people, a footprint of 181 m^2 and 577 m^3 in volume. Its reverberation time is $T_{20} = 0.93$ s, resulting in a Schroeder-frequency of $f_{\text{Schroeder}} \approx 80$ Hz. The lecture hall was chosen because its dimensions and most of the acoustic surface properties are well documented from previous works [11, 12].

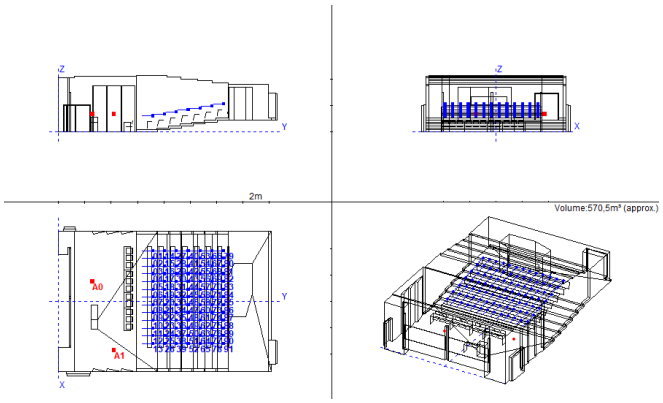


Figure 1: Floor plans of the examined lecture hall from the simulation model with source (red) and receiver (blue) positions.

At first, a simulation model using the software CATT Acoustic [13] was created (fig. 1). Next, the model was calibrated with a spatially sparse, conventional reverberation measurement with two source and 15 receiver positions. Then, a high resolution simulation was performed with two source positions and one receiver position at each seat in the audience area. For every source-receiver combination a four channel B-format RIR was exported, giving a total of $N_{\text{source}} \cdot N_{\text{receiver}} \cdot N_{\text{channel}} = 2 \cdot 91 \cdot 4 = 728$ impulse responses. In MATLAB, the evaluation of the RIRs was implemented using parts of [14]. Next, the

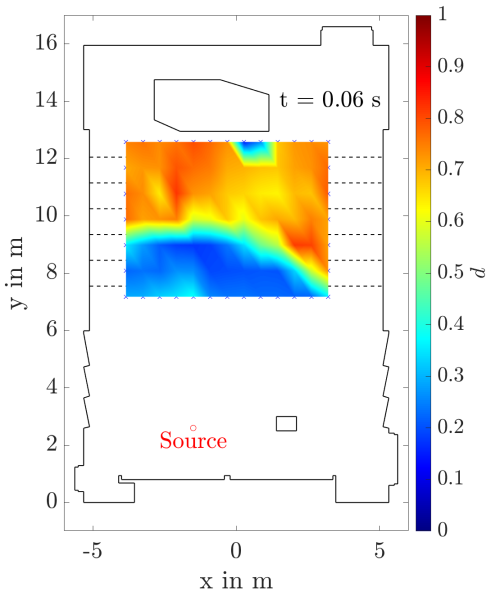


Figure 2: Diffuseness estimate d at time $t = 0.06$ s over the floor plan of the lecture hall.

high resolution measurement was performed with a tetrahedral four channel microphone and an omnidirectional measurement speaker at the same source and receiver positions as stated previously in the simulation model. This results in a spatial resolution of $\Delta x = 0.59$ m and $\Delta y = 0.9$ m, respectively, at a height of $z = 1.2 \ldots 2.1$ m following the steps of the auditorium (fig. 1, upper left). For excitation, an exponential sine sweep with a FFT degree of $N_{\text{FFT}} = 19$ was generated using ITA Toolbox [15, 16], resulting in a sweep duration of approx. 10 s. The measurement was repeated three times for each source-receiver combination.

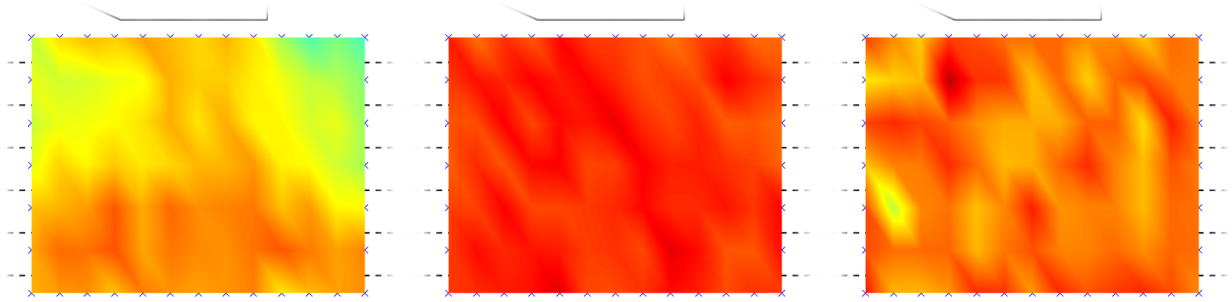
Measurement Results

Figure 2 shows d over the floor plan of the lecture hall. It was calculated according to eq. (4) for every measurement and time step using a rolling rectangular window of $\Delta t = 30$ ms for the computation of the covariance matrix. In order to fill the gaps between the actual receiver positions, linear interpolation was applied.

The first wave front originating from the omnidirectional point source is clearly visible in fig. 2. Since this is pure direct sound, the diffuseness estimate d reaches values close to zero. The drop in d at the upper edge of the heatmap is due to a measurement error and of no further concern. For better visibility, the images in fig. 3 are cropped to the audience area. The scaling of the colours and dimensions remain the same as in fig. 2. They are snapshots from an animation[1] and represent the most interesting points in time during the decay process.

After the initial incidence of the direct sound the field strives for a diffuse state. But, as shown in fig. 3a, this is not a homogeneous process. Depending on location a more or less diffuse state is reached with different time constants, resulting in some slower inclining spots of d. The reason for this is the naturally uneven distribution

[1] https://vimeo.com/586675748



(a) $t = 0.10\,\mathrm{s}$ **(b)** $t = 0.17\,\mathrm{s}$ **(c)** $t = 0.65\,\mathrm{s}$

Figure 3: Diffuseness estimate d at various times t over the floor plane of the lecture hall, cropped to the audience area.

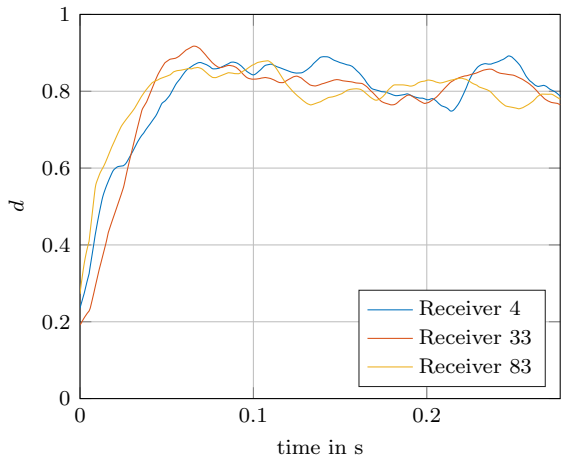
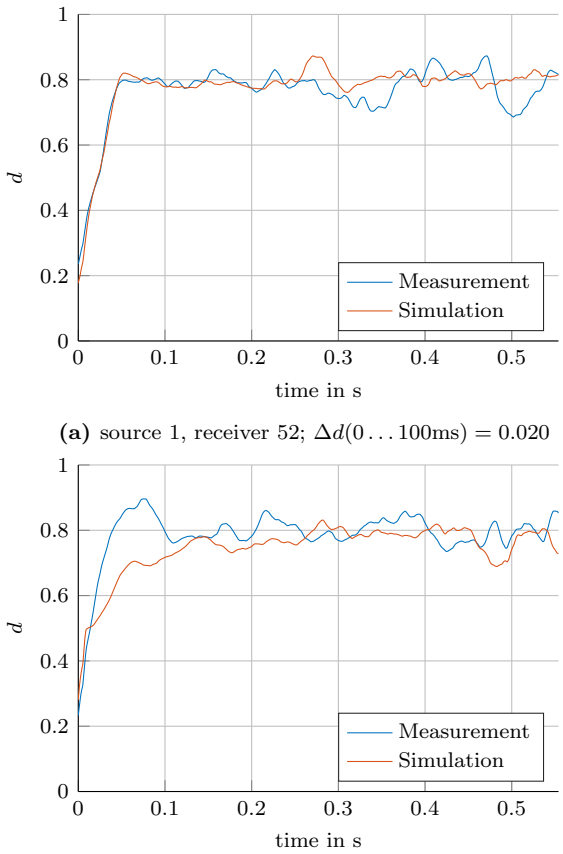


Figure 4: Delay compensated diffuseness estimate d (smoothed) at various receiver positions and source 1 over time.

of early reflections due to the geometry of the room. This can be also observed in fig. 4, which shows the course of d at three random receiver positions and source position 1. Since the receivers have different distance to the sound source, the plots are delay compensated, where the point $t = 0\,\mathrm{s}$ is the time of direct sound incidence. The resurgence of d immediately afterwards has a characteristic curvature that depends on the distribution and density of the early reflections.

At time $t = 0.17\,\mathrm{s}$, as depicted in fig. 3b, an almost homogeneous state of diffuseness is reached. This is the point at which the transition from the geometric to the statistical domain of room acoustics takes place, to put it in classical terms. The reflections become so dense that they can no longer be distinguished from each other, marking the start of the "diffuse reverberation".

However, this highly diffuse state does not remain until the end of the decay process. As fig. 3c shows for the lower values of d, there is a locally fluctuating decrease of incoherence and/or isotropy in the late reverberation. This could be due to a double decay process as described by Berzborn et al. [4]. Nevertheless, it is hard to determine to which point d shows actual diffuseness fluctuations in the decay process and does not just indicate directionality of some randomly occurring sound events in the quiescent noise. This is certainly a point for further investigation as to whether it is possible to clearly identify the end of the decay process with respect to d.



(a) source 1, receiver 52; $\Delta d(0 \ldots 100\mathrm{ms}) = 0.020$

(b) source 1, receiver 82; $\Delta d(0 \ldots 100\mathrm{ms}) = 0.143$

Figure 5: Delay compensated diffuseness estimate d (smoothed) at the source and receiver position with best (fig. 5a) and worst match (fig. 5b) between measurement and simulation.

Comparison: Measurement and Simulation

In order to collect more data and gain more knowledge on a possible application in measurement techniques it would be desirable to observe the diffuseness estimate d in as many different rooms and measurement positions as possible. An efficient way to do so is to use conventional room acoustic simulation instead of extensive measurement series.

In fig. 5 two calculations of d for each measurement and simulation are depicted. Within all source-receiver combinations, these two examples represent both the best (fig. 5a) and the worst match (fig. 5b) between measure-

ment and simulation with mean deviation $\Delta d_{\min} = 0.020$ and $\Delta d_{\max} = 0.143$, respectively within the first $T = 100\,\text{ms}$. This mean deviation is given by

$$\Delta d = \frac{1}{T} \int_{t=0}^{T} |d_{\text{meas.}}(t) - d_{\text{sim.}}(t)| dt \tag{5}$$

Overall, measurement and simulation are in good alignment regarding d. The average deviation is $\overline{\Delta d} = 0.067$ with median $m = 0.065$ and 95% confidence interval $C_{95\%} = [0.061, 0.072]$ over all 182 source-receiver data sets and $T = 100\,\text{ms}$. The initial resurgence right after the direct sound, that might be of interest for a closer qualitative description of early reflections, is reproduced clearly for the majority of the simulations. Even if the absolute value of d after the direct sound differs from the corresponding measurement, the slope of the respective curves is very similar. Due to the limited number of simulated sound rays (setting: 50000 rays) and the partly unknown simulation parameters - especially scattering coefficients - a certain deviation was expected.

For the evaluation of the early reflections, which occur individually and mostly clearly separated from each other, conventional simulation seems to be an appropriate means for further investigations regarding the diffuseness estimate d.

Conclusion

In this paper the suitability of a diffuseness measure such as the covariance matrix eigenvalue estimate d for room acoustic characterisation was examined. The results of a spatially high resolution measurement campaign using a tetrahedral microphone array showed, that there is a lot of potential for new measurement techniques or as an supplement to existing ones to it. Because of the basic assumption of homogeneity and diffuseness without further investigation and the vague definition, the use of such a microphone array, or any other spaced microphone array giving directional room impulse responses would be very beneficial in order to gain a better understanding of the sound field. The ongoing discussions about, for example, absorption coefficient measurements in the reverberation chamber show that there is still much to be done in this field of research.

For applied room acoustic measurement techniques the diffuseness estimate d could be useful for the qualification of early reflections. Even though there are room acoustic measures like C_{80} (Clarity) or D_{50} (Definition) meant for the evaluation of early reflections, those measures only consider the energetic quantity of early reflections in a certain time frame, but not their distribution within that frame. The characteristic curvature of the diffuseness estimate depending on location might be a tool to do so.

References

[1] De Vries, D. et. al.: Spatial fluctuations in measures for spaciousness. The Journal of the Acoustical Society of America (2001), 110, 947-954

[2] International Organization of Standardization: ISO 354:2003 - Measurement of sound absorption in a reverberation room. Geneva, 2003

[3] Del Galdo, G. et al.: The diffuse sound field in energetic analysis. The Journal of the Acoustical Society of America (2012), 131, 1241-1251

[4] Berzborn, M. et al.: On the Directional Propoerties of Energy Decay Curves. Proceedings of the 23rd International Congress on Acoustics (2019), 4043-4050

[5] Witew, I. and Vorländer, M.: Wave Field analysis in concert halls using large scale arrays. Proceedings of the Institute of Acoustics (2018), 40/3, 319-326

[6] Merimaa, J. and Pulkki, V.: Spatial impulse response rendering I: Analysis and synthesis. Journal of the Audio Engineering Society (2005), 53, 1115-1127

[7] Ahonen, J. and Pulkki, V.: Diffuseness estimation using temporal variation of intensity vectors. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (2009), 285-288

[8] Pulkki, V. et al.: Parametric Time-Frequency Domain Spatial Audio. John Wiley & Sons, 2017

[9] Epain, N. and Jin, C.: Spherical Harmonic Signal Covariance and Sound Field Diffuseness. IEEE/ACM Transactions on Audio, Speech, and Language Processing (2016), 24, 1-1

[10] Gover, B. N. et al.: Measurements of directional properties of reverberant sound fields in rooms using a spherical microphone array. The Journal of the Acoustical Society of America (2004), 116, 2138-2148

[11] Behr, S.: Raumakustiksimulationen mit Auralisationen unter Verwendung von CATT-Acoustic. Master's project, TU Graz (SPSC).

[12] Frischmann, C. et al.: Akustische Bestandsaufnahme der TU-Hörsäle Inffeldgasse. Bachelor's thesis, TU Graz (SPSC).

[13] Room acoustic simulation software CATT Acoustic, URL: `https://www.catt.se/`

[14] Spherical Array Processing Toolbox, URL: `https://github.com/polarch/Spherical-Array-Processing`

[15] Bomhardt, R. et al.: The ITA-Toolbox: An Open Source MATLAB Toolbox for Acoustic Measurements and Signal Processing. DAGA Conference Paper (2017), 222-225

[16] ITA Toolbox, URL: `https://git.rwth-aachen.de/ita/toolbox`