HAART: a New Impulse Response Toolbox for Spatial Audio Research

Dale Johnson[1], Alex Harker[2], and Hyunkook Lee[1]

[1] Applied Psychoacoustics Lab, University of Huddersfield, West Yorkshire, HD1 3DH, United Kingdom dale.johnson@hud.ac.uk, h.lee@hud.ac.uk

[2] CeReNeM, University of Huddersfield, West Yorkshire, HD1 3DH, United Kingdom a.harker@hud.ac.uk

ABSTRACT

This engineering brief describes a new, open source code library named HAART (Huddersfield Acoustical Analysis Research Toolbox). HAART simplifies the measurement and analysis of multi-channel impulse responses (IRs). For the purposes of this engineering brief the code library is compiled as a set of Max objects that form a prototype program in Max. This program is able to perform the acquisition, manipulation and analysis of IRs using subjective and objective measures described in acoustics literature. HAART is also able to convolve IRs with audio material and, most importantly, able to binaurally synthesize virtual, multichannel speaker arrays over headphones, negating the need for multichannel setups when out in the field. The code library is freely available from:
http://www.hud.ac.uk/research/researchcentres/mtprg/projects/apl/

1. INTRODUCTION

HAART is a set of modular tools designed to facilitate the acquisition, analysis and binaural auralisation of multi-channel impulse responses that embeds and expands the HISSTools software library [1] It is ultimately intended to be a complete, self-contained program available for both Mac OS and Windows, and other platforms in the future, accompanied by a modular codebase for other researchers to adapt to future requirements. However, the software presented at this time is a Max implementation/prototype of the program. Both the application and library code are completely open source and written in C++.

1.1. Aim and Scope

HAART aims to address issues faced with the analysis and auralisation of multi-channel impulse responses whilst out in the field. The equipment used when carrying out IR measurement sessions commonly includes microphones, speakers, a laptop, an audio interface and headphones. Normally, one would have to acquire the IRs first, edit them and then manually analyse them and convolve them all within separate applications. This can make the entire process inefficient. As IR acquisition is a precise and painstaking process, it is necessary for results to be immediately available so that the acquired IR can be validated.

The proposed tool conducts the acquisition, analysis and binaural synthesis within one interface, thus providing researchers with an efficient workflow in room acoustics and spatialisation research. HAART acts as an expansion to the HISSTools library by adding functions for the analysis of spatial/binaural attributes and multichannel virtual loudspeaker simulation. The

program is designed to facilitate multichannel situations in which the user wishes to capture multiple groups of sources / receivers (e.g. multiple microphone setups or loudspeaker configuration) within a single session, either through simultaneous or sequential capture. The user interface is designed to facilitate this level of flexibility in setup, whilst also ensuring ease of use. Importantly, because HAART is aimed at users in the field, it is possible to convolve existing audio material with multi-channel IRs and, via the use of head related impulse responses (HRIRs), synthesize the result over headphones so that users can anticipate how the result will sound over multichannel loudspeaker setups.

1.2. Workflow

The basic workflow is as follows. Once multichannel impulse responses are acquired, they are manipulated as desired (normalisation, time alignment, D/R ratio variation, Ambisonics/virtual microphone decoding, etc.). The interface then enables the user to virtually monitor the loudspeaker playback of the IR-convolved sound through convolution, as well as to analyse the spatial characteristics of the IRs objectively based on the standard parameters defined in the room acoustics

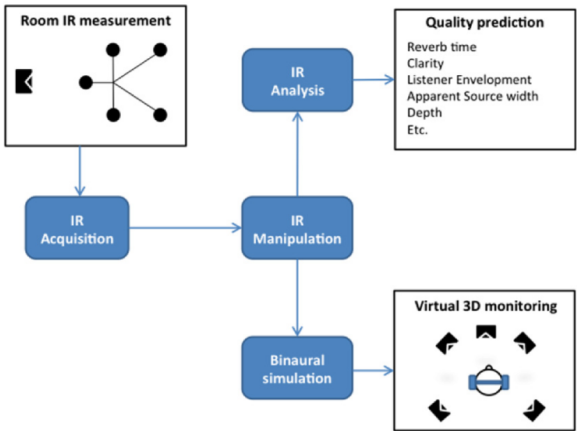

Fig.1. Basic workflow of HAART

literature. Therefore, one can have both the subjective and objective ways of monitoring acquired IRs, and changes in microphone configuration can be made accordingly to achieve target IR characteristics.

2. SYSTEM OVERVIEW

2.1. Acquisition

The IR acquisition process is provided by the HISSTools software library. Multichannel response can be acquired using various excitation methods including the exponential sine sweep (ESS) or maximum length sequence (MLS). Recordings of the excitation signals are de-convolved immediately after capture to form the impulse response using appropriate inverse filter and spectral equalisation techniques, such as those proposed by Farina [2]. At this stage IRs are tagged with metadata about how the impulse response was acquired (e.g. microphone technique, position, spacing, type etc). It is worth noting that users are not limited to using IRs acquired from HAART alone and can skip the acquisition stage if they wish to provide their own IRs that have been acquired using other tools.

2.2. Metadata

HAART makes use of metadata in human readable XML format that is designed to aid the user throughout the measurement process and act as a record in complex measurement scenarios in which there are many channels/measurements. The user specifies sources and receivers in terms of positionable groups of loudspeakers or microphones that are subsequently treated as a unit within the software. Settings correspond to well-known formats (e.g. ORTF, Decca Tree, B-format, binaural dummy head, etc. for microphones, or stereo, 5.1, 7.1 etc, for loudspeakers), with adjustable parameters for distance, spacing and position. Fully customizable configurations are also possible.

As well as acting as a readable record of the session, metadata guides normalization and editing (which is applied in relevant groups), as well as data display and selection during the analysis procedure.

2.3. Manipulation

After acquisition, the IR can be manipulated to achieve:

- IR Normalization (in relevant groups).
- Time alignment via the use of onset detection.
- Trimming of IRs
- Decoding and conversion from B-Format and Mid-Side signals.

2.4. Analysis

After onset detection, trimming and decoding has been performed, the impulse response can be analysed using relevant acoustic parameters found in the existing literature, including the parameters documented in ISO 3382-1 [3]:

- Reverberation Time (T30, T20, EDT)
- Sound Strength (G)
- Clarity (C)
- Definition (D)
- Centre Of Gravity
- Interaural Cross-correlation Coefficient (e.g. IACCE3, IACCL3)
- Interchannel Cross-correlation Coefficient (ICCC)
- Lateral Energy Fraction (e.g. Early LF, Late LF, Lateral G)
- Early/Late Support
- Front-Back Energy Ratio
- Direct to Reverberant (D/R) Energy Ratio

In addition to the predefined analysis parameters HAART allows the user to define their own values for dividing the direct, early and late sound paths of the IR for calculating further energy ratio and cross-correlation based parameters. Multiband filtering is available for parameter calculation, with options for octave bands, 1/3 octave bands and ERB critical bands provided.

2.5. Auralisation

Acquired impulse responses can be convolved with existing recorded material and auralised over multichannel systems for verification. However as recording equipment is becoming more portable, and many engineers wish to hear their results without requiring immediate access to multichannel listening environments, it is desirable for engineers to listen to their results over headphones. This requires binaural auralisation using convolution. HAART is able to simulate multichannel auralisation over headphones by convolving audio material with each captured IR channel, followed by a HRIR that maps the IR convolution results onto the virtual speaker setup for headphone listening. HRIRs that simulate various multichannel speaker setups will be provided. If the IR is inherently binaural, the HRIR convolution can be bypassed.

3. INTERFACE

The interface is designed to balance simplicity and ease of use whilst allowing a high degree of flexibility. To this end, each module in the program is independent, with its own interface, so that the user is presented with only a limited and relevant amount of information at any given time.
For the acquisition stage, the user chooses the excitation method they will be using and enters data about microphone/loudspeaker setups, via a series of presets, which can be customized for non-standard configurations. The user can test prior to acquisition to ensure no clipping or spurious anomalies occur. Metadata about the setup is stored with the IRs for later use.

After the acquisition stage is complete, the program presents the IR manipulation stage where the user can view the IR waveforms to determine any necessary corrections, alongside applying relevant editing and decoding. Most importantly this stage allows for onset detection in order to automatically determine the relevant sections of the IRs for analysis. Edits can be verified and corrected visually, and rendered to disk if desired.

In the analysis stage, the user can select to perform any of the provided acoustic analysis functions on specific IRs from the captured set, producing onscreen readout or file export. As certain acoustic parameters are only applicable to certain setups, (e.g., IACC is only useful for binaural recordings) access to these is handled using the relevant metadata. Users can also perform analysis on subregions and subsets of multichannel IRs, thus deriving analysis data beyond the conventional used ISO measures, making HAART a potentially useful tool for researchers.

Depending on the analysis parameters, the results can be displayed as a table of values where each parameter is matched with the corresponding channel and filter band result. These results can also be viewed in the form of a graph. For time varying analysis such as cross- correlation, the results are represented as a graph. The results can be exported either as a text file or an image.
During both the manipulation and analysis stages, the user has the option to convolve existing audio material with the IR. The tool provides them with different playback modes for either loudspeaker arrangements or headphones. Included in the tool is a simple file browser for selecting between different existing audio materials from a single directory in order to quickly load them into HAART for auralisation.

4. SUMMARY AND FUTURE PLAN

To summarise, HAART is designed to provide users with an easy way of measuring, analysing and auralising impulse responses in a simple and easy to use environment. It allows researchers to obtain results quickly, allowing them to make changes easily without requiring access to a multichannel speaker setup via the use of binaural synthesis.

The future of HAART includes extra functionality to analyse the IRs further such as observing the interchannel time and level differences and their effect on stereophonic localisation.
HAART will ultimately be made available as a self- contained program for Mac OS and Windows, with a library codebase that can be integrated into other applications either as a plugin or an extension, or used to create related applications for any platform. The current version of HAART and its code library is freely available through the website below:
http://www.hud.ac.uk/research/researchcentres/mtprg/projects/apl/

5. ACKNOWLEDGEMENTS

The authors would like to thank Prof. Liza Lim and Prof. Monty Adkins at the Centre for Research in New Music (CeReNeM) for supporting the project. This work was funded by the University of Huddersfield, Grant Ref. URF213-01/1014, and the Engineering and Physical Sciences Research Council (EPSRC), Grant Ref. EP/L019906/1.

6. REFERENCES

[1] Harker, A., & Tremblay, P. A. (2012), *The HISSTools impulse response toolbox: convolution for the masses.* In: ICMC 2012: Non-cochlear Sound. The International Computer Music Association, pp. 148-155. ISBN 9780984527410

[2] Farina, A. (2000). Simultaneous measurement of impulse response and distortion with a swept-sine technique. *108th Audio Engineering Society Convention, Preprint 5093.*

[3] British Standards Institution. (2012). ISO 3382-1:2009: Acoustics Measurement of room acoustic parameters: Part 1: Performance spaces. London: BSI.