Vibratory and Acoustical Factors in Multimodal Reproduction of Concert DVDs

Sebastian Merchel and M. Ercan Altinsoy

Chair of Communication Acoustics, Dresden University of Technology, Germany
sebastian.merchel@tu-dresden.de

Abstract. Sound and vibration perception are always coupled in live music experience. Just think of a rock concert or hearing (and feeling) a church organ sitting on a wooden pew. Even in classical concerts kettle-drum and double bass are sensed not only with our ears. The air-borne sound causes seat vibrations or excites the skin surface directly. For some instruments (e.g. an organ) structure-borne sound is transmitted directly from the instrument to the listener.

If concert recordings are played back with multimedia hi-fi systems at home, these vibratory information is missing in the majority of cases. This is due to low reproduction levels or to the limited frequency range of conventional loudspeakers. The audio signal on todays DVDs contains an additional channel for low frequency effects (LFE), which is intended for reproduction using a subwoofer. The generation of tactile components is still very restricted. An enhancement of such a systems might be possible using an electrodynamical shaker which generates whole body vibration (WBV) for a seated person.

This paper describes a system implementing this approach. The generation of a vibrotactile signal from the existing audio channels is analyzed. Different parameters during this process (amplitude of the vibration, frequency range) are examined in relation to their perceptual consequences using psychophysical experiments.

Keywords: Multimodal Music Reproduction, Whole Body Vibration, Audiotactile Concert Perception.

1 Introduction

Measurements in real concert situations confirm the existence of whole body vibrations. If a bass drum is hit or the double bass plays a tone the perceived vibrations are noticeable. Nevertheless, in most cases the concert visitor will not recognize the vibrations as a separate event. The vibrotactile percept is integrated with the other senses (e.g. vision and hearing) to one multi-modal event.

Experiencing a concert, the listener expects vibrations, even if he is not aware of it all the time. These expected vibrations are missing in a traditional multimedia reproduction setup. According to Jekosch [1] the perceived quality of an entity (e.g. a reproduction system) results from the judgment of the perceived characteristics of an entity in comparison to its desired/expected characteristics.

M.E. Altinsoy, U. Jekosch, and S. Brewster (Eds.): HAID 2009, LNCS 5763, pp. 119–127, 2009.

If in the reproduction situation the vibratory component is missing, there might be a loss of perceived quality, naturalness or presence of the concert experience. To look at it from the other side: The perceived quality of a conventional reproduction system might be improved by adding whole body vibrations. This study focuses on whole body vibrations for a seated person, like they are perceived in a classical chamber concert.

Unfortunately there is no vibration channel in conventional multimedia productions. Therefore it would be advantageous if the vibration signal could be generated using the information stored in the existing audio channels. This might be reasonable, since the correlation between sound and vibration is naturally high in everyday situations. The questions in focus of this study are:

1. Is it possible to generate a vibrotactile signal using the existing audio channels of a conventional 5.1 surround recording?
2. Up to which frequency should the WBV be reproduced?
3. Is there an ideal amplitude for the reproduction of WBV?

Previous studies are primarily concerned with the perception of synchrony between acoustical and vibratorical stimuli ([2], [3], [4], [5],[6]). Walker et al. [7] investigates the tactile perception during reproduction of action oriented DVD movies. In Merchel et al. [8] a pilot experiment is described, which aims at investigating the influence of WBV on the overall quality of the reproduction of concert DVDs. This paper describes an extended experiment with 20 subjects. The focus is on different vibration parameters like amplitude and frequency range of the vibration.

2 Experiment

2.1 Stimuli

The stimuli should include instruments for which low frequency vibrations and sounds are expected. Another criteria was that the stimuli represent typical concert situations for both classical and modern music. To place the subject in a standard multimedia reproduction context, an accompanying picture from the DVD can be projected. The video sequence shows the stage, the conductor or the individual instrumentalists while playing. The participant in the experiment should have enough time to become familiar with the stimulus. Thus a stimulus length of 1.5 minutes was chosen. The following sequences were selected:

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

The sum of the three frontal channels and the LFE channel was used to generate the vibration signal (see Figure 1). Two low pass frequencies were implemented

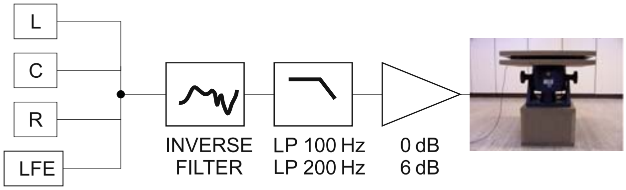

Fig. 1. Generation of the vibration signal using four out of six DVD audio channels. The transfer characteristics of the vibration chair was compensated and the signal was filtered with a variable low pass as well as variable amplification.

($f1 = 100\,\text{Hz}$, $f2 = 200\,\text{Hz}$), which seemed suitable for the chosen stimuli. The low pass filter was a steep Butterworth filter with 10th order. Two test persons adjusted the amplitude of the vibration, until it was just noticeable for all frequencies. This amplitude is further referred to as $a1 = 0\,\text{dB}$. To study the influence of the vibration amplitude, an additional signal with an $a2 = 6\,\text{dB}$ amplified vibration was generated. The peak value of the vibration at the surface of the seat in vertical direction was measured using an Endevco triaxial seat pad accelerometer. The peak vibration was measured between 0.25 and $0.60\,\text{m/s}^2$, depending on the reproduced sequence.

2.2 Setup

Figure 2 shows the used setup for reproduction of surround recordings according to ITU [9]. It was build in front of a silver screen for video projection. Five Genelec 8040A loudspeakers and a Genelec 7060B subwoofer were used. In addition vertical whole body vibrations have been reproduced using a self build electrodynamical vibration seat. The transfer characteristic of the shaker loaded with a seated person has been measured using an Endevco triaxial vibration pad. This frequency response depending on the individual test person is called the Body Related Transfer Function (BRTF) [10]. All stimuli have been compensated for the transfer characteristic of the seat in vertical direction by using inverse filters in MATLAB.

2.3 Crosstalk

There was also crosstalk between the different systems. The subwoofer excited some seat vibrations. They were measured to be below $0.5\,\text{mm/s}^2$. The peak vibrations reproduced with the shaker reached from 250 to $600\,\text{mm/s}^2$. This is factor 1000 above the sound induced vibrations. Thus it was concluded, that crosstalk is uncritical.

The vibration seat itself radiated some sound. It was measured at the hearing position to be 40 dB below the signal reproduced by the loudspeakers. The loudspeaker generated sound pressure level was approximately 68 dB(A).

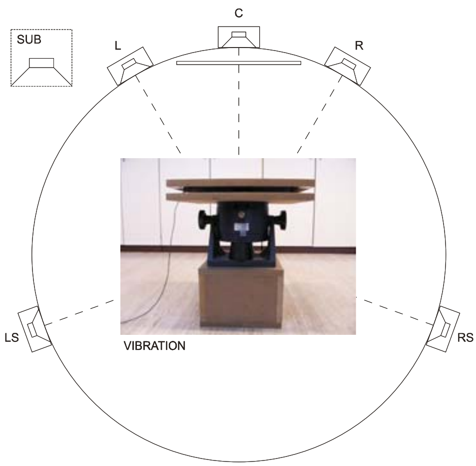

Fig. 2. Setup with six loudspeakers according to ITU [9]. An additional shaker was used to reproduce vertical whole body vibration.

2.4 Subjects

20 Subjects participated voluntarily in this experiment (15 male and 15 female). Most of them were students between 20 and 46 years old (mean 23 years) and between 48 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 18 and ranged from 1 to 100. The highest number was reported from one subject, who played guitar in a band. The preferred music styles were rock music (15 subjects), classic (9 subjects), pop (6 subjects), jazz (3 subjects) and 12 subjects preferred other genres in addition.

2.5 Experimental Design

Each subject had to judge 20 stimuli, five for each music sequence. All versions of one sequence were played on after the other, always starting with the no vibration condition. The remaining four combinations of low pass frequency ($f1 = 100\,\text{Hz}$, $f2 = 200\,\text{Hz}$) and amplitude ($a1 = 0\,\text{dB}$, $a2 = 6\,\text{dB}$) had been randomized between subjects using a balanced latin square, a williams square. The final presentation order is illustrated in Table 1.

Before starting with the experiment the subjects had to do a training with two stimuli to get familiar with the task and the stimuli range. The used stimuli was the first 1.5 minutes from Bizet - Carmen (Prelude), which is a classical composition with kettledrum and contrabass. After or while listening to the concert reproduction, 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 [11]. Figure 3 shows the used questionnaire. In addition presence and naturalness had to be evaluated by means of a five point Rohrmann scale [12].

Table 1. Order of presentation of all factor combinations to the first four subjects. The first stimuli in one sequence block is always without vibration. The other four combinations are randomized using a williams square. The presentation order of the music sequences are also randomized between subjects using a williams square.

Sequence	**Bach**					**Verdi**					**BMG**					**Dvorak**				
Vibration Amplitude	**no**	**a1**		**a2**		**no**	**a1**		**a2**		**no**	**a1**		**a2**		**no**	**a1**		**a2**	
Vibration Lowpass Freq.	**no**	**f1**	**f2**	**f1**	**f2**	**no**	**f1**	**f2**	**f1**	**f2**	**no**	**f1**	**f2**	**f1**	**f2**	**no**	**f1**	**f2**	**f1**	**f2**
Subject 1	1	2	3	5	4	6	8	9	7	10	16	19	20	18	17	11	15	12	14	13
Subject 2	11	13	14	12	15	1	4	5	3	2	6	10	7	9	8	16	17	18	20	19
Subject 3	16	19	20	18	17	11	15	12	14	13	1	2	3	5	4	6	8	9	7	10
Subject 4	6	10	7	9	8	16	17	18	20	19	11	13	14	12	15	1	4	5	3	2
...																				

Fig. 3. Questionnaire to evaluate the overall quality of the concert experience

3 Results and Discussion

For statistical analysis the evaluation values were interpreted as numbers on a linear scale from 1 to 5. Data was checked for normal distribution with the KS-test. A multifactorial analysis of variance was carried out. The averaged results for the overall quality evaluation are plotted in Figure 4 with mean and 95% confidence intervals for all 20 stimuli. It can be seen that the influence of the different vibration parameters (low pass frequency $f1 = 100\,\text{Hz}$, $f2 = 200\,\text{Hz}$ and amplitude $a1 = 0\,\text{dB}$, $a2 = 6\,\text{dB}$) on the overall quality judgement is relative small.

Although reproduction with vibration is judged better than reproduction without vibration in most cases. This is illustrated in Figure 5. The quality evaluations for all vibration versions have been averaged for each music sequence. A t-test for paired samples showed very significant differences on a 1% significance level for all music sequences, except Dvořák. This music composition is very calm with gentle contrabass. Whole body vibrations might not be expected for this kind of stimuli. For all other sequences the vibration reproduction improved the perceived quality of the concert experience significantly. There was no significant influence of preferred music style on the evaluation of the reproduction.

There was no significant overall preference for a specific vibration amplitude or low pass frequency.

The interaction between the factors *low pass frequency* and *sequence* was significant on a 5% significance level. An interaction diagramm is plotted in Figure 6. It can be seen that for the sequences from Bach, Verdi and Dvořák it is preffered to reduce the frequency range for whole body vibrations to 100 Hz. If frequencies

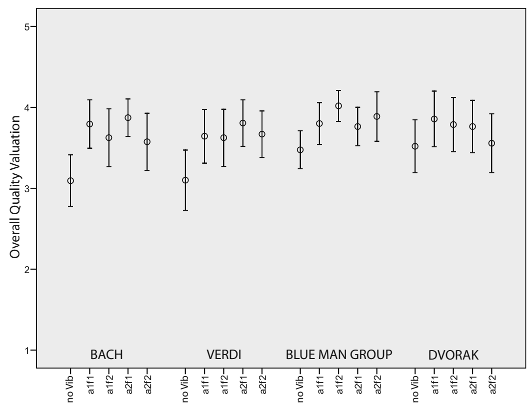

Fig. 4. Mean overall quality evaluation of all 20 participants with 95% confidence intervals

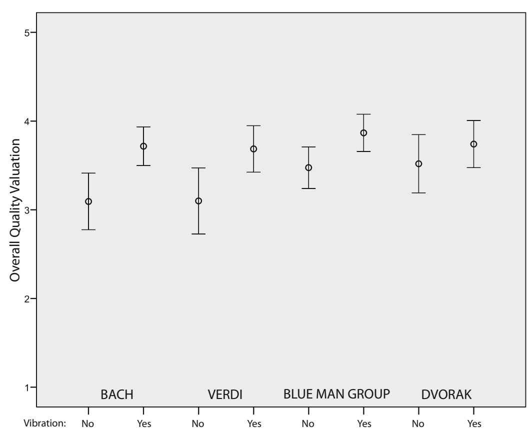

Fig. 5. Comparison of overall quality evaluation for reproduction with and without accompanying vibration plotted with 95% confidence intervals. It can be seen that reproduction with vibration ist judged better.

up to 200 Hz are reproduced the judgement is slightly worse. This might be due to a prickling sensation that is induced through higher frequency content. In addition no strong whole body vibrations might be expected in the range from 100 to 200 Hz for organ, kettledrum or contrabass. Contrary for the Blue Man Group sequence a 200 Hz low pass is favored.

Figure 7 helps to understand this result. The figure shows spectrograms which plot the frequency content (mono sum of L, C, R and LFE channel) over time for all four music sequences. For the church organ in Bachs Toccata in D minor (top left) the keynote and overtones are apparent. Note that the keynote remains below 100 Hz. In the Dvořák (bottom left) and Verdi (bottom right) sequence

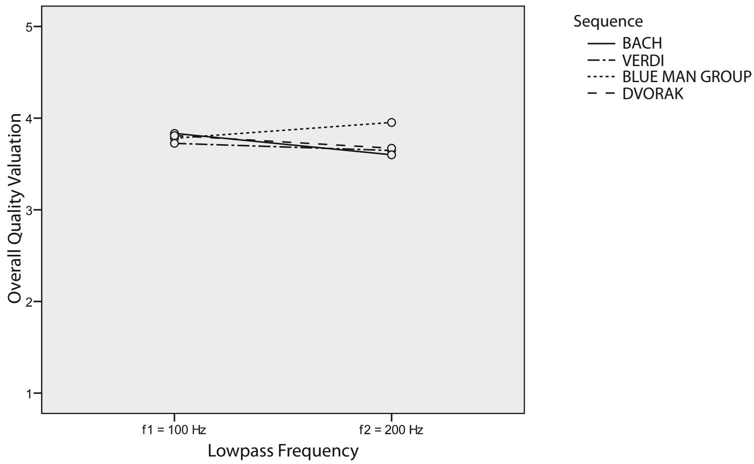

Fig. 6. Diagram showing the significant interaction between low pass frequency and sequence. The lower low pass frequency ($f1 = 100\,\text{Hz}$) is judged better in all cases except for the sequence Blue Man Group (BMG).

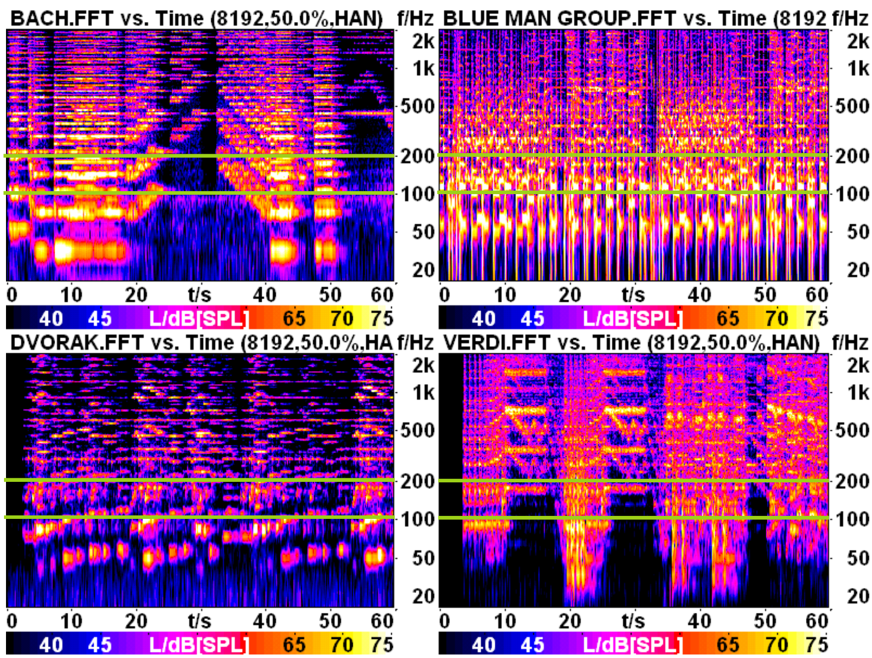

Fig. 7. Spectrograms of the mono sums for all four music sequences. The low pass frequencies (100 Hz and 200 Hz) are plotted with solid lines.

the keynotes remain as well below 100 Hz. In contrast in the Blue Man Group sequence (top right) the kick drum and the alternating keynote pattern of the bass guitar can be seen up to 200 Hz . If this signal is now low pass filtered with 100 Hz the frequency content of the electric bass keynote pattern is clipped. Thus the alternating bass guitar pattern is only partly reproduces as whole body vibration. This might be the reason for slightly worse quality judgments of this stimuli.

It can be concluded that the ideal low pass frequency is dependent on the particular music sequence.

The vibration amplitude had no influence on the overall quality judgments of the music sequences. This is surprising, since the just noticeable difference for vibration amplitude is approximately 1.5 dB [13]. The used amplitude difference of 6 dB was clearly noticeable. Still no preference for a specific amplitude was found. In an additional experiment, the subjects were asked to adjust the vibration amplitude for the same music sequences to an optimal level. A markerless infinite rotary knob (PowerMate, Griffin Technology) was used to avoid any visual cues. The results varied between subjects with a standard deviation of approx. 6 dB. This indicates a broad range of preferred vibration amplitudes for musically induced whole body vibrations. If one subject had to adjust the vibration amplitude for one stimuli repeatedly, the within subject standard deviation was approx. 4 dB. This again indicates no strong preference for a specific amplitude.

4 Summary

This paper investigates a reproduction method for whole body vibrations, which are generated from audio recordings.

- The perceived overall quality of concert DVD reproduction can be improved by adding vertical whole body vibrations.
- The vibration signals for the selected sequences could be generated by low pass filtering the audio sum signal.
- The ideal low pass frequency depends on the specific music content.
- The preference for a specific whole body vibration amplitude varies.

5 Outlook

This study uses a broad band calibrated whole body vibration reproduction system. The influence of an applicable reproduction solution in real life situations needs to be investigated. Further research is necessary to evaluate the general perception of whole body vibrations. The generation of the vibration signal using an audio recording is promising. However more complex processing than low pass filtering seems necessary. There are different approaches which are investigated at the moment.

References

1. Jekosch: Voice and speech quality perception - Assessment and Evaluation. Springer, Berlin (2005)
2. Altinsoy, Blauert, Treier: Intermodal effects of non-simultaneous stimulus presentation. In: Proceedings of the 7th International Congress on Acoustics, Rome, Italy (2001)
3. Daub, Altinsoy: Audiotactile simultaneity of musical-produced whole-body vibrations. In: Proc. of the Joint Congress CFA/DAGA, Strasbourg, France (2004)
4. Martens, Woszczyk: Perceived Synchrony in a Bimodal Display: Optimal Delay for Coordinated Auditory and Haptic Reproduction. In: ICAD, Sydney, Australia (2004)
5. Walker, Martens, Kim: Perception of Simultaneity and Detection of Asynchrony Between Audio and Structural Vibration in Multimodal Music Reproduction. In: AES 120th Convention, Paris, France (2006)
6. Altinsoy: Auditory-Tactile Interaction in Virtual Environments. Shaker Verlag, Aachen (2006)
7. Walker, K., Martens, W.L.: Perception of audio-generated and custom motion programs in multimedia display of action-oriented DVD films. In: McGookin, D., Brewster, S. (eds.) HAID 2006. LNCS, vol. 4129, pp. 1–11. Springer, Heidelberg (2006)
8. Merchel, Altinsoy: 5.1 oder 5.2 Surround - Ist Surround taktil erweiterbar? In: DAGA, Dresden, Germany (2008)
9. ITU-R BS.775-1: Multichannel stereophonic sound system with and without accompanying picture (1992)
10. Altinsoy, Merchel: BRTF - Body Related Transfer Functions for Whole-Body Vibration Reproduction Systems. In: DAGA, Rotterdam, Netherlands (2009)
11. ITU-T P.800: Methods for objective and subjective assessment of quality (1996)
12. Rohrmann: Empirische Studien zur Entwicklung von Antwortskalen für die sozialwissenschaftliche Forschung. Zeitschrift für Sozialpsychologie 9, 222–245 (1978)
13. Bellman, Remmers, Mellert: Basic Experiments on the Perception of Vertical Whole-Body Vibrations. VDI-Tagung Humanschwingungen, Darmstadt (2004)