AURALIZATION AS A TOOL FOR EVALUATING A MUSICAL INSTRUMENT

Matthew Boucher[1], *David Pelegrin-Garcia*[2], *Bert Pluymers*[1], *and Wim Desmet*[1]

[1] Dept. of Mech. Eng.
[2] Dept. of Physics & Astronomy
KU Leuven, Heverlee, Belgium
`matt.boucher@kuleuven.be`

ABSTRACT

When a musician is considering the purchase of a new instrument, the test room often does not have ideal acoustics, and a musician rarely has the opportunity to test the new instrument in a concert or recital hall before purchase. However, one of the most important criteria when choosing a musical instrument is its sound in a representative performance environment. If recorded in a relatively anechoic environment, the dry sound of an instrument can be merged with the acoustic reflections of a performance environment through a process known as auralization. The technique of auralization, therefore, allows the musician to evaluate the sound of the instrument without being influenced by non-musical cues, such as quality or knowledge of manufacturing. By reproducing the auralized sound, a musician can listen to his/her own playing as if it were played in a concert hall, for example, with the auralized sound quality being one aspect of total quality. Auralization techniques can then be a service offered to musicians by the instrument maker at the time of purchase in addition to testing in standard factory showrooms.

1. INTRODUCTION

At the time of purchasing a musical instrument, a musician starts with several instruments to test. By repeatedly playing typical repertoire on each instrument, unsuitable instruments are identified and removed from the pool of possibly suitable instruments. Eventually, one instrument is judged to be the best, and a happy musician goes home with a new instrument.

While this process seems simple enough, there are many factors that contribute to the decision process. For example, when choosing a tuba, it is suggested that intonation, tone, response and dynamics are all important playing criteria and that the evaluation may be different for low, middle and high tones [1].

As a further complication, non-musical cues, such as touch, can affect one's perception of sound during playing [2]. It has been shown that pianists can determine which piano they are playing even when their vision and hearing are intentionally impaired [3], indicating that the mechanical response of piano keys is an identifying factor of quality.

Another non-musical cue can be simple knowledge of how the instrument was made or materials of construction. Smith's experiment showed that a copper trombone bell was not distinguished from brass bells of varying thickness by a group of 10 professional trombonists during double blind tests. However, distinguishable playing characteristics were attributed to the copper bell during non-blind tests [4].

This provides motivation to separate the evaluation of sound from complicating factors such as touch and knowledge of fabrication. A simple way to do this is to record the playing and judge the instruments only based on a listening test without any visual or tactile clues. However, the listening test is then only valid to judge the instrument's sound in a particular recording room. Of more concern is that musicians, whether consciously or subconsciously, change their performance based on their acoustical environment [5].

For these reasons, isolating the sound of a musical instrument from non-musical cues and room effects is desired. This can be achieved through a technique known as auralization. The technique consists of recording a sound in a room with sound absorbing walls and then adding the perceivable effects of another acoustic environment. An added benefit is to judge the sound of the instrument in different musical settings [6].

In this work, the basics of the auralization techniques are outlined, and their application to the evaluation of a basstuba (in F) is demonstrated. Two excerpts of different musical styles, common to the tuba repertoire, are chosen and auralized. They are taken from: (1) *Vocalise*, Op. 34, No. 14 by Sergei Rachmaninoff, a lyrical, singing melody originally written for soprano or tenor voice that has been transcribed for various instruments, including the tuba and (2) the *Hungarian March* from Hector Berlioz's *The Damnation of Faust*, which is played at a higher dynamic level and requires a harder, march-like articulation.

Applying auralization in this context, the intent is to isolate the sound of the instrument so that it may be evaluated in the absence of non-musical cues. This work is intended to aid musicians and musical instrument manufacturers in the evaluation of a musical instrument at the time of purchase.

2. AURALIZATION

The basic process of auralization consists of recording a sound source in a room with sound absorbing materials on its walls and then blending this "dry" sound with the response of a particular performance environment. It is most often associated with the evaluation of concert halls, which has been the subject of lots of research [7]. Even though the usual goal is to compare the quality of different concert halls, the method is equally valid to evaluate different sources of sound in the same concert hall, which would be valuable in the case of a musician testing
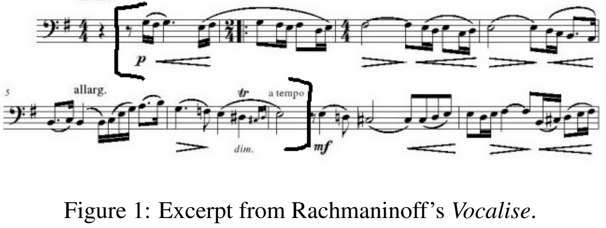
Figure 1: Excerpt from Rachmaninoff's *Vocalise*.
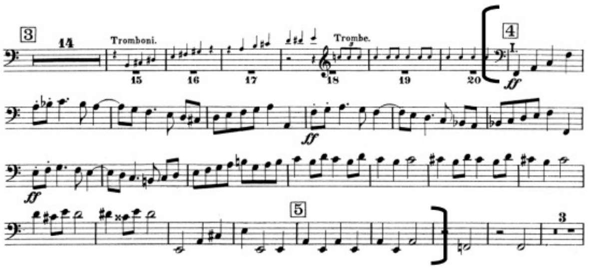
Figure 2: Excerpt from Berlioz's *Hungarian March.*

several instruments for purchase.

New challenges are presented in this case, of course, namely, that the same music should be reproduced exactly on each instrument. However, as is the case in most recording sessions, several takes are allowed, so it is assumed that experienced musicians are able to reliably reproduce high quality performances on different instruments. Any differences that remain are attributed to the characteristics of the instrument.

In this section, the main steps of the auralization technique are outlined: 2.1 the making of dry recordings, 2.2 taking into account room and listener effects of a typical musical environment, 2.3 adding these effects to the dry recordings and 2.4 listening and evaluation. The process is applied to the musical excerpts shown in Figs. 1 and 2.

2.1. Anechoic recording

An anechoic (*i.e.*, without echo) room is an acoustically dry room that is specifically designed for performing sensitive acoustic experiments. Such experiments include measuring the directivity of loudspeakers, human hearing tests and noise radiation measurements of machines. Although not achievable in practice, a perfectly anechoic room absorbs all sound energy at its walls at all frequencies, imitating an infinite space. This means that a listener inside an anechoic room in the presence of a sound source only hears the direct sound from the source and not the reflected sound from the walls of the anechoic room.

When a musical instrument is played in a normal room, the sound heard by a listener contains the sound of the musical instrument and the sound reflected from the walls of the room. The differences between a normal room and an anechoic room are illustrated in Figs. 3 and 4. If a musical instrument is played inside an anechoic room, only the sound starting from the source and traveling directly to the microphone is present; any reflected sound from the walls is absent.

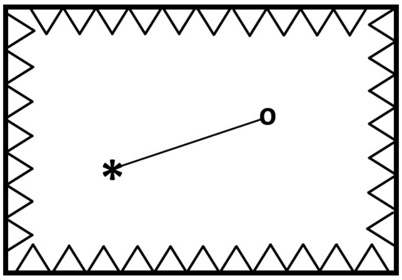

Figure 4: Only the direct sound is present in an anechoic room. The reflections are eliminated by the triangular wedges, which are made of absorbing materials.

The recording setup in the anechoic room of the KU Leuven Dept. of Physics & Astronomy is shown in Fig. 5. Two microphones are placed in front of the player at $+/-45^\circ$ and one microphone directly behind the player. All three microphones were placed at a distance of 150 cm. The two front microphones were Behringer ECM8000 1/2" microphones, and the one behind was a Brüel & Kjaer type 4192 1/2" pressure-field microphone with measurement amplifier type 2160. A fourth microphone, a miniature Sennheiser MKE2 (4.8 mm), was placed on-axis at the opening of the bell using tape and string. The audio interface was a Roland UA-1010 Octa-Capture, and the Reaper Digital Audio Workstation was used, which allowed multitrack recording [8].

It is important to note that the choice of microphones has an impact on the quality of the recordings. For example, large diaphragm microphones (1") are more sensitive than microphones with smaller diaphragms (1/2" or below). While this allows the recording of more faint sound pressure levels (SPL), they can not be used to record very high SPLs, such as those arising at fortissimo passages played with the tuba at short distances. For auralization purposes, therefore, it is important to maximize the dynamic range of the recording, *i.e.*, the difference between the strongest signal peaks and the noise floor. Moreover, the frequency response of the microphones should be flat.

The directivity of the musical instrument should also be considered when making anechoic recordings, as it plays a role in the placement of the recording microphones. The directivity of a musical instrument describes how the low, middle and high tones radiate in different directions. Fortunately, due to the bass nature of the tuba, its sound radiates rather evenly in all directions over most of its playing range [9, 10], and the recordings at the three microphone positions are quite similar. Other instruments with more complicated directivity patterns (*e.g.*, violin or flute [9]) may need a more complex arrangement of micro-

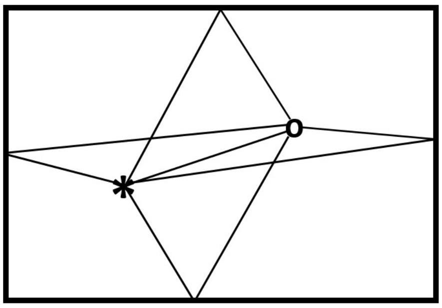

Figure 3: The sound in a normal room comes from the source ($*$) and reflections from the walls and arrives at the listener (**o**).

Figure 5: Recording setup in anechoic room using 3 microphones.

RIR	all room effects including the positions of the source and listener, materials on walls, size and shape of room
HRIR	all effects due to the presence of a listener including shape of upper body, head and ear
BRIR	room and listener effects contained in RIR and HRIR

Table 1: Summary of different impulse responses.

phones [11]. Other sounds like mechanical noise or breathing from the player may also have a more complicated directivity pattern. With the intention of evaluating the auralized sound of the tuba, these additional sounds should be ignored as much as possible.

Under these conditions, anechoic recordings of a tuba performing the excerpts found in Figs. 1 and 2 were obtained, eliminating the room acoustics feedback given to the musician [1]. The next step in the auralization process is to determine the effects a real musical environment has on the sound of the instrument. This is done by using impulse responses.

2.2. Impulse responses

Once an anechoic recording is obtained, the relevant, perceivable aspects of listening to this sound as if it were performed on the stage of a concert hall must be added. This is done by using impulse responses that contain the effects of the room and listener's presence. There are three types of impulse responses that are relevant to the auralization of a musical instrument, the RIR, HRIR and BRIR, which are explained below.

A room impulse response (RIR) is the resulting sound in a room after a very short sound. In classical measurements of the RIR, the sound is recorded after the popping of a balloon. Short impulsive sounds are used, because they excite all frequencies evenly. This is why most acousticians clap their hands when they enter an acoustic space. They are performing a very crude measurement of the room impulse response.

The impulse response of a room depends on the size and shape of the room, the objects inside the room and the materials on its walls. If someone adjusts the position of reflecting panels on stage or if curtains are drawn or closed, all of these things change the RIR.

Even if the wall treatments and other objects do not change, a room impulse response is still dependent upon the source and receiver locations. At an opera, an RIR changes if the singer walks across the stage or if the listener changes seats. Therefore, the room impulse response contains all the sound effects that are determined by the room and the positions of the sound source and listener.

The RIR describes how the sound travels through a room and arrives at the location of a listener. However, this measurement is from a microphone, a physically different situation than if the sound were measured at the listener's eardrums. The presence of a human listener affects the sound waves arriving at the listener's location. For these reasons, a head-related impulse response (HRIR) is needed.

The HRIR gives the relationship between the sound at the eardrum of the listener and the sound as measured by a microphone in place of the listener. The two sounds are different, because the shape of the body, head and ear all affect the incoming sound. Sound coming from the left and arriving at the right ear must travel around the head, resulting in a different sound compared to a direct line of travel, likewise for the sound coming from the right, arriving at the left ear. Therefore, the HRIR contains all the information related to the physical presence of a human listener.

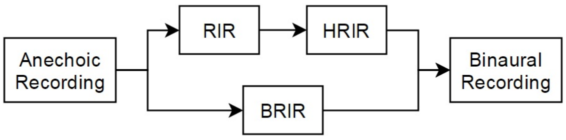

Figure 6: How to make a binaural recording.

The combination of both the RIR and HRIR gives the binaural room impulse response (BRIR), which contains the information of how a sound source on stage travels through the concert hall to the listener's location and arrives at the listener's eardrum. The different impulse responses are summarized in Table 1.

It should be noted that there are several databases of impulse responses available for academic use. Room impulse responses of different rooms are available [12, 13]. Head-related impulse responses can also be found [14, 15, 16].

The anechoic recording and the binaural room impulse response are the two main building blocks of the auralization process. How to combine them into a binaural recording is described next.

2.3. Processing

Adding room and head effects to an anechoic recording consists of a convolution process of the recorded signal with a binaural room impulse response. Convolution is a mathematical process of combining two sequences. Specifically, the sound at a particular time in the first sequence (the anechoic recording) results in a large amount of echoes that are determined by the impulse response. Each instance in time of the anechoic recording gives rise to a similar sequence of echoes. Finally, all the echoes overlap and are added together to get the total effect.

In digital audio workstations (DAW), the convolution process is similar to adding artifical reverb. The reverb in this case is the reflections measured inside a real concert hall. In addition to Reaper, the already mentioned DAW used to make the anechoic recordings, Audacity [17], Cubase [18] and Wavosaur [19] are other popular DAWs.

It is usually necessary to use a reverb or convolution plug-in with a DAW, and there are many available. Two examples are Freeverb (opensource) [20] and SIR2 (license required) [21]. They are provided as Virtual Studio Technology (VST) plug-ins that are compatible with most DAWs. Available from CATT-Acoustic, GratisVolver is available as a free stand-alone convolver [22].

While DAWs are quite user friendly, the convolution algorithm used is not always specified, and computation time-saving measures are often employed, which reduce the accuracy of the convolution process. For these reasons, Octave [23] is used in this work for the convolution process. Octave is an open-source software used primarily for advanced numerical computations. With the execution of only a few commands, an anechoic recording can be read, processed and listened to. The process of obtaining binaural recordings is summarized in Fig. 6.

---

[1] It is recognized that an anechoic environment may itself be a form of feedback that may influence the musician. However, it has been reported that musicians are able to quickly adapt and perform well in this acoustically foreign environment [11].

By convolving the anechoic recordings with binaural room impulse responses, binaural recordings of both musical excerpts are calculated. The anechoic recordings as well as the binaural recordings using one chamber hall and one concert hall are available at `people.mech.kuleuven.be/~u0086891/Data`. Audio files for all four microphones are included.

2.4. Listening and evaluation

If this process of creating binaural recordings were repeated for other tubas, the binaural recordings could be compared. The musician trying out a new musical instrument could listen to Instrument A, playing several contrasting pieces of music in a few different musical environments. Next, Instrument B could be auralized under the same conditions. In this way, the musician could judge the musical characteristics of the instrument independent of several biasing factors, such as mechanical response and knowledge of manufacturing techniques.

Even though the sound has been successfully separated from any non-musical cues, determining the best sound is still quite involved. Therefore, the subjective evaluation is left to the musician purchasing the instrument.

On the other hand, if there are several evaluators, perhaps several representatives of the musical instrument manufacturer, other factors should be taken into consideration. Specifically, attributes have to be clearly defined, and some training is often needed. Returning to the suggestions found in [1], there are many attributes to consider. Intonation is obviously related to the frequency at which the instrument plays relative to accepted standards, but evaluating the tone of an instrument may not be as straightforward. A bright or a dark tone may be preferred, and this may differ for different musical styles. Even more problematic is that people have a range of opinions of what bright means and what dark means.

Once the attributes are well-defined, these can be judged and ranked for the different recordings using standardized blind listening tests like MUSHRA [24] or paired comparisons. The results of the listening tests can give customers and manufacturers more objective information when evaluating the auralized sound of the instruments.

3. DISCUSSION

It is possible to approximate the experience of playing an instrument in a concert hall by using real-time auralization. This technique is essentially identical to the offline convolution described in section 2.3. Instead of using a recording, the microphone input during live performance is directly processed by the convolution plugin, and the output is played through headphones worn by the player. In order to use this technique, headphones should be of high quality, open type and with a rather flat frequency response. The time delay between the instant a sound is played and the time for the sound to be processed by the audio interface, *i.e.*, the system's latency, should be lower than 5 ms. In addition, the early samples of the BRIRs should be trimmed, accounting for the latency of the sound card and the propagation of sound between the instrument and the microphone. Despite the benefits of this technique, it brings back the effects of visual and tactile cues in judging the value of an instrument.

A critical requirement for accurate auralizations is the use of an anechoic room. Since construction of an anechoic room is quite expensive, it is not expected that these techniques can be immediately applied to the evaluation of musical instruments. However, anechoic rooms are standard facilities at acoustics research laboratories. Therefore, auralization of musical instruments can be a point of collaboration between research institutions and instrument makers.

4. CONCLUSIONS

This work applies the basics of auralization techniques to the evaluation of a musical instrument. The main concepts, including anechoic recordings, impulse responses and data processing, are described. Auralization is intended to be of use for musicians and instrument makers at the time of purchasing a new instrument. This method successfully isolates the sound perception from the non-musical cues experienced by a musician when testing an instrument. A musician can record several instruments and then listen to the binaural recordings in order to independently evaluate the playing quality of the instrument.

5. ACKNOWLEDGEMENTS

The research of Matthew Boucher is supported by the European Commission and the Marie Curie Initial Training Networks GRESIMO (No. 290050) and BATWOMAN (No. 605867), the Research Fund KU Leuven and the Fund for Scientific Research - Flanders (F.W.O.). David Pelegrin-Garcia gratefully acknowledges F.W.O.-V for supporting this research by financing his postdoctoral grant No. 1280413N. Alexander Lindau and the Audio Communication Group of TU Berlin are recognized for providing binaural impulse response data.

6. REFERENCES

[1] R. Bobo, *Choosing a tuba*, `http://www.rogerbobo.com/musical_articles/tubashopping.shtml`, 2006.

[2] D. M. Campbell, "Evaluating musical instruments," *Physics Today*, vol. 67, no. 4, pp. 35–40, 2014.

[3] A. Galembo, "Acoustics of extreme ranges of piano, part 2...," *Piano Tech. J.*, vol. 55, pp. 14–17, 2012.

[4] R. Smith, "The effect of material in brass instruments: a review," in *Proceedings of the Institute of Acoustics*, 1986, vol. 8.

[5] Z. Shärer Kalkandjiev and S. Weinzierl, "The influence of room acoustics on solo music performance: an empirical case study," *Acta Acust united Ac*, vol. 99, no. 3, pp. 443–441, 2013.

[6] M. Vörlander, *Auralization*, Springer, 1st edition, 2008.

[7] T. Lokki, "Tasting music like wine: Sensory evaluation of concert halls," *Physics Today*, vol. 67, no. 1, pp. 27–32, 2014.

[8] Reaper, `http://www.reaper.fm/`.

[9] J. Meyer, *Acoustics and the Performance of Music*, Springer, 5th edition, 2009, Translated by Uwe Hansen.

[10] PTB, `http://www.ptb.de/en/org/1/17/173/richtchar.htm`.

[11] J. Pätynen, V. Pulkki, and T. Lokki, "Anechoic recording system for symphony orchestra," *Acta Acust united Ac*, vol. 94, pp. 856–865, 2008.

[12] Aachen Impulse Response Database, `http://www.ind.rwth-aachen.de/en/research/tools-downloads/aachen-impulse-response-database/`.

[13] R. Stewart and M. Sandler, "Database of omnidirectional and b-format impulse responses," in *Proc. of IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Dallas, Texas, 2010.

[14] MIT Sound Media Lab, `http://sound.media.mit.edu/resources/KEMAR.html`.

[15] UC Davis, `http://interface.cipic.ucdavis.edu`.

[16] IRCAM, `http://www.ircam.fr/equipes/salles/listen/download.html`.

[17] Audacity, `http://audacityteam.org/`.

[18] Steinberg Media Technologies GmbH, `http://www.steinberg.net/en/products/cubase/start.html`.

[19] Wavosaur, `www.wavosaur.com`.

[20] Freeverb3, `http://freeverb3vst.osdn.jp/`.

[21] Christian Knufinke, SIR Audio Tools, `http://www.siraudiotools.com/` `http://www.knufinke.de/sir`.

[22] CATT-Acoustic, `http://www.catt.se/`.

[23] S. Hauberg J. W. Eaton, D. Bateman and R. Wehbring, *GNU Octave version 4.0.0 manual: a high-level interactive language for numerical computations*, 2015.

[24] Institute of Sound Recording, `http://iosr.surrey.ac.uk/software/`.