APPLYING ANECHOIC RECORDINGS IN AURALIZATION

Tapio Lokki
Dept. of Media Technology,
Helsinki University of Technology
P.O.Box 5400, 02015 TKK, Finland
Tapio.Lokki@tkk.fi

Jukka Pätynen
Dept. of Media Technology,
Helsinki University of Technology
P.O.Box 5400, 02015 TKK, Finland
Jukka.Patynen@tml.fi

ABSTRACT

Auralization needs anechoic recordings as stimuli. Earlier, a single instrument or speech has been applied in auralization studies, but nowadays the whole symphony orchestra can also be used as stimuli. We have recently published anechoic symphony orchestra recordings for all researchers at http://auralization.tkk.fi. In this paper we discuss about these recordings in more detail and we give advices how to use these recordings. In addition, some indications for the direction of future studies on auralization are presented.

1. INTRODUCTION

Anechoic symphony orchestra recordings are needed when high quality auralization of a concert hall model is performed. In particular, when auralizing the modeling results to public audience, it would be convenient to apply symphonic music. In literature, only a few auralization studies with orchestra music are reported [1, 2, 3, 4]. The reasons for the small number of studies are probably the lack of anechoic recordings and the lack of appropriate modeling software. However, nowadays the suitable recordings are available [5] and commercial software (such as Odeon[1] and CATT-Acoustics[2]) supports multiple sources. So, in near future more studies should be prepared.

The recent recordings [5] were performed in particular for auralizations in mind. The microphone setup as well as music repertoire were chosen so that many different concert hall models could be auralized with reasonable music. The recorded excerpts of music were the following:

- *W. Mozart* (1756-1791), a soprano aria of *Donna Elvira* from the opera *Don Giovanni*, Act II, Scene III
- *L. van Beethoven* (1770-1827), Symphony no. 7, I movement, bars 1-53
- *A. Bruckner* (1824-1896), Symphony no. 8, II movement, bars 1-61
- *G. Mahler* (1860-1911), Symphony no. 1, IV movement, bars 1-85

These four short passages of music represent different eras of classical music and the size of the orchestra varies between them. The Mozart's aria represents classical style and has a vocal soloist. The introduction part of Beethoven's 7th symphony has big chords in the beginning with the whole orchestra, and many slow string crescendos with which the acoustics of a hall, in particular, can be evaluated. On the other hand Bruckner and Mahler are great examples of works which require large orchestras. Musical texture of Bruckner's music is quite conventional while Mahler is really complex music. Both of them are good samples for auralization demonstrations since these works are rather well known.

In this paper we discuss about the mentioned recordings, in particular some practical issues when applying these recordings in auralization. The next section discusses about noise reduction and Sec. 3 overviews the methods used to take into account the radiation pattern of an instrument. In Sec. 4 more detailed practical suggestions are given in addition to a brief overview of the directivities of different instrument groups. Finally, some opinions on the required future work are given before the conclusions.

2. NOISE REDUCTION

The applied microphones in the anechoic recordings were large-diagram condenser microphones with very low self noise (Røde NT1-A with self noise of 5 dB(A)). Thus, the individual sound tracks contain only very little noise regarding that the full dynamic range could not be utilized, since the gains were kept equal for all instruments. However, when all tracks are composed together as the full orchestra, the noise is audible. In other words, level of uncorrelated noise increases with stacked tracks. The situation is comparable to an orchestra recording with dozens of microphones with equal gains adjusted to capture the sound of the cymbals.

A simple and straightforward way to reduce noise is to apply a noise gate[3] to each individual track. A noise gate takes the noise out from the parts where no signal is present. With individual tracks such processing might be audible as a "pumping" of background noise, but with a stack of all tracks, when the gates work at different times, the possible artifacts are not audible. Moreover, the parameters of the noise gate are suggested to be selected in a way that the gate is opened and closed slowly. If possible, the gate should open well in advance before the signal. This can be accomplished with various audio processing software. The recorded tracks can be processed with an automatic gate easily, since the level difference of the quietest instrument sound is well above the noise floor.

3. DIRECTIVITY OF SOUND SOURCES

An essential part of high quality auralization is the use of proper directivities of sound sources. Natural instruments are far from omnidirectional point sources, which are often used in room acous-

[1] http://www.odeon.dk
[2] http://www.catt.se
[3] There are different techniques to implement noise gates, see e.g. http://en.wikipedia.org/wiki/Noise_gate

tics modeling. At least two different approaches for taking into account the directivity have been proposed: directional filtering [6, 7, 8] and so called multi-channel method [9, 4].

In the directional filtering technique the signal recorded from one direction (with one microphone) is processed with filters which are fitted to measured directivity data. This approach can be intricate, since there is only one recording direction and the main radiation direction varies between instruments. However, if the recording direction is the main radiation direction of the instrument, this method is good. Alternatively the sound should be recorded from the same position that is taken as the reference direction for the directivity data. Currently, there is directivity data available [10, 11] in de-facto standard format (namely Common Loudspeaker Format) and this data can be applied when designing directional filters.

Another method, called multi-channel auralization, comprises of representing a sound source in auralization as a point source, but this point emits different anechoic signals to different directions [9]. Thus, anechoic recordings have to be made with multiple microphones simultaneously and in auralization all these recorded signals are applied. In this technique, the directivities of instruments—in which the radiation is a function of the played note—are inherently modeled in this context more accurately than with the directional filtering technique. However, the number of signals to be applied in rendering might be very large. For example, 20 channels per instrument and 50 instrument positions results in 1000 signals to be rendered in the auralization process. However, such a big number of channels are probably not needed, as suggested in [9, 4]. On the other hand, the computational load could be possibly decreased by combining nearby sound sources before applying them in auralization.

Both methods described above have their pros and cons, and more research is needed to find the best possible solution to model sound sources and their directivity characteristics in auralization. The previously accomplished anechoic orchestra recordings provide suitable material for comparative studies.

4. APPLYING ANECHOIC RECORDINGS IN HIGH QUALITY AURALIZATIONS

In this section, some practical advices are given when our symphony orchestra recordings [5] are applied. The discussion of the directivities of instruments is based on [12, 13].

4.1. Woodwinds

Based on informal listening [14] we suggest that in auralization studies only four point sources are needed for woodwinds, one for each instrument group. All the parts are then applied in these points. Naturally, all parts could be applied in separate positions, but with the cost of increased rendering workload.

The overall directivity of all woodwinds can be described as follows. At low frequencies the directivity is more or less omnidirectional. At mid frequencies the instruments radiate sound from a few first open finger holes (and from the mouthpiece in the flute) and the directivity pattern is quite heavily depending on the played note. At high frequencies, almost all sound is emitted from the bell of the instrument. Therefore, the clarinet and the oboe direct the middle frequencies mostly to the forward region and the high frequencies to front and down. The bassoon radiates higher frequencies to up and left. With the flute the directivity is particularly complex, as the mouthpiece must be considered as a separate sound source unlike with the reed instruments. However, at frequencies over 2 kHz the sound is radiated strongly in the direction of the open end. Based on our measurements [13] it seems that multi-channel method should be applied in directivity modeling, as shown by Otondo and Rindel [9]. However, the directional filtering method is not far from the optimal and it could be an option too in full orchestra auralizations. It should also be considered that the excited frequency band with woodwinds does not span over the whole audible bandwidth. Thus, the design of directivity filters can be concentrated on frequencies where there are most radiated sound energy, e.g., between 100–8000 Hz with all woodwinds.

4.2. Brass

Similar approach to combine one instrument group to one point source can be considered for brass. However, if there are more than three instruments in one section, more point sources should be used. This is often the case with the French horns, as large sections of the instruments are characteristic especially for late Romantic compositions such as Mahler's or Bruckner's symphonies.

The directivity pattern of brass instruments is not depending on the played note. The sound is always emitted from the bell, and the higher the frequency, the more directive the instruments are. At low frequencies, when the wavelength is larger than the diameter of the bell, all brass instruments are fairly omnidirectional. Towards higher frequencies the "sound beam" comes narrower being very narrow already at 2 kHz. For instance, the difference of the trombone sound level between the front and above directions can be nearly 20 dB. This behavior makes the brass instruments an ideal object for modeling the directivity with the filtering approach without prominent loss in the accuracy.

4.3. Percussion Instruments

The percussion instruments can be modeled with point sources, although some of them (e.g., gran cassa and timpani) are exceptionally large instruments. In the published recordings [5] all instruments are on separate tracks except timpani where both of them are naturally on the same track.

The directivities of percussion instruments are quite tricky. For example, a tam-tam (i.e. a big gong), has a time-dependent radiation pattern. Some modes are excited or attenuated faster than others and for this reason the directivity changes when the instrument vibrates freely after hitting it. In addition, the cymbals are typically moved apart and rotated after a loud strike in order to provide a longer audible decay to the listeners. These properties suggest that directivity can not be conveniently modeled with time-invariant directional filters, it should rather be modeled with multi-channel technique. The same principle applies to the timpani as well. As they commonly appear by pairs or in larger numbers, the directivity is affected by each others. Therefore the directivity is not very reasonable to model with filters.

4.4. Strings

Strings require the most attention before they can be applied in auralization. Due to the lack of resources only one instrument per part was recorded for each string section. This means that this one recording should be multiplied so that the result sounds like a section instead of a loud violin, for example. While in some pieces we managed to record several takes, there is still not enough non-identical tracks to have a nice section sound without any signal processing.

Much of the familiar sound of strings in a symphony orchestra is created by factors outside the actual instruments. The variations in the synchronization and tune of players combined with a large number of differently positioned sources (thus different reflection paths in the hall) all contribute to the so called chorus effect [12]. Some of the prominent details such as altering the tuning and timing are addressed in the following.

The authors are not aware of published papers citing the statistical values of the collective synchronization and tune of orchestra players. The optimal way to multiply strings has been studied by Lokki [1] and Vigeant et al. [4]. The former study applied the Pitch Synchronous Overlap Add (PSOLA) method to modify the spectrum of one recorded signal. The results suggest that PSOLA algorithm did not make any difference to non-processed copies of one instrument, if a section was modeled with one point source for each musician. Vigeant *et al.* [4] have utilized time shifting, where prime number-valued delays up to 23 ms are applied to string instrument recordings. Such a value has been deducted from a study dating back to 1950's. However, they did not separately study the section sound itself, since the main study concerned different orchestra layouts. Therefore, it is hard to tell how good a section sound can be achieved with this technique.

By ad-hoc listening and comparison to concert hall recordings we suggest a combined pitch and time shift to be applied to the copied tracks in order to achieve "chorus" effect typical for a string section. The following equations can be applied for creating different sets of pitch and time shifts for parallel tracks in a repeatable manner.

$$P = 2p_{max} * \text{normcdf}(x, 0, \sqrt{N}) - p_{max}, \tag{1}$$

$$T = t_{max} * sin(x), \tag{2}$$

where P is the amount of pitch shifts in cents and T is time shift in ms. N denotes the number of required track copies and p_{max} and t_{max} are the maximum mean differences of pitch and time shift, respectively. Each pitch and time shift is determined by the variable x, which receives N equally spaced samples between $[-2\sqrt{N}, 2\sqrt{N}]$. Normcdf() refers a cumulative distribution function (e.g. in Matlab) at point x of a normal distribution with zero mean and variance $\sqrt{N}$.

The result of this formulation is shown in Fig. 1, where the thick lines showing continunous shift amounts are sampled from 6 points each. Pitch shifts are applied with a combination of a phase vocoder and resampling in Matlab. Actual time shifts are selected by quantizing the shifts to the closest prime number in samples.

The largest number of generated copies of instrument tracks has been 9, and the formulas do not function properly for excessively large number of copies. However, in our experimental auralizations this method has been found to provide natural multiplication of string instrument tracks.

4.4.1. Directivity of string instruments

The directivity of the string instruments can be generalized with similar description than the woodwinds. The violin and the viola are very omnidirectional up to around 500 Hz, after which the radiation is directed to the frontal area of the player. The violoncello

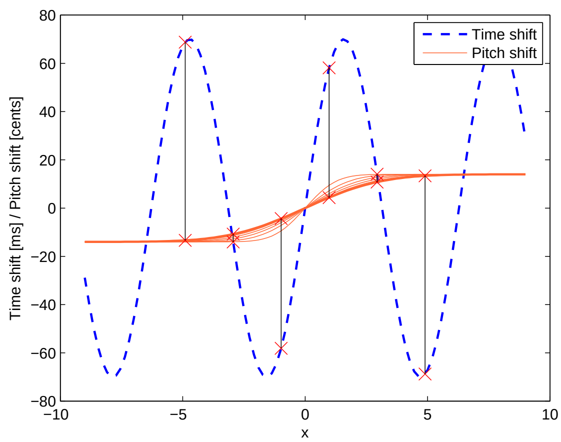

Figure 1: Illustration of the devised functions for modifying string instrument tracks with pitch and time shifts with 6 copies. Series of thin solid lines shows the different pitch shift behavior from $N = 1...6$. Each shift (marked with a cross) combination applied to separate copies is connected with a line. $p_{max} = 70, t_{max} = 14$.

and the contrabass have similar properties, but only at lower frequencies. In overall, the directivity of the strings is very complex due to a large number of excited vibrational modes. The directivity also changes with the played note in the manner of the woodwinds [13]. However, the differences between radiation directions are not nowhere near as strong as for instance with brass instruments.

In a real orchestra there are usually at least 10 player in the first and second violins. From the auralization point of view, the neighboring players can be assumed to be in slightly different positions. Therefore the combined directivity of these positions from a point source is nearer to an omnidirectional pattern. From this it can be possibly concluded that modeling a string section with few sources the importance of correct directivity is somewhat decreased and could be implemented with directivity filters.

4.5. The proposed orchestra layout for auralizations

Based on the discussion presented in this section, we propose a layout of 30 point sources to represent a full symphony orchestra in auralization. The layout is illustrated in Fig. 2 and sound source positions are from 1 to 2 meters distances. Note that source positions are not at equal distances in purpose, since such positioning might cause artifacts, such as comb filters. Based on our previous studies and facts on human capabilities to perceive multiple simultaneous sound streams, we are quite convinced that there is no audible difference between using only 30 point sources and a point source for every single instrument. Unfortunately, we cannot yet prove this with proper listening test results.

5. FUTURE WORK

In this paper, we have tried to give our opinion on how auralization with anechoic signals of a full orchestra should be performed. Some presented solutions are only "educated guesses" which are based on our previous experiences as musicians and as auralization

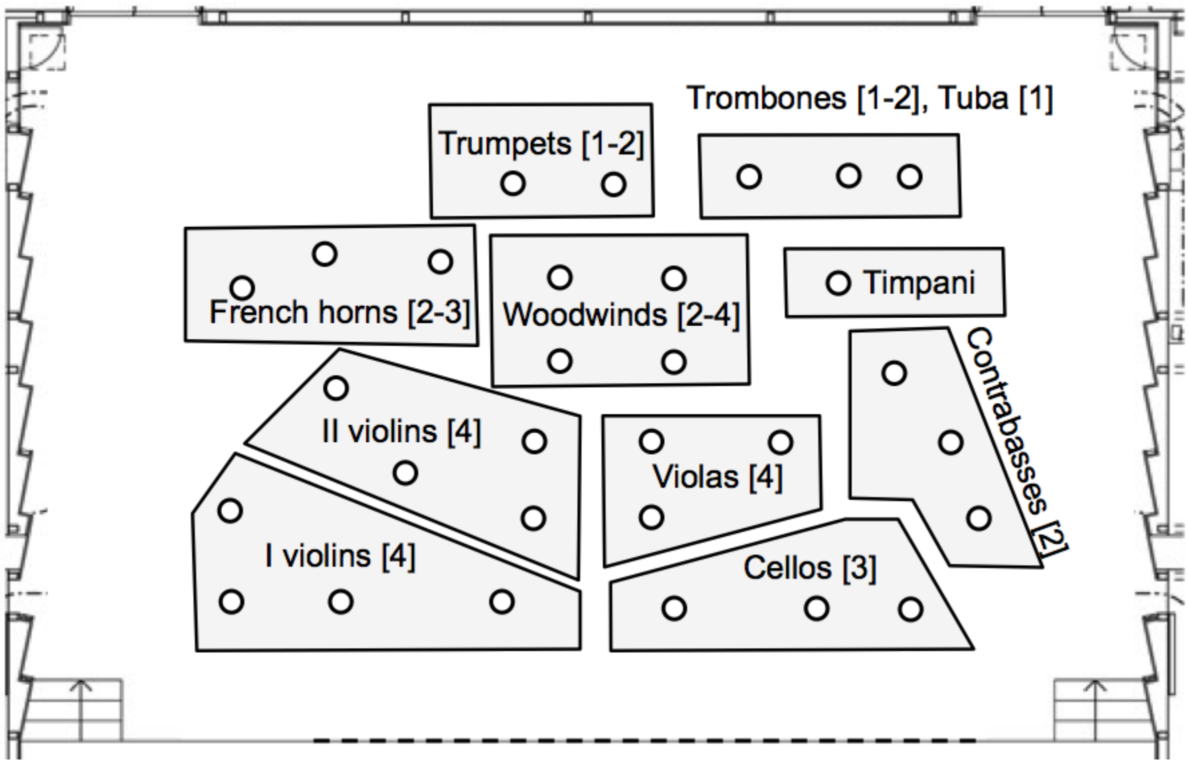

Figure 2: A symphony orchestra representation with point source on the stage of a concert hall. The number in brackets indicate the number of tracks applied in one point source. With such a representation the large orchestra (consisting of about 100 musicians) could be modeled with 30 point sources.

researchers. Naturally, these solutions should be carefully studied to show which of them are correct. Thus, at least the following studies could be done with the anechoic recordings [5] and directivity data [11].

- The audible differences of directional filtering and multichannel methods when the full orchestra is applied in auralization?
- The minimum number of required point sources to represent the full orchestra?
- The optimal way to multiply one recorded violin to create rich section sound?
- Movement of orchestra players, does it have an audible effect?
- Music stands and musicians, should they be modeled also?
- Could an orchestra be modeled with a surface source (as in [4])? What is the anechoic signal for a surface source? The directivity of a surface source?

6. CONCLUSIONS

In this paper we have given suggestions how the anechoic orchestra recordings should be used in auralization. Some practical methods to reduce noise from recordings as well as to take into account the directivities of musical instruments are discussed. Finally, we propose an orchestra layout with 30 point sources to represent a full size symphony orchestra in auralization.

7. ACKNOWLEDGMENTS

The research leading to these results has received funding from the Academy of Finland, project no. [119092] and the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. [203636].

8. REFERENCES

[1] T. Lokki, "How many point sources is needed to represent strings in auralization?," in *the International Symposium on Room Acoustics (ISRA'2007)*, Seville, Spain, September 10-12 2007, Paper P11.

[2] M. C. Vigeant and L. M. Wang, "Investigations of multichannel auralization technique for various orchestra arrangements, with phase-shifted string sections," in *the International Symposium on Room Acoustics (ISRA'2007)*, Seville, Spain, September 10-12 2007, Paper S36.

[3] M. C. Vigeant, L. M. Wang, and J. H. Rindel, "Investigations of multi-channel auralization technique for solo instruments and orchestra," in *the 19th International Congress on Acous-*

tics (ICA'2007), Madrid, Spain, September 2-7 2007, Paper RBA-15-004.

[4] M. C. Vigeant, L. M. Wang, and J. H. Rindel, "Investigations of orchestra auralizations using the multi-channel multi-source auralization technique," *Acta Acustica united with Acustica*, vol. 94, no. 6, pp. 866–882, November/December 2008.

[5] J. Pätynen, V. Pulkki, and T. Lokki, "Anechoic recording system for symphony orchestra," *Acta Acustica united with Acustica*, vol. 94, no. 6, pp. 856–865, November/December 2008.

[6] J. Huopaniemi, M. Karjalainen, V. Välimäki, and T. Huotilainen, "Virtual instruments in virtual rooms - a real-time binaural room simulation environment for physical models of musical instruments," in *Proc. Int. Computer Music Conf. (ICMC'94)*, Aarhus, Denmark, Sept. 1994, pp. 455–462.

[7] M. Karjalainen, J. Huopaniemi, and V. Välimäki, "Direction-dependent physical modeling of musical instruments," in *Proc. 15th Int. Congr. Acoust. (ICA'95)*, Trondheim, Norway, June 1995, pp. 451–454.

[8] L. Savioja, J. Huopaniemi, T. Lokki, and R. Väänänen, "Creating interactive virtual acoustic environments," *J. Audio Eng. Soc.*, vol. 47, no. 9, pp. 675–705, 1999.

[9] F. Otondo and J. H Rindel, "A new method for the radiation representation of musical instruments in auralizations," *Acta Acustica united with Acustica*, vol. 91, no. 5, pp. 902–906, 2005.

[10] Physikalisch-Technische Bundesanstalt, "Directivities of musical instruments," Available at http://www.ptb.de/en/org/1/17/173/richtchar.htm, accessed Jan. 12, 2009.

[11] J. Pätynen, "Directivities of orchestra istruments for auralization," in *EAA Symposium on Auralization*, Espoo, Finland, June 15-17 2009.

[12] J. Meyer, "The sound of the orchestra," *J. Audio Eng. Soc.*, vol. 41, no. 4, pp. 203–213, Apr. 1993.

[13] J. Pätynen and T. Lokki, "Directivities of the symphony orchestra instruments," Unpublished manuscript, 2009.

[14] J. Pätynen, S. Tervo, and T. Lokki, "A loudspeaker orchestra for concert hall studies," in *The Seventh International Conference On Auditorium Acoustics*, Oslo, Norway, October 3-5 2008, Institute of Acoustics, pp. 45–52.