

INTRODUCTION

Perceptible whole-body vibrations, which show a strong correlation with the sound, have been measured in real concert venues [1, 2, 3]. Taking into account the rising threshold of vibration perception towards higher frequencies, these vibrations correspond basically to a low-passed version of the audio signal. If such subtle vibrations are added during the reproduction of music recordings, e.g., by using a vibration seat, the perceived quality of the concert experience increases [4, 5]. In the above-mentioned experiments, a precisely calibrated vibration actuator was applied, capable of reproducing frequencies from 10 Hz to 200 Hz. In practical applications, smaller and cheaper vibration actuators would be beneficial. However, these shakers are usually limited to a small frequency range around a resonance frequency. The question rises if such simple vibration reproduction systems can be used for music reproduction? To answer this question, the following experiment investigates the effect of compressing the vibratory frequency range in the context of music perception.

For a plausible multi-sensory concert experience, it is important that input from all sensory systems is integrated into one unified percept. Therefore, the delay between different sensory inputs is an important factor. Several studies have been concerned with temporal aspects between acoustical and vibrational stimuli [6, 7, 8, 9, 10]. It can be summarized that auditory-tactile asynchrony detection seems to depend on the reproduced signal. Impulsive content is obviously more prone to delay between modalities. Because music often contains transients, the delay between sound and vibration will be aligned to 0 ms in the following. However, for a real-time implementation of audio-generated vibration reproduction, a slight delay seems to be tolerable.

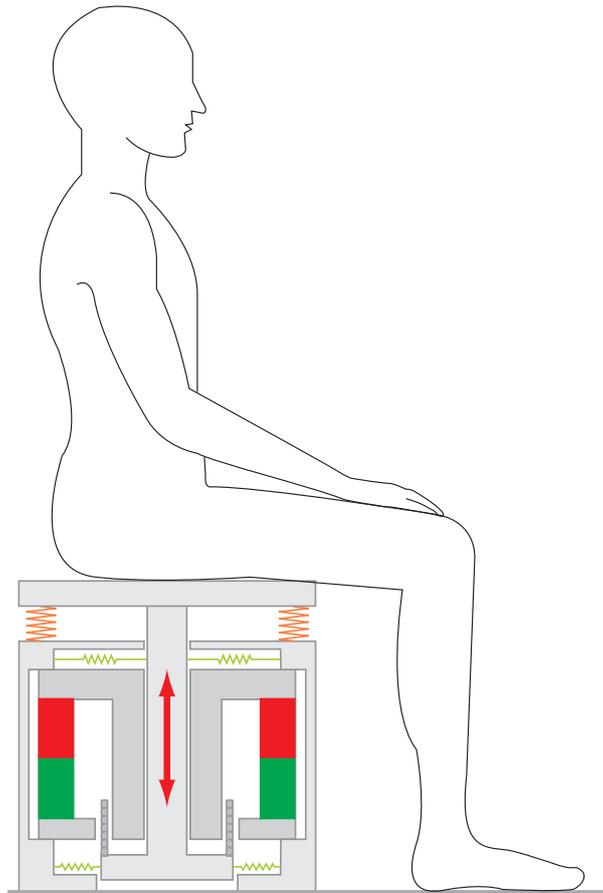


FIGURE 1: Vibration chair with an electro-dynamic exciter.

SETUP

In this study, surround recordings have been played back using a 5.1 loudspeaker setup according to ITU-R BS.775-1 [11]. Additional vibrations have been rendered using a custom vibration chair with an electro-dynamic exciter (RFT Messelektronik Type 11076). Seat vibrations were generated vertically as shown in Figure 1. Subjects were asked to sit on a flat, hard, and wooden seat (46 cm x 46 cm) with both feet flat on the ground. The transfer characteristic of the vibrating chair was strongly dependent on the individual person. This phenomenon is referred to as the body-related transfer function (BRTF) [12]. The BRTF of each subject was individually monitored and equalized during all experiments. The transfer functions were measured using a vibration pad (B&K Type 4515B) and a Sinus Harmonie Quadro measuring board, and compensated using inverse filters in Matlab.

SUBJECTS

20 Subjects participated voluntarily in this experiment (14 male and 6 female). Most of them were students between 20 and 55 years old (mean 24 years) and between 58 and 115 kg (mean 75 kg). All stated to have no known hearing or spine damages. The average number of self reported concert visits per year was 9. Two subjects were members in a band. The preferred music styles were manifold, ranging from rock and pop to classic and jazz.

STIMULI AND EXPERIMENTAL DESIGN

In order to represent typical concert situations for both classical and modern music, four 5.1 surround sequences were selected from music DVDs, which include low frequency content. A stimulus length of approximately 1.5 minutes was chosen in order to ensure that the participants had sufficient time to become familiar with a stimulus before giving their quality judgements. The following sequences were selected:

- Bach, Toccata in D minor (*church organ*)
- Verdi, Messa Da Requiem, Dies Irae (*kettledrum, contrabass*)
- Dvořák, Slavonic Dance No. 2 in E minor, op. 72 (*contrabass*)
- Blue Man Group (BMG), The Complex, Sing Along (*electric bass, percussion, kick drum*)

The first sequence, Toccata in D minor, is a well known organ work, which will be further simply referred to as 'BACH'. An exemplary spectrogram of the first 60 s is plotted in Figure 2. Rising and falling succession of notes, covering a broad frequency range, can be seen. Additionally, steady state tones with a rich overtone spectrum dominate the composition. Strong seat vibrations would be expected in a real church for this piece of music [2]. The sequence from 'DVORAK' is a calm orchestral piece, dominated by bowed and plugged strings. Contrabasses and cellos continuously generate low frequencies, however, with low level. In the 'VERDI' composition, impulsive fortissimo parts with a concert bass drum, a kettledrum and tutti orchestra alternate rapidly with parts which are dominated by the choir, bowed instruments and brass winds. The sequence is dominated by strong transients. The fourth sequence is a typical pop music example. It is performed by the Blue Man Group, which will be further abbreviated as 'BMG'. The sequence is characterized by the heavy use of drums and percussions.

The concert recordings were played back to each participant using the setup described above. In order to generate a vibration signal from these sequences, the sum of the low-frequency effects

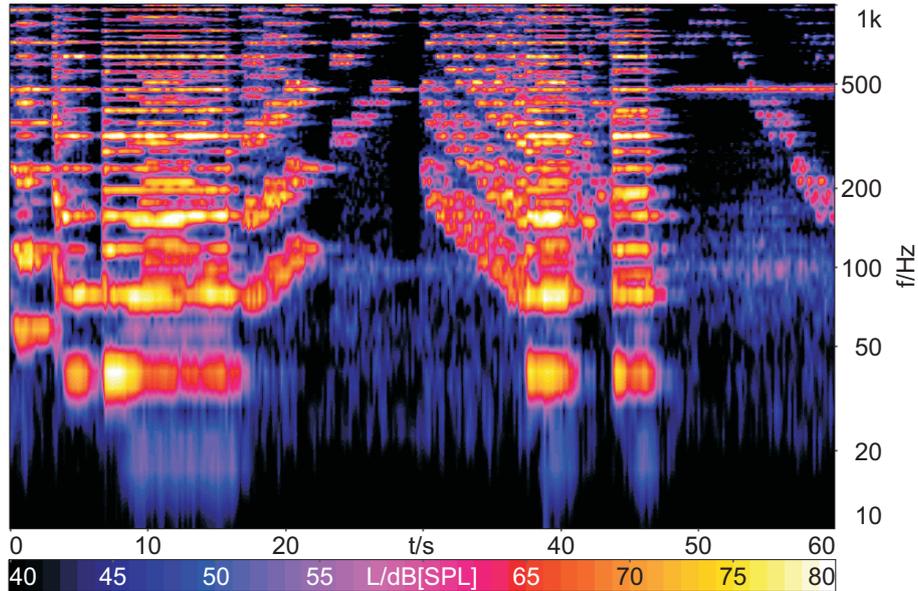


FIGURE 2: Spectrogram of the mono sums for 60 s from the sequence BACH. Fast Fourier transforms were calculated with 8192 samples using 50% overlapping Hanning windows.

(LFE) channel and the three respective frontal channels (left, right and center) was calculated. There was no low-frequency content in any of the surround channels. Figure 3 shows the corresponding signal processing chain. Sinusoidal tones at 20 Hz, 40 Hz, 80 Hz and 160 Hz were generated using Pure Data. The frequencies were selected to span a broad frequency range and to be clearly distinguishable, taking into account the tactile frequency resolution [13]. These simple signals have been further multiplied with the envelope of the low-passed audio signal. An envelope follower was implemented, which calculated the RMS amplitude of the input signal using successive analysis windows. Hanning windowing was applied and the window size was set to 1024 samples, which corresponded approximately to 21 ms, in order to avoid smearing of impulsive signal content. The period for successive analysis was half the window size.

Additionally, the low passed audio signal was reproduced directly via the vibration seat as reference condition.

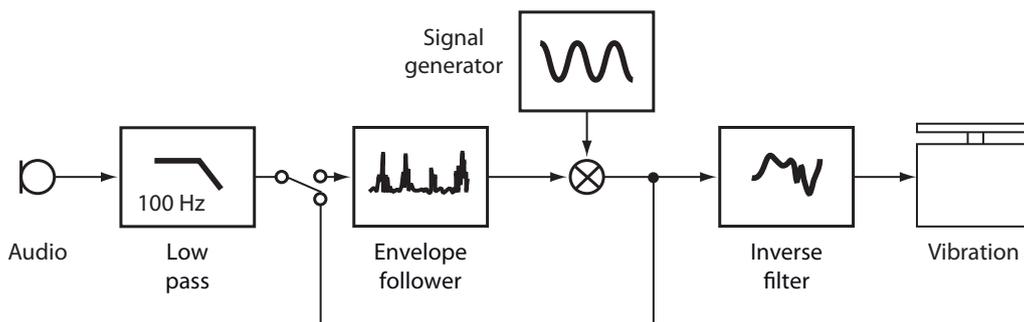


FIGURE 3: Signal processing to generate vibration signals from audio sum. The envelope of the low pass filtered audio signal was extracted and multiplied with sinusoids at 20 Hz, 40 Hz, 80 Hz and 160 Hz. Alternatively, the low passed signal was routed directly to the inverse filter and the vibration actuator.

The vibration intensities were initially adjusted, so that the peak acceleration levels reached 100 dB and were thus clearly perceptible. However, if such a reproduction system would be implemented at home, the vibration level could be varied easily. Additionally, the perception threshold can vary heavily between subjects [14]. Therefore, each subject was asked to adjust the vibration

amplitude individually to the preferred level. This was usually achieved within the first 5 s to 10 s of a sequence. Subsequent, the subject had to judge the *overall quality of the concert experience* using a quasi continuous scale. Verbal anchor points from bad to excellent have been added similar to the method described in ITU-T P.800 [15]. Figure 4 shows the used rating scale.



FIGURE 4: Rating scale for evaluation of the overall quality of the concert experience.

In order to prevent dissatisfaction, the subject could interrupt the current stimuli, as soon as he/she was confident with his/her judgment. The required time varied between subjects from 30 s to usually no more than 60 s.

RESULTS AND DISCUSSION

For statistical analysis the individual quality ratings were interpreted as numbers on a linear scale from 0 to 100, with 0 corresponding to ‘bad’ and 100 to ‘excellent’. Data was checked for sufficient normal distribution with the Kolmogorov-Smirnov test (KS-test). A two-factorial analysis of variance (ANOVA) with repeated measures was carried out using IBM SPSS Statistics, which also checks for homogeneity of variances. The two factors were the played music *sequence* and the applied *treatment*. Averaged results for the overall quality evaluation are plotted in Figure 5 with mean and 95% confidence intervals. The quality ratings for the concert reproduction without vibration are shown on the left for comparison.

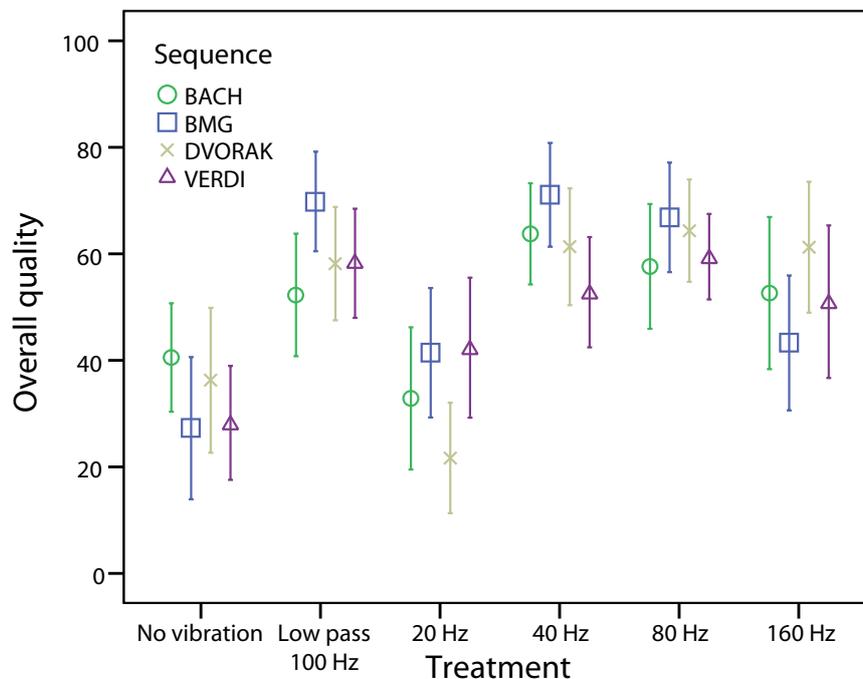


FIGURE 5: Mean overall quality evaluation for reproduction using different vibration generation approaches. For comparison, the ratings for the condition without vibration are plotted on the left.

The plot shows that the overall quality of the concert experience increase if low passed vibrations were added ($p < 1\%$). This confirms results from earlier studies [4, 5]. Poor quality ratings were achieved with the 20 Hz condition. No significant difference was found between the 20 Hz vibration and the no-vibration condition. Subjects indicated that the 20 Hz vibration was too low in frequency and did not fit with the audio content. In contrast, 40 Hz and 80 Hz seemed to fit well. They were judged better than reproduction without vibration on a very significant level ($p < 1\%$). No complaints about a mismatch between sound and vibration were noted. The resulting overall quality was judged comparable to the low-pass conditions. Average differences to the no-vibration condition were between 29 for the 40 Hz vibration and 18 for WGN and the 160 Hz vibration. Interestingly, even the 160 Hz vibration resulted in fair quality ratings. However, a trend towards worse judgements compared to the 80 Hz condition can be seen ($p \approx 11\%$). A much stronger effect was expected, because this vibration frequency is relatively high and tingling effects may occur. There was some disagreement between subjects, which can be seen from slightly larger confidence intervals for this condition.

SUMMARY

The results of this study show that concert reproduction with vibration was judged better than reproduction without vibration in most cases. As expected, the low-pass condition resulted in good quality ratings. However, even strong compression of the frequency range can result in good reproduction quality. This enables the use of small and cheap vibration actuators in the context of music reproduction. However, some simple signal processing is necessary in order to extract the envelope of the original vibration signal. For the tested sequences, amplitude modulated sinusoids at 40 Hz and 80 Hz worked well.

REFERENCES

- [1] S. Merchel and M. E. Altinsoy, "Der Konzertsaal bebt - Vibroakustische Messungen in der Dresdner Semperoper", in *Proceedings of DAGA 2012 - 38th German Annual Conference on Acoustics* (Darmstadt, Germany) (2012).
- [2] S. Merchel and M. E. Altinsoy, "Music-Induced Vibrations in a Concert Hall and a Church", To appear in *Archives of Acoustics* 1 (2013).
- [3] C. L. Abercrombie and J. Braasch, "Perceptual Dimensions of Stage-Floor Vibration Experienced During a Musical Performance", in *Proceedings of Audio Engineering Society Convention 129* (2010).
- [4] S. Merchel and M. E. Altinsoy, "5.1 oder 5.2 Surround - Ist Surround taktil erweiterbar?", in *Proceedings of DAGA 2008 - 34th German Annual Conference on Acoustics* (Dresden, Germany) (2008).
- [5] S. Merchel and M. E. Altinsoy, "Vibratory and Acoustical Factors in Multimodal Reproduction of Concert DVDs", in *Haptic and Audio Interaction Design* (Springer) (2009).
- [6] M. E. Altinsoy, "Perceptual aspects of auditory-tactile asynchrony", in *Proceedings of the Tenth International Congress on Sound and Vibration* (Stockholm, Sweden) (2003).
- [7] M. Daub, "Audiotactile simultaneity perception of musical-produced whole-body vibrations", in *Proceedings of CFA/DAGA* (2004).
- [8] W. L. . Martens and W. Woszczyk, "Perceived synchrony in a bimodal display: Optimal intermodal delay for coordinated auditory and haptic reproduction", in *Proceedings of ICAD* (Sydney, Australia) (2004).

- [9] M. E. Altinsoy, *Auditory-Tactile Interaction in Virtual Environments* (PhD Thesis, Shaker Verlag) (2006).
- [10] K. Walker, W. L. Martens, and S. Kim, "Perception of Simultaneity and Detection of Asynchrony between Audio and Structural Vibration in Multimodal Music Reproduction", in *Proceedings of Audio Engineering Society Convention 120* (Paris, France) (2006).
- [11] ITU-R BS.775-1 International Telecommunication Union, "Multichannel stereophonic sound system with and without accompanying picture", (1992).
- [12] M. E. Altinsoy and S. Merchel, "BRTF - Body Related Transfer Functions for Whole-Body Vibration Reproduction Systems", in *Proceedings of NAG/DAGA* (Rotterdam, The Netherlands) (2009).
- [13] S. Merchel, M. E. Altinsoy, and M. Stamm, "Just-Noticeable Frequency Differences for Whole-Body Vibrations", in *Proceedings of Internoise* (Osaka, Japan) (2011).
- [14] S. Merchel, A. Leppin, and M. E. Altinsoy, "Hearing with your Body: The Influence of Whole-Body Vibrations on Loudness Perception", in *Proceedings of ICSV - 16th International Congress on Sound and Vibration* (Kraków, Poland) (2009).
- [15] ITU-T P.800 International Telecommunication Union, "Methods for Objective and Subjective Assessment of Quality", (1996).