



Audio Engineering Society Conference e-Brief

Presented at the International Conference on
Immersive and Interactive Audio
2019 March 27 – 29, York, UK

This Engineering Brief was selected on the basis of a submitted synopsis. The author is solely responsible for its presentation, and the AES takes no responsibility for the contents. All rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Audio Engineering Society.

Multichannel Impulse Response Measurements in MATLAB: An Update on ScanIR

Julian Vanasse, Andrea Genovese, and Agnieszka Roginska

New York University - Music and Audio Research Lab

Correspondence should be addressed to Andrea Genovese (genovese@nyu.edu)

ABSTRACT

This document illustrates new important feature additions to the ScanIR impulse response measurement tool for MATLAB. ScanIR is an academic acoustic measurement software which streamlines the process of recording and analyzing different types of impulse responses for scientific purposes. The main changes to the new version of the software regard the possibility of storing and reading measurements in SOFA format, the possibility of interfacing with a rotating step-motor platform and the computation of relevant acoustic metric parameters from measured impulse responses.

1 Introduction

The response of any linear time-invariant system can be measured via broadband impulse signals. Room acoustic reflections, reverberation or the interference filtering due to the head and outer ear can thus be modeled by an impulse response. First released in 2011 [1], ScanIR is a general Finite Impulse Response (FIR) measurement program for MATLAB which permits the capture of acoustic impulse responses using a variety of measurement signals, such as sine sweeps [2], maximum-length sequence (MLS) [3] or Golay code [4]. ScanIR's main measurement modes include single-channel IRs, stereo HRIR/BRIR, and multichannel measurement.

A pressing issue in the field of acoustic measurement is the use of standards for storing measurement metadata. In an effort to further collaborative research and otherwise encourage the sharing of acoustic measurements we have adopted the Spatially Oriented Format

for Acoustics (SOFA) file-type [5] as an available destination format for measurements made in ScanIR. The SOFA data-structure encapsulates a measurement session's audio data as well as information pertaining to the system measured such as the source and receiver position and angle information.

Additionally, the change in position of sources relative to a measurement array is an important element of the process to fully characterize directional room impulse responses (such as DRIRs or HRIRs in the field of spatial audio). The most rudimentary method to achieve these ends are human intervention in the form of a researcher manually changing the system or source orientation. This is time consuming and leaves open the opportunity for human imprecision. With this in mind, ScanIR now includes the option to add an ARDUINO-controlled step-motor to one's setup to rotate the capture apparatus after each measurement is made.

ScanIR is intended for different sorts of academic users of varying domain expertise. We now provide reference room-acoustics and 3D-audio analytical metrics within the tool to enhance the data collection process. These metrics can prove useful for quick feedback about the space of recording and allow non-experts to access an estimation of the acoustic parameters relative to the measurement.

2 Spatially Oriented Format for Acoustics

2.1 Background

The SOFA file-type is a container developed to standardize the exchange of multichannel and multi-measurement acoustic capture [5]. Head-related impulse responses (HRIRs) and directional room impulse responses (DRIRs) are composed of a series of measurements made from a number of positions relative to one or more microphones. A SOFA file includes fields associating each measurement with a direction, or position, along with additional metadata regarding the room type and author information. In the generalized technical language of the SOFA file-type, a measurement is made through a setup containing *emitters*, devices which broadcast the measurement signal, and *receivers*, devices which capture the response of the acoustic environment.

The adoption of this format as an available final destination file-type for a ScanIR measurement session was informed by the readily available SOFA programming interfaces for MATLAB, Octave and C++, which further facilitates the exchange of data in this format. The numerical container relies on netCDF-4 (Unidata), a set of software libraries supporting the creation and sharing of scientific data. SOFA supports a number of measurement conventions, the convention of interest for measurements made in ScanIR being the ‘General-FIR’.

2.2 SOFA Implementation

In the source directory included with the most recent release of ScanIR are the two functions `sofa2scanIR()` and `scanIR2sofa()`. These functions convert without loss between the previously used `.mat` format and the adopted SOFA file-type, and vice versa. They are embedded in the program

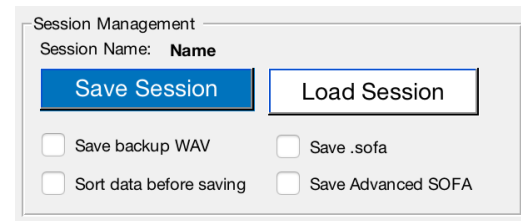


Fig. 1: Added output selection interface.

user-interface (see Fig. 1) but may also be used to bulk convert previous ScanIR measurement sessions to the new file-type.

A session of finite impulse response (FIR) measurements is stored in a multidimensional array of size $[M \times R \times N]$ where M is the number of measurements, R is the number of receivers and N is the length of the measurement (in samples). This structure is stored in the field `Data . IR`. Each measurement has a corresponding source position vector entry in a matrix of vectors stored in the `SourcePosition` field with coordinate type (Cartesian, horizontal, etc) stored in the `SourcePosition_Type` field. The data fields between the previously used output format and the SOFA standard are not in total correspondence. To ensure loss-less conversion and ensure compatibility some ScanIR specific metadata is written to the `GLOBAL_Comments` field of the resultant `.sofa` file.

The ‘Advanced SOFA’ checkbox in Fig. 1 permits the ScanIR user to enter additional metadata for storage in the output file. When checked, as the program saves the measurement session the dialog window featured in Fig. 2 will become visible allowing the user to enter and save information regarding the database, author contact and make additional comments.

3 Rotating Motor Platform

3.1 ARDUINO Motor

ScanIR is flexible to a variety of uses. Certain types of measurements such as directivity patterns for reproduction and capturing equipment, directional responses, head-related-impulse-responses, etc., depend on the polar angle between source and receiver. To address the need of experimenters of repeating measurements at different angle rotations, an automated rotating step-motor

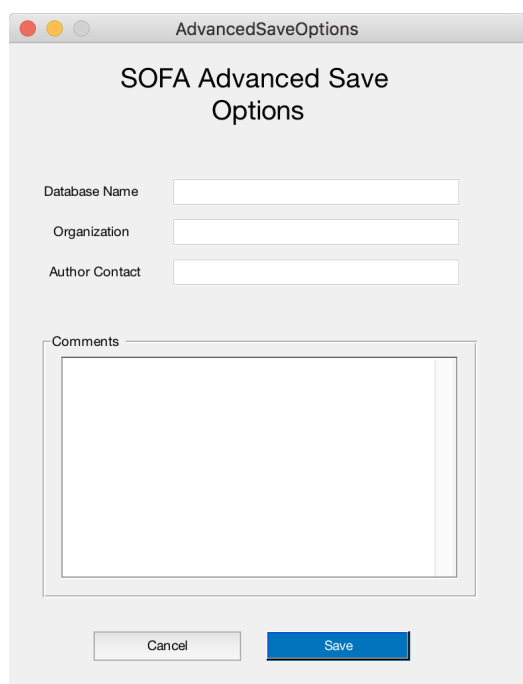


Fig. 2: Advanced SOFA save options.

functionality feature has been added to the ScanIR program. This feature system was first implemented in [6] under the name of ARM (Automatic Rotating Motor). The ARM components allow either source or receiver to rotate on axis through remote control, for a specified number of step increments.

This optional system requires an *ARDUINO UNO Rev3* micro-controller, an *Adafruit Motor Shield V2*, and the *ARDUINO MATLAB* support package. The feature works with a range of compatible stepper motors of various step resolutions. The use of the ARM mount system was first tested for the measurements described in [6, 7], which respectively involved, a directivity pattern measurement for a 1st order ambisonics microphone and a HRIR dataset for 200 azimuth positions around the head. The stepper motor used in those studies was the *NEMA23 3A* which featured 200 steps for 1.8° per step, although any motor device interfaceable with the listed components is compatible with the implementation.

3.2 Step-Motor usage

ScanIR automatically checks for the presence of an ARDUINO micro-controller device and allows the user to

input a range of rotation in degrees for the step-motor. It is possible to choose a number of steps for a specified step-size (based on the available step-grid resolution) or alternatively choose a target angle of arrival. The measurement signal of choice is automatically emitted at each step rotation. If no step-motor is detected at start-up, the user has a chance to re-establish connection within the session by clicking on the "Connect ARDUINO" button. It is suggested to clamp the rotating motor firmly to a support as the vibrations of the step movement may displace the emitter/receiver device.

4 Analysis Metrics

In the fields of room acoustics and immersive audio, scientists and engineers often refer to specific analysis metrics when providing parametric descriptions of impulse responses. These metrics can be used to describe the character of an acoustic response in various forms. ScanIR now provides to the user a number of reference metrics in addition to the impulse response data. While expert users may want to look into computing their customized acoustic analysis based on their own specifications, these metrics can be useful for general analysis purposes and data-collection monitoring.

These parameters are displayed within the GUI in either single-channel form, for each input channel, or in multi-channel average form, and can be accessed after recording from the `.mat` file output generated after saving the measurement session. Formulas are taken from both domain-related books of classic acoustic theory and recent publications about novel algorithms. For more context on the meaning and interpretation of these parameters, please refer to [8, 9].

Reverberation Time One of the most common parameters in regards to acoustic impulse responses, is the reverberation time. Usually referred to T_{60} , this parameter indicates the time it takes for the sound energy to decay by 60 dB within the space of recording. The parameter is usually calculated from an *Energy Decay Curve* (EDC) calculated using the Schroeder integration method [10]. The EDC is defined as the tail integral of the squared impulse response h at time t :

$$\text{EDC}(t) \triangleq \int_t^{\infty} h^2(\tau) d\tau$$

As this represents the total signal energy remaining in the impulse response at time t , we can extract the T_{60}

value, in seconds, from the \log_{10} of the EDC, as simply the t value at which the decay reaches -60dB . In practice, this metric is often miscalculated as it necessitates a recording performed with a microphone with a high dynamic-range and with enough headroom from the environmental noise-floor. To address this possible issue, the metrics of T_{30} and T_{20} are calculated on the EDC over a smaller decay range in which the measurement is more likely to be accurate. A regression line is then used to extrapolate the T_{60} reverberation time value along the fitted line of decay [11]. The ISO 3382 [12] standard specifies a minimum decay range of 45 dB for T_{30} (with regression range from -5dB to -35dB) and a minimum decay range of 35dB for T_{20} (with regression range from -5dB to -25dB). Having a regression range starting at -5dB reduces the interfering influence of early signal fluctuations caused by reflections and thus providing a better estimate for T_{60} [13].

Similarly, the Early Decay Time (EDT) is computed from the EDC using a regression line going from 0dB to -10dB . The EDT is a psychoacoustically relevant metric, useful for room acoustic design. Since the perception of reverberation is non-linear, sharp decreases in early energy decay are often the cause for the space being perceived as “drier” than it really is [14].

Energy Ratios Another common IR metric is the *Direct-to-Reverberant-Ratio* (DRR). The DRR is often used to understand the amount of reverberation in comparison to the direct sound path of the source to the microphone and it can be an indicator of “wetness” and apparent source distance. The DRR is given as the \log_{10} of the ratio of the direct energy of the IR - which is the energy in between the direct path onset up to 2.5 ms after - to the late energy [15]:

$$\text{DRR} = 10 \log_{10} \left(\frac{\int_0^{2.5\text{ms}} h^2(t) d(t)}{\int_{2.5\text{ms}}^{\infty} h^2(t) d(t)} \right)$$

Other similar parameters are the *Definition*, or *Deutlichkeit*, (D_{50}) and the *Clarity Index* (C_{50}/C_{80}). The D_{50} is reported as an indicator of sound intelligibility and consists in the ratio of the early energy within the first 50 ms of the IR, to the total signal energy.

$$D_{50} = \frac{\int_0^{50\text{ms}} h^2(t) d(t)}{\int_0^{\infty} h^2(t) d(t)}$$

“Clarity” is instead used as an indicator of sound “brilliance” and is computed as the log of the ratio between the energy in the early part of the IR - up to 50 milliseconds for speech (C_{50}) or 80 milliseconds for music (C_{80}) - to the energy in the later part of the IR.

$$C_{50} = 10 \log_{10} \left(\frac{\int_0^{50\text{ms}} h^2(t) d(t)}{\int_{50\text{ms}}^{\infty} h^2(t) d(t)} \right)$$

Reflection parameters Two additional IR parameters are given by the *Initial Time Delay Gap* (ITDG) and the *Central Time* (T_s). The ITDG relates to room dimensions or wall proximity and consists in the time in milliseconds passing between the first direct onset and the first environmental reflection arriving to the receiver. The T_s parameter is instead related to the first momentum of the IR and corresponds to the “gravity center” of the whole impulse response, thus reflecting the balance between clarity and reverberation. Larger values of T_s mean that the energy is more distributed over time, giving a perceptual impression of a larger reverberant space [13].

$$T_s = \frac{\int_0^{\infty} t \cdot h^2(t) d(t)}{\int_0^{\infty} h^2(t) d(t)}$$

HRIR parameters These parameters are accessible when using the HRIR/BRIR measurement mode and relate to the relationship between the left and right channels of the impulse response as descriptors of perceptual spatial cues. At lower frequencies, the *Interaural Time Delay* (ITD) is the dominant localization cue embedded within an HRIR filter pair. The ITD is the difference between the direct path onset arrival times between the left and the right channels [16]. The ITD is here calculated from low-passed ($<3\text{ kHz}$) minimum-phase HRIRs as the max displacement value of the interaural cross correlation [17]. Differently, in the high-frequency region, the *Interaural Level Difference* becomes the dominant cue for spatial localization. A simple measure of the ILD is given by the dB power difference taken from the band-passed ($2\text{-}4\text{ kHz}$) HRIRs [18]:

$$\text{ILD} = 10 \log_{10} \left(\frac{\int h_L^2(t) d(t)}{\int h_R^2(t) d(t)} \right)$$

Finally, the *Inter-Aural Cross-correlation Coefficient* (IACC), a measure of interaural dissimilarity, is taken from the band-passed HRIRs (>500 Hz) as:

$$\text{IACC} = 1 - \max_{\tau} \left| \frac{\int h_L(t)h_R(t + \tau)d(t)}{\sqrt{\int h_L^2(t)d(t) \cdot \int h_R^2(t)d(t)}} \right|$$

Where τ is a time interval typically chosen to be from -1 ms to +1 ms where the maximum correlation is sought. The IACC takes a value from 0 to 1 which reports the amount of channel dissimilarity, corresponding to increasing perceptual “spaciousness” [13].

5 Other changes

To address one common issue with the past version, the tool now allows a user to themselves decide which audio input/output interface to connect to, instead of using the system default. In ScanIR version 2.0 the user is able to select a single device acting for both signal emission and measurement, from a drop-down menu. Built-in options, aggregate devices, or external interfaces (suggested) can be used.

A BRIR (Binaural Room Impulse Response) measurement mode setting has been added to allow for longer measurements to be recorded, while also maintaining the convenient presets of the HRIR measurement mode. BRIRs are reverberant HRIR filters which capture the room reflections using a binaural microphone array.

6 Download

The ScanIR v2.X matlab tool can be downloaded at the NYU Immersive Audio Group Github page <https://github.com/NYU-ImmersiveAudio/ScanIR>. The git includes the full user manual and more information about the tool. Please note that the tool has been extensively tested on MacOS Sierra and MATLAB R2018b. The use of other operating systems or older MATLAB versions may possibly lead to errors or bugs. The authors encourage the users to point out any bug, or desired features, at the contact email address given in this e-brief.

For reference about the pre-existing features of ScanIR, please refer to the publication accompanying the first release [1] and the user manual. All publications and links are available on the Music and Audio Research Lab website (<https://steinhardt.nyu.edu/marl/>).

6.1 Installation Requirements

ScanIR is built upon version 3 of PsychToolbox [19] which needs to be installed to MATLAB ¹. If the presence of the Signal Processing Toolbox is detected, the program is able to compute all the analytical metrics, otherwise some wont be shown. For using the optional rotating platform feature, please download and install the MATLAB ARDUINO support package².

The SOFA-API for MATLAB/OCTAVE is required to use the SOFA optional saving format [20]. Version 1.02 of the API is already included in the ScanIR release. If a different version is desired, it is possible to download and place the new API parent folder in the `api` folder and delete the older ones.

6.2 Future Work

Version 2.1 of ScanIR is planned to be released in 2019. The main additions will include patches for Windows 10, separate input and output device selection and updated analysis metrics, allowing sub-band analysis for octave-bands. Further releases will include windowing options to allow users to further process the signals as desired. ScanIR is currently a MATLAB tool tested on MacOS. However, a Python version might be discussed at a future stage to enhance use and compatibility with different systems.

7 Acknowledgements

The authors would like to thank all the previous contributors to the ScanIR source code and the students who used and tested the tool in recent years, providing important feedback leading to bug fixes. A particular thanks to Gabriel Zalles and Charlie Mydlarz who first implemented the ARM ARDUINO feature and to Braxton Boren who published the first version and provided guidelines for the developments in place.

¹<http://psychtoolbox.org/>

²<https://www.mathworks.com/hardware-support/arduino-matlab.html>

References

- [1] Boren, B. and Roginska, A., “Multichannel impulse response measurement in matlab,” in *Audio Engineering Society Convention 131*, Audio Engineering Society, 2011.
- [2] Farina, A., “Simultaneous measurement of impulse response and distortion with a swept-sine technique,” in *Audio Engineering Society Convention 108*, Audio Engineering Society, 2000.
- [3] Schroeder, M. R., “Integrated-impulse method measuring sound decay without using impulses,” *The Journal of the Acoustical Society of America*, 66(2), pp. 497–500, 1979, doi:10.1121/1.383103.
- [4] Foster, S., “Impulse response measurement using Golay codes,” volume 11, pp. 929 – 932, 1986, doi:10.1109/ICASSP.1986.1168980.
- [5] Majdak, P., Iwaya, Y., Carpentier, T., Nicol, R., Parmentier, M., Roginska, A., Suzuki, Y., Watanabe, K., Wierstorff, H., Ziegelwanger, H., et al., “Spatially oriented format for acoustics: A data exchange format representing head-related transfer functions,” in *Audio Engineering Society Convention 134*, Audio Engineering Society, 2013.
- [6] Zalles, G., Kamel, Y., Anderson, I., Lee, M. Y., Neil, C., Henry, M., Cappiello, S., Mydlarz, C., Baglione, M., and Roginska, A., “A Low-Cost, High-Quality MEMS Ambisonic Microphone,” in *Audio Engineering Society Convention 143*, Audio Engineering Society, 2017.
- [7] Genovese, A., Zalles, G., Reardon, G., and Roginska, A., “Acoustic perturbations in HRTFs measured on Mixed Reality Headsets,” in *Audio Engineering Society Conference: 2018 AES International Conference on Audio for Virtual and Augmented Reality*, Audio Engineering Society, 2018.
- [8] Kuttruff, H., *Room acoustics*, Crc Press, 2016.
- [9] Blauert, J., *Spatial hearing: the psychophysics of human sound localization*, MIT press, 1997.
- [10] Schroeder, M. R., “New Method of Measuring Reverberation Time,” *The Journal of the Acoustical Society of America*, 37(3), pp. 409–412, 1965, doi:10.1121/1.1909343.
- [11] Hak, C., Wenmaekers, R., and Van Luxemburg, L., “Measuring room impulse responses: Impact of the decay range on derived room acoustic parameters,” *Acta Acustica united with Acustica*, 98(6), pp. 907–915, 2012.
- [12] ISO International Standard ISO/DIS, *ISO/3382-2:2008: Acoustics — Measurement of room acoustic parameters — Part 2: Reverberation time in ordinary rooms*, pub-ISO, 2008.
- [13] Schroeder, M., Rossing, T. D., Dunn, F., Hartmann, W., Campbell, D., and Fletcher, N., “Springer handbook of acoustics,” 2007.
- [14] Howard, D. M. and Angus, J., *Acoustics and psychoacoustics*, Focal press, 2017.
- [15] Zahorik, P., “Direct-to-reverberant energy ratio sensitivity,” *The Journal of the Acoustical Society of America*, 112(5), pp. 2110–2117, 2002.
- [16] Begault, D. R. and Trejo, L. J., “3-D sound for virtual reality and multimedia,” 2000.
- [17] Katz, B. F. and Noisternig, M., “A comparative study of interaural time delay estimation methods,” *The Journal of the Acoustical Society of America*, 135(6), pp. 3530–3540, 2014.
- [18] Boyd, A. W., Whitmer, W. M., and Akeroyd, M. A., “Recording and analysis of head movements, interaural level and time differences in rooms and real-world listening scenarios,” *ISRA 2013*, 2014.
- [19] Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., et al., “What’s new in Psychtoolbox-3,” *Perception*, 36(14), p. 1, 2007.
- [20] “SOFA (Spatially Oriented Format for Acoustics),” <https://www.sofaconventions.org/>, Online; accessed May 2018.