

Der Einfluss von Vibrationen auf die Konzertwahrnehmung

Sebastian Merchel, M. Ercan Altinsoy

Lehrstuhl für Kommunikationsakustik, TU Dresden, Deutschland, Email: sebastian.merchel@tu-dresden.de

Abstrakt

Die Wahrnehmung von Musik ist kein ausschließlich auditives Phänomen. So können zum Beispiel visuelle, aber auch vibratorische Aspekte in einem Konzert eine entscheidende Rolle spielen. Dieser Beitrag untersucht den Einfluss von Vibrationen auf die Qualität eines Konzerterlebnisses. Deutlich spürbare Schwingungen der Körperoberfläche treten bei Rock- oder Orgelkonzerten auf. Die Körperoberfläche wird dabei direkt durch den Luftschall oder indirekt durch Körperschall zum Mitschwingen angeregt. Auch während eines Sinfoniekonzertes im klassischen Konzertsaal sind überschwellige Vibrationen nachweisbar. Diese können einen Einfluss auf wahrnehmbare Merkmale des Konzerterlebnisses, wie die Präsenz oder Umhüllung, haben. Das Fehlen der Vibrationen, zum Beispiel bei Wiedergabe einer Konzertaufnahme zu Hause, kann die wahrgenommene Konzertqualität beeinflussen. Um diesen Einfluss zu untersuchen, wurden Schall und Sitzvibrationen im Wahrnehmungsexperiment separat kontrolliert. Es zeigt sich, dass Vibrationen für die Qualität eines Konzerterlebnisses von Bedeutung sind. Dabei ist es möglich diese mit Hilfe verschiedener Ansätze aus dem Audiosignal zu generieren.

Einleitung

Misst man während eines Konzertes die Schwingungen in Konzertsälen und Kirchen, so liegen diese teils deutlich über der Wahrnehmungsschwelle [1, 2, 3]. Die Schwingungen an Sitzoberflächen oder Fußboden korrelieren dabei meist stark mit dem Schall am Hörort. Unter Berücksichtigung der ansteigenden Fühlschwelle für die Wahrnehmung von Vibrationen bei höheren Frequenzen, entsprechen diese Schwingungen in erster Näherung dem tiefpassgefilterten Audiosignal. Fügt man nun solche leichten Vibrationen während der Wiedergabe von Musikaufnahmen hinzu, z.B. unter Verwendung eines Vibrationssitzes, dann steigt die wahrgenommene Qualität des Konzerterlebnisses [4, 5]. Um solche Experimente durchzuführen, wurden individuell kalibrierte Schwingungserzeuger verwendet, welche Frequenzen von 10 Hz bis 200 Hz wiedergeben konnten. Für eine praktische Anwendung wären kleinere und billigere Shaker wünschenswert. Diese arbeiten jedoch meist nur in einem eingeschränkten Frequenzbereich um eine Resonanzfrequenz. Es stellt sich die Frage ob solche einfacheren Vibrationssysteme bei der Wiedergabe von Musik verwendet werden können. Um diese Frage zu beantworten, soll in dieser Studie der Einfluss der Kompression des vibratorischen Frequenzbereichs auf die Konzertqualität untersucht werden.

Damit das multisensorische Konzerterlebnis als plausibel beurteilt wird, ist es wichtig dass die Informationen der unterschiedlichen Wahrnehmungsmodalitäten zu einem einheitlichen Perzept integriert werden. Dabei spielt die Verzögerung zwischen akustischem und vibratorischem Stimuli eine bedeutende Rolle. Mehrere Studien haben sich mit solchen zeitlichen Aspekten auseinandergesetzt [6, 7, 8, 9, 10]. Altinsoy [9] ermittelte beispielsweise die maximale Verzögerung von Ganzkörperschwingungen gegenüber einem Audiosignal, sodass diese noch als Synchron wahrgenommen wurden. Für die von ihm verwendeten Signale erhielt er tolerierbare Verzögerungszeiten zwischen 47 ms und 58 ms. Zusammenfassend lässt sich sagen, dass die Erkennung audio-taktiler Asynchronität vom wiedergegebenen Signal abzuhängen scheint. Verzögerungen zwischen den Modalitäten werden für impulsive Signale offensichtlich am besten erkannt. Da Musik oft transiente Komponenten enthält, werden im Folgenden Schall und Vibrationen synchron dargeboten. Dennoch scheint eine leichte Verzögerung tolerierbar zu sein, was von Interesse ist wenn Vibrationen in Echtzeit aus einem Audiosignal erzeugt werden sollen.

Setup

In dieser Studie wurden Surround Aufnahmen über ein konventionelles 5.1 Lautsprecher-setup nach ITU-R BS.775-1 [11] wiedergegeben. Zusätzlich wurden mit Hilfe eines angepassten Schwingstuhles (elektrodynamischer Erreger basierend auf Shaker RFT Messelektronik Type 11076) Vibrationen erzeugt. Wie in Abbildung 1 dargestellt wurde eine flache Holzoberfläche (46 cm x 46 cm) in vertikaler Richtung angeregt. Die Höhe des Stuhles wurde angepasst, so dass die Oberschenkel der Probanden auf der Sitzoberfläche auflagen und beide Füße flach auf dem Boden standen. Die Übertragungsfunktion eines solchen Vibrationssitzes hängt dabei stark von der Versuchsperson ab. Sie wird körperbezogene Übertragungsfunktion oder body-related transfer function (BRTF) genannt [12]. Die BRTF wurde im Folgenden für jede Versuchsperson individuell gemessen und mittels inverser Filter in Matlab kompensiert. Die Messung erfolgte mit Hilfe eines Sitzkissens mit Beschleunigungsaufnehmern (B&K Type 4515B) und der Messkarte Sinus Harmonie Quadro.

Versuchspersonen

An diesem Experiment nahmen 20 überwiegend studentische Versuchspersonen teil (14 männlich und 6 weiblich). Das Alter der Probanden lag zwischen 20 und 55 Jahren (Mittelwert 24 Jahre) und 58 und 115 kg (Mittelwert 75 kg). Alle gaben an keine bekannten Gehör- oder Rückenschäden zu haben. Nach eigenen Angaben besucht

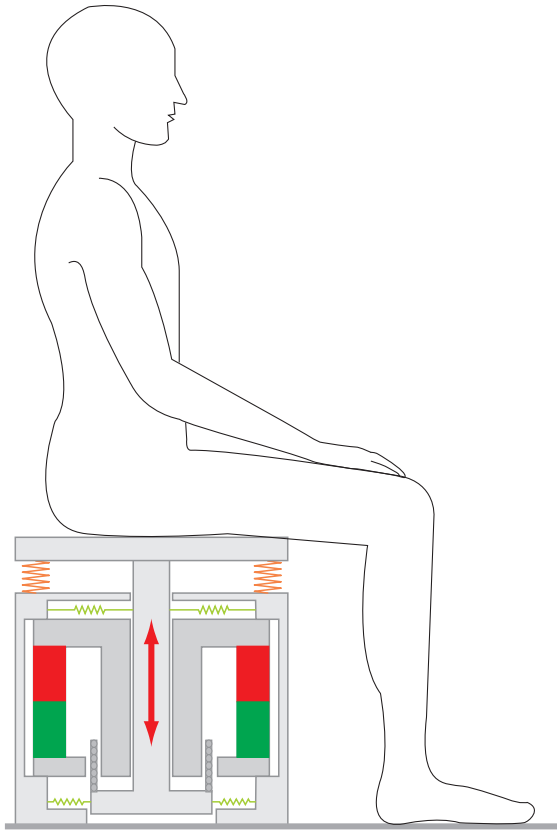


Abbildung 1: Vibrationssitz mit elektro-dynamischer Anregung.

jeder Proband im Mittel 9 Konzerte pro Jahr. Zwei Versuchspersonen spielten Instrumente in einer Band. Die bevorzugten Musikrichtungen waren vielfältig, von Rock und Pop bis zu Klassik und Jazz.

Stimuli und Versuchsdesign

Vier Konzertmischungen wurden ausgewählt, welche sowohl klassische als auch moderne Musik repräsentieren. Die Stimuluslänge wurde mit 1.5 Minuten relativ lang gewählt, um sicherzustellen, dass die Versuchspersonen genug Zeit zur Verfügung hatten, um mit der Sequenz vertraut zu werden. Die folgenden Sequenzen wurden ausgewählt:

- Bach, Toccata in d-Moll (*Orgel*)
- Verdi, Messa Da Requiem, Dies Irae (*Pauke, Kontrabass*)
- Dvořák, Slawischer Tanz Nr. 2 in e-Moll, op. 72 (*Kontrabass*)
- Blue Man Group (BMG), The Complex, Sing Along (*E-Bass, Trommeln*)

Die erste Sequenz, Toccata in d-Moll, ist ein bekanntes Stück für Orgel, welches im Folgenden verkürzt als BACH bezeichnet werden soll. Abbildung 2 zeigt exemplarisch das Spektrogramm der ersten 60s. Gut zu erkennen sind ansteigende und abfallende Tonleitern, welche ein breites Frequenzspektrum überdecken. Zusätzlich dominieren langanhaltende Akkorde mit einem reichen Obertonspektrum die Komposition. Für

dieses Musikstück werden starke Sitzvibrationen in einer echten Kirche erwartet [2]. Die Sequenz DVORAK ist ein ruhiges Orchesterstück, welches durch gestrichene und gezupfte Streicher dominiert wird. Kontrabässe und Cello erzeugen kontinuierlich tiefe Frequenzanteile, allerdings mit geringem Pegel. In der Komposition von VERDI wechseln impulsive fortissimo Passagen mit einer großen Konzerttrommel, einer Pauke und vollbesetztem Orchester mit Passagen, welche durch Chor, Streicher und Blechbläser dominiert werden. Die Sequenz ist durch starke Transienten gekennzeichnet. Der vierte Musikausschnitt ist ein typisches Popmusikbeispiel. Es ist eine Sequenz der Blue Man Group Show, welche im folgenden als BMG abgekürzt wird. Der Stimulus ist gekennzeichnet durch die starke Verwendung von Trommeln und Perkussionsinstrumenten.

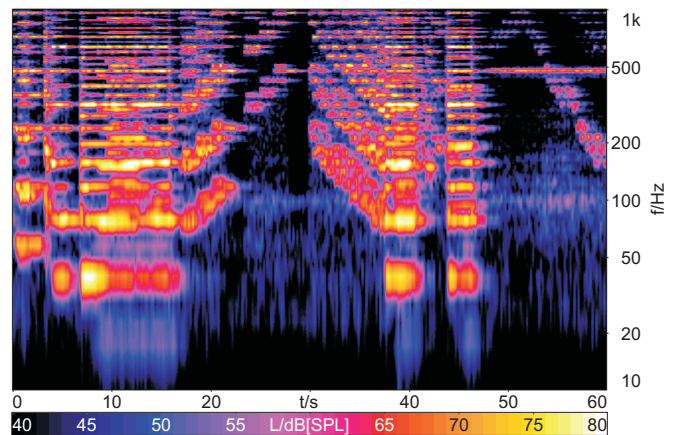


Abbildung 2: Spektrogramm der ersten 60s für die Sequenz BACH (Monosumme).

Die Signalverarbeitungskette zur Generierung von Vibrationen aus dem Audiosignal ist zur Veranschaulichung in Abbildung 3 dargestellt. Zuerst wurde die Summe der vorderen drei Kanäle und des LFE-Kanals (low-frequency effects) gebildet. Die Surroundkanäle enthielten in bei keinem Musikstück tieffrequenten Signalanteile. Anschließend wurde dieses Signal tiefpassgefiltert und dessen Einhüllende bestimmt, indem der Effektivwert für aufeinanderfolgende Analysefenster berechnet wurde. Als Fensterfunktion kamen kurze Hanning-Fenster mit einer Fenstergröße von 1024 Abtastwerten (ca. 21 ms) zum Einsatz, um impulsive Anteile im Signal nicht zu verwischen. Aufeinanderfolgende Fenster überlappten dabei jeweils um eine halbe Fensterlänge. Die resultierende Einhüllende konnte nun mit Sinustönen unterschiedlicher Frequenz multipliziert werden. Die Auswahl der Ausgangsfrequenzen deckte dabei ein breites Frequenzspektrum ab (20 Hz, 40 Hz, 80 Hz und 160 Hz). Zusätzlich wurde als Referenz das ausschließlich tiefpassgefilterte Signal wiedergegeben.

Die Vibrationsintensität wurde zu Beginn so eingestellt, dass die Spitzen des Beschleunigungspegels Werte von 100 dB erreichten. Damit lagen diese deutlich über der Wahrnehmungsschwelle für Ganzkörperschwingungen. Allerdings lässt sich der Pegel eines solchen Systems, z.B. in einer Heimkinoanwendung, leicht ändern. Außerdem können sich die individuellen Wahrnehmungsschwellen

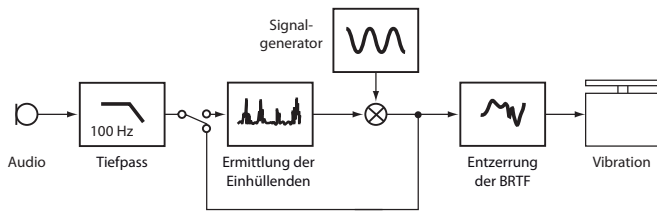


Abbildung 3: Signalverarbeitungskette für die Erzeugung von Vibrationen aus der Summe des Audiosignals. Die Einhüllende der Monosumme wurde ermittelt und mit verschiedenen Sinustönen (20 Hz, 40 Hz, 80 Hz und 160 Hz) multipliziert. Alternativ wurde das tiefpassgefilterte Audiosignal ohne weitere Bearbeitung für die Wiedergabe entzerrt und über den Vibrationssitz ausgegeben.

deutlich unterscheiden [15]. Daher wurde jeder Proband aufgefordert die Vibrationsintensität auf einen angenehmen Wert anzupassen. Dieser Vorgang dauerte meist nicht länger als 5 s bis 10 s zu Beginn einer Sequenz. Anschließend sollten die Versuchsteilnehmer die Gesamtqualität des Konzerterlebnisses (overall quality of the concert experience) auf einer quasikontinuierlichen Skala von mangelhaft (bad) bis ausgezeichnet (excellent) bewerten. Die verwendete Bewertungsskala ist in Anlehnung an ITU-T P.800 [16] in Abbildung 4 dargestellt.

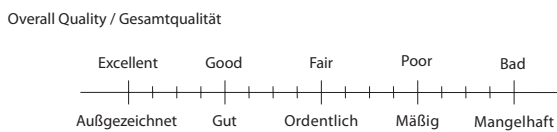


Abbildung 4: Bewertungsskala für die Gesamtqualität des Konzerterlebnisses.

Um Unzufriedenheit vorzubeugen, konnten die Probanden den aktuellen Stimuli abbrechen sobald sie mit ihrem Urteil zufrieden waren. Die dafür benötigte Zeit variierte von Versuchspersonen zu Versuchsperson zwischen 30 s und üblicherweise nicht mehr als 60 s.

Ergebnisse und Diskussion

Für die statistische Auswertung wurden die Qualitätsurteile als Zahlen auf einer linearen Skala von 0 bis 100 interpretiert. Dabei entspricht 0 dem Attribut "schlecht" und 100 dem Attribut "ausgezeichnet". Für die statistische Auswertung wurde IBM SPSS Statistics verwendet. Die Daten wurden mit Hilfe des Kolmogorov-Smirnov-Test (KS-Test) auf hinreichende Normalverteilung überprüft. Anschließend konnte die zweifaktorielle Varianzanalyse (ANOVA) mit Messwiederholungen angewendet werden. Die zwei Faktoren waren die Musiksequenz (Sequenz) und die angewendete Signalverarbeitungsmethode zur Generierung der Vibrationen (Methode). Die Mittelwerte und 95% Konfidenzintervalle der Gesamtqualitätsurteile sind in Abbildung 5 zu sehen. Die Qualitätsurteile für Konzertwiedergabe ohne Vibration sind auf der linken Seite als Referenz abgebildet.

Es ist deutlich erkennbar, dass die Qualität des Konzerterlebnisses ohne Vibrationen durch die Wiedergabe

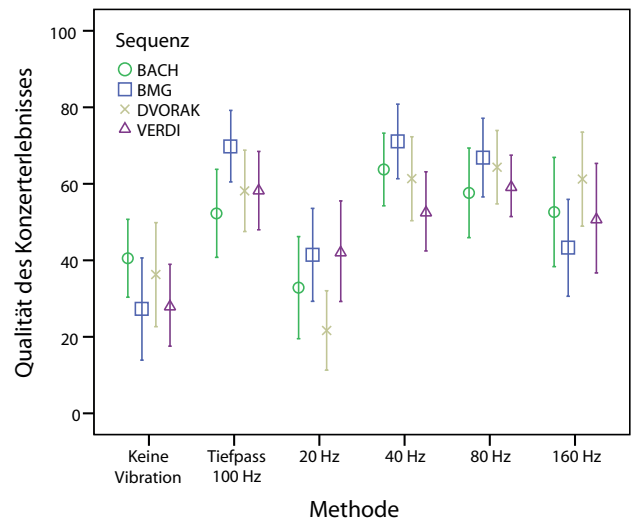


Abbildung 5: Gemittelte Gesamtqualitätsurteile des Konzerterlebnisses und 95% Konfidenzintervalle für unterschiedliche Methoden der Vibrationserzeugung. Links sind zum Vergleich Urteile für konventionelle Wiedergabe ohne Vibrationen dargestellt.

tiefpassgefilterter Vibrationen (100 Hz) verbessert werden kann ($p < 1\%$). Dies bestätigt Ergebnisse aus vorangegangenen Experimenten [4, 5]. Bei Verwendung der Einhüllenden muss zwischen unterschiedlichen Vibrationsfrequenzen unterschieden werden. Schlechte Beurteilungen erhielt das 20 Hz Anregungssignal (kein signifikanter Unterschied zur Kondition ohne Vibrationen). Die Versuchspersonen bemängelten das die Schwingungsfrequenz zu niedrig sei und nicht zum Audiosignal passe. Es ist nicht überraschend, dass unpassende Vibrationen eine Verschlechterung des Konzerterlebnisses hervorrufen können. Im Gegensatz dazu scheinen die Vibrationsgrundsignale mit 40 Hz und 80 Hz passender zu sein. Sie wurden signifikant besser bewertet verglichen mit der Kondition ohne Vibrationen ($p < 1\%$). Das resultierende Qualitätsurteil ist in beiden Fällen vergleichbar mit der Bewertung der breitbandigen Tiefpass-100-Hz-Kondition. Der mittlere Unterschied des Qualitätsurteils für den 40 Hz Ansatz zur Kondition ohne Vibrationen betrug 29 Skaleneinheiten. Interessanterweise wurde auch der 160 Hz Ansatz gut bewertet. Im Vergleich mit der 80 Hz Kondition ist jedoch ein Trend zu schlechterer Beurteilung sichtbar ($p \approx 11\%$). Für die relativ hohe Vibrationsfrequenz von 160 Hz wurde allerdings ein wesentlich schlechteres Ergebnis erwartet, da sie kribbelnd und unangenehm wahrgenommen werden kann. Die etwas größeren Konfidenzintervalle für diese Kondition zeigen, dass die Versuchspersonen in diesem Punkt uneinig sind.

Zusammenfassung

Die Ergebnisse dieser Studie haben gezeigt dass die Konzertqualität bei gleichzeitiger Wiedergabe von Vibrationen in vielen Fällen besser bewertet wird als ohne Vibrationen. Wie erwartet liefert das tiefpassgefilterte Audiosignal für die Vibrationsgenerierung gute Ergebnisse. Allerdings konnte der Frequenzbereich des

Vibrationssignals zusätzlich stark komprimiert werden ohne das die wahrgenommene Qualität des Konzerterlebnisses abnahm. Unter Anwendung der entsprechenden Signalverarbeitung ermöglicht dies die Verwendung von preisgünstigen schmalbandigen Vibrationserregern. Für die untersuchten Musikstücke erwiesen sich amplitudenmodulierte Sinustöne mit 40 Hz und 80 Hz als sinnvoll.

Literatur

- [1] S. Merchel and M. E. Altinsoy, “Der Konzertsaal bebte - 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] I. R. Summers, *Tactile Aids for the Hearing Impaired* (Whurr, London) (1992).
- [14] S. Merchel, M. E. Altinsoy, and M. Stamm, “Just-Noticeable Frequency Differences for Whole-Body Vibrations”, in *Proceedings of Internoise* (Osaka, Japan) (2011).
- [15] 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).
- [16] ITU-T P.800 International Telecommunication Union, “Methods for Objective and Subjective Assessment of Quality”, (1996).