Technische Universität Dresden

Perceptual Model-Driven Authoring of Plausible Vibrations from User Expectations for Virtual Environments

Dipl.-Ing.

Robert Rosenkranz

der Fakultät Elektrotechnik und Informationstechnik der Technischen Universität Dresden

zur Erlangung des akademischen Grades

Doktoringenieur
(Dr.-Ing.)

genehmigte Dissertation

Vorsitzender:	Prof. Dr.-Ing. Andreas Richter
Gutachter:	Prof. Dr.-Ing. habil. Ercan Altinsoy
Gutachter:	Prof. Seungmoon Choi, Ph.D.
Tag der Einreichung:	02.11.2020
Tag der Verteidigung:	12.10.2022

Preface

The presented thesis was created during my work at Dresden University of Technology, Faculty of Electrical and Computer Engineering, Chair of Acoustic and Haptic Engineering. This work was supported by German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) AL1473/5-1 and as a part of Germany's Excellence Strategy – EXC 2050/1 – Project ID 390696704 – Cluster of Excellence "Centre for Tactile Internet with Human-in-the-Loop" (CeTI) of TU Dresden.
First and foremost i want to greatly thank my scientific supervisor Prof. Ercan Altinsoy for inspiring my research on the plausibility illusion and for supporting me with stimulating discussions. I want to also thank Prof. Seungmoon Choi for kindly agreeing to be the second reviewer of my thesis. Furthermore, i would also like to thank the collegues of the Chair of Acoustic and Haptic Engineering for further in-depth discussions, especially Sebastian Merchel and Jürgen Landgraf. Moreover, i want to thank the students Daniel Schecker, Sebastian Gruschwitz, Martin Wilberg and Sebastian Metzner supervised by me for their support in creating parts of the reproduction setup and stimuli utilized in this work.
Likewise, i am grateful to my student assistants Leonie Tribukait and Lukas Haack for their support in conducting some or the perceptual studies as well as Fabian Koller for implementing some of the software functionalities required for stimulus creation.
Last but not least, i want to thank my parents for their unconditional support and for proofreading my publications. Finally, i want to thank my wife, my children, and the rest of the family for their support and understanding.

Contents

Abstract

One of the central goals of design is the creation of experiences that are rated favorably in the intended application context. User expectations play an integral role in tactile product quality and tactile plausibility judgments alike. In the vibrotactile authoring process for virtual environments, vibration is created to match the user's expectations of the presented situational context. Currently, inefficient trial and error approaches attempt to match expectations implicitly. A more efficient, model-driven procedure based explicitly on tactile user expectations would thus be beneficial for authoring vibrations. In everyday life, we are frequently exposed to various whole-body vibrations. Depending on their temporal and spectral properties we intuitively associate specific perceptual properties such as "tingling". This suggests a systematic relationship between physical parameters and perceptual properties. To communicate with potential users about such elicited or expected tactile properties, a standardized design language is proposed. It contains a set of sensory tactile perceptual attributes, which are sufficient to characterize the perceptual space of vibration encountered in everyday life. This design language enables the assessment of quantitative tactile perceptual specifications by laypersons that are elicited in situational contexts such as auditory-visual-tactile vehicle scenes. However, such specifications can also be assessed by providing only verbal descriptions of the content of these scenes. Quasi identical ratings observed for both presentation modes suggest that tactile user expectations can be quantified even before any vibration is presented. Such expected perceptual specifications are the prerequisite for a subsequent translation into physical vibration parameters. Plausibility can be understood as a similarity judgment between elicited features and expected features. Thus, plausible vibration can be synthesized by maximizing the similarity of the elicited perceptual properties to the expected perceptual properties. Based on the observed relationships between vibration parameters and sensory tactile perceptual attributes, a 1-nearest-neighbor model and a regression model were built. The plausibility of the vibrations synthesized by these models in the context of virtual auditory-visual-tactile vehicle scenes was validated in a perceptual study. The results demonstrated that the perceptual specifications obtained with the design language are sufficient to synthesize vibrations,

which are perceived as equally plausible as recorded vibrations in a given situational context. Overall, the demonstrated design method can be a new, more efficient tool for designers authoring vibrations for virtual environments or creating tactile feedback. The method enables further automation of the design process and thus potential time and cost reductions.

Zusammenfassung

Eines der zentralen Ziele des Designs von Produkten oder virtuellen Umgebungen ist die Schaffung von Erfahrungen, die im beabsichtigten Anwendungskontext die Erwartungen der Benutzer erfüllen. Gegenwärtig versucht man im vibrotaktilen Authoring-Prozess mit ineffizienten Trial-and-Error-Verfahren, die Erwartungen an den dargestellten, virtuellen Situationskontext implizit zu erfüllen. Ein effizienteres, modellgetriebenes Verfahren, das explizit auf den taktilen Benutzererwartungen basiert, wäre daher von Vorteil. Im Alltag sind wir häufig verschiedenen Ganzkörperschwingungen ausgesetzt. Abhängig von ihren zeitlichen und spektralen Eigenschaften assoziieren wir intuitiv bestimmte Wahrnehmungsmerkmale wie z.B. “kribbeln”. Dies legt eine systematische Beziehung zwischen physikalischen Parametern und Wahrnehmungsmerkmalen nahe. Um mit potentiellen Nutzern über hervorgerufene oder erwartete taktile Eigenschaften zu kommunizieren, wird eine standardisierte Designsprache vorgeschlagen. Sie enthält eine Menge von sensorisch-taktilen Wahrnehmungsmerkmalen, die hinreichend den Wahrnehmungsraum der im Alltag auftretenden Vibrationen charakterisieren. Diese Entwurfssprache ermöglicht die quantitative Beurteilung taktiler Wahrnehmungsmerkmale, die in Situationskontexten wie z.B. auditiv-visuell-taktilen Fahrzeugszenen hervorgerufen werden. Solche Wahrnehmungsspezifikationen können jedoch auch bewertet werden, indem der Inhalt dieser Szenen verbal beschrieben wird. Quasi identische Bewertungen für beide Präsentationsmodi deuten darauf hin, dass die taktilen Benutzererwartungen quantifiziert werden können, noch bevor eine Vibration präsentiert wird. Die erwarteten Wahrnehmungsspezifikationen sind die Voraussetzung für eine anschließende Übersetzung in physikalische Schwingungsparameter. Plausible Vibrationen können synthetisiert werden, indem die erwarteten Wahrnehmungsmerkmale hervorgerufen werden. Auf der Grundlage der beobachteten Beziehungen zwischen Schwingungsparametern und sensorisch-taktilen Wahrnehmungsmerkmalen wurden ein 1-Nearest-Neighbor-Modell und ein Regressionsmodell erstellt. Die Plausibilität der von diesen Modellen synthetisierten Schwingungen im Kontext virtueller, auditorisch-visuell-taktiler Fahrzeugszenen wurde in einer Wahrnehmungsstudie validiert. Die Ergeb-

nisse zeigten, dass die mit der Designsprache gewonnenen Wahrnehmungsspezifikationen ausreichen, um Schwingungen zu synthetisieren, die in einem gegebenen Situationskontext als ebenso plausibel empfunden werden wie aufgezeichnete Schwingungen. Die demonstrierte Entwurfsmethode stellt ein neues, effizienteres Werkzeug für Designer dar, die Schwingungen für virtuelle Umgebungen erstellen oder taktiles Feedback für Produkte erzeugen.

Table of Contents

List of Abbreviations

1-NN	1-nearest-neighbor
AM	amplitude modulated
ANOVA	analysis of variance
FIR	finite impulse response (filter)
ICP	integrated circuit piezo
IIR	infinite impulse response (filter)
JNDF	just noticeable difference in frequency
JNDL	just noticeable difference in level
k-NN	k-nearest-neighbor
LFE	low frequency effects/enhancement, audio channel containing low frequency signals
MDS	multidimensional scaling
NP	non Pacinian (channel)
PCA	principal component analysis
RA	rapidly adapting (mechanoreceptor)
RMS	root mean square
SA	slowly adapting (mechanoreceptor)
SL	sensation level
STFT	short-time Fourier transform
SVM	support vector machine
WBV	whole-body vibration
WGN	white Gaussian noise

1 Introduction

1.1 General Introduction

One of the central goals of design is the creation of experiences that are judged favorably in the intended application context [1]. Tactile design strategies are applied in many fields. In vibrotactile product design, the intrinsic device vibration or additional feedback vibration is being perceptually optimized until the user expectations are sufficiently fulfilled. For example, in the vehicle industry, the noise vibration harshness departments are occupied with controlling perceptual properties of whole-body vibration (WBV) intrinsically produced by various excitations required for the cars' functionality. The goal is to meet user expectations on the vibration quality regarding comfort or sportiveness. Vibrotactile feedback is frequently utilized in vehicles to convey feedback to the user, e. g. for driving assistance systems. The design of such feedback should match the expected tactile properties for the application context enabling preferably intuitive communication of the desired information about the vehicle's state. Vibrotactile design is also utilized for other devices such as handhelds.

In the vibrotactile authoring process for virtual reality like applications, vibration is created to match the expectations on the situational context presented. In the last decade, there has been remarkable progress in reproduction technologies for virtual reality. Besides high-resolution head-mounted display, also tracking technologies or tactile reproduction devices as well as performant computing systems for rendering the virtual environment have been seeing increasing adoption. The technological progress has led to broader utilization of virtual reality in many fields [2]. Virtual environments offer many advantages. A user can be immersed safely into an arbitrary environment, whose properties can be precisely controlled. These properties made virtual reality ideal for the field of research and development, which was among its first users. This includes virtual reality applications with WBV i.e. vibration presented via a seat. Virtual reality enables simulators for flight training or vehicle driving without exposing the user to real danger. Furthermore, they are being utilized in e-commerce [3] e. g. in the form of virtual test drives as offered by AUDI car brand or for online shopping enabling tactile interaction with products such as clothing. Virtual environments are being utilized in the health sector in the form of virtual therapy [4] as for exposure therapy treatment of phobias. In entertainment virtual environments that included vibrotactile elements were first limited to amusement parks. Now, vibration reproduction is increasingly applied to cinemas for example by DBox Technologies, providing a more immersive experience due to the added tactile modality [5]. Affordable head-mounted displays brought virtual environments even to the homes of many consumers sometimes accompanied by vibration reproduction.

With tactile virtual reality technology becoming ubiquitous, creating tactile virtual environments that users interact with like real environments comes into focus. Mel Slater [6] suggests that two illusions are contributing to this goal: the place illusion and the plausibility illusion. The place illusion refers to the users' sense of presence in the virtual environment and is influenced by the degree to which the technical capabilities as display resolution or latency enable immersion. The plausibility illusion is created by the content of the depicted environment matching the user expectations. While the place illusion received much interest from researchers, the plausibility illusion has attracted much less attention.

Obviously, user expectations play an integral role in product quality and plausibility judgments alike. However, their role is often implicit, e. g. when users check design proposals or prototypes for agreement with their expectations without explicitly stating their expectations first. Thus, content needs to be created by anticipating the vibration expected from the user. Attempting to transfer vibration from a real environment to virtual environments would require the creation of enormous databases of each situation. This is inflexible, inefficient, and only possible for virtual environments with a real counterpart. Producing vibration with complex models of the physical excitation process is more flexible, but requires modeling of each process. Both methods assume accurate playback of the vibration signals potentially requiring cost-intensive reproduction systems.

The main shortcoming of these approaches is the assumption that user expectations resemble the vibration occurring in the specific real environment, despite users potentially never having experienced it. It is often assumed that sound and vibration are physically coupled and thus also perceptually coupled. However, simply generating WBV from the LFE audio channel of a DVD is perceived as much less realistic than manually authoring a suitable vibration signal [5]. When presenting a basketball game in a virtual environment, [7] vibration for each ball impact was presented to the audience despite it not being present in the auditorium. The experience was judged more realistic than without the vibration. The results from [8] indicate that the recorded vibration is not necessarily the most plausible. It has been shown that the perception of contradictory stimuli in different modalities is dominated by the most convincing cue for the context [9]. Thus, the elicited tactile perceptual properties of vibration can differ from the expected perceptual properties in the virtual environment. This suggests that the composition of the tactile user expectations is more complex than previously assumed. This explains the necessity of design experts anticipating user expectations or of trial and error approaches iteratively approximating user expectations. A systematic, explicit approach to determine the expected tactile properties of a situational context simulated in the virtual environment which would elicit the plausibility illusion is still missing.

Tactile user expectations are rarely explicitly assessed in free elicitation tasks. Furthermore, they are mostly accompanied by hands-on evaluations. Laypersons cannot communicate about vibration utilizing engineering terms such as level or frequency directly. Instead, they communicate with associations such as “tingling”. Such assessments of free associations are mostly non-standardized, complicating their analysis. Unfortunately, there is no universally agreed tactile design language available yet [10] that enables standardized communication with future users about their tactile experiences. To generate vibration susequently, it is furthermore necessary to obtain suitable vibration parameters from such perceptual assessments. However, this semantic gap between the physical engineering space of vibration and the perceptual space has not been bridged yet.

1.1 Objectives of the Thesis

The shortcomings discussed above are an obstacle impeding a universally valid, efficient tactile design process that ensures the fulfillment of user expectations. The lack of a systematic approach prevents further automatization and thus potential time and cost reductions. A standardized vibrotactile design language could enable the communication about vibrotactile experiences and expectations in a systematic way. This could allow the quantification of vibrotactile perceptual properties. A systematic relationship between the elements of the vibrotactile design language and vibration parameters would enable the generation of vibrations with expected perceptual properties. This would be beneficial for both the tactile design of products and virtual environments. This vision defines the objectives of this thesis. Therefore, this work investigates the feasibility of quantifying user expectations on WBVs directly in laypersons terms. The goal is to explicitly assess vibrotactile user expectations on WBV with the help of a tactile design language, enabling the translation into plausible vibration. To achieve this vision, three main obstacles need to be overcome. First, a set of tactile perceptual attributes, which sufficiently characterize the perceivable variation of WBV and their direct relationship to physical vibration parameters need to be identified. Second, it needs to be demonstrated that the expected perceptual property profile of a context can be quantified utilizing this tactile design language. Third, it needs to be shown that the model-based synthesis produces plausible vibrations for this context from the expected perceptual profile.

1.2 Structure of the Thesis

In this chapter, the motivation and the goal of this work were elaborated. **Chapter 2** will outline the findings, which are relevant to the thesis. The different levels of the tactile perceptual process are presented. Furthermore, the perception of virtual environments is discussed. **Chapter 3** attempts to join previous findings to derive a research concept for the stated goal. **Chapter 4** describes the virtual environment reproduction system with a focus on the development of the WBV reproduction system. Furthermore, the test setup for the perceptual studies to be conducted in this work is documented. In **Chapter 5** the first main objective of assessing a tactile design language consisting of a set of tactile sensory-perceptual attributes will be investigated. Subsequently, the relationship between physical vibration parameters and these sensory tactile perceptual attributes will be examined. **Chapter 6** builds onto Chapter 5 by utilizing the tactile design language for the quantification of tactile expectations on a set of vehicle contexts. In **Chapter 7** models will be built from the relationships uncovered in Chapter 5 that can translate expected tactile perceptual profiles, as quantified in Chapter 6, into vibration parameters. Subsequently, the synthesized vibrations will be compared to recorded vibrations regarding the perceived plausibility in their multimodal context. **Chapter 8** will summarize the thesis with an overview of the results, a description regarding their novelty, and an outlook on further research questions suggested by the findings.

2. Tactile Perception in Real and Virtual Environments

In the previous section, the aim of this work was set to the creation of a model that enables the translation of tactile expectations of a situational context into plausible vibration. This aim requires an explicit assessment of expected tactile sensations. Furthermore, it requires knowledge about the controlled elicitation of the illusion of plausibility. Therefore, it is necessary to analyze prior research for understanding on how to approach this goal. In the first section of this chapter, an overview of tactile perception as a multilayered process will be provided with a focus on WBVs i.e. vibrations introduced to seated subjects. The second section will present a summary of the perception of virtual environments with an emphasis on the illusion of plausibility. The third section will discuss different authoring approaches that attempt to elicit the plausibility illusion.

2.1 Tactile Perception as a Multilayered Process

As humans, we perceive our everyday life environment through our five senses. The encountered stimuli form a base layer on which a multilayered tactile perceptual process occurs. The somatosensory system enables us to judge the qualities of physical contact with the environment [11], e.g. vibration. The temperature of the touched object can be perceived by thermoreceptors. Pain can be registered by the nociceptors. The position of the body and the limbs is perceived by the proprioceptors. The mechanoreceptors are sensitive to stretching, pressure, and vibration and are thus most relevant for the given focus. The neural excitation patterns produced by the receptors in response to the stimuli encountered in the environment subsequently elicit a percept. The capabilities of the receptors, as investigated by psychophysics, determine e. g. the perception threshold and difference thresholds and thus the potential range and resolution of the percepts elicited by vibrations. This can be interpreted as a second layer of abstraction that represents the theoretical capability of conveying tactile information. In everyday life, some types of vibration are encountered frequently and thus we have learned to intuitively associate specific attributes, e.g. “tingling” to them. Such sensory tactile perceptual attributes [12] can be seen as a third layer of abstraction representing the interpretation of the percept directly elicited by vibration. Depending on sensory tactile perceptual properties, an emotional response, e.g. arousal or pleasantness can be produced [12], [13]. This can be interpreted as a fourth layer of abstraction, the affective layer. Each of the four layers and their relevance to the goal of this work will be discussed in detail in this section.

2.1.1 Physical Layer

Technological developments have greatly influenced our everyday lives. Starting from ships, to horse carriages, proceeding to trains, planes, and ultimately to automobiles, we experience WBVs directly in the vehicle or indirectly by transmission into buildings. Today, several million people in the Federal Republic of Germany are exposed to vibrations through transport in occupational and non-occupational settings. Furthermore, vehicles (e. g. trains) as well as stationary machines (e.g. generators) can also produce vibration transmitted into buildings thus resulting in more human exposure [14]. On the physical layer, vibrations have been investigated intensively, and thus well-established measures for characterizing vibrations from an objective perspective are available. Griffin [15] classifies vibration typically encountered into six classes of excitation patterns (see Figure 2.1.).

One of the most frequently used measures of characterization is the magnitude of the oscillatory motion. The magnitude can be either quantified as displacement, velocity, or acceleration with acceleration being the preferred unit as it is most convenient to measure with widely available measurement instrumentation [15]. As the acceleration of oscillatory motion is varying over time, the root mean square acceleration provides an average measure of the vibration magnitude that humans are exposed to. Since perceivable vibrations can vary by orders of magnitude, the level of vibrations is a more convenient measure of magnitude. The reference for the vibration level is 1 $\mu m/s^2$. However, besides a magnitude dependence, human response to vibration is also frequency-dependent [15]. Therefore, spectra or spectrograms are utilized to

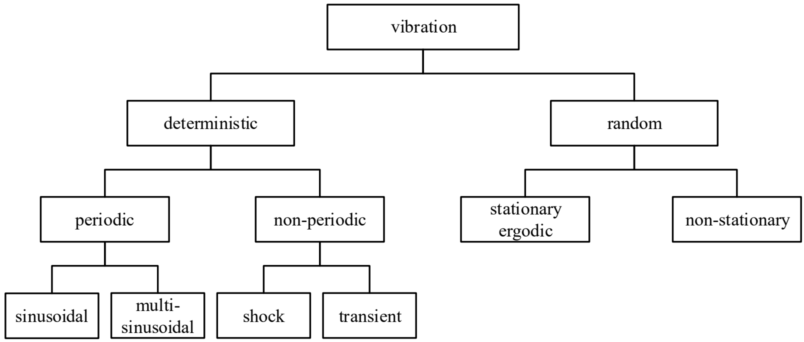

Figure 2.1.
Classification of vibration according to Griffin (adopted from [15]).

characterize vibration regarding their frequency content on the physical layer. From a workplace safety perspective, it is obvious that also the direction of vibration can influence the human response to vibration. This is evidenced by different thresholds for fore-and-aft, lateral, and vertical vibration [16].
The information about vibration available on the physical layer is usually only understandable by experts with prior knowledge in this field. If human response to vibration is not of interest, this layer offers an objective description of vibrations as a basis for analysis. However, for the perceptual quality of a product or the perceived realism of a virtual environment the human response obviously does play a role and thus no inference can be drawn from the information available on this layer. The design of vibration for virtual environments on the physical layer would need to rely on mappings from discrete scene type to vibrations and thus have to omit perceptual aspects that include preferences of users.

2.1.2 Mechanoreceptor Layer

Most research on the human response to vibration has focused on understanding tactile perception from a psychophysical perspective. The psychophysical approach attempts to identify the joint capabilities and limits of the somatosensory system that mediate the perception of vibration. The goal of psychophysical studies lies in the identification of vibration properties which make them perceivable by humans. Furthermore, it investigates the capability of humans to discriminate vibration varying in certain physical vibration properties. The somatosensory system comprises cutaneous receptors, i.e. the sense of touch as well as proprioception, i.e. the sense of position and movement [17].

2.1.2.1 Anatomy and Physiology of Mechanoreception

The anatomy of the glabrous skin suggests four mechanoreceptors types and their associated afferent nerve fibers forming four psychophysical channels [18] through which humans can perceive touch and vibration. These mechanoreceptors are Merkel's Receptors (slowly adapting (SA)-I fiber, Non-Pacinian (NP) III channel), Ruffini corpuscle (SA-II fiber, NP II channel), Meissner’s corpuscles (rapidly adapting (RA)–I fiber, NP I channel), and Pacinian corpuscle (Pacini channel (PC) fiber, P channel). Each receptor signals over its associated afferent nerve fiber type, thus mediating four psychophysical channels detecting and encoding information about the vibration stimulus and ultimately resulting in a sensory event at the somatosensory cortex in the perceiver's brain [11]. An overview of the different capabilities of the four channels is provided in Table 2.1.

The properties relevant to vibration perception of the channels caused by their underlying functional mechanisms are described by [11] and summarized in the following. The NP III channel is mediated by Merkel neurite complexes and SA-I afferent nerve fibers. This channel is more sensitive than the other channels in the frequency range of 0.2 Hz to 2 Hz. Stimulation of this channel

Table 2.1.
Vibrotactile psychophysical channels mediated by their underlying mechanoreceptors and associated afferent nerve fiber types determining the frequency range at perception threshold level and the characteristic sensations produced according to [11].

channel	receptor type	afferent nerve fiber type	frequency range at threshold	associated sensation
NP III	Merkel receptor	SA-I	0.4-2 Hz	“pressure”
NP II	Ruffini corpuscles	SA-II	100-500 Hz	“buzz-like”
NP I	Meissner corpuscles	RA-I	2-40 Hz	“flutter”
P	Pacini corpuscles	PC	40-500 Hz	“vibration”

elicits the sensation of “pressure”. The NP II channel is mediated by Ruffini end organs and the associated SA-II afferent nerve fibers. This channel is less sensitive than the other channels. It produces a “buzz-like” sensation in the frequency range of 100 Hz to 500 Hz. The NP I channel is mediated by Meissner corpuscles and RA I afferent nerve fibers. This channel is dominating the perception threshold in the frequency range of 2 Hz to 40 Hz. Stimulation of this channel is associated with the sensation of “flutter”. Finally, the P channel is mediated by Pacinian corpuscles and the associated rapidly adapting PC fibers. The perception threshold in the frequency range from 40 Hz to 500 Hz is dominated by this channel. Stimulation of this channel elicits the sensation of “vibration”. Each channel is characterized by a different neural code in response to stimulation of different frequencies or magnitudes. While each of the four channels dominates the perceptual threshold in a specific frequency range, the total range of each of the channels mostly exceeds this range, as shown by [18]. The four channels are typically stimulated simultaneously and thus the information mediated by these channels is ultimately integrated into a unified percept. Therefore, only the combination of the four channels can explain the entire perceptual capabilities [11].

While cutaneous receptors are most relevant for receiving information in the intermediate to the high-frequency range [15], WBVs i.e. vibrations perceived by seated subjects often include low-frequency vibrations that can e.g. cause movement of the limbs. In such cases proprioception, which is mediated by receptors in the muscles (muscle spindles), joints, or tendons (Golgi tendon organ) can contribute to the perception of oscillatory motion [15]. Therefore, proprioception also likely contributes to the entirety of the vibrotactile perceptual capabilities. The joint capabilities are reflected in the model proposed by [11] and shown in Figure 2.2. Since it is difficult to stimulate one of the channels in isolation, usually the joint capabilities of vibrotactile perception are studied in psychophysical research. Thus, the effect of changes in vibration stimulation on changes in the elicited percept can only be studied for joint stimulation.

2.1.2.2 Assessment of the Tactile Perceptual Dimensions on the Mechanoreceptor Layer

The limited capabilities of vibrotactile perception define the boundaries of the perceptual space in which percepts can be elicited by vibration, e.g. the perceivable vibration level range is determined by the perceptual threshold. Furthermore, the dimensions of the perceptual space are influenced by the limited capabilities of vibrotactile perception to resolve changes in physical vibration properties (e. g. level) through changes in the elicited percept (e.g. perceived intensity).

There are three general experimental approaches to investigating the perceptual space of vibration: the classification of vibration, the semantic differential method, and the similarity estimation method [19]. Of the three methods, only the classification method and the related similarity estimation method remains on the mechanoreceptor layer to reveal underlying perceptual dimensions according to which vibration can be discriminated. In contrast to the semantic differential approach, no explicit rating criteria in the form of perceptual attributes are required for these approaches, since the vibration stimuli are only rated for their similarity or dissimilarity. By applying a multidimensional scaling (MDS) the perceptual dimensions underlying these ratings can be identified. However, the dimensions of the perceptual space of vibration acquired by this method need to be interpreted as being implicit, since the approach does not reveal, why subjects perceive specific stimuli as being similar. Therefore, it is often attempted by researchers to map explicitly interpretable labels manually onto the uncovered dimensions. In the following an overview is given over prior research conducted according to this method.

The perceptual space of sinusoidal vibration was investigated by [20]. They presented vibrations in the frequency range from 40 Hz to 250 Hz and a level range from 30 dB to 40 dB sensation level (SL) to subjects at the hand. An MDS based on dissimilarity judgments revealed two perceptual dimensions: a dimension associated with low vibration frequency and a dimension associated with high vibration frequency.

The perceptual space of amplitude-modulated (AM) sinusoidal vibration presented at the hand was examined by [21]. For this study, the level was kept constant at 30 dB (SL) as well as the carrier frequency at 150 Hz while the modulation frequency was varied from 0 Hz to 80 Hz. Participants judged the

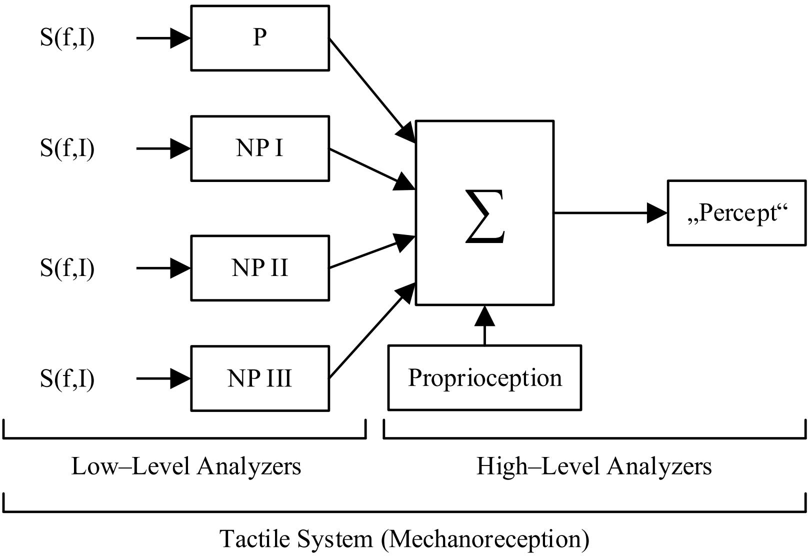

Figure 2.2.
Integration of low-level psychophysical channels as well as proprioception into an unified high-level percept of a vibration stimulus S with frequency f and intensity I (adapted from [11]).

dissimilarity and again an MDS was applied, revealing two perceptual dimensions. One dimension is associated with modulation i.e. dissimilarity from unmodulated sinusoidal vibration, while the other dimension is associated with envelope frequency. They hypothesize that amplitude-modulated sinusoidals with modulation frequencies from 2 Hz to 160 Hz likely elicit similar percepts as pure sinusoidals with a comparable frequency.
Another study by this group [22] presented two superimposed sinusoidal signals of different frequencies. The frequency range of the sinusoids was 50 Hz to 320 Hz. Again, MDS was applied to dissimilarity judgments confirming a perceptual dimension associated with low-frequency vibration and a perceptual dimension associated with high-frequency vibration. The low-frequency component was more salient, i.e. it dominated the resulting percept in comparison to the high-frequency component.
Besides sinusoidal vibration, also sawtooth and square vibrations presented over a knob stimulating fingers were investigated by [23]. Vibration signal frequency vas varied over the range from 3 Hz to 25 Hz and vibration level

at two magnitudes with about 6 dB difference. MDS revealed two dimensions: one associated with frequency and one with signal shape. Sinusoidal, sawtooth, and square vibration presented with three different actuators at the finger were also investigated by [24]. The device utilized did not change the uncovered perceptual dimensions as the vibration introduced to the finger was similar across the three devices.

Also, the perceptual space of regular and irregular rhythmic patterns of successive impulses was investigated [25]. The resonance frequency of the pulses presented over a piezo-mounted hand-held touch screen was 200 Hz or 300 Hz. An MDS based on dissimilarity judgments revealed two perceptual dimensions explaining the perceived difference in these stimuli: "even" – "uneven" and "low amplitude" – "high amplitude". Vibration Frequency was not associated with an additional dimension possibly due to the small frequency range.

Overall each of these studies uncovered a fraction of the potential tactile perceptual dimensions due to only varying few parameters of the utilized vibration and not representing the range of vibrations encountered in everyday life. However, the required pairwise comparisons of this method grow rapidly with an increasing number of stimuli, thus limiting the practical feasibility of this approach [19].

The advantage of investigating the perceptual space on the mechanoreceptor layer with the similarity approach is that no rating criterion besides similarity is necessary for the rating of the stimuli. This facilitates the task [19] because the researcher does not need to define relevant attributes in advance. Such investigations can enable the assessment of the perceived relative difference between two stimuli by the designer of tactile feedback. However, this advantage is also inherently tied to the disadvantage of this approach. It only shows that changes in vibration effect changes in the percept, but it does not hint at the semantic meaning associated with the revealed perceptual dimension of vibration, requiring the researcher to interpret them [19], [23]. Thus, the approach does not uncover what has been labeled explicitly verbalizable subjective sensations [26], useable for laypersons to communicate about vibrations. Therefore, it cannot reveal the intuitively conveyed perceptual properties elicited by vibration [23].

2.1.2.3 Investigations on the Tactile Perceptual Dimensions of Whole-body Vibration on the Mechanoreceptor Layer

Many of the especially salient percepts depending on vibration level or frequency have been investigated in detail by various psychophysical methods [27]. Such properties are e.g. the perceptual threshold and the just noticeable difference thresholds for level as well as frequency or perceived intensity [26]. Some of these characteristic properties for WBVs are summarized due to their relevance to the goals of this work.

The lower end of the perceivable level range of vibrations is limited by the perceptual threshold. The perceptual threshold differs depending on the location of the vibration introduction [16], which might be explained by the receptor density varying across different parts of the skin. Similarly, to the auditory threshold [28], it is frequency-dependent, as the sensitivity of the psychophysical channels is also frequency-dependent [18]. The perceptual threshold of vertical WBVs was investigated by several studies. For recumbent i.e. lying subjects the perceptual threshold of vertical vibration exposure was determined to fall into the level range from 60 dB to 70 dB in a frequency range from 1 to 100 Hz [29]. For seated subjects, the perceptual threshold of vertical vibration in the frequency range of 5 Hz to 200 Hz was measured by [30]. For rigid seats, a threshold between 83 dB and 88 dB was found. An extended frequency range from 2 Hz to 315 Hz of vertical WBV presented at a contoured, rigid seat was examined by [16]. They found the perceptual threshold rising from 83 dB at 2 Hz to 95 dB at 315 Hz. The perceptual threshold of high-frequency vertical WBV from 100 Hz to 500 Hz was investigated by [31] for subjects seated on a rigid plate. Their data suggests that the perceptual threshold rises steeply from 85 dB at 100 Hz to 122 dB at 500 Hz. The perceptual threshold according to [16] and [31] is shown in Figure 2.3.

The influence of acceleration level on perceived intensity has been one of the most investigated physical vibration parameters, e.g. by [15]. The majority of these studies utilize the method of magnitude estimation. In this method, a stimulus with a fixed level is provided as an anchor against which the perceived intensity is to be rated. Thus, the resulting intensity curves are a measure of relative intensity. The experimental design of magnitude estimation was also applied by [32] for the investigation of the perceived intensity of

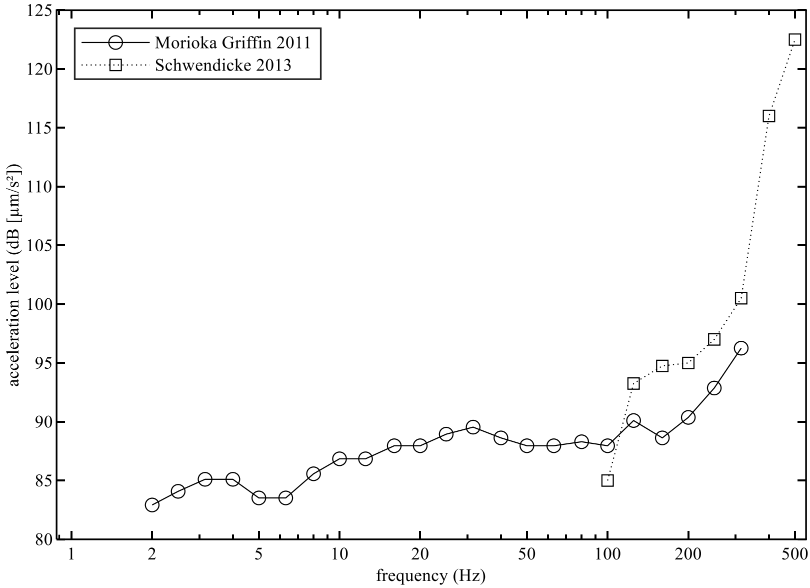

Figure 2.3.
The perceptual threshold of vertical WBV according to [16] and [31].

vertical sinusoidal WBV in the frequency range of 10 Hz to 200 Hz. Based on these results equivalent perceived intensity contours were calculated in steps that correspond to a doubling in perceived intensity each. The almost constant distance between all the equal intensity contour pairs suggests a linear relationship between acceleration level and perceived intensity. Furthermore, the equal intensity contours are approximately parallel to the perceptual threshold. These equal intensity contours are shown in Figure 2.4. Bandlimited white Gaussian noise (WGN) vibration seems to be perceived with a similar intensity compared to sinusoidal vibration, if they have an equal root mean square (RMS) acceleration level and are in a similar frequency range judging by the similar discomfort they elicit [33].
Another important property of tactile perception of intensity is the capability to distinguish two vibration stimuli differing only in level. Investigations of the just noticeable difference in level (JNDL) of vertical WBV have been conducted by multiple researchers. The JNDL reported by [34] at frequencies

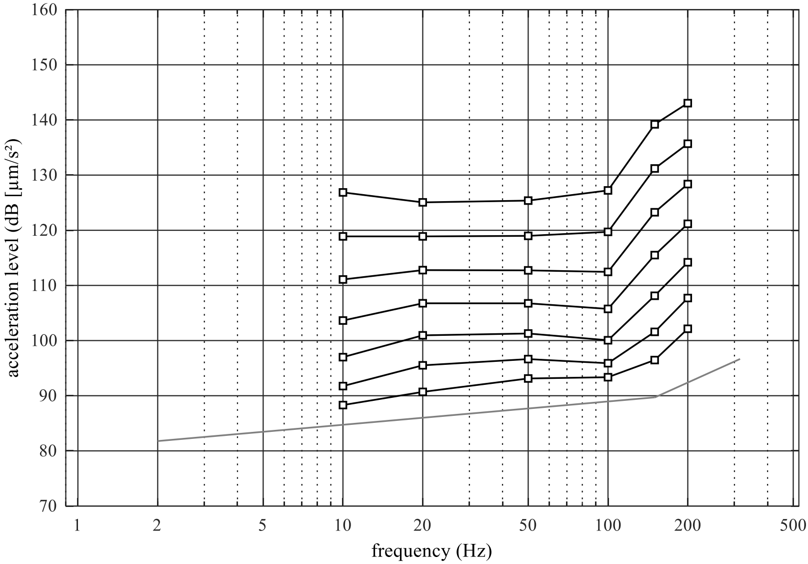

Figure 2.4.
Contours with equivalent perceived intensity of vertical WBV (black) based on magnitude estimation data and two part linear fit of the perception threshold (grey). An upward shift to the next equal intensity contour corresponds to a doubling in perceived intensity (adapted from [32]).

of 5 Hz and 20 Hz at vibration levels of 100 dB and 114 dB is approximately 1 dB providing a lower bound for the ability of subjects to distinguish level increments. A slight decrease of JNDL with increasing vibration levels was observed. An extended frequency range of 5 Hz to 50 Hz was examined by [30], revealing a JNDL of 1.5 dB, which was approximately constant over frequency.

Besides the investigations of the influence of frequency on the perceptual threshold, the just noticeable difference in frequency (JNDF) for vertical WBV also received some attention. The JNDF in the frequency range of 5 Hz to 40 Hz was investigated by [30]. The results suggest that the positive JNDF can be approximated depending on the reference frequency by $34\ f_{ref} - 1.25\ Hz$ in the direction of increasing frequency. A higher frequency range of 20 Hz to 90 Hz was examined by [32]. The positive JNDF as a fraction of the reference frequency was found to increase from approximately 35 % at 20 Hz to 80 % at 90 Hz. Overall, this implies that the capability of tactile perception of WBV to discriminate sinusoidal stimuli with different frequencies, i.e. the frequency selectivity depends highly on the frequency range. The frequency selectivity might be further reduced by masking effects, i.e. low frequency components leading to an increase in the perceptual threshold for high frequency vibration [35].

2.1.2.4 Advantages and Disadvantages of Investigations on the Mechanoreceptor Layer

While the physical layer provides an objective description of physical vibration properties, the mechanoreceptor layer provides insight into the limited capabilities of vibrotactile perception in resolving these properties. This might be interpreted as the theoretical limits for conveying information via vibration. A transfer rate of around 6 bits per second is suggested as a quantification of information transfer capability [36]. Similarly, to the physical layer such information is mostly valuable to experts e. g. for the selection of discriminable i.e. non-redundant vibration feedback. However, it provides no information about the intuitively conveyed information and thus explicit communication with laypersons about the sensations is not possible on this layer.

2.1.2.5 Applications enabled by Investigations on the Mechanoreceptor Layer

This layer of abstraction is most suitable for encoding new information to be delivered as tactile icons, or tactons [37] via tactile feedback e.g. in vehicles [38]. However, the intended mapping between vibration and newly encoded information about the meaning of the feedback needs to be learned by users, especially if the tactons are arbitrarily assigned to the information that are to be conveyed. Furthermore, it has been suggested by [37] for the auditory domain that relying on paired comparisons for judging the ability to differentiate of two stimuli might lead to on overestimation of the reliably achievable performance in absolute identification of the feedback. One application that relies on this abstraction layer is tactile hearing aids that utilize the tactile channel for conveying information to persons with hearing loss [39], [40]. Another application enabled by this layer of abstraction is tactile codecs [26]. Since indiscriminable stimulus variation is redundant from a perceptual point of view, the codec can reduce the required amount of bandwidth for the transmission of vibration. However, in contrast to e.g. mp3 for the auditory modality, there is no established perceptual codec for vibration, yet [41]. For the goal of designing vibration for virtual environments, this layer of abstraction does not enable the communication with laypersons about their preferences. Therefore, designing vibrations on this layer would likely also rely on mappings from discrete scene type to vibration. The only difference might be a reduction in stimulus set size, due the elimination of perceptually indiscriminable, i.e. redundant stimuli pairs.

2.1.3 Sensory Layer

On the mechanoreceptor layer, perception is understood as a bottom up process, i.e. vibration excites our tactile receptors that convert the stimulus into neural signals that are processed in the somatosensory cortex. While this view enables the investigation of the theoretically transferred information, it cannot explain how that information might be interpreted by the perceiver, i.e. the intuitive semantic meaning. In agreement with this assessment, [26] suggest that psychophysical properties do not cover every aspect of tactile perception. On the sensory layer of abstraction, vibrations encountered frequently in everyday life are seen as a carrier of information about the environment, in the form of verbalizable sensory-perceptual attributes associated to the percepts [42]. Thus, this layer focusses on the relationship between vibrations and their elicited sensory-perceptual attributes such as “tingling”, which are directly related to the physical properties of the vibration signal.

2.1.3.1 Ecological Approach to Perception

For the visual domain, Gibson argued that visual perception can not only be understood from the physics of light, the anatomy and physiology of the eye as well as the brain [43]. He argues that perception is not only a bottom up process, i.e. a processing of sensory inputs in response to stimuli. Instead it is rather a top down process that extracts invariants from the stimuli. Similarly, Gaver argues for the auditory domain that bottom up processing, as understood by psychoacoustics, does not capture every aspect of auditory perception [44], [45]. He divides the perception of sound into two strategies: musical listening (listener focusses on the experience of the sounds themselves) and everyday listening (listener focusses on the information conveyed by the sound producing event). He suggests that everyday life sounds are interpreted by laypersons as a carrier of information about the environment, especially in open ended description tasks. He refers to this as the ecological approach to auditory perception. The physical properties of the sound producing event influence the spectral temporal properties of the sound and thus its audible source attributes. For example, an approaching car's engine emits a sound with specific temporal and spectral properties. These properties will elicit specific sensory-perceptual attributes such as “humming” that convey information about the sound source. This implies an intuitive relationship between physical sound properties and elicited sensory-perceptual attributes acquired from experiences in everyday life. Based on this relationship it is possible to synthesize sounds with specific sensory-perceptual properties and or analyze sounds for their elicited properties [45]. The synthesis needs to be only accurate enough, to elicit the relevant perceptual attributes.

2.1.3.2 Perceptual Categorization

Gaver describes ecological perception as being an interdisciplinary approach between psychophysics and cognitive science [44]. A topic at the center of cognitive science is the perceptual process of categorization [46]. A free verbal elicitation task to find sensory-perceptual attributes might indeed be interpreted as a categorization task. Rosch [47] provides the following definition for categorization:

> “The world consists of a virtually infinite number of discriminably different stimuli. One of the most basic functions of all organisms is the cutting up of the environment into classifications by which nonidentical stimuli can be treated as equivalent."

This suggests a generalization process in which two stimuli, despite being discriminable in a paired comparison, might be perceptually equivalent because they are assigned to the same category. Harnard argues that categorization may be a general principle of perception, which might be explained by neural networks representing categories in the human brain [48]. Furthermore, these semantic categories can be interpreted as a bridge between individual sensory experiences and collective representations through a shared language [49]. Thus, sensory-perceptual attributes are suitable as a means of communicating with laypersons about their percepts.

It is well known that categorization can occur at different levels of abstraction depending on the inclusiveness of the category [47]. At the most commonly used level of categorization, the basic level, the conveyed information is maximized while the number of categories is minimized. Categories at the superordinate level share less common properties and should thus be too general to cover all the perceptual differences of interest. Categories at the subordinate level share more common properties and should thus be too correlated or more redundant than desired. Categories of vibration elicited in [50] obviously seem to be part of different levels of categorization: “cobblestone-like”, “shaking”, “vibrating”, with “shaking” likely being a compromise between explanatory power and generalizability. In contrast to laypersons, experts have been shown to use the subordinate level as frequently as the basic level [51]. This emphasizes the necessity of using laypersons for the elicitation of perceptual attributes, if their perspective is of interest.

In linguistic categorization, nouns are typically understood as entities belonging to a category, while adjectives are understood as qualities of these entities. Sensory properties refer to a specific value of such a quality [52]. However, as far as sensory properties are concerned this general statement may be reevaluated since nouns may refer to general qualities (operating as categorical membership criteria) and adjectives may therefore refer to a certain value or quantification of a given quality, property or criteria. Membership in categories of natural language can be misunderstood as being dichotomous, but [53] provide a list of evidence that it is rather continuous. Furthermore, they demonstrate that adjectives such as sensory-perceptual attributes also show this property.

2.1.3.3 Assessment of the Tactile Perceptual Dimensions on the Sensory Layer

The ecological approach to perception might also be applied to the tactile modality. In our everyday lives, we are frequently exposed to all kinds of vibrations such as at the seat of a moving vehicle. These vibrations convey information about the environment, e. g. the operating mode of the vehicle, or the condition of the road. Depending on the temporal and spectral properties of the encountered vibration, specific sensory-perceptual attributes are elicited, such as “tingling” or “shaking”. In comparison to psychophysical investigation of vibrotactile perception, an ecological perspective was rarely followed in previous studies. While language might be a promising way to assess vibration properties, a systematic analysis of laypersons communication about the sensory-perceptual space of vibration using natural language is lacking and thus no set of sensory-perceptual attributes sufficiently describing the perceptual space has been agreed upon [10], [54].

Besides investigating the perceptual space on the mechanoreceptor layer of abstraction, the perceptual space can also be investigated on the sensory layer of abstraction. However, in contrast to the former investigations, the latter investigations utilized explicit verbal labels for the characterization of the sensory tactile perceptual properties. The semantic differential approach applied to investigate the perceptual space on the sensory layer of abstraction is also described by [19]. In a first step, a set of sensory-perceptual attributes is selected either by an expert from a dictionary or by conducting free elicitation tasks with subjects while presenting vibration. In a second step, a set of stimuli is defined that should be rated according to the selected attributes. On the mechanoreceptor layer, only a relative rating of vibrations for their similarity is utilized. In contrast, on the sensory layer, the attributes are absolutely rated on a Likert scale for their suitability in describing vibrations in a second step. In a final step, a factor analysis or principal component analysis (PCA) can be conducted to identify fully or partially redundant attributes. In the following an overview is given over prior research conducted also according to this method.

A simple approach to assess the perception of vibration on the sensory layer is to simply map physical event labels or situation labels to their specific vibration [55], [56]. However, this has the disadvantage that such labels are not necessarily generalizable across different situational contexts. Therefore, it is more meaningful to investigate the perceptual space according to the elicited sensory-perceptual attributes.

When touching textures, the mechanoreceptors might be stimulated similarly as they are stimulated when exposed to vibration. In [57] 33 tactile sensory-perceptual attributes were assembled from a dictionary. A small stimulus set of five texture stimuli was presented to subjects' forearms and all 33 attributes were rated on a five-point scale. The results suggest four dimensions of the sensory-perceptual space of textures: roughness, slip, firmness and pile. The dimensions of the perceptual space of textures found by many studies were aggregated by [19]. He suggests four dimensions: “rough” - “smooth”, “hard” - “soft”, “cold” - “warm” and “friction”.

Also vibration stimuli produced directly by various active principles were studied. While the first part of the study described in the previous section on sinusoidal vibration presented at the hand [20], was conducted without any explicit verbal labels, the second part involved the rating of 13 adjective pairs in Korean language. These sensory-perceptual attribute pairs were mapped onto the two-dimensional perceptual space. They suggest to label the low frequency dimension “slow” - “fast” or “vague” - “distinct” and the high frequency dimensions “thick” - “thin” or “heavy” - “light”.

A free elicitation task was conducted by [58] with sinusoidal vibration of 16 Hz and 250 Hz reproduced with an ultrasound transducer array causing acoustic radiation pressures at subjects hands. They manually summarized the descriptions of test subjects experiences to 14 categories depending on the stimulus frequency (16 Hz Stimulus/250 Hz Stimulus): “puffs” - “breeze”, “pulsing” - “flowing”, “soft” - “dense”, “coming and going” - “constant”, “pointed” - “dispersed”, “weak” - “strong”, “prickly” - “tingling”.

A library of 120 vibrotactile effects was assembled by [59] with the goal of aiding designers in the selection of feedback with controlled perceptual properties. These effects varying mostly in signal envelope were presented over a wristband actuator with a frequency range from 200 Hz to 300 Hz. Subjects rated emotional and sensory tactile perceptual attributes relevant to vibration extracted from the touch dictionary presented by [57]. The resulting database consisted of mappings from each stimulus to ratings of each attribute. In their subsequent study [60], they investigated these mappings for underlying perceptual dimensions. They suggested 4 dimensions of sensory-perceptual attributes for their stimulus space: complexity (e.g. “regular”), continuity (e.g. “continuous”), roughness (e.g. “smooth”), and duration (e.g. “long”). However, due to their mostly non-systematic variation of vibrations they did only identify a discrete mapping between their vibration effects in their database and the sensation associated with them. In order to overcome the constraint of having to rely on a limited set of vibration items from a database and their discrete, manually rated perceptual profiles, [12] suggested to tune the vibration items according to physical tuning parameters to continuously shift the perceptual attributes ratings. They selected 10 basis rhythmic vibration patterns from their database and modified their tempo or energy. However, a different relationship between the physical vibration parameters and elicited attributes depending on the basis vibration was observed. This suggests that the relationship observed for one vibration cannot be generalized to all vibrations.

2.1.3.4 Investigations on the Relationship between Whole-body Vibration and Sensory-perceptual Attributes

A generalized relationship between physical vibration parameters and perceptual attributes would enable a more flexible selection of vibration from perceptual attribute ratings. Therefore, some studies systematically varied vibration parameters while investigating the ratings of the elicited sensory-perceptual attributes. In [50] sinusoidal stimuli from 1 Hz to 50 Hz were presented to subjects as WBV at an individually selected level and associations were collected. These associations ranged from specific terms such as “cobblestone road” to attributes as “shaking” to very general terms as “vibration”. Some attributes were rated in a semantic differential. A factor analysis identified three components: attributes describing uniform vibration, attributes describing low frequency vibration and attributes which are independent from frequency. In [61] a wider frequency range of 8 to 200 Hz was presented. Stimuli were not limited to sinusoidal vibrations. However, parameter variation in amplitude modulated sinusoidal vibrations and in bandlimited WGN vibrations was very limited. All stimuli were presented at 10 dB above SL. After a free elicitation task, the suitability of all found perceptual attributes was rated. Instead of identifying perceptual dimensions by a factor analysis, eight attributes were manually selected to represent the perceptual space: “rattle”, “beating”, “bumpy”, “shaky”, “humming”, “ticking”, “wavy”, and “booming”. [62] presented vibrations recorded from 14 cars in idle mode. 22 perceptual attributes were elicited by those stimuli. Factor analysis revealed 3 components: “comfort”, “regularity” - “impulsiveness”, “tonality”.

2.1.3.5 Advantages and Disadvantages of Investigations on the Sensory Layer

In comparison to the investigations of tactile perceptual dimensions on the mechanoreceptor layer where the implicit dimensions need to be interpreted, the investigation on the sensory layer offers direct interpretability of the revealed dimension due to the loadings of explicit sensory attributes on the components of a PCA or the factors of a factor analysis. In addition to the mechanoreceptor layer of abstraction, which provides insights on the capabilities of the tactile perceptual system in resolving variation in physical vibration parameters, the sensory layer of abstraction provides insights on the explicit interpretation of these variations in the form of sensory tactile perceptual attributes. Furthermore, it enables communication with laypersons about their percepts elicited by vibration. Vibration designers are ultimately interested in the judgments of future users which are typically non-experts. In contrast to the physical and the mechanoreceptor layer of abstraction, communication about the sensory layer of abstraction requires no prior training, because laypersons should be familiar with sensory-perceptual attributes from their everyday life experiences.
As for the mechanoreceptor layer, the utilized stimulus set determines the revealed dimensions of the perceptual space. Each of the presented studies investigated only a fraction of the range of vibrations encountered in everyday life. Thus, it is unclear, whether all relevant perceptual dimensions have been identified yet. Identification of all relevant sensory-perceptual attributes in the vast amount of verbal descriptions of vibration is difficult, since the required effort in rating the attributes rises with each additional attribute. If none of the attributes loading onto a potential dimension are included in the semantic differential, this dimensions will not emerge from the factor analysis. Unsurprisingly, the attributes rated in the semantic differentials of the previous studies are often arbitrarily limited.

Identification of tactile sensory-perceptual attributes, representing the dimensions of the tactile perceptual space is the prerequisite of investigating the relationship between physical vibration parameters and these attributes. However, by arbitrarily choosing vibration instead of systematically varying their parameters, it is difficult to draw inferences regarding a universal relationship between physical vibration parameters and their elicited attribute ratings. Databases of discrete vibration items and their associated absolute attribute ratings, might be a tool for designers and engineers for the selection of vibration by their sensory-perceptual rating profiles. However, the utility of such databases is limited because only the discrete vibration items in the database can be selected. Studying the relationship between physical vibration parameters and the ratings of the sensory-perceptual attributes might produce models, which would enable a much more flexible selection of vibration with known sensory-perceptual properties. They might enable a translation of perceptual specifications consisting of sensory tactile perceptual attribute ratings into physical vibration parameters.

2.1.3.6 Applications enabled by Investigations on the Sensory Layer

This layer is suitable to design feedback, intuitively conveying sensory tactile perceptual properties intrinsically encoded in vibration. With the help of models, an automatic translation of perceptual specifications collected from laypersons might become feasible. Selecting vibration according to such a procedure would ensure the controlled elicitation of the desired perceptual properties by the vibration of a product or a virtual reality application. Thus, this layer of abstraction seems promising regarding the goal set out for this work.

2.1.4 Affective Layer

While the sensory layer of abstraction provides an interindividually shared and thus mostly objective description of perceptual properties of vibration applicable across situational contexts, the designer ultimately wants to effect the emotional response to the potential users. The user should perceive the product or virtual environment as e. g. “exciting” or “pleasurable”. The emotional response to the elicited sensory tactile perceptual properties[1] is indeed suggested by [12], [13] to be a superordinate layer of abstraction.

2.1.4.1 Assessment of the Tactile Perceptual Dimensions on the Affective Layer

The assessment of the dimensions of the emotional perceptual space of vibration is conducted in a similar method compared to the sensory-perceptual space. Besides investigating the sensory tactile perceptual space, the emotional perceptual space of textures was also investigated by [57]. They assembled 14 emotional attributes from a dictionary. Again, a small stimulus set of five texture stimuli was presented to subjects' forearms and all 14 attributes were rated on a five-point scale. The results suggest two emotional dimensions of perceptual space of textures: comfort and arousal. Similarly, [60] also investigated the emotional perceptual dimensions of vibration besides sensory-perceptual dimensions for their stimulus set. They suggested three dimensions of emotional perceptual attributes for their stimulus space: agitation (e.g. “urgent”), liveliness (e.g. “lively”), and strangeness (e.g. “strange”).

2.1.4.2 Investigations on the Tactile Perceptual Dimensions of Whole-body Vibration on the Affective Layer

The mapping from physical parameters of vibration to these emotional dimensions was also investigated. Starting from 10 basis rhythmic vibration patterns from their database, [12] modified their tempo or energy to study the effect on emotional attributes “agitation”, “liveliness” and “strangeness” depending on the presented vibration. A linear relationship between actuator output energy and agitation and liveliness is suggested. Similar to the sensory attributes rated in their study, a different relationship between the physical vibration parameters and elicited attributes depending on the basis vibration was observed. This suggests that the relationships for the basic vibration cannot be generalized to all vibrations.

The relationship between amplitude modulated sinusoidal vibration in a frequency range from 60 Hz to 300 Hz, five amplitude steps over a range of approximately 6 dB and a duration form 50 ms to 2000 ms was investigated by [63]. The emotional attributes valence and arousal were rated by subjects on a -100 to +100-point scale for each stimulus. The arousal rating increased mostly on stimulus duration and amplitude, varying across two thirds of the arousal scale range. Valence increased from about -25 points to 25 points with increasing frequency. This suggests, the stimuli were neither perceived as very positive or negative. One possible explanation might be found in the context dependency of valence i.e. preference judgments. No situational context for subjects was provided beyond the laboratory context, possibly hinting at the neutral ratings.

From the domain of sound quality judgments, it is indeed known that preference judgments are influenced by the situational context [64]. This may be obvious from the example of ocean waves and tire noise of cars being spectrally and temporally quite similar on the physical layer of abstraction. However, their “pleasurable” ratings are likely very different, with ocean waves being perceived as much more pleasant than tire noise. Such an example might also be found for the tactile domain, with the touch of a loved one being perceived as surely more pleasant than the touch of stranger. Overall, this implies that a universal mapping between physical vibration parameters and affective perceptual attributes is likely difficult to obtain, because they depend on the situational context.

2.1.4.3 Advantages and Disadvantages of Investigations on the Affective Layer

Compared to the sensory layer of abstraction, the affective layer of abstraction offers mostly similar advantages and disadvantages. Explicit communication is also possible on this level. But the affective perceptual dimensions depend on the initial attributes and stimuli utilized for the investigation. If the stimulus set is not representative of everyday life experiences with vibrations, it would be unclear whether all relevant dimensions have been uncovered yet. However, while the sensory layer might be useful to provide a context independent description of perceptual properties of vibration, a description on the affective layer would also be dependent on the situational context.

2.1.4.4 Applications enabled by Investigations on the Affective Layer

As a consequence of the context dependency, a tactile designer relying solely on affective attribute ratings would have to utilize different models for different situational contexts, increasing the effort to create such models enormously. Thus, for the goal of this work, i.e. synthesizing vibration fulfilling user expectation, the affective layer is likely not preferable.

2.2 Perception of Virtual Environments

Virtual reality has been fascinating humans by enabling the interaction with a seemingly endless range of environments, in and out of reach of everyday life experience. They offer the advantage of immersing the user safely into an arbitrary environment, whose properties can be precisely controlled. The generation of virtual environments has been enabled by increasingly capable rendering, tracking, and reproduction systems. This technological progress has facilitated broader utilization of virtual reality in many fields [2]. Such applications include research and development, training, shopping [3], therapy [4], and entertainment [5].

Eventually, understanding and thus controlling the involved illusions enabling virtual realities became a topic of research interest. What are the illusions that lead to the user perceiving the virtual environment and interacting with it just like with real environments? Slater suggests the separation into the two orthogonal factors place illusion and plausibility illusion to explain different aspects of the illusions [6]. The place illusion explains the sense of presence perceived by the user immersed in to the virtual environment. The plausibility illusions explain the overall credibility of the experienced environment. Both illusions will be discussed in detail in the following two sections.

2.2.1 The Place Illusion

Slater [6] provides definitions for immersion, presence and the place illusion which are all related to the qualia of the “sense of being there”, as defined by

[65]. A qualia cannot be communicated and can only apprehended by direct experience [66]. These terms will be summarized below.

2.2.1.1 Definition of Place illusion and Related Terms

The technical capabilities of the virtual environment such as display (e. g. resolution, field of view), the tracking technology (e. g. latency), and rendering (e. g. framerate) determine the quality of the virtual reality experience. These capabilities constrain the sensorimotor contingencies [67], i.e. the actions that can result in a meaningful change in stimuli presented to the different modalities. For example, a head movement shifting the field of view by a specific angle can exceed the boundaries of the display breaking immersion. Thus the stimuli from the real instead of the environment would be perceived. Slater suggests to delineate this “system immersion” referring to valid actions enabled by the technical capabilities of the virtual environment from the “immersion response” from a user i.e. the qualia of being there [68].
The term presence is applied to wide range of phenomena in an often ambiguous way [69], but it generally includes the qualia of the sense of being there [65]. This qualia simply happens and requires no deliberate attention as e. g. in desktop computer games [6].
In order to avoid confusion, the term place illusion is defined as strictly relating to the “sense of being there” by [6]. Place illusion can only arise in the boundaries of system immersion and depends on the behavior of the user testing these boundaries. Thus it is likely binary, i.e. the boundaries are encountered or they are not encountered [6]. Interindividual differences can be explained by user behavior. One user might be moving more carefully and thus never run into any system immersion limits. Another user might extensively probe the boundaries and would thus encounter the system immersion limits much more frequently, i.e. he would perceive stimuli from outside the virtual reality breaking the place illusion [70].

2.2.1.2 Measurement of the Place Illusion

The investigations of the place illusion often necessitate the quantification i.e. measurement of the place illusion occurring in users of the virtual environments. Slater argues that place illusion is a qualia and thus it has to be measured indirectly via questionnaires or behavioral observation in order to compare the interaction with the virtual environment to the interaction with the real environment [6]. A presence questionnaire is presented by [71] in order to quantify individual differences as well as differences in characteristics of different virtual environments. Another questionnaire was presented by [72]. However, [72] obtained similar ratings of presence for real environment and a virtual environment with a low-resolution head-mounted display. They argue that the qualia of being there is specific to the system immersion enabled by a specific virtual reality system and are thus not comparable via questionnaires between different system immersions. The validity of presence questionnaires is further questioned by [73], because there is no methodological proof that the questionnaire assesses the mental state of the user experienced place illusion. Instead, it is suggested to observe the behavior and physiology of users in the virtual environment, e. g. if they avoid virtual obstacles [6].

2.2.2 The Plausibility Illusion

In contrast to the place illusion, the plausibility illusion has received much less attention. If the place illusion is concerned with how the virtual environment is perceived, the plausibility illusion would be concerned with what is perceived, i.e. the content of the virtual environment [6]. This illusion is discussed in detail below, too.

2.2.2.1 Definition of the Plausibility Illusion

The term plausibility illusion refers to the illusion of the events, over which the user has no direct control, in the virtual environment perceived as being real [6]. Slater suggests that the illusion arises from the events occurring in the scene and the corresponding sensations felt by the users having a high correlation, without physically correct reproduction being necessary. However, the question about the correctness of the judgement, regarding the sensation felt during the event, arises. He hypothesizes that prior experience of such events would likely shape the expectations of the sensation. For example, if visual stimuli in virtual environment depict a person driving on a cobblestone road, the tactile stimuli should match the expected tactile properties typically experienced in such a situation. The mechanisms underlying the illusion of plausibility or verisimilitude have not received much attention yet. The definition of [6] suggests it can be understood as a measure of coherence of the stimuli encountered dependent on the expectations of the depicted virtual environment.

2.2.2.2 Authenticity and Plausibility

Eliciting the place illusion seems relatively easy in the light of technological advances facilitating virtual realities. But how can the content of a virtual environment be designed in a way that the encountered stimuli match expected properties? There are two general approaches to the generation of such stimuli: the authentic approach and the plausible approach.
The authentic approach attempts to elicit percepts in a specific virtual environment which are identical to percepts elicited in the corresponding specific real environment [74]. Often this naive approach attempts a complex, physically accurate reproduction of the real environment that cannot be differentiated in an A/B comparison. The implicit assumption of the approach is that the properties of the percepts elicited by stimuli encountered in the real environment are a heuristic of the properties expected in the real environment. Furthermore, it is assumed that e.g. sound and vibration are physically coupled that auditory and tactile perception should by coupled as well. However, it is unclear, whether this assumption holds. It is often the case that humans are confronted with conflicting cues perceived in different modalities. In the real environment the multimodal percept of specific situational contexts might be dominated by one modality [9], thus dominating the experience and ultimately the expectations of such contexts. If the stimuli of the different modalities in the virtual environment are based on recordings of the real environment, a mismatch between expected perceptual properties and the elicited perceptual properties would occur. This would likely impair the plausibility illusion. There are hints that such effects might also involve the tactile modality. For example, in a basketball game, the guests in the stadium can perceive ball impacts visually and auditorily. However, the vibration caused by the ball impacts is attenuated below the tactile perceptual threshold of the

guests in the stadium. The authentic reproduction of vibrations would thus not include any tactile stimuli. When transferring the basketball game into a multimodal, virtual environment, [7] added vibration for each ball impact, resulting in a much more plausible experience than without vibration. Furthermore, the results from [8] indicate that the recorded vibration is not necessarily the most plausible vibration.

The plausible approach attempts to elicit percepts that are perceived as having occurred in a comparable environment. Thus it focusses directly on the expectations of such an environment. In comparison to the authentic approach, the plausible approach has many advantages. The scope of this definition also includes environments unknown to the user, since it is sufficient to hold expectations on such an environment, for the plausibility illusion to occur. Expectations might also be influenced by media exposure. It seems to be generally accepted by laypersons that spaceships should emit a sound in vacuum, despite them never having experienced space flight. It has indeed been demonstrated that expectations on bang-like sounds are shaped by media exposure [75]. If such a bang does not include an audible decay as ubiquitous in movies, it will not be perceived as being bang-like. Therefore, the plausible approach can also aid in vibration design, where no corresponding real environment exists and thus vibration recording is impossible. Furthermore, if it is sufficient to elicit the expected perceptual properties, the requirements on the properties of the reproduction system might be lower. The plausible approach is likely to avoid the case of conflicting sensory cues, since it builds directly onto the expectations of a situational context. This suggests that the plausible approach to authoring vibration might be preferable to elicit the plausibility illusion. However, this requires understanding mechanisms underlying the plausibility illusion, especially regarding the contribution of expectations.

2.2.2.3 Investigations on the Plausibility Illusion

The plausibility illusion was investigated conceptually and practically by [76] in a series of experiments. He introduces the term coherence as measure reasonable circumstances, i.e. how well the encountered virtual environment agrees with prior knowledge about such an environment, e.g. regarding environmental coherence of physical interaction with objects, or regarding the

own body. Subsequently, the coherence factors behavior of other humans, behavior of the own body, physical behavior of objects, and appearance of the environment, were reduced in complexity over three levels each. Participants started at the lowest level for each factor could improve each coherence factor one level of a time. The experiment assessed, which factors subjects prioritized. The results suggest that body coherence has the highest contribution to overall plausibility, followed by scenario coherence and finally by physical coherence and human behavior coherence. However, the properties relevant to the coherence factors are not investigated quantitatively, offering little cues for the design of these factors.
The mechanisms underlying the plausibility illusion were also analyzed und investigated by [77] for the auditory perception of room properties. He divides perceptual measurements into sensory measurements, evaluation measurements, and proprioceptive measurements. Sensory measurements refer to psychophysical measurements to characterize the functional qualities of perception. Evaluation measurements such as quality or plausibility judgments are driven by top down perceptual processes, since the elicited percept is compared to an inner reference. Proprioceptive measurements refer to the perception of the self, e. g. the qualia of “being there”, i.e. the place illusion. The perceptual quality of a product is a measure of the match between the elicited characteristics and the expected characteristics with respect to the totality of relevant perceptual features [1]. Accordingly, [77] hypothesizes that also the plausibility illusion arises from a perceptual process measuring the adequacy of the elicited perceptual object against an inner reference. The inner reference is formed by expectations which are based on prior experiences. The prior experience arises from the perceptual objects encountered in everyday life, which are aggregated into a system of abstract classes with similar perceptual properties, likely related to the perceptual process of categorization. The inner reference for a situational context presented in a virtual environment is inferred from these systems of classes. Thus, a plausibility judgment is a search in this system of classes for a class whose properties are similar to the properties of the perceptual object elicited in the situational context. However, he assumes that the elicited perceptual object and the inner reference are difficult to investigate separately, because the assessment of a plausibility judgment in response to a stimulus would always involve both. Providing an example for the tactile domain, the perceptual properties (e.g.

“tingling”) elicited by car vibration in a virtual environment would be compared to the perceptual properties of multiple classes (e.g. driving over cobblestone, driving over tarmac). The more similar the elicited perceptual properties are to the properties of one of these classes, the more plausible the vibration would be perceived.

Based on his analysis of the mechanisms of the plausibility illusion, [77] suggested that the plausibility can be interpreted as a similarity judgment between the properties of the elicited perceptual object and the properties of the inner reference. Thus he generalized this similarity judgment into a distance measure between the elicited perceptual object and the inner reference in a perceptual space, created by the relevant perceptual properties. In order to test this hypothesis, [77] conducted a series of experiments regarding the auditory perception of room acoustical properties. Participants rated the plausibility of a speech sample folded with 30 room impulse response of four room classes according to their plausibility for the verbally communicated room classes (small chamber, medium room, hall, and cathedral), resulting in 30 plausibility judgments for each room class. Subsequently, also similarity judgments were collected for all combinations of the 30 stimuli on a 10-point scale. A MDS was utilized, extracting two dimension of the perceptual space underlying these similarity judgments. Each stimulus is represented by a point in the resulting two-dimensional perceptual space. This enabled the calculation of four room class centroids from the stimuli associated with each room class. Subsequently, the Euclidean distance between the stimuli and each class centroid was calculated in the two-dimensional perceptual space. Finally, for each room class, the Euclidean distance was compared to the plausibility rating of each stimulus. The results suggest that the plausibility of an elicited perceptual object (depending on the room reflections and decay) decreases with increasing Euclidean distance from the expected room class centroid in the perceptual space. From their results it can be concluded that plausibility seems to be proportional to the inverse Euclidean distance between an elicited perceptual object and an expected perceptual object in the perceptual space.

2.2.2.4 Measurement of the Plausibility Illusion

There is no established measure for plausibility yet. Similarly, to the place illusion, physiological as well as psychological metrics have been investigated for the assessment of the plausibility illusion. Heart rate, skin conductivity and skin temperature were investigated as a predictor for the plausibility illusion and the place illusion by [76]. For this purpose, the system immersion (associated with place illusion) and the coherence (associated with plausibility illusion) were varied on two levels each for an interactive virtual environment were participants threw balls. System immersion was varied by reducing or increasing the field of view and by adding or removing tactile feedback. Coherence was varied, by adding or removing the law of physics for the simulation of the ball. Furthermore, it was varied by providing subjects a narrative prior to the experiment and subsequently adhering to it or violating it. While the first experiment failed to show any differences in the physical metric, the second experiment with larger differences in coherence showed a negative correlation of heart rate with coherence (Bayesian posterior probability 87.1 %). Skin temperature and conductivity demonstrated inconclusive results. However, these results suggest that physiological measures have a large variance, making it difficult to detect smaller significant changes.
An alternative forced choice method to detect whether a virtual stimulus is less plausible than an original stimulus was suggested by [78]. The method requires the successive presentation of a real stimulus and a simulated stimulus. Participants are forced to select the stimulus that they perceive to be the real stimulus. By evaluating the ratio between correct and incorrect choices, the threshold for plausibility can be determined. They utilize this method for evaluating which of two binaural simulation methods has a better performance. However, this method has the disadvantage that the simulation needs to be fairly close to reality, since subjects would always be able to detect the difference otherwise. Furthermore, such a method is not necessarily representative for a real world application scenario, where the reference of the real stimulus is rarely available for comparison. In the framework of authenticity and plausibility laid out in [74], it seems that the method of [78] is rather assessing authenticity, since it compares a specific percept elicited by the real

environment to a specific percept elicited by the virtual environment. However, in a judgment of plausibility the percept elicited in the virtual environment is rather compared to the expectations of the user, without any concrete reference being available for comparison.

Besides physiological measures, [76] also investigated questionnaires for the measurement of the plausibility illusion with the experimental setup described previously. The Wittmer-Singer presence questionnaire (PQ) [71] as well as the Slater-Usoh-Steed presence score (SUS) [72] were utilized. These questionnaires were originally intended to be used for measuring the place illusion. Unsurprisingly, there was little evidence (Bayesian posterior probability of SUS 61.7 % and PQ 71.3 %) for differences in overall ratings of these questionnaires caused by the large differences in coherence of the virtual reality scenario. But when only the single item "naturalness" contained in these questionnaires was analyzed, good evidence (Bayesian posterior probability 86.9 %) was found.

However, [77] argues that "natural", as well as "realistic" or "authentic", refer to a specific outer reference, i.e. the percept elicited in the specific real environment should be identical to the percept elicited in the specific virtual environment. He suggests that plausibility extends this definition by referring to a percept elicited in the virtual environment, which could have believably been elicited in any comparable real environment. Therefore, he suggests to utilize "plausible" instead of "natural" for questioning subjects about the plausibility illusion. In an experiment, where subjects rated the plausibility of room reflections regarding a specific room size on a 100-point scale. The participant's judgments reach almost 100 for reflections matching room size and decline towards zero for reflections not matching room size. These results suggest that the method is suitable in identifying small to large differences in the elicited plausibility illusion.

The plausibility illusion was assessed by rating the attribute "plausible" by only a few studies. The attribute "plausible" was utilized to assess the plausibility illusion of virtual handshakes by [79]. The attribute "verisimilitude", which is a synonym for plausibility, was rated on a 6 point Likert scale for WBVs of passing trains by [80]. When the level of these vibrations was reduced from the level encountered in the real environment below the perceptual threshold, the plausibility ratings declined. Furthermore, when the audio-visual to tactile delay was increased or decreased, the plausibility ratings declined, too. Similar observations were assessed by [81], when applying low pass or high pass filters to WBVs encountered in vehicles. However, the validity of such assessments of plausibility can also be inferred from the conceptual similarity between quality and plausibility [77]. Both are a measure of the match between the elicited characteristics and the expected characteristics with respect to the totality of relevant perceptual features [1]. Rating scales are very frequently utilized to assess the perceptual quality of products [82], suggesting an agreement on the validity of such an approach. This suggests that measuring the plausibility with ratings scales is likely also a valid method for measuring plausibility.

2.3 Approaches for the Authoring of Vibrations

In the vibrotactile authoring process for virtual reality like applications, vibration is created with the goal of feeling “realistic” in the situational context presented. There are multiple approaches to the problem, which will be classified, according to the layers of tactile perception suggested in section 2.1.

2.3.1 Approaches on the Physical Layer

Vibration authoring on the physical layer can be mostly classified as the authentic approach. This can be realized by recording the vibration in a real environment and reproducing it in a physically accurate way in the real environment. However, this requires access to the real environment and measurement hardware such as acceleration sensors for recording the vibration, which is not always easy or even possible. The theoretically infinite variations of vibration encountered in scenes, quickly increases the required effort for recording vibration for a larger domain of scenarios to a practically infeasible level. An alternative to recording is the physical modeling of the excitation processes as well as the transfer functions. While this might be feasible for simple structures, models of increasingly complex real structures are often decreasingly accurate while requiring an increasing amount of computational

resources for e.g. boundary element or finite element simulations. This impedes the real time performance for the simulator.
If only optical or acoustical recordings were conducted, vibration might be reconstructed from the signals of the other modalities with an algorithm. This was attempted by [83] for optical recordings of helmet cameras e.g. for driving on a bicycle or motorcycle. However, the question is whether sufficient information for vibration synthesis is even contained in the video signal. For example, such an approach is limited by the low framerate of the video, determining the sampling frequency of vibration and thus the maximum frequency to be reconstructed. Furthermore, such an approach is dependent on the positioning of the camera. This also implies that the transfer function of camera mount will influence the vibration introduced into the camera and thus contained in the optical recording. This transfer function might be different from the transfer function of the shoulder neck section [30]. Apart from that, pitch of the head might lead to the same shift in video images as vertical movement. Indeed, [83] suggests that not all parameters can be estimated from the video alone, requiring additional heuristics. The results of the optical reconstruction algorithm suggest little improvement in the perceived realism of visual-tactile compared to visual only. Subjects remarked that they expected high frequency vibration for the motor cycle, which was not extracted by the algorithm. Based on the content potentially benefitting from vibration, [84] suggests the division into slow and fast point of view movement, discrete and continuous object movement, impulses and vibration from e.g. explosions, and context movement, e.g. vibration following a step on the gas pedal of car. While reconstruction is especially suitable for slow motion, the other classes likely require understanding of the semantic and spatiotemporal content of the scene. [85] suggests that such reconstruction can work for scenes with simple semantics but are difficult for more complex semantics.
Vibration may also be reconstructed from acoustical recordings. For example, vibration was reconstructed from the LFE channel of concert DVDs by [5]. Again, the sample frequency of the LFE channel is limited to 240 Hz and thus the maximum extractable vibration frequency is 120 Hz. Furthermore, frequencies below 20 Hz are inaudible and thus likely not included in the LFE channel. The results from this study suggest that reconstructed vibration is perceived as much less realistic than manually coded vibration.

Overall, the approaches on the physical layer are oblivious to the semantics of the scene and to the expectations of the users. Furthermore, without incorporating knowledge from the perceptual domain, the potential of such approaches for eliciting the plausibility illusion needs to validated. The general aim of approaches on the physical layer of abstraction is an authentic reproduction. The advantages and disadvantages of such an approach were discussed in section 2.2.2. Authentic reproduction that does not take into account tactile receptor capabilities would require physically accurate reproduction, which is likely to be difficult to achieve with simple reproduction systems. For many virtual environments, such as situations depicted in movies, no corresponding real environment is available as a reference and thus the “correct” vibration is unknown. Furthermore, an authentic reproduction is not necessarily the most plausible reproduction.

2.3.2 Approaches on the Mechanoreceptor Layer

Approaches to authoring content for virtual environments on the mechanoreceptor layer are mostly not feasible. If a perceptual codec would exist, it might be used to eliminate perceptually identical vibration of a database with recorded vibration. However, it might be suitable for augmented reality, where new information might be encoded in vibration, which requires taking into account receptor capabilities for encoding such vibration. In contrast, for authoring vibration for virtual environments the information intuitively encoded in vibration, i.e. the conveyed perceptual properties need to be known in order to incorporate them into the authoring process.

2.3.3 Approaches on the Sensory Layer

The sensory layer abstraction enables the incorporation of the user's perception into the authoring process, since vibration is interpreted as a carrier of intuitively elicited information. On this layer of abstraction users can communicate about the perceptual attributes anticipated to be elicited in a situa-

tional context presented in a virtual environment. However, the difficulty encountered on this layer, is the translation from the perceptual space into the engineering space.
The simplest method is a manual trial and error approach. The designer generates a vibration, assesses the perception in the situational context and iteratively improves the vibration signal. Such an approach is usually done in the auditory domain by Foley artists. With increasing experience, the designer will learn the underlying relationship between vibration and the elicited perceptual properties. Such an approach typically ensures favorable user judgments, as confirmed by the manual authoring of vibration for movies [5]. Unfortunately, this method is not very efficient requiring substantial time and thus cost for each new virtual environment.
In order to avoid repeating the manual vibration authoring process for similar situational contexts, it has been suggested to create databases with vibration effects [55]. The designer might use such a tool by simply querying the database of a suitable situation. For each situation, a e.g. rain drops falling or a buzzing motor, a vibration effect is created. However, this is still quite inefficient, since a new effect needs to be added for each new situational context. Furthermore, it is inflexible, since it does not take into account nuances in these situational contexts.
If the database does not characterize vibrations by scene descriptions, but by sensory-perceptual attributes elicited by the vibration, the efficiency may be greatly increased. Since vibration encountered in similar situational contexts would likely elicit similar attribute ratings, redundant assessment of vibration can be avoided. A designer can utilize such databases by querying a sensory-perceptual profile consisting of attribute ratings of each relevant sensory tactile perceptual attribute. Such an approach is presented by [60] for wrist vibrations with the intention of aiding product designers, who have similar goals. However, for each discrete vibration item entered into the database, a perceptual profile needs to be obtained by having users rate the new vibration. Furthermore, it is important to compensate the transfer function of the vibration reproduction utilized for the ratings, to ensure that the discrete mapping can be generalized to other vibration reproduction system.
Similar to the selection of vibration from databases on the physical layer of abstraction, the utility of selecting discrete vibration items by their discrete tactile sensory-perceptual profiles is limited. Ideally, the designer would like

to define tactile sensory-perceptual profiles on a continuous scale and to translate them into continuous vibration parameters. However, such a tool would require models of the relationship between physical vibration parameters and tactile sensory-perceptual attributes. The formalization of the plausibility illusion as a similarity judgment between the elicited perceptual object and the expected perceptual object in the sensory-perceptual feature space [77] offers a theoretical basis for such models. If users are able to reflect about the expected sensory tactile perceptual properties, it might even be possible to crowdsource the assessment of user expectations on the web by simply providing, e.g. verbal scene descriptions of the scenes for which vibration is to be synthesized. The designer might use the tool as following. For example, the vibration anticipated for a car driving over a cobblestone road would be rated by potential users as feeling very shaking. Subsequently, the model translates the sensation into engineering parameters. Since such a synthesis approach based on user expectations incorporates the semantic information of the scene, it would ensure the elicitation of the plausibility illusion. Unfortunately, models taking into account all relevant sensory tactile perceptual properties describing the range of everyday life vibration do not exist yet. The approach most similar to this authoring category was presented by [86]. For dual frequency sinusoidal vibration, they determined the relationship between amplitude of each frequency component and perceived intensity as well as perceived roughness. Based on this relationship, vibration with controlled sensory tactile perceptually properties were generated. Instead of assessing the intensity and roughness anticipated by the users for a specific situational context, they based their synthesis on auditory loudness and roughness extracted from an audio signal of that context. However, the plausibility (vibration matching audio-visual content) performance of the perception level vibration synthesis was worse than the direct audio signal synthesis for movies and music. Furthermore, the necessity of using different models for different situational context types (games vs. music) demonstrates that finding a universal relationship for the translation from the auditory to the tactile domain is difficult. Overall, this highlights the fact that the expected sensory tactile perceptual properties need to be explicitly assessed to avoid mismatching sensory cues and thus to benefit fully from such perceptual models.

2.3.4 Approaches on the Affective Layer

Theoretically, vibration authoring might be conducted on the affective layer, since the ultimate goal is often to affect the emotions of potential users. Such a method has been attempted by [12]. Their results suggested that the identified mappings between the physical tuning parameters and the affective attributes are dependent on the base vibration. Since especially the pleasurable dimension is likely context dependent as argued for the auditory domain by [64], it would require situational context specific models, limiting their practical usability of such authoring.

2.4 Summary

In the introduction chapter it was laid out that plausibility judgments as well as quality judgments compare the elicited properties to the expected properties. The goal of this work was defined as systematically fulfilling user expectations on WBV in virtual environments. Thus, it is necessary to communicate with users about their perception or expectations of vibration. First, tactile perception was interpreted as multi-layered process. Each layer was examined according to prior research for its potential in facilitating this goal. On the physical layer, a fully objective description of vibration can be obtained. However, communication about physical vibration properties requires expert knowledge. Laypersons rather rely on their perception of vibration to communicate about vibration. but the physical layer ignores perceptual factors. Thus, this layer of abstraction is impractical for communication with laypersons about the perception and expectation of vibrations.
On the mechanoreceptor layer, the capabilities of tactile receptors in resolving changes in vibration are taken into account. On this layer of abstraction the perceptual space of vibration was examined, providing insights into potentially important perceptual dimensions. However, no interpretation of these dimensions by laypersons is provided by such studies. Therefore, this layer of abstraction is impractical for communication with laypersons about the perception and expectation of vibrations as well.

On the sensory layer, the interpretation of these perceptual dimensions by users is taken into account. Thus, this layer of abstraction is promising for the defined goal of this work, as it enables explicit and quantitative communication with laypersons about elicited and expected perceptual vibration properties (e. g. “tingling”). However, prior studies on this layer of abstraction each only mapped a fraction of the perceptual space, due the utilized stimulus sets covering only a fraction of everyday life exposure to vibrations. Especially for WBVs, these dimensions have not been identified yet. The disadvantage of communication with users in free association tasks are the often disparate descriptions about the perceptual dimensions. In order to efficiently communicate with laypersons on this layer, it is necessary to standardize the utilized attributes in a way that they sufficiently represent all relevant perceptual dimensions.

On the affective layer, the affective dimensions of vibrations such as arousal or comfort are in focus. While communication about affective properties with laypersons is surely possible, such ratings are likely context dependent. Thus, identical vibration may elicit different affective ratings in different situational contexts. When attempting to find physical vibration parameters that elicit specific perceptual attributes, a separate model would be required for each situational context Therefore, this layer is also impractical for the scope of this work.

Subsequently, perception of virtual environments was examined and two factors place illusion and plausibility illusion were outlined. Ignoring intraindividual differences, place illusion mainly depends on how the virtual environment is perceived, which is directly influenced by the technical capabilities of the reproduction, tracking and rendering, etc. The plausibility illusion is created by the content of the depicted environment matching the user expectations. The plausibility illusion was formalized as a similarity judgment between elicited perceptual object and an inner reference, i.e. an expected perceptual object in a n-dimensions perceptual space. By minimizing the Euclidean distance between the elicited and the expected perceptual object it is possible to maximize the plausibility illusion [77].

Previous approaches of authoring tactile content for virtual environments have mostly ignored user expectations. Such approaches either attempted to build huge inflexible databases of vibrations for each situational context of interest or to extract a suitable vibration signal from the optical or acoustical signal. However, since the semantics of the scene is not taken into account, these approaches cannot guarantee that a plausible vibration is produced. Furthermore, it was shown that the expected tactile perceptual object of a situational context can be different from the elicited tactile perceptual object, e.g. in the case of a virtual basketball game.

An universal approach to authoring plausible vibration that is centered on user expectations is still lacking. The formalization of plausibility by [77] might enable such an approach. However, the implicit dimensions of the perceptual space extracted with an MDS offer no insights on their interpretation by laypersons, as argued in section 2.1.2. Without perceptual attributes associated with these dimensions, no communication about the inner reference from the system of classes involved in the plausibility judgment according to [77] is possible. Thus, the position of the inner reference in the perceptual space cannot be assessed and the abstraction of plausibility as a distance measure by [77] offers little practical benefit. Finding suitable perceptual properties e.g. by semantic differentials, which are a reasonable abstraction of classes shared across individuals, is an open topic of research according to [77].

In order to utilize this abstraction of plausibility for vibration design, it is therefore necessary to identify relevant tactile perceptual properties, which are interindividually shared across classes (i.e. situational contexts) and across individuals. However, the ecological approach to perception suggests that everyday life vibration intuitively elicits perceptual properties, which are explicitly verbalizable. It is very likely that prior experience of vibration encountered in everyday life across many situational contexts is indeed shared by the population to a large extent, and thus similar classes as well as perceptual properties. Based on the sensory layer of tactile perception and the formalization of plausibility, a research concept will be laid out in the following.

3. Research Concept

3.1 Research Questions

In the course of the first chapter, the importance of a systematic approach to fulfilling user expectations for tactile product design as well as for creating plausible vibration for virtual environments was elaborated. In the second chapter existing approaches were analyzed. It was expounded that each approach has specific limitations. Generating vibration from audio or video signals is not universally possible. Creating vibration databases of all existing situations is inefficient. Furthermore, it only works if there is a corresponding real environment, and if expected vibration properties of the virtual environment do not differ from vibration properties in the real environment. Various findings of research fields with relevance to the goal of vibration design were analyzed in chapter two. By combining these conceptually related findings to understanding perception, a research concept for the creation of a systematic, user-centered approach to tactile design for plausible vibration in virtual environments is developed in the following section.

3.1.1 Foundations of the Research Concept

Mel Slater suggested that two factors contribute to users interacting with virtual environments like real environments: the place illusion and the plausible illusion [6]. While the place illusion and the related concept of immersion rather depends on the presentation of the stimuli, the plausibility illusion depends on the content of the scene matching the user expectations. There are two general approaches to matching user expectations: the authentic approach and the plausible approach [74]. The authentic approach attempts to elicit a perceptual object in the virtual environment which is perceived to be identical to the perceptual object elicited in the real environment. Thus, in a paired comparison test, where users can rely on their tactile working memory, they would not perceive a difference. However, in most applications, the real environment cannot be presented in a short succession after the virtual environment or it cannot be presented at all. Thus, the plausible approach only attempts to elicit a perceptual object which might have occurred in a comparable real environment. Therefore, the user needs to rely only on his expectations, formed by previous experiences retrieved from long term memory, to decide whether the presented vibration is plausible or not.

Kuhn-Rahloff [77] argues that such a plausibility judgment can be interpreted as a similarity judgment. The elicited perceptual object is compared to the recalled inner reference defined by expectations on a specific context i.e., a scene in an n-dimensional perceptual feature space. However, utilizing the proposed relationship for plausible vibration design requires perceptual properties that are valid across contexts and individuals. The properties need to be valid across situational contexts i.e. scenarios to provide a benefit compared to the database approach i.e. assessing the vibration of the scenario individually. The properties need to be valid across individuals to enable an objectively valid assessment. Such perceptual properties represent a shared meaningful level of abstraction of intraindividual inner references that makes the interindividually comparable. However, he only hypothesized about such properties in the domain of sounds.

It seems likely that verbalizable perceptual properties enable communication about them and thus make them interindividually shared. To build upon this hypothesis, the verbalizable perceptual properties of the tactile perceptual space need to be identified. Most studies that investigated the vibrotactile perceptual space aimed at assessing the potential of the tactile modality for encoding and transmitting arbitrarily defined new information (see applications in section 2.1.2). Therefore, they focused on the tactile ability to distinguish vibrations differing in various physical parameters. The ecological approach contrasts this by assuming information to already be encoded in everyday life stimuli. The ecological approach to perception might be useful to take on this task. This ecological approach to perception was first described by Gibson for the visual modality [43]. It was applied to the auditory modality by Gaver [44], [45]. He argues that in everyday life sounds are a carrier of information about the environment.

In many everyday life situations, as e. g. in vehicles also vibration occurs. The driver of a vehicle is exposed to specific vibration depending on road conditions (tarmac road, cobblestone road, ...) or properties of the car (accelerating, idling, ...). Depending on the vibration properties, certain perceptual attributes like “bumpy” or “tingling” are elicited. Therefore, the ecological approach might also be applied to the tactile modality in which vibration is interpreted as a carrier of information about the environment. The approach implies the existence of a set of verbalizeable perceptual attributes. It also implies that these attributes are intuitively associated with everyday life vibration excitation patterns suggesting a direct relationship to physical vibration parameters. An exhaustive investigation of the vibrotactile perceptual space from the ecological perspective as well as on the the relationship between physical vibration parameters and tactile perceptual attribute ratings is still missing (see section 2.1.3).

3.1.2 Research Concept

Combining these findings, a research concept was developed. The goal is to demonstrate that plausible whole-body vibration can be universally synthesized from expected tactile perceptual properties. To achieve this goal three main problems had to be solved. The interplay of these tasks is shown in Figure 3.1.

First, a set of tactile perceptual attributes needs to be identified that are inter-individually shared and that can sufficiently represent the tactile perceptual space. Furthermore, it is necessary to assess whether there is a systematic relationship between physical vibration parameters and the suitability of these perceptual attributes in describing vibrations.

Second, the perceptual attribute set needs to be utilized to quantitatively rate tactile vibration properties of the provided situational context. It needs to be demonstrated that the expected tactile perceptual attribute profile of the inner reference of a context can be quantified by presenting a multimodal scene of this context with vibration. If the situational context is communicated without presenting vibration it should result in a similar attribute profile of the inner reference for this context.

Third, utilizing the previously identified relationship between vibration and the elicited perceptual attributes, a model can be created. This model should be able to translate an expected perceptual attribute profile of a context into

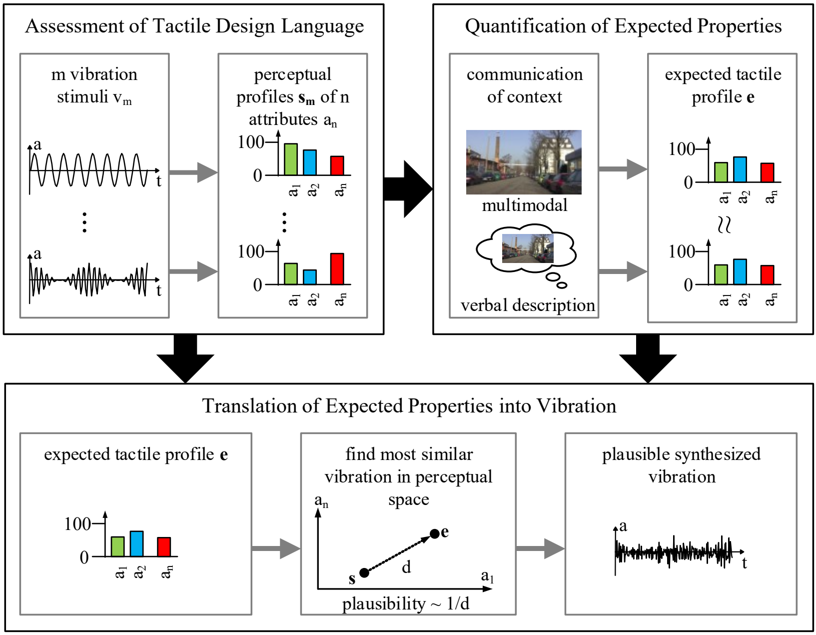

Figure 3.1.
Overview on the Research Concept

vibration parameters. Finally, it needs to be demonstrated that by eliciting the expected tactile perceptual attributes, the synthesized vibration will be perceived as plausible, thus validating the approach.

3.2 Limitations

The scope of this work will be clarified here to delimit investigations to be conducted in the course of this work. The research will focus on the haptic modality. The auditory and visual modality will only be utilized to communicate the situational context in the virtual environment. Furthermore, it will concentrate only on the vibrotactile subset of the haptic modality. The perception of vibration is often subdivided depending on the location of the introduction [15]. The focus will be set on subjects seated on vibrating surfaces i.e. WBV in contrast to hand-arm, feet, or finger vibration. Thus, spatial parameter variation will be excluded and only spectral and temporal parameters of the vibrations will be examined. The division into different locations of introduction is partly motivated by work safety research, which is not relevant due to vibration levels utilized in this work being below exposure limits [87]. However, the location of introduction also defines the reproduction system to be used since there exisits no single reproduction system for all locations of excitation.

In everyday life, WBV is most frequently encountered in situations involving vehicles. Therefore, the model validation will be conducted with scenes involving various vehicle situations. However, the relationship between perceptual attribute ratings and vibration parameters will be explored from a more general point of view, facilitating the transfer to other domains. WBVs can be subdivided into vertical, lateral, and fore-aft vibration, depending on the direction of excitation [87]. The perceptual threshold for vertical WBV is lower than for lateral and fore-aft vibration [16]. Apart from the static acceleration of vehicles in curves or of vehicles gaining and reducing speed, vertical vibration is more dominant due to the vehicle traveling orthogonally over surface irregularities. Indeed, [14] suggests for moving vehicles that vibration intensity of fore-aft and lateral vibration only amount to 50 % to 70 % of the value of vertical vibration. Due to the heightened sensitivity and the higher amplitudes usually observed, vertical WBV can be considered as potentially more relevant compared to lateral or fore-aft vibration and were thus selected as the focus of this study.

Due to the novel approach to tactile design, the emphasis was put on its general feasibility. Thus only segments of scenes were selected that would elicit quasi constant perceptual attribute ratings. Longer scenes with larger changes in perceptual attribute ratings over time would require corresponding changes in vibration parameters throughout the scene. Such scenes would need to be treated as a set of successive segments with constant perceptual attribute ratings. From these profiles, a set of vibrations could be synthesized. Subsequently, the transitions would need to be treated e. g. by applying crossfading to eliminate artifacts. Thus, longer scenes with dynamically changing perceptual attribute ratings as e. g. required for interactive simulations would be

potentially feasible. However, the optimization of this scenario should be investigated in a successive study after the general feasibility of the approach is demonstrated by this study.
Furthermore, only vibration at and above 1 Hz will be considered. Constant acceleration is produced by motion. When acceleration frequency increases above 0 Hz, vibration occurs. Vibration below 0.5 Hz can lead to motion sickness [15], especially when not presented synchronously with vibration. Thus, exact synchronization of movement with the visual stimuli would be required in that case.

4. Development of the Experimental Setup

The investigations conducted in this work required a suitable WBV reproduction system. Furthermore, it required additional capabilities in presenting acoustic and optic stimuli for conveying auditory and visual context in the virtual environment. The multi-modal measurement laboratory [88] of the Chair of Acoustic and Haptic Engineering provides such capabilities (see Figure 4.1). It consists of a separate control room and a visually and auditorily neutral room with concealed reproduction systems to facilitate immersion into the virtual environment. In section 4.1 the optic, acoustic, and vibration reproduction systems are described in detail. The software required for presenting multimodal stimuli and for conducting perceptual studies is outlined in section 4.2.

4.1 Hardware

4.1.1 Optical Reproduction System

The optical reproduction system provided a visual context for the tactile stimuli. Visual stimuli were presented with an Epson EMP-TW980 projector. It supports 1080p Full-HD resolution with 1920 x 1080 pixels and a brightness of 1200 ANSI-Lumen. For acoustical reasons, the device is encapsulated in a sound proof housing. Thus, the projector noise does not interfere with the audio reproduction. The image is projected onto an acoustically transparent screen with a width of 260 cm resulting in a diagonal of 118 inches at an aspect ratio of 16:9. The distance between the subjects seated on the motion platform and the screen is 340 cm. The image is centered on the visual axis of the test subject.

Figure 4.1.
Simulator for the presentation of visual, auditory, and tactile virtual environments.

4.1.2 Acoustical Reproduction System

The acoustic reproduction system provided auditory context for the tactile stimuli. The auditory stimuli were presented with a wave field synthesis system. Wavefield synthesis is an optimal reproduction system for virtual environments compared to classical reproduction systems such as stereo systems or headphones. In contrast to headphones reproduction, it does not require wearing a foreign object and thus there is no susceptibility of in head localization of sound sources potentially breaking immersion. Stereo or 5.1 setups rely on the phantom source effect to convey sound source direction. The phantom source effect is only stable in a relatively small listening position, thus requiring tracking of the listener position. Wavefield synthesis system recreates the wave field emitted by a sound source, thus enabling localization mostly independent from the listener position. Furthermore, it enables the reproduction of focused sound sources that can be positioned in front of the speaker plane.

The utilized wave field synthesis system was produced by ISONO (Barco). The system can be controlled from a MATLAB interface. Audio signals and sound source position streams of up to 32 sources can be provided to this interface. Audio signals are transferred from an RME HDSP MADI sound card and position streams are transferred over Ethernet to the IOSONO renderer. The renderer calculates the signals sent over 464 channels to the IOSONO amplifiers and subsequently to the 464 loudspeakers. The 464 speakers consist of 4 subwoofers, 116 mid-range, and 348 high range speakers with a distance of 6 cm. The mid-range and high range speakers are placed circularly at ear height on each wall. They are hidden from the view of the subject behind acoustically transparent panels to facilitate immersion. The superimposition of the speaker sound sources as provided by the renderer creates a virtual focused sound source at the desired position.

The wave field synthesis system was utilized for the presentation of scenes containing auditory stimuli in chapter 6 and chapter 7. The auditory stimuli were obtained in the process of recording the vehicle scenes with two integrated circuit piezo (ICP) microphones placed on-ear axis of the driver approximately 30 cm to the ears (see section 6.1.2). To reflect the recording setup in the vehicle, the recorded signals were presented as focused sound sources placed on-ear axis of the subjects in 30 cm distance. The reproduction system was calibrated at the subjects position on the seat on the motion platform. The transfer function was equalized to a flat frequency response. The reproduced sound pressure level was calibrated to the recorded sound pressure level.

4.1.3 Whole-Body Vibration Reproduction System

4.1.3.1 Requirements on the Reproduction System

The investigations conducted in this work attempt to cover the perceivable frequency and level range of WBV occurring in everyday life. Constant acceleration is produced by motion. When acceleration frequency increases above 0 Hz, vibration occurs. Vibration below 0.5 Hz can lead to motion sickness [15], especially when not synchronous with visual stimuli. Thus, a frequency of 1 Hz was selected as a lower frequency limit, clearly separating static motion from oscillatory motion i.e. vibration. The perceptual threshold is rising steeply above 300 Hz as shown in section 2.1.2. Very high-frequency WBV above 500 Hz can theoretically be perceived. However, vibration occurring in everyday life is unlikely to reach above the perceptual threshold over 500 Hz, as confirmed by the vehicle measurements in section 6.1.3. Therefore, a frequency of 500 Hz was selected as the upper-frequency limit for the investigations. The lower limit of the level range is determined by the perceptual threshold for vertical WBV (see section 2.1.2). The upper-level range is determined by the exposure limits for one-hour exposure to vibration [87] conforming to acceptable vibration levels. The reproduction system needed to cover this large frequency and level range. Unfortunately, no single reproduction system was available for this purpose. Thus, a reproduction system was created by combining a hydraulic motion platform suitable for low-frequency reproduction with an electrodynamic shaker for high-frequency reproduction. This system is shown in Figure 4.2.

4.1.3.2 Electrodynamic Shaker Subsystem

Electrodynamic shakers are well suited for vibration reproduction because of their favorable frequency response especially in the required mid to high-frequency range up to 500 Hz. Vibration signals are transferred from an RME HDSP MADI sound card to an RME M-32 digital-analog converter. Subsequently, the signal is amplified by an Alesis RA 150 amplifier. The analog signal is reproduced by an RFT Messelektronik ESE201 Type 11075 electrodynamic shaker, which is placed below the Recaro Pole Position seat made

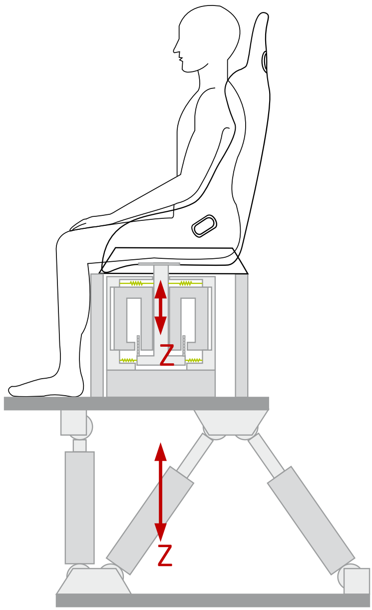

Figure 4.2.
Schematic of the combined reproduction system for the presentation of vertical WBVs consisting of a motion platform and an electrodynamic shaker.

of glass-fiber reinforced plastic. The voice coil of the shaker is attached rigidly to the horizontal seat surface enabling vertical vibration reproduction. The default foam rubber on seat surface was replaced by a thin piece of fabric to avoid vibration attenuation in the higher frequency range. The shaker placement is shown in Figure 4.3.

For each of the experiments, participants were instructed to sit comfortably in the seat with their thighs resting on the seat surface and to keep their sitting posture for the duration of the experiment. It is well known that interindividual differences in body parameters such as weight and height influence the dynamic interaction of the excitation system with the participant. [89] stressed the need to calibrate vibration reproduction systems in the actual tactile experiment condition. Differences of up to 10 dB have been measured in the body related transfer function for different persons as documented by [90]. Thus, it was necessary to compensate these interindividual differences to ensure controlled vibration presentation for the experiments conducted in

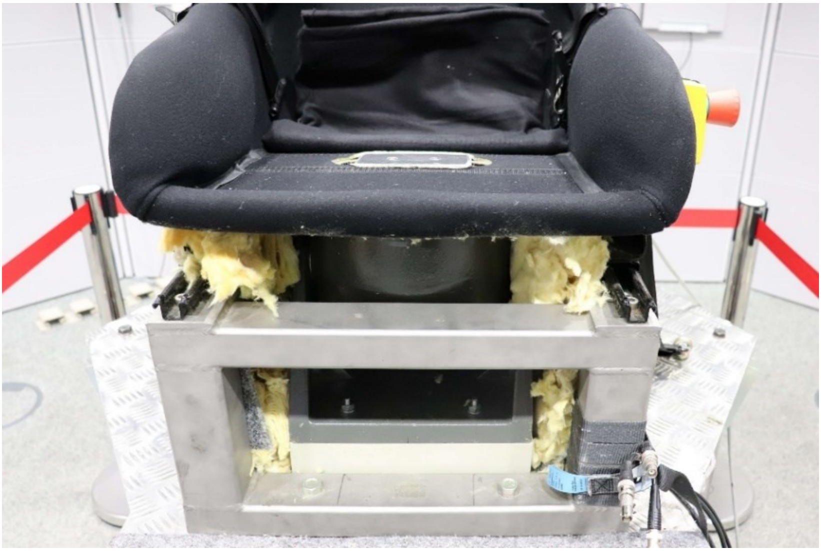

Figure 4.3.
Electrodynamic shaker placed below a glass-fiber reinforced plastic seat with a voice coil actuator coupled to the seat surface.

the course of this work. The transfer function between the input voltage and acceleration occurring at the seat surface was measured with a B&K 4515B seat pad accelerometer. The sensor was connected to a Zodiac DataRec 4 measurement frontend. The measurements were conducted in the application HEAD Recorder and were imported into MATLAB. From the measurements, inverse finite impulse response (FIR) filters were constructed. Each vibration signal was filtered using the *filtfilt* command for zero-phase filtering, thus compensating the interindividual differences in the transfer functions within a 2 dB range, which amounts to approximately the JNDL of WBVs (see section 2.1.2).

The shaker can be operated at a maximum RMS current of 4 amperes which limits the maximum acceleration. The acceleration level for the maximum current was measured with a sweep signal at each frequency. The resulting maximum acceleration curve is shown in Figure 4.4. It is obvious that the output is not sufficient for the low-frequency range, exceeding the sensation threshold only by 10 dB at 7 Hz and above. The transfer function of the shaker

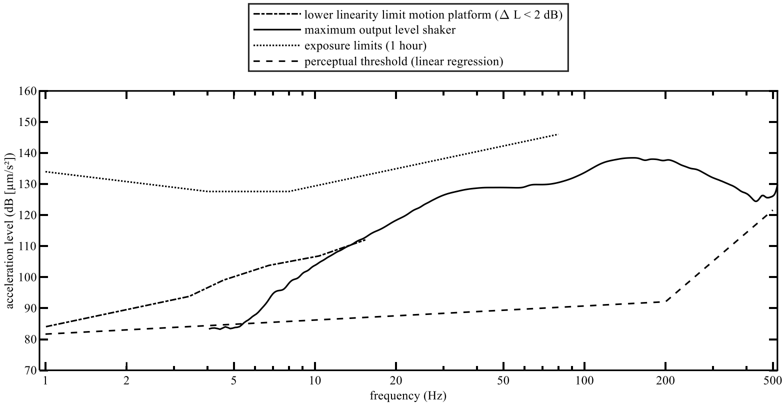

Figure 4.4.
Maximum output of the electrodynamic shaker and linearity limits for the motion platform with perceptual threshold and exposure limits for orientation.

is linear in the level range from the perceptual threshold to the maximum output level with a maximum deviation of 2 dB.
Due to the vertical stiffness of the seats' surface being much higher than the horizontal stiffness crosstalk of the vertical excitation to the lateral or transversal axis was negligible. Only slight harmonic distortion at the second harmonic was observed for sinusoidal excitation, which was at least 25 dB below the base frequency level. When interpreting the potential impact of the crosstalk, it should be noted that the perceptual threshold is approximately 10 dB lower for vertical WBV compared to lateral or transversal WBV [16]. Furthermore, a 10 dB reduction in vibration level is perceived to be less than half as intensive [32].

4.1.3.3 Hydraulic Motion Platform Subsystem

Due to their high elongations, hydraulic motion platforms are especially suitable for the low-frequency range. Therefore, a hydraulic Stewart motion platform was utilized for this range. The platform was manufactured by Oelhydraulik Hagenbuch. It is driven by six hydraulic actuators that are powered by a pump. The pump is placed in a separate room, which is acoustically insulated from the laboratory to attenuate the operating noise. The platform

can present motion and vibration in six degrees of freedom. It supports a maximum vertical elongation of 407.6 mm. The peak acceleration is limited to one g. The platform has an active control system, which is interfaced over a .NET API. Position values can be sent over Ethernet to the active control system, which renders the position to the six hydraulic actuators.
Many valve controlled actuators have a small dead zone, i.e. the actuator position output remains constant for very small input position changes. This property needs to be taken into account when reproducing low amplitude narrowband i.e. sinusoidal vibration, for which a superimposition of multiple frequency components doesn't produce a total position change which is relatively large in comparison to the dead zone. Controlling this property for the utilized motion platform, the transfer function was determined with sweep signals in 6 dB steps in the acceleration level range from the exposure limits to the perceptual threshold. The same measurement setup as for the shaker was utilized but with a Kistler 8305B10 capacitive acceleration sensor placed on the seat's surface. A nonlinearity threshold was defined for each level step at the frequency, where the deviation from the transfer function of the highest input level step was more than 2 dB (approximately the JNDF of vibration) below the expected value. The nonlinearity threshold curve is obtained from these thresholds (see Figure 4.4). The frequency and level range above this threshold can be presented by the motion platform system while the range below this threshold should be presented by the electrodynamic shaker.
In the utilized frequency range of the motion platform, interindividual differences in the transfer function were below 1 dB due to the active control system compensating weights of persons varying interindividually. Therefore, no transfer function compensation was necessary for this subsystem.
For near-threshold vertical vibration crosstalk to lateral and transversal acceleration was observed, which was close to the acceleration level of the vertical excitation. However, this property can be avoided by utilizing the shaker for near-threshold vibration level except for the frequency range below 5 Hz. Only slight harmonic distortion at the third and fourth harmonic was observed for sinusoidal excitation, which was at least 15 dB below the base frequency level. Restating the points brought up in the discussion of the electrodynamic shaker, a 10 dB reduction in vibration level is perceived to be less than half as intensive [32]. Furthermore, the perceptual threshold is approximately 10 dB lower for vertical WBV compared to lateral or transversal WBV with the

difference decreasing towards low frequencies [16]. When interpreting the results for the low frequency and low-level range, crosstalk should potentially be taken into account.
For the electrodynamic shaker, it was possible to input the acceleration signal directly. However, the motion platform required an elongation i.e. position signal instead of an acceleration signal. Since the acceleration is the second derivative of the elongation, it is possible to obtain the elongation signal by double integration of the acceleration signal. Double numeric integration was implemented utilizing the Simpson's rule of integration. Potentially present static acceleration needed to be removed for the acceleration signal to be reproducible in the limited movement space of the platform. A Chebyshev Type II high pass filter with a passband frequency of 1 Hz, a passband ripple below 1 dB, and 60 dB attenuation was applied to the acceleration signal using the MATLAB *filtfilt* command for zero-phase filtering. Chebyshev Type II filters were selected due to their steep edges and smooth passband at the cost of stopband ripple.

4.1.3.4 Properties of the Aggregate System

The maximum output of the electrodynamic shaker and the nonlinearity threshold of the hydraulic motion platform define the lower frequency and upper frequency limit of each subsystem. These limits are both dependent on the acceleration level as visible in Figure 4.4. Due to the negligible crosstalk of the electrodynamic shaker, this reproduction system was always utilized if possible, and the motion platform was utilized to cover the remaining level and frequency range.
For a defined narrowband stimulus level and frequency the reproduction system was selected depending on the maximum output of the electrodynamic shaker. Thus, for vibrations at 10 dB above the perceptual threshold, the motion platform was utilized below 4 Hz. For vibrations at 36 dB above the perceptual threshold the electrodynamic shaker was utilized above 15 Hz. For broadband stimuli such as bandlimited WGN or recorded vehicle acceleration, simultaneous utilization of both reproduction systems was necessary. The separation frequency between both systems was assessed by determining the fraction of the stimulus acceleration spectrum below the maximum output of the shaker. Subsequently, the separation frequency defined the passband

frequency of an infinite impulse response (IIR) high pass filter and an IIR low pass filter with 60 dB attenuation and below 1 dB passband ripple. By utilizing the MATLAB *filtfilt* command the acceleration signal was split into a low-frequency component reproduced by the motion platform and a high-frequency component reproduced by the electrodynamic shaker using zero-phase filtering. The superimposition of both reproduction systems resulted in the desired acceleration signal.

In the previous section, the reproduction accuracy was described from a general perspective for each reproduction system that can easily be applied to e.g. sinusoidal signals. More complex signals based on vibration measurements were reproduced on both systems simultaneously. In addition to that, they needed to be prepared for playback on the motion platform. The controlled reproduction of each vehicle vibration scene utilized in this work was important and therefore reproduction was individually verified. For stationary scenes, the maximum spectral differences between suprathreshold recorded vibration and reproduced vibration were always below 6 dB in each frequency sub-band (see Figure 4.5). For scenes with impulse-like vibration the maximum difference between recorded peak vibration and reproduced peak vibration was at or below 4 dB (see Figure 4.6).

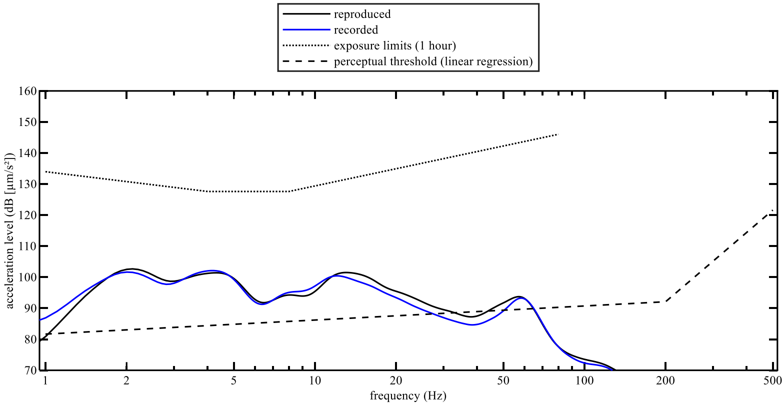

Figure 4.5.
Spectrum of the recorded and reproduced acceleration of driving on a cobblestone road at 30 kph with perceptual threshold and exposure limits for orientation.

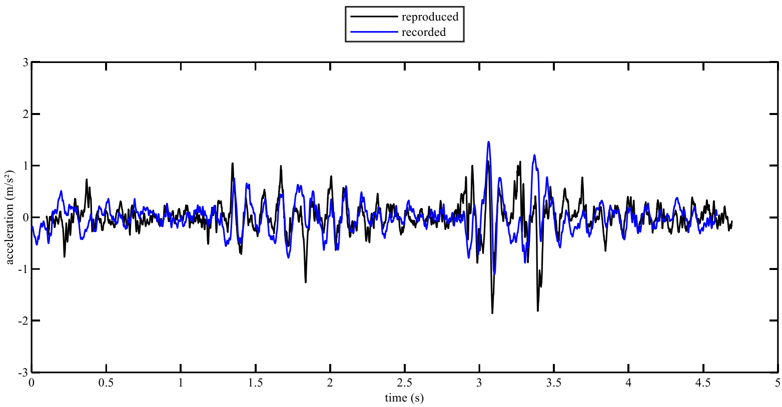

Figure 4.6.
Time signal of the recorded and reproduced acceleration of driving over a manhole at 30 kph.

4.2 Software

4.2.1 Combination of Reproduction Systems for Unimodal and Multimodal Presentation

For conducting the perceptual studies, it was necessary to present tactile as well as multimodal (auditory, visual, tactile) stimuli. No standard software was available for this purpose due to the non-standard reproduction systems. Therefore, a multimodal scene player was required that enabled the creation of multimodal virtual environments utilizing the hardware described in the previous section. Due to the scope of this work being on non-interactive scenes the simulation was limited to passive playback. At first, a multimodal scene format was defined as a MATLAB *struct* data type, which contained links to individual files for each modality. Optical signals were stored as MPEG-4 video files. Acoustical signals were stored as 32-bit stereo wave files. Vibratory signals were stored as comma-separated value files separately for the motion platform and the electrodynamic shaker.
Subsequently, a multimodal scene player was implemented in MATLAB to playback the scene format. It provided a central control to all reproduction systems enabling a synchronous reproduction of the signals. Facilitating this task, the player was split into two main functions: loading the signals associated with the various modalities and starting of the reproduction systems synchronously. The loading of the signal files preceded the playback to avoid relative delays being introduced by the loading itself. Optical signals were loaded to a windows media player ActiveX element. Acoustical signals were uploaded to the wave field synthesis server. Shaker signals were simply loaded into the working memory to be available for playback via the soundcard. Motion platform signals were uploaded to its control server.
After all the separate signal files finished loading, a callback was triggered to initiate the actual playback. The video and shaker reproduction were started on the control PC, the motion platform reproduction was started on its corresponding control server and the wave field synthesis reproduction was started on its corresponding control server as well. The main task of this method was to ensure that the playback is perceived to be synchronous. The auditory tactile asynchrony threshold was determined by [91]. For complex car stimuli

consisting of WBV and sound recordings, audio delays within the range of -47 ms to 63 ms are perceived to be synchronous. For the auditory visual-tactile scene to be perceived as synchronous these limits need to be adhered to. The auditory-visual asynchrony threshold was determined by [92]. For complex speech stimuli, an audio delay within the range of -131 ms to 258 ms is perceived to be synchronous.
Three potential sources for delays were identified. Due to the distributed architecture of the reproduction system, the playback on the motion platform server might run out of synchronization with the playback on the wavefield server. The remaining total delay between the first reproduction system start and the last reproduction system start was below 20 ms. However, the stimuli utilized in this thesis had a duration of approximately 10 seconds only. Therefore, a delay spread beyond the synchronicity thresholds was not an issue. Static delays were measured and subsequently compensated. Small dynamic delays, i.e. delays varying from playback to playback were also observed. The preemptive scheduling of the windows 7 operating system was identified as the main source of dynamic delays. The problem was reduced by setting the process priority of the player to real-time before the starting procedure was initiated.

4.2.2 Conducting Perceptual Studies

One of the core tasks of this thesis was the assessment of perceptual judgments of the various experiments. Therefore, a framework was implemented in MATLAB, enabling efficient handling of perceptual studies. A brief overview of the core features is provided. First, an experiment configuration is set up. Various experiment types as free association task or semantic differential can be selected. Furthermore, a set of scenes (in the previously defined scene format) is provided, which will be presented in the experiment. At the beginning of each experiment, the experiment configuration is loaded. Subsequently, the participant is registered for the experiment configuration and his or her age and sex are collected and an individual stimulus randomization is generated. As mentioned before, the transfer function of each subject was assessed, an inverse filter of the individual transfer function generated and applied to all vibration stimuli. In the evaluation phase, one stimulus was rated per trial. First, the participant played back the stimulus at least one time. Subsequently, the perception of the subject was assessed. For the free interview experiment, associations could be entered as a comma-separated list. A semantic differential was used to assess selected perceptual properties on rating scales. The chosen quasi-continuous Rohrmann scale [93] (see Figure 5.11) was implemented as a slider attached to a scale image to rate the suitability of each perceptual property in describing the presented stimulus. The slider knob was only shown after a click on the slider to avoid a rating bias towards the initial knob position. At the beginning of each trial, the slider knob was hidden again.

5. Assessment of a Sensory Tactile Design Language for Characterizing Vibration

The goal of this thesis is to demonstrate that plausible whole-body vibration can be universally synthesized from expected tactile perceptual properties. The basis for the proposed approach is a vibrotactile design language that enables communication about the relevant tactile perceptual properties. [1]

5.1.1 Design Language Requirements

No universally agreed upon tactile design language exists yet [10]. There are several requirements for a vibrotactile design language that could facilitate the synthesis of vibration. Vibrotactile design would be far simpler if non-experts could communicate in engineering terms such as level or frequency directly. Only relying on expert judgments does not necessarily result in perceptual properties preferred by the majority of future users. However, potential users are mostly laypersons which are not able to communicate these engineering parameters without prior training [99]. Thus, vibrations need to be described with a vocabulary understandable to laypersons. According to the ecological approach to perception, vibration occurring in everyday life are a carrier of information about the environment. Therefore, there should be associations intuitively and frequently elicited by vibration by all persons including laypersons. Only such perceptual attributes would enable laypersons to explicitly and efficiently communicate about their vibrotactile experiences since no prior explanation is required.
If different laypersons use different associations, then it becomes difficult to generalize the perceptual properties into properties preferred on average. Frequently used, attributes maximize the potential to be interindividually understandable. However, they should not only be understandable across different

[1] Parts of this work were presented in [93], [94], [95], [96], [97] and [98] by Rosenkranz et al.

individuals but also across different situational contexts while still maximizing conveyed information (see section 2.1.3). Therefore, the perceptual properties should be mostly independent of the situational context i.e. applicable for describing a wide range of scenes with vibration while still accounting for differences between vibrations of different scenes. On the one hand, very specific associations or metaphors (e. g. “cobblestone road”) have only a narrow applicable domain and would not offer much insights about most situations outside their domain. On the other hand, if such attributes are too general (e.g. “vibration”), they would be unsuitable for describing specific perceptual properties.
Another requirement is the attribute's suitability in enabling translation into physical vibration parameters. Attributes relating exclusively to other, non-vibrotactile modalities provide little information about vibration parameters. It has been suggested by [13] and [12] that tactile perceptual attributes can be understood as a multilayer structure with engineering parameters mapping onto sensory-perceptual attributes mapping onto affective attributes. Affective attributes in general and e. g. “pleasantness” in particular are likely influenced by the situational context. Sound quality is known to be influenced by situational context [64]. This suggests that identical vibration might be judged as pleasant in one context and as unpleasant in another context. Furthermore, affective attributes might differ between groups of users as e. g. indicated by different preferences of sports car drivers and limousine drivers with regards to sound level judged as pleasant. For such attributes, it is therefore difficult to find universal relationships to physical vibration parameters independent from situational context. Affective attributes can be regarded as an emergent property of lower-level sensory-perceptual attributes and the situational context. Sensory-perceptual attributes such as “tingling” have a direct relationship to the perceptual properties elicited by vibration and are thus more likely independent from the situational context and of user group preference. Thus, they are more suitable for forming a tactile design language from which vibration parameters can be derived universally i.e. in the form of models.
To translate physical vibration parameters, the design language needs to enable explicit, quantitative communication about vibrotactile perceptual properties. Therefore, psychophysical knowledge about how similar a stimulus is to a chosen reference is not sufficient since it reflects only implicit perceptual

dimension (see section 2.1.2). The ecological approach to perception (see section 2.1.3) implies that from our everyday life experiences we have learned to associate specific explicitly verbalizeable tactile perceptual properties (e. g. “tingling”) with specific vibration intuitively. It also implies that the range of everyday life experiences of laypersons with vibration provides an inner reference against which the perceptual properties elicited by vibration can be absolutely quantified.

Furthermore, the tactile design language should enable efficient and effective communication about sensory-perceptual properties. Therefore, it should not contain redundant attributes while still explaining the majority of variance in sensory-perceptual properties. The effectiveness of the design language depends on whether its elements sufficiently explain all prominent sensory tactile perceptual properties. The efficiency of the design language depends on whether it contains only necessary elements for explaining all prominent sensory tactile perceptual properties. Since the required effort of profiling perceptual properties rises with the number of attributes to be rated synonymous, redundant attributes should be avoided. Thus, the set of sensory-perceptual attributes should be compact to minimize the required effort of profiling perceptual properties.

5.1.2 Method to Assess the Design Language

There are three approaches to investigating the tactile sensory-perceptual space according to [19]. In the classification method, subjects categorize a set of stimuli into perceptually similar groups, with each stimulus belonging to one group without explicit attributes attached. The result is only nominal scaled judgements, i.e. non-nuanced judgments of each stimulus about implicit properties. Thus the requirement of finding a set of attributes which does enable explicit quantification of sensory tactile properties cannot be achieved with this method.

In the similarity estimation method, participants rate the pairwise similarity of the stimulus set. Based on these similarity judgments MDS can reveal the underlying, implicit perceptual dimensions (see section 2.1.3). However, labels need to be manually and subjectively attached to the implicit dimensions

[23]. Thus, this approach does not produce attributes for explicit communication with laypersons. The inherently relative similarity judgments can also not be directly transformed to mappings between vibration and absolute attribute ratings. Thus, this approach is not suitable for finding a tactile design language fulfilling the stated requirements.
Based on the semantic differential method an approach was developed by [99] for the auditory domain for finding attributes with relevance to quality judgments of vehicle sounds. The approach adopted to the tactile domain by [62] for finding quality attributes for vehicle idle mode vibration. They named the approach "multi-sequential systematic approach". The first step is a free association task to collect the subject's associations on a set of stimuli. Subsequently, individually elicited attributes are checked for interindividual understandability. The subset of understandable attributes is reduced by the attributes that subjects judge to be unsuitable for describing the stimuli in general. Afterward, ratings of the remaining attributes are obtained for each stimulus. A PCA aids in the elimination of redundant attributes. Finally, the completeness of the attribute set is validated by comparing similarity judgments of stimulus pairs to rating differences of stimulus pairs. If attribute ratings can explain all the differences, the attribute set is complete.

5.1.3 Goals of this Chapter

This chapter will assess a sensory tactile design language according to the stated requirements for representing the domain of WBV with a focus on vehicle vibration. It will adapt the approach suggested by [62] to the goals of this thesis. In Figure 5.1 an overview of the single steps is shown, which are described below in detail. Previous studies have not investigated a WBV stimulus set sufficiently representing everyday life experience with vibration. Thus, the first step will be the selection of a stimulus set of WBV representing the variation encountered in everyday life, suitable for eliciting all relevant sensory-perceptual attributes. Due to potentially infinite vibration stimuli, a method needs to be developed to reduce the necessary number of stimuli to elicit all relevant perceptual attributes.

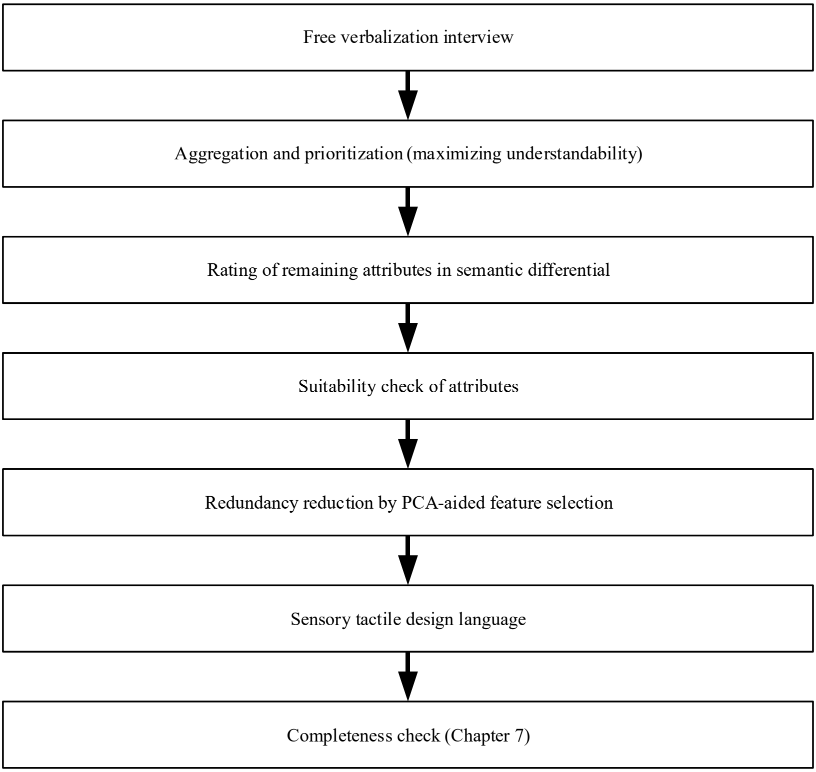

Figure 5.1.
Overview over the successive steps to assess the sensory tactile design language.

Since it is not only interesting which attributes are elicited by specific vibrations but also how well the attributes describe these vibrations, each attribute will be rated for each stimulus. The results can be utilized for three tasks. First, an implicit suitability check can be conducted for each attribute by examining rating differences between stimuli. If an attribute does not show differences between stimuli, it is not suitable for deriving vibration parameters. Second, the attribute ratings enable a PCA to eliminate redundant attributes and thus to determine a compact sensory-perceptual attribute set forming the design language. Third, the attribute ratings provide the basis for building the synthesis model in chapter 7.

A completeness check similar to [62] would require the assessment of similarity ratings of all stimuli pairings, requiring an enormous number of similarity judgments. However, completeness can also be validated in the course of the synthesis model validation in chapter 7. If the attribute set is complete, then it should be possible to obtain ratings only of these attributes for a vibration of any situational context and translate the ratings into a vibration. If synthesized vibration and recorded vibration both elicit the same perceptual attribute ratings and if they are perceived as equally plausible, then the attribute set is complete with regards to describing the sensory-perceptual properties of vibration.

5.2 Tactile Stimuli

As argued about the shortcoming of the semantic differential method in the previous section, the coverage of the stimuli set will affect the attributes elicited. If the goal is to elicit a sufficient set of attributes that can explain the perceptual properties of WBV, then potentially the majority of WBVs occurring in everyday life needs to be presented. Obviously, it is difficult if not outright impossible to record vibration of potentially infinite variations. Therefore, the first section of this chapter will develop a method to reduce the infinite amount of stimuli to a finite stimulus set while ensuring all sensory-perceptual properties are reflected by the elicited attribute set.

5.2.1 Generalization into Excitation Patterns

In everyday life, WBV i.e. vibration perceived by seated subjects is predominantly encountered in situations involving vehicles i.e. ships, aircraft, trains, and cars. However, recording all potential variations of vibration encountered in such situations would require tremendous effort which makes this strategy infeasible. According to the ecological approach to tactile perception, vibration is a carrier of information about the environment. Therefore, vibration with certain temporal spectral properties will elicit certain perceptual properties verbalizable in the form of sensory tactile perceptual attributes. Gaver

[44] argues that the ecological approach to perception falls into the intersection of psychophysics and cognitive psychology. One of the major research areas of cognitive psychology that is concerned with the associations elicited by stimuli, is the process of categorization. Rosch [47] provides a definition of this perceptual process:

> “The world consists of a virtually infinite number of discriminably different stimuli. One of the most basic functions of all organisms is the cutting up of the environment into classifications by which nonidentical stimuli can be treated as equivalent.”

This suggests that it is not necessary to assess all WBVs potentially occurring in everyday life. Instead only one vibration signal for each sensory attribute that needs to be elicited to enable description of all sensory-perceptual properties is required. Furthermore, the vibration presented does not need to be perceptually identical and thus not physically identical to vibration encountered in everyday life. Especially for an elicitation task the presented vibration only needs to be sufficiently similar to elicit the sensory-perceptual property, since laypersons can likely only come up with intuitive associations for vibrations similar to vibrations previously encountered in everyday life.
Gaver [44] suggests to generalize ecological stimuli according to their excitation process for the investigation of the elicited perceptual attributes. Griffin [15] classifies vibration typically encountered in seven classes (see Figure 5.2 in section 2.1.1). The focus of this work lies in vibration segments that would elicit quasi constant sensory-perceptual attribute ratings (see section 3.2). Four of the seven categories fit into the scope of this thesis: sinusoidal, multi-sinusoidal, shock, and stationary ergodic. Arguably in our everyday lives, we most frequently encounter WBV in cars. The four categories can be found in car vibration indeed as evident from Figure 5.2.
Sinusoidal vibration is produced by periodic mechanical processes e. g. regular road structures such as cobblestones with the same diameter. Multiple correlated periodic excitation processes produce acceleration signals with an

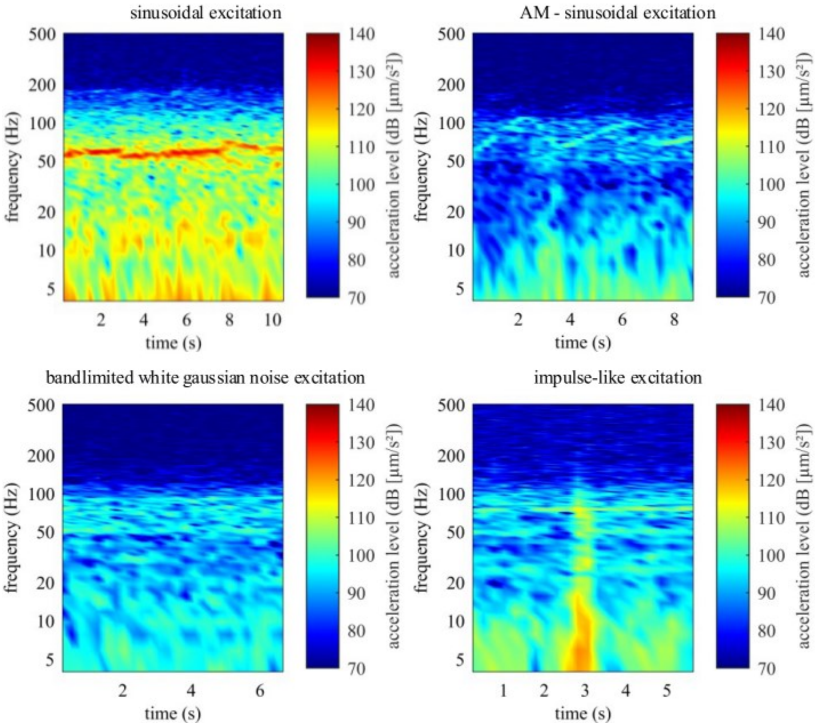

Figure 5.2.
Spectrograms (short-time Fourier transforms (STFT): 8,192 samples, 50 % overlapping Hann windows) of the vertical acceleration of four different vehicle scenes exhibiting four excitations process: sinusoidal, AM-sinusoidal, bandlimited WGN and impulse-like vibration.

envelope that shows regular temporal variation i.e. modulation, e.g. motor orders of combustion engines. Thus, this excitation process can be represented by amplitude modulated (AM) vibration. Broadband vibration such as stationary stochastic vibration is caused by a superimposition of uncorrelated sources. Impulse-like vibration is caused by shocks when e.g. driving across singular road irregularities such as manholes. If the single event is considered in isolation, it can also potentially be described by a quasi-constant perceptual profile.

It was considered to also include the superimposition of multiple vibration signals. The frequency selectivity of tactile receptors with a JNDF of about

30 % or more [30], [100] and masking effects [35], [101] limit the ability of subjects to resolve spectrally more complex structured vibration. However, if two vibration components are separated further in frequency, it would need to be considered to include superimpositions of multiple vibration signals. [22] investigated such a superimposed excitation pattern consisting of a low frequency and a high-frequency sinusoidal vibration. They demonstrated that the low-frequency component dominates the overall perception. This might suggest a saliency effect, in which the percept elicited by one vibration component dominates the overall percept of a superimposed component. Therefore, no superimpositions of multiple vibration signals were included.
The required stimulus set should cover scenes with vibrations of each of the four categories. Simply recording one scene of each of the four categories will not ensure to elicit all perceptual properties, since also the variation in these categories needs to be taken into account. However, the abstraction of vibration into excitation processes enables the systematic variation of the parameters defining the excitation process.

5.2.2 Definition of Parameter Values of the Excitation Patterns

Previously assessed psychophysical properties of tactile perception enable the definition of parameter ranges of perceivable WBV ranges encountered in everyday life for each of the four excitation process. Therefore, a finite stimulus set can be constructed while ensuring all sensory-perceptual properties are reflected by the attribute set elicited by these stimuli.

5.2.2.1 Sinusoidal Excitation Pattern

The sinusoidal excitation pattern shown in Figure 5.2 can be generalized according to the formula:

$$a(t) = A \sin(2\pi f t) \quad \textbf{5.1}$$

$$a(t) = 10^{-6} 10^{\frac{SL(f)+3dB}{20dB}} \sin(2\pi f t) \quad \textbf{5.2}$$

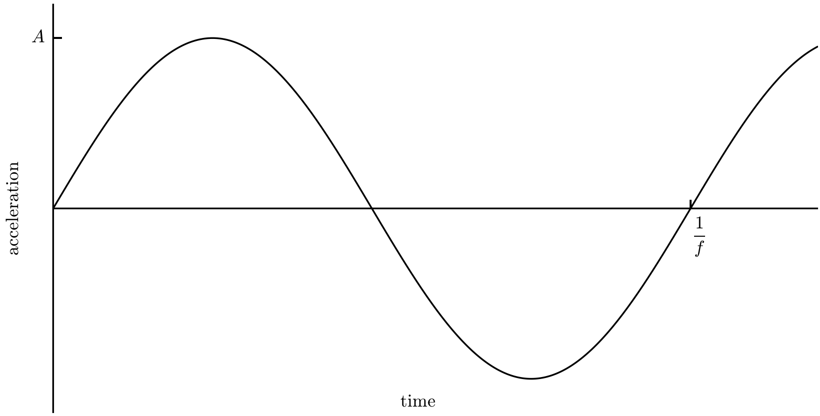

Figure 5.3.
Sinusoidal excitation pattern.

The resulting signal (see Figure 5.3) is characterized by the peak acceleration amplitude A represented by RMS sensation level SL and frequency f.
Figure 5.4 provides an overview of the sinusoidal stimuli. The level range is limited by the perceptual threshold [16] towards lower acceleration levels The exposure limits [86] for one-hour exposure with one hour being a typical session duration for participants to provide a reasonable upper boundary for everyday life WBV. The perceptual threshold of vertical WBV is frequency-dependent which is comparable to the auditory threshold [28]. Equal intensity curves of vertical WBV are approximately parallel to the perceptual threshold [102]. To enable comparisons of attribute ratings at different frequencies but the same perceived intensity in section 5.3, it is reasonable to define the vibration level relative to the perceptual threshold. Indeed, many psychoacoustic studies define the stimulus level as sensation level (SL) [28] relative to the threshold of hearing. The perceptual threshold of vertical WBV was approximated as a two-part linear regression based on the data of [16] for frequencies below 315 Hz and on [31] for frequencies above 315 Hz to account for a steeper rise towards higher frequencies beginning at 315 Hz.
The utilization of average perceptual thresholds for the definition of stimulus levels has the advantage of presenting identical stimuli for each participant. The disadvantage lies in the perceptual threshold varying across participants which potentially leads to some stimuli being below the perceptual threshold

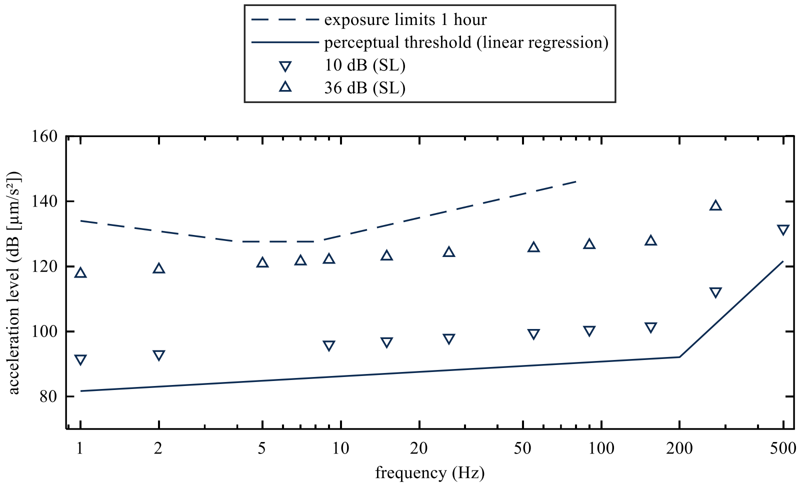

Figure 5.4.
Selected parameter value combinations for the sinusoidal excitation pattern.

for some of the participants. Therefore, 10 dB (SL) were selected as a compromise between covering the lower level range and stimuli being below the perceptual threshold for only a small fraction of participants. A level of 36 dB (SL) which was just below the one-hour exposure limits [86] was chosen. The JNDL for vertical WBV of about 1 dB provides a lower bound for meaningful level increments for the stimuli [34]. However, filling the interval between 10 dB (SL) and 36 dB (SL) in twice the JNDL increments for clearly distinguishable stimuli would have produced very many stimuli. Previous studies for WBV only presented one level [50], [61]. In combination with insights gained from a small preliminary test, it was concluded that small level differences would not likely change whether an attribute is elicited at all. Therefore, only two levels were selected with 10 dB (SL) representing a low perceived intensity and 36 dB (SL) representing a high perceived intensity.
The lower limit of the frequency range for vibration is about one Hz, since below this frequency oscillator motion will converge to non-oscillatory motion. The upper limit is determined by the perception threshold rising steeply above 315 Hz [31]. Thus, it is reasonable to select an upper limit of 500 Hz, above which everyday life WBV is likely below the threshold, as confirmed by the vehicle measurements in section 6.1.3. The JNDF for vertical WBV of

approximately 30 % provides a lower bound for meaningful frequency increments for the stimuli [30], [100]. For frequencies below 20 Hz, the JNDF is lower than 30 % suggesting a higher resolution than above 20 Hz. To create clearly distinguishable stimuli frequency intervals of about double the JNDF were selected. The frequencies of the stimuli were selected in such a way that they extended the stimuli range of the study [61]. Since reproduction in the limits defined in section 4.1.3 could not be achieved for the combinations 7 Hz and 9 Hz with 10 dB (SL) as well as 500 Hz at 36 dB (SL), they were omitted.

5.2.2.2 Amplitude Modulated Sinusoidal Excitation Pattern

The multi-sinusoidal category shown in Figure 5.2 can be represented by an AM-sinusoidal excitation pattern according to the formula:

$$a(t) = A\,(1 + m\cos(2\pi f_m t))\cos(2\pi f t) \tag{5.3}$$

$$a(t) = 10^{-6} 10^{\frac{SL(f)+3dB}{20dB}} \left(1 + m\cos(2\pi f_m t)\right) \cos(2\pi f t) \tag{5.4}$$

The resulting signal (see Figure 5.5) is characterized by the peak acceleration amplitude A represented by RMS sensation level SL, carrier frequency f, modulation frequency $f_{m,}$, and modulation index m.
To enable comparisons to the stimuli of the sinusoidal excitation pattern, parameter values of the AM-sinusoidal stimuli were selected accordingly. The RMS levels of the AM-sinusoidal stimuli were selected to be identical to the levels of sinusoidal stimuli with the same frequency. The carrier frequency values were selected to be identical to the frequency of the sinusoidal stimuli. AM-sinusoidal excitation patterns are characterized by a carrier frequency modulated with a relatively lower modulation frequency. Therefore, the modulation frequency values were not identical for all carrier frequencies but instead were chosen to always be a fraction of the carrier frequency. As a consequence, the carrier frequencies of 1 Hz, 2 Hz, 5 Hz, and 7 Hz were dropped. 2 Hz was selected as the lower limit for a modulation frequency to be clearly distinguishable from unmodulated vibration. With increasing modulation fre-

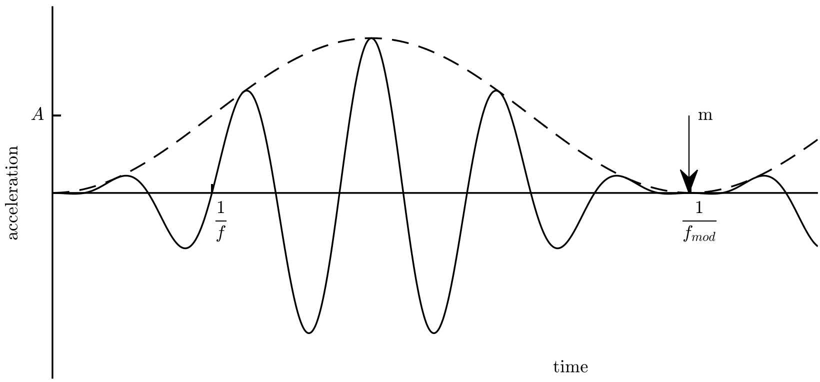

Figure 5.5.
AM-Sinusoidal excitation pattern.

quency, AM-sinusoidal stimuli will also become increasingly similar to sinusoidal stimuli [21]. Therefore, 26 Hz was selected as the upper limit for a modulation frequency to be clearly distinguishable from unmodulated vibration. In general modulation frequency values were selected to be identical to the carrier frequency values, but with modulation frequency being a fraction of carrier frequency in particular. As argued for the modulation frequency, a modulation depth of 0 would be equivalent to sinusoidal excitation. For AM-sinusoidal stimuli to be clearly distinguishable from sinusoidal stimuli a modulation depth value of one was chosen.

Figure 5.6 provides an overview of the AM-sinusoidal stimuli for the high SL. In a preliminary experiment, all combinations of carrier and modulation frequency were assessed for their similarity. To keep the stimulus set compact, the stimuli with carrier frequencies of 55 Hz, 275 Hz, and 500 Hz were omitted since they were perceived to be very similar to the remaining stimuli with the carrier frequencies of 9 Hz, 15 Hz, 26 Hz, 90 Hz, and 155 Hz.

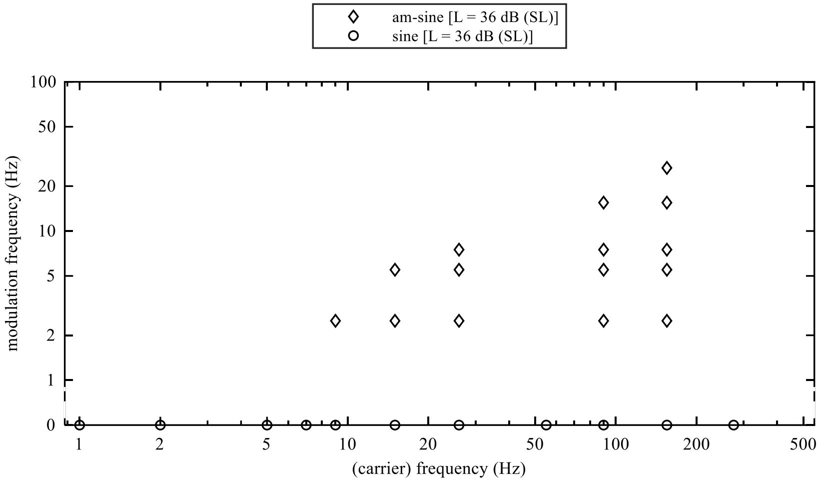

Figure 5.6.
Selected parameter value combinations for the AM-sinusoidal excitation pattern for 36 dB (SL).

5.2.2.3 Bandlimited White Gaussian Noise Excitation Pattern

The stationary ergodic stochastic excitation pattern shown in Figure 5.2 can be represented by bandlimited WGN. The resulting signal (see Figure 5.7) is characterized by the RMS sensation level SL, center frequency f_c, and bandwidth f_b. To enable comparisons of RMS acceleration level to the stimuli of the sinusoidal excitation pattern, parameter values of the narrowband noise stimuli were selected to 10 dB (SL) and 36 dB (SL) accordingly. Most studies on perception thresholds were conducted with sinusoidal vibration, but no data was available for bandlimited noise. A study investigating perceived discomfort [33] suggests that the perceived intensity of sinusoidal vibration is comparable to the perceived intensity of band-limited noise with the same RMS level with both vibrations falling into the same frequency range. Therefore, the perception threshold at the center frequency of the bandlimited noise vibration was inferred from the perception threshold of sinusoidal vibration. The bandwidth and the center frequency were varied in combination. The frequency range of sinusoidal signals was selected from 1 Hz to 500 Hz, implying that the maximum bandwidth should be about 500 Hz to span the per-

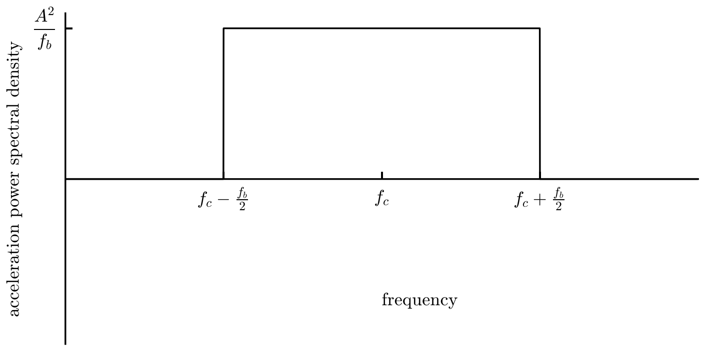

Figure 5.7.
Bandlimited WGN excitation pattern.

ceivable range. However, the perception threshold rises steeply above approximately 315 Hz would likely results in the high-frequency components falling below the perception threshold above 315 Hz. Therefore, 400 Hz was selected as the maximum bandwidth at a center frequency of 201 Hz as a compromise between covering the whole frequency range and not falling partly below the perception threshold. The bandwidth was halved four times to 200 Hz, 100 Hz, 50 Hz, and 25 Hz while with noise with 25 Hz bandwidth still being clearly distinguishable from sinusoidal stimuli in the respective frequency range. For a chosen bandwidth, the center frequency was determined in such a way that the resulting stimuli would not overlap. The RMS SL was kept constant for the different bandwidths.

Figure 5.8 provides an overview of the bandlimited WGN stimuli for the high SL. Also for this excitation pattern, a preliminary experiment with all parameter combinations of center frequency and bandwidth was conducted to assess the stimuli for their similarity. To keep the stimulus set compact, the stimuli with center frequencies of 100 Hz and below in the frequency range above 200 Hz were omitted since they were perceived to be very similar to the 200 Hz bandwidth stimulus in the frequency range above 200 Hz in a preliminary study.

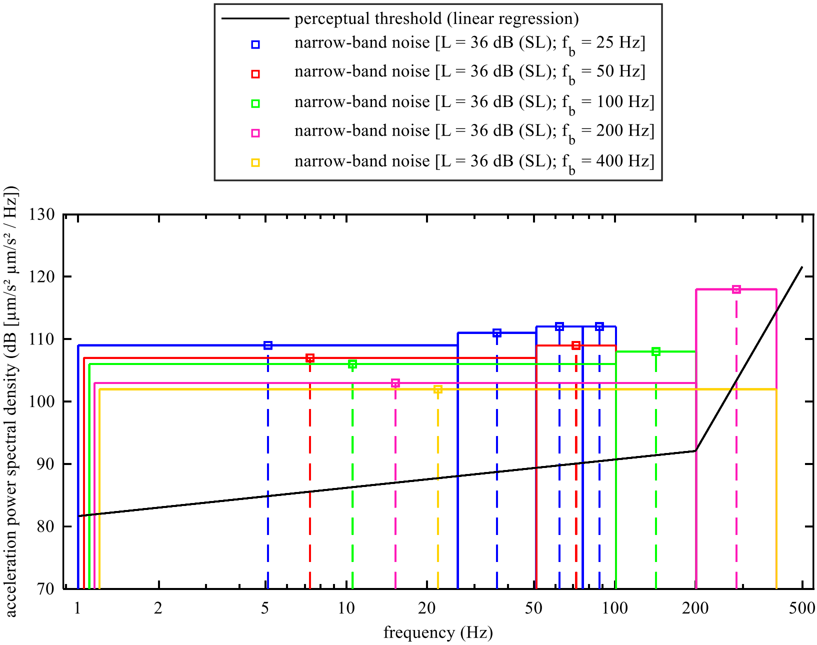

Figure 5.8.
Selected parameter value combinations for the bandlimited WGN excitation pattern for 36 dB (SL).

5.2.2.4 Impulse-like Excitation Pattern

The fourth excitation pattern is shock excitation. In the domain of vehicles encountering transient excitation in everyday life, such excitation can be produced by the vehicle driving over road irregularities such as manholes. However, the impulse like excitation is not transmitted directly to the driver. The suspension of the vehicle can be considered as a mass-spring-damper system. Such a system produces damped oscillation as a response to a transient excitation. According to [103] it is possible to generalize this behavior as a sudden onset of a sinusoidal vibration with a decay:

$$a(t) = \begin{cases} 0 & \text{if } t < 0 \\ A_0 e^{-\alpha t} sin(2\pi f t) & \text{if } t \geq 0 \end{cases} \quad 5.5$$

The resulting signal (see Figure 5.9) is characterized by the peak acceleration amplitude Â represented by RMS sensation level SL, resonance frequency f, and decay rate α. The peak acceleration Â can be determined from the initial acceleration A_0 with the following equation [103]:

$$\hat{A}_{peak} = A_0 e^{-\frac{\alpha}{2\pi f}\arctan(\frac{2\pi f}{\alpha})}\sin(\arctan(\frac{2\pi f}{\alpha})) \qquad \mathbf{5.6}$$

To facilitate comparisons to sinusoidal excitation with the same frequency as the resonance frequency, the RMS SL was defined as if the impulse was a non-decaying sinusoidal signal with a peak acceleration A corresponding to the peak acceleration Â of the impulse. Thus, the impulse-like signal can be generated from RMS SL with:

$$a(t) = \begin{cases} 0 & \text{if } t < 0 \\ 10^{-6}\frac{m}{s^2}10^{\frac{SL(f)+3dB}{20dB}}\dfrac{e^{-\alpha t}\sin(2\pi f t)}{e^{-\frac{\alpha}{2\pi f}\arctan\left(\frac{2\pi f}{\alpha}\right)}\sin\left(\arctan\left(\frac{2\pi f}{\alpha}\right)\right)} & \text{if } t \geq 0 \end{cases} \qquad \mathbf{5.7}$$

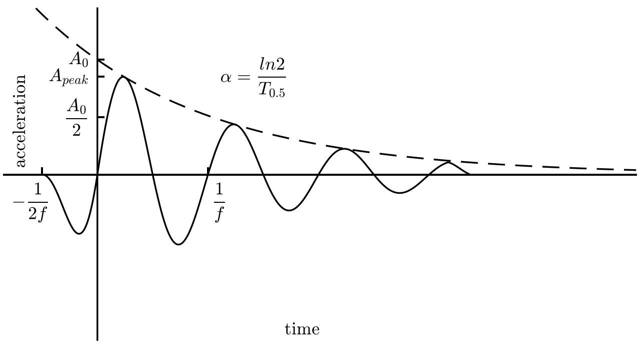

Figure 5.9.
Impulse-like excitation pattern.

The impulse-like stimulus also required a short linear fade in for half of a period of the resonance frequency of the stimulus to enable correct presentation on the reproduction system.
Figure 5.10 provides an overview of the impulse-like signal for the high SL. The resonance frequency values were selected to be identical to the frequency of the sinusoidal stimuli to enable comparison to sinusoidal stimuli. However, the resonance frequencies of 1 Hz and 2 Hz could not be reproduced for the high SL so 3 Hz was introduced as the lowest resonance frequency instead. The initial peak acceleration instead of the total signal RMS was selected to characterize the vibration, because of the decaying characteristic of this excitation pattern. Due to the decay, the signals were intrinsically limited in duration. Tactile receptors are integrating the energy over a short period [29]. If the vibration duration is below about one second, then the perception threshold rises and the perceived intensity decreases. Therefore, 30 dB (SL) was selected as the lower peak level of the stimuli to remain clearly perceivable. The upper level was shifted to 42 dB (SL).
The WBV recordings of impulse-like events in section 6.1.3 were analyzed for typical decay constants. The observed decay constant α range was increased from 2 s^{-1} to 8 s^{-1} to include the behavior of a highly damped (α = 8 s^{-1}) and a weakly damped (α = 2 s^{-1}) resonance system. Multiple successive Impulses separated by a fixed duration pause were also included in the experiment since such vibration can be encountered on certain road types such as motorways with concrete plates. The beginning of the pause was determined from the decaying sinusoidal falling below the perception threshold. The end of the pause corresponded to the beginning of the next impulse. The pause was varied from zero seconds to one second. Longer pauses can be considered as a succession of multiple impulse-like events instead of an aggregate event.
In a preliminary experiment, all combinations of resonance frequency with decay constants were assessed for their similarity. To keep the stimulus set compact for this excitation pattern as well, the stimuli with the resonance frequencies of 7 Hz, 15 Hz, 55 Hz, 155 Hz, and 500 Hz were omitted also here since they were perceived to be very similar to the remaining stimuli with the resonance frequencies of 9 Hz, 26 Hz, 90 Hz, 275 Hz in a preliminary experiment (see 5.2.2.2). Due to limitations of the reproduction system, the decay constant of the impulse-like stimulus with a resonance frequency was

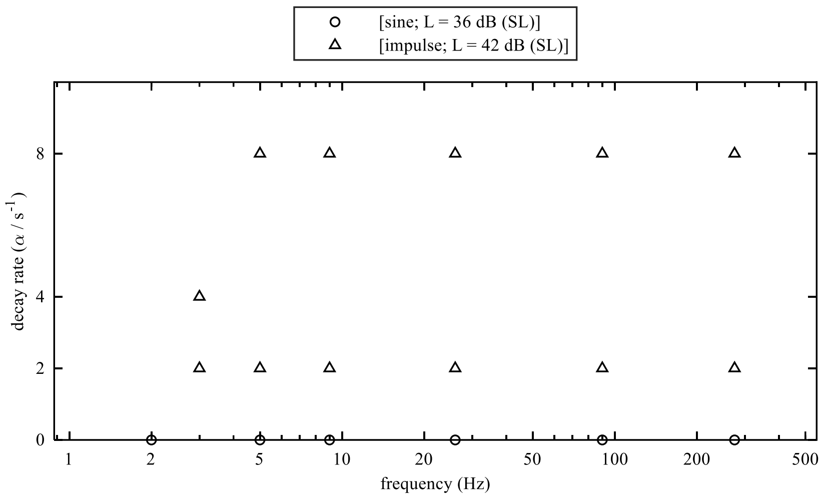

Figure 5.10.
Selected parameter value combinations for the shock excitation pattern for 42 dB (SL).

reduced from 8 s^{-1} to 4 s^{-1}. The goal was to investigate the difference between multiple impulses and single impulses in general. Thus, not all resonance frequencies were investigated, but only 5 Hz and 90 Hz.

5.2.3 Generation of the Stimuli

All stimuli were generated in MATLAB according to the formulas and parameter value combinations documented in the previous section. The bandlimited WGN stimuli were generated with the *randn* function in MATLAB. Subsequently, a Chebyshev Type II bandpass filter with passband frequencies according to the bandwidth and center frequency of the stimulus, a passband ripple below 1 dB and 60 dB attenuation was applied to the acceleration signal using the MATLAB *filtfilt* command for zero-phase filtering. Chebyshev Type II filters were selected due to their steep edges and smooth passband at the cost of stopband ripple. Finally, the RMS of the stationary stochastic vibration was scaled to the defined SL RMS. To enable smooth reproduction, the non-transient scenes were faded in linearly at the beginning of the stimulus and faded out linearly at the end of the stimuli for 300 ms.

The transient stimuli also required a short linear fade in for half of a period of the resonance frequency of the stimulus. No fade-out was required due to the decaying characteristic. The stimulus duration for the non-transient excitation patterns was 9 seconds to 10 seconds. The stimulus duration for the transient excitation pattern was 2 seconds to 3 seconds. While duration was suggested to be a perceptual dimension of vibration [60], it is often externally constrained by the duration of the scene for which vibration is to be synthesized in chapter 7. Thus, the parameter duration was not varied for this investigation.

5.2.4 Summary

By generalizing the WBV occurring in everyday life into excitations patterns, it is possible to systematically vary their parameters in the limits of the tactile receptors. Since the generalized vibration will be perceptually equivalent i.e. perceptually very similar to WBV actually encountered, all relevant sensory-perceptual properties should potentially be elicitable by the representative stimulus set. The final stimulus set including all four selected excitations patterns consisted of a total of 99 WBV stimuli. The structure of the stimulus set needs to be balanced if missing out characteristics of underrepresented, less dense data should be avoided [19]. In the representative stimulus are 21 sinusoidal stimuli, 30 AM-sinusoidal stimuli, 22 bandlimited noise stimuli, and 26 impulse-like stimuli. Since the number of potential parameter combinations is not the same across the four excitation patterns there is some small variation in the total amount of stimuli per excitation pattern. But overall no excitation pattern is dominating the stimuli structure which can thus be interpreted as sufficiently balanced.

5.3 Assessment of the most relevant Sensory Tactile Perceptual Attributes

After the definition of the representative stimulus set in the previous section, the first step was to conduct a free association task to find attributes elicited by these stimuli. The goal was to extend the findings of [50], [61] gained by investigating a small stimulus set consisting mostly of periodic WBV with only one SL with the representative stimulus. Therefore, all stimuli were presented to laypersons and each association mentioned was collected. The vast amount of elicited attributes requires the identification of a representative attribute subset suitable for enabling the subsequent semantic differential analysis for accurately assessing redundancy of attributes. An alternative but less precise method of aggregating redundant attributes is developed based on a thesaurus containing synonyms and antonyms. Selection of the most frequently mentioned attributes will maximize understandability and maximize the likelihood of including all relevant attributes in the sensory tactile design language while minimizing the attribute set further.

5.3.1 Experimental Design

All stimuli consisting of four excitation pattern subsets described in section 5.2 were presented randomized in separate blocks. Due to their similarity, sinusoidal and AM-sinusoidal stimuli were aggregated in the first block. The bandlimited WGN stimuli were presented in the second block. The third block consisted of impulse-like stimuli. All tactile stimuli were presented in the auditory-visual-tactile virtual environment described in chapter 4.
The general instruction for the free association task was to mention all attributes characterizing the presented vibration. According to the stated requirements on the sensory tactile design language (see section 5.1.1) participants were further instructed not to describe specific situations e. g. “cobblestone road” in which they encountered similar vibrations, since such descriptions would be difficult to transfer across different situational contexts. Similarly, they were instructed to avoid general descriptions matching all oscillating motion such as “vibrating”. Another instruction derived from the requirements was to avoid affective terms such as “annoying” and attributes relating exclusively to other modalities.
Such a free association task is often difficult and [54] argues that this is also the case for tactile perceptual properties. Indeed, in a preliminary test, some subjects reported difficulties. It has been shown that productive vocabulary size is often smaller than receptive vocabulary size [104]. Thus, providing a

list with potentially suitable words can facilitate the task, since subjects do not need to rely exclusively on their receptive vocabulary. Attributes mentioned in previous studies with similar goals [50], [61], [105] were candidates for inclusion in such a list. However, handing a list of attributes can bias the results towards terms included in the list. Since the occurrence of attributes in the previous study can be understood as evidence for their relevance in describing vibration, a potential bias towards attributes on the list was deemed acceptable. Therefore, after coming up with associations themselves, subjects were optionally handed a list that consisted of potential attributes found by the previous studies. The time to think about associations for one stimulus was not limited and varied by participant. Typically, it did not exceed one minute. Participants could play the stimulus again as often as necessary. The experiment was split into two sessions of approximately one hour each in order to maintain participants' attention.

5.3.2 Participants

Sample sizes comparable to the sample size of 17 for the successful elicitation study by [99] were selected for the free association task. All subjects were German native speaking laypersons. The study was conducted with the understanding and written consent of each participant. For the (AM-)sine stimuli 19 subjects (12 male, 7 female) with an average age of 35 years (23 to 71 years) took part in the experiment. For the noise stimuli, 18 subjects (12 male, 6 female) with an average age of 35 years (23 to 71 years) took part. For the impulse stimuli, 18 subjects (12 male, 6 female) with an average age of 35 years (23 to 71 years) participated in the experiment.

5.3.3 Results

The total numbers of attributes found per block are displayed in Table 5.1 depending on the filtering or aggregation applied. In the first block, a total of 117 unique German attributes were mentioned for sinusoidal and AM-sinusoidal stimuli. The second block produced 102 unique attributes for bandlimited WGN stimuli. In the third block, 88 attributes were found for impulse-

Table 5.1.
The number of occurrences for each excitation pattern group before and after each successive reduction step.

	Number of Occurrences		
Cumulative Filtering Step	(AM)-Sine	Bandlimited Noise	Impulse-like
none	117	102	88
affective, connotative, non-tactile removed	98	87	80
attributes with same word-stem removed	93	80	72
less frequent synonyms and antonyms merged	39	35	42
infrequent attributes discarded	17	14	19

like stimuli. A lot of the associations were found for more than one excitation pattern. Associations which were not in the focus of the study according to section 5.1.1 were removed. Therefore, affective terms, scene content descriptions, general indications of the presence of oscillating motion were discarded. This resulted in 98, 87 and 80 unique attributes for the three blocks (see Table 5.1).

5.3.4 Aggregation and Prioritization

It is obvious that a large number of sensory tactile perceptual attributes would not represent a useful tool for communication efficiently about the perceptual properties of WBV. The goal of the followed multistep method (see section 5.1.2) is to explain relevant variance in the perceptual space with a minimum number of perceptual dimensions represented by potentially redundant perceptual attributes. To reduce the attributes elicited by the free association task for the semantic differential without impairing its outcome, two approaches might be followed. First, if attributes were mentioned very rarely, they would likely not refer to a salient perceptual property shared by all laypersons and thus would likely be less relevant for explaining variance in the perceptual space. Second, as evident by consulting a thesaurus, natural language contains many synonyms and antonyms i.e. redundancy. Thus, such redundant attributes might be removed. For each of the three excitation pattern groups, these steps were applied separately. The order, in which such reductions are conducted can potentially affect the final set of perceptual attributes. If two low frequently occurring attributes e.g. “even” and “uneven” referring to the same perceptual dimension would simply be omitted due to their low occurrence, it would potentially lead to a failure to identify this dimension in the semantic differential stage. Therefore, the thesaurus aggregation was conducted before the omission of low frequently occurring perceptual attributes.

The core step of the identification of the elements of the sensory tactile design language is identifying the perceptual dimension of the vibrotactile perceptual space with the semantic differential method. The purpose of the semantic differential is to accurately assess redundant attributes by investigating correlations between ratings of the attributes directly in the perceptual domain. As evident by consulting a thesaurus, natural language contains many synonyms and antonyms i.e. redundancy. However, the required number of ratings across all stimuli for each attribute is linearly dependent on the number of attributes. If an attribute is listed as synonym or antonym in the thesaurus, it would be an indication of approximate, heuristic redundancy indirectly in the language domain. However, the semantic differential method is likely more accurate in identifying redundant attributes than the heuristic thesaurus approach. Thus, reduction in the language domain should not be applied if two attributes are not clearly redundant or if an attribute is among the most frequently mentioned attributes suggesting high relevance. Therefore, it was only applied to reduce the number of attributes to a small enough set making the semantic differential stage feasible.

In the first step of the aggregation, synonymous attributes with the same word stem were aggregated by e.g. by replacing “tingle” with “tingling” in the dataset. In the second step synonyms and antonyms were aggregated. To enable this step all attributes occurrences were counted. Subsequently, the algorithm iterated through the attributes by ascending order. For the attribute selected in the iteration step, the synonyms or antonyms were extracted from the machine-generated thesaurus provided by the online database Wortschatz [106]. The synonyms and antonyms were then searched in the list of attributes occurring more frequently than the selected attribute selected in the current iteration step. If a match was found, the occurrences of the selected attributes were replaced by the synonymous or antonymous attribute occurring more frequently. For example, if “intense” had occurred 10 times and “weak” 20

times, then the occurrence of “intense” would have been replaced by “weak” in the dataset. This resulted in 39, 35, and 42 attributes for the three excitation pattern groups (see Table 5.1).
After aggregating the attributes, their overall relevance for explaining the tactile perceptual objects elicited by the stimuli was analyzed. An attribute's low global occurrence might have suggested its irrelevance, but it could have been very relevant for only a few stimuli. Therefore, the relevance was analyzed locally i.e. per stimulus and globally i.e. over all stimuli. The local per stimulus relevance threshold was set to 15 %. If an attribute was not mentioned by at least 15 % of the participants for at least one stimulus, it fell below this threshold. The global relevance threshold was set to 2 %. If an attribute was not mentioned in at least 2 % of the total judgments assessed in the experiment (number of participants times the number of stimuli), it fell below this threshold. If an attribute fell below any or both of the thresholds, it was discarded. The omission of rarely occurring attributes also maximizes the understandability of the tactile design language by laypersons by only including the remaining high frequently mentioned attributes. This resulted in 17, 14, and 19 attributes for the three excitation pattern groups (see Table 5.1).

5.3.5 Summary

A free association task conducted on the selected stimulus set elicited a great number of attributes. After discarding synonyms and antonyms identified by a thesaurus and low frequently occurring attributes for each excitation pattern group, only a fraction of the attributes remained. Since the attributes were largely overlapping for the excitation pattern groups, the intersection of the attributes of the three groups was created. This resulted in 21 attributes which are shown in Table 5.2.
To translate the German attributes into corresponding English attributes, an English native speaker who is a professional bilingual language expert was consulted. The suggested translations of each attribute were validated together with the expert by demonstrating him live vibrations that were previously rated low or high by subjects of the semantic differential experiment (see section 5.4). The thesaurus aggregated and occurrence prioritized attribute set enabled the subsequent semantic differential stage. The omissions

Table 5.2.
Overview of the thesaurus aggregated and occurrence prioritized attributes for each of the three excitation pattern groups. Attribute numbers below the occurrence thresholds are displayed in brackets.

Attribute (Translation)	Attribute (German)	Number of Occurrences (AM)-Sine	Bandlimited Noise	Impulse-like
weak	schwach	178	102	100
trembling	wackelnd	143	49	176
jolting	schlagend	136	32	243
bumpy	holprig	124	70	85
buzzing	summend	113	14	102
pulsating	pulsierend	108	(4)	104
tingling	kribbelnd	108	48	103
calm	ruhig	101	10	41
humming	brummend	99	17	54
rattling	ratternd	92	18	18
grinding	rauschend	78	22	(0)
shaky	rüttelnd	70	27	161
shuddering	zittrig	54	19	26
throbbing	wummernd	47	(6)	77
up and down	auf und ab	32	11	72
uniform	gleichmäßig	30	33	(0)
decaying	abklingend	(0)	(0)	40
fading	nachschwingend	(0)	(0)	35
soft	weich	17	(16)	30
ticking	tickend	(17)	(1)	17
repetitive	wiederholend	(5)	(3)	14

maximized understandability and minimized the likelihood of excluding relevant attributes for the sensory tactile design language.

5.4 Identification of the Attributes forming the Design Language

The thesaurus aggregated and occurrence prioritized attribute set was constructed in the previous section, to contain relevant attributes for sufficiently but possibly redundantly explaining variance in perceptual space of WBV. To eliminate the remaining redundancy, a semantic differential analysis was conducted. In this analysis, all remaining 21 perceptual attributes were rated for each stimulus. An analysis of variance (ANOVA) was conducted for each attribute to identify attributes explaining no differences in the representative stimulus set and which would thus be redundant. Finally, a PCA was conducted to identify the remaining redundancy and to enable feature selection to represent the sensory tactile perceptual dimensions. These steps were necessary to ensure the fulfillment of the compactness requirement on the tactile design language by producing a minimal set of perceptual attributes.

5.4.1 Experimental Design

The representative stimulus set was presented again to laypersons to rate each of the 21 attributes for each of the 99 stimuli in a semantic differential test. The enormous amount of required judgments necessitated the split into two times two blocks of approximately one hour each to keep participant's attention. The 21 perceptual attributes were split into groups of three attribute triples which were shown simultaneously on the semantic differential graphical user interface (see section 4.2.2). A repeated measures design was chosen i.e. each subject rated each stimulus and each attribute to facilitate the comparison between the ratings. All tactile stimuli described in section 0 were presented in the tactile virtual environment described in section 4.1.3 again.
According to the ecological approach to perception (see section 2.1.3) we have learned to intuitively associate specific explicitly verbalizeable tactile perceptual properties (e. g. "tingling") with specific vibration. Therefore, instead of providing an external anchor against which perceptual attributes were to be rated against, participants were assumed to be able to rate the attribute against their inner reference formed by their previous everyday life experiences with WBV. Thus, they were explicitly instructed to compare the presented vibration against the range of their experiences with WBV from everyday life according to the provided perceptual attribute.
A quasi-continuous 100-point rating scale was utilized on which subjects could indicate their judgments for the semantic differential. The scale was

chosen to be unipolar, since no clear antonyms existed for many of the attributes such as “up and down” or “tingling”. A verbal rating scale with equidistant verbal labels at five tick marks was used to facilitate the rating of the perceptual attribute relative to the range of previous experiences [92]. Depending on the rating task the verbal labels on such a scale differ. There are different verbal anchors for judgments of probability, quality, agreement, and intensity. For attribute ratings, the intensity anchors are most suitable. Selecting the endpoint labels of the scale will influence how often the extremes are selected [92]. Since the goal of the task was to reflect the intensity of the perceptual attributes in relation to the range of everyday life experiences, extreme endpoints were preferred. The verbal anchors (translated from German) “not at all” (“nicht”) was displayed at 0, “slightly” (“wenig”) at 25, “moderately” (“mittel”) at 50, “very” (“ziemlich”) at 75, and “extremely” (“sehr”) at 100 as suggested by [107]. The rating scale is shown in Figure 5.11.
As argued in section 5.2.2 the selection of SL in combination with individually varying perception thresholds might lead to some stimuli falling below the individual perceptual threshold for a small fraction of participants. In order not to confuse subjects, they could skip the attribute rating, if they could not perceive the vibration. A short training phase preceded the experiment in which participants were familiarized with the rating scale by providing some example stimuli from the subsequent test. Participants were allowed to repeat the stimulus, but rarely did so. The rating time of an attribute triple per stimulus was not limited, but typically took 10 seconds. The large number of trials

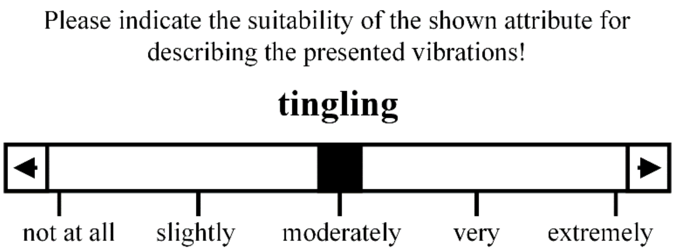

Figure 5.11.
A verbal rating scale with verbal anchors at 0, 25, 50, 75, and 100 points was utilized to assess the attribute ratings for each stimulus.

required the splitting of the rating task into four separate sessions of one hour each to maintain participants' attention.

5.4.2 Participants

For each session 28 to 30 German native speakers (70 % male, 30 % female) with an average age of 35 years (16 to 74 years) took part. The study was conducted with the understanding and written consent of each participant.

5.4.3 Results

In section 5.2.2 it was expected that some 10 dB (SL) stimuli would fall below the individual perception threshold of some participants. In fact, some low SL stimuli were not perceivable in up to 8 % of all ratings. However, simply rejecting the "not perceivable" attribute ratings would introduce a bias since they still implicitly contain information about perceptual properties. For example, if a stimulus was rated as "weak" by the majority of participants and "not perceivable" by one subject, it implies that the stimulus is at the extreme end of the range of potential everyday life vibration for that subject. If the implicit extreme judgment of the participant would be dropped, a bias of the average attribute rating towards the middle of the scale would be introduced. To derive correct conclusions about an attribute's rating on an ecological level rather than on an individual level, "not perceivable" should be replaced by the extreme values of the scale. Such a replacement had to be conducted at not more frequently than for 8 % of the participants for some of the 10 dB (SL) stimuli. For attributes that are positively correlated with level, a stimulus converging towards the perception threshold will elicit an attribute rating converging to zero. Therefore, the small fraction of "not perceivable" ratings were replaced with "not" i.e. zero ratings for these attributes. For attributes that are anticorrelated with level ("weak", "soft", and "calm"), a stimulus converging towards the perception threshold will elicit a rating converging to one hundred. Therefore, the small fraction of "not perceivable" ratings were replaced with "extremely" i.e. one hundred ratings for these attributes.

The experiment resulted in a dataset consisting of ratings of each of the 21 attributes for each of the 99 stimuli for each participant. However, not all 21 attributes are necessarily required to form the design language if e. g. one attribute is correlated with another attribute. According to the stated requirement on the design language of compactness (see section 5.1.1), redundancy needs to be removed to enable efficient communication. Therefore, a PCA or factor analysis can be conducted on the matrix of mean ratings of each stimulus for each attribute. A PCA was preferred due to the focus lying on the selection of explicit non-redundant features i.e. attributes for the design language and not necessarily on the extraction of underlying features, i.e. implicit perceptual dimensions. The analysis suggested that four dimensions can explain 91 % of variance. The first factor explains 46% of the variance, the second factor explains 23 % of the variance, the third factor explains 12 % of the variance, and the fourth factor explains 10 % of the variance. Including an additional factor would only explain 3 % additional variance and was thus rejected. Table 5.3 shows the attribute loadings onto each factor.
This data set also enabled the investigation of each attribute's suitability in describing vibration. An attribute can likely only meaningfully describe vibration if rating differences are observable between stimuli of the representative stimulus set. Therefore, a repeated-measures ANOVA was conducted over the 99 stimuli for each of the 21 attributes. It demonstrated a highly significant effect of stimulus ($p < 0.001$) for every attribute. If there are differences between stimuli for a given attribute it can be concluded that the attribute can be used to potentially describe differences between WBVs. Therefore, all 21 attributes are generally suitable as an element of the sensory tactile design language for WBV.

5.4.3.1 Attributes Describing Vibration Level

The first component consists of two subgroups. The first group had high positive loadings while the second group had high negative loadings. The perceptual attributes “weak”, “calm”, and “soft” belong to the first subgroup with highly negative loadings. In Figure 5.12 the mean ratings and 95 % confidence intervals of the attribute “weak” are shown for sinusoidal vibration. Stimuli with a low level of 10 dB (SL) have a high rating and stimuli with a high level of 36 dB (SL) have a rating of approximately 50 to 60 scale points

Table 5.3.
Principal Component Analysis (varimax rotated) for the remaining 21 attributes with attributes loading higher than 0.6 or lower than -0.6 shown in bold.

Attribute		Component				Subgroup
Translation	German	1	2	3	4	
bumpy	holprig	**0,94**	-0,26	0,1	0,09	1b)
buzzing	summend	-0,19	**0,94**	0,13	0	2
calm	ruhig	**-0,76**	-0,4	-0,23	-0,31	1a)
decaying	abklingend	0,06	-0,05	0,01	**0,96**	4a)
fading	nachschwingend	0,29	-0,12	0,06	**0,91**	4a)
grinding	rauschend	-0,01	**0,84**	-0,16	-0,38	2
humming	brummend	0,24	**0,9**	0,16	-0,1	2
jolting	schlagend	0,58	0,05	0,47	**0,6**	4a)
pulsating	pulsierend	0,47	0,17	**0,84**	0,1	3
rattling	ratternd	**0,82**	0,42	0,21	-0,2	1b)
repetitive	wiederholend	0,18	-0,05	**0,89**	-0,34	3
shaky	rüttelnd	**0,95**	-0,14	0,1	0,14	1b)
shuddering	zittrig	**0,93**	0,13	0,16	-0,04	1b)
smooth	weich	**-0,7**	-0,41	-0,19	-0,28	1a)
throbbing	wummernd	**0,73**	0,2	0,53	0,21	1b)
ticking	tickend	0,14	0,18	**0,92**	0,19	3
tingling	kribbelnd	0,01	**0,95**	0,19	0,08	2
trembling	wackelnd	**0,9**	-0,33	0,03	0,19	1b)
uniform	gleichmäßig	-0,15	0,16	0,52	**-0,67**	4b)
up and down	auf und ab	**0,8**	-0,38	0,03	0,22	1b)
weak	schwach	**-0,82**	-0,36	-0,25	-0,25	1a)

less. Compared to the influence of level, the stimulus frequency has only little influence with ratings increasing from low to high frequencies by approximately 20 points. This might be attributed to the linear approximation of the perception threshold on which the SL of the stimuli is based on (see section 2.1.2). Thus, it can be concluded for sinusoidal vibration that this subgroup is mainly utilized to communicate about the stimulus level.

Applying modulation to the sinusoidal vibration doesn't change the attribute rating of “weak”. The ratings of “weak” for impulse like vibration with a peak level of 42 (SL) are similar to sinusoidal vibration with an RMS level of 36 dB (SL) despite their differences in level (see Figure 5.13). This is likely caused by the decaying characteristic of the impulse-like vibration compared

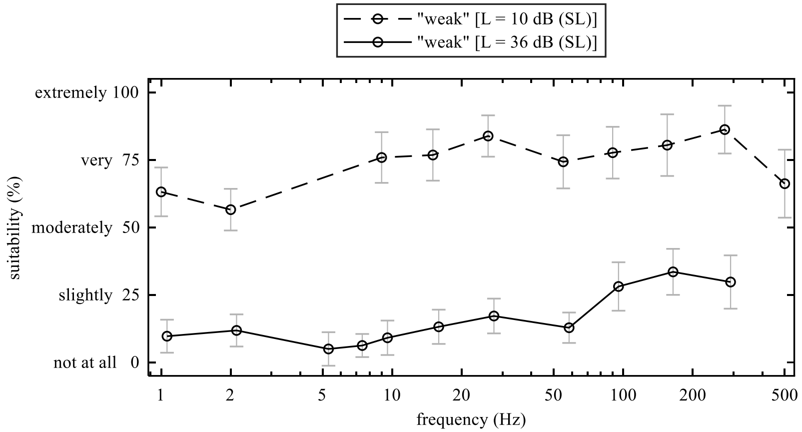

Figure 5.12.
Mean ratings and 95 % confidence intervals for the attribute "weak" for sinusoidal vibration with a level of 10 dB (SL) and 36 dB (SL).

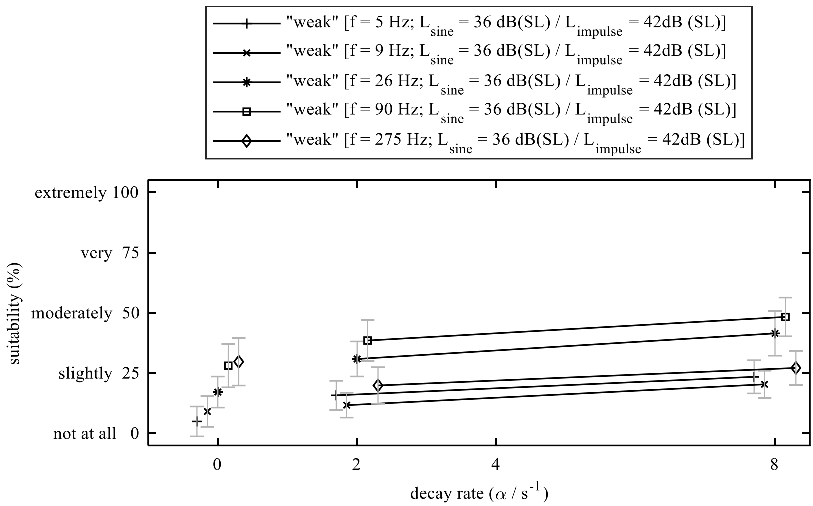

Figure 5.13.
Mean ratings and 95 % confidence intervals for the attribute "weak" for sinusoidal vibration with an RMS level of 36 dB (SL) and impulse like vibration with a peak level 42 dB (SL).

to the constant level of the sinusoidal vibration. When comparing the stimuli of the stochastic excitation pattern to stimuli with the sinusoidal excitation pattern the attribute ratings are also very similar if both stimuli have the same RMS and are in the same frequency range. However, some small differences for WGN vibration with a bandwidth of 200 Hz or 400 Hz compared sinusoidal vibration in the corresponding frequency range might be caused by the perception threshold rising over the frequency range of the narrowband noise (see section 5.2.2). There is a high correlation between the vibration parameter SL and the attribute “weak” (Spearman’s $\rho = -0.739$, $p<0.01$). Summarizing the observations, the first factor is used mainly for describing the vibration level across the four excitation patterns.

5.4.3.2 Attributes Describing Low Vibration Frequency

The perceptual attributes “up and down”, “bumpy”, “rattling”, “shaky”, “trembling”, and “shuddering” belong to the second subgroup with highly positive loadings. The attributes “throbbing” and “jolting” load on this component only slightly less than the previously mentioned attributes. The rating patterns of “throbbing” or “jolting” are also showing similarities to attributes loading onto the third component or the forth component. In Figure 5.14 the attribute “up and down” shows the typical ratings patterns demonstrated by

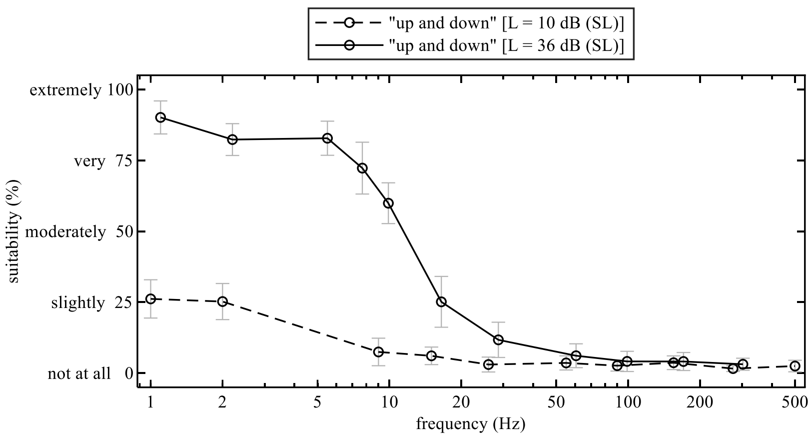

Figure 5.14.
Mean ratings and 95 % confidence intervals for the attribute “up and down” for sinusoidal vibration with a level of 10 dB (SL) and 36 dB (SL).

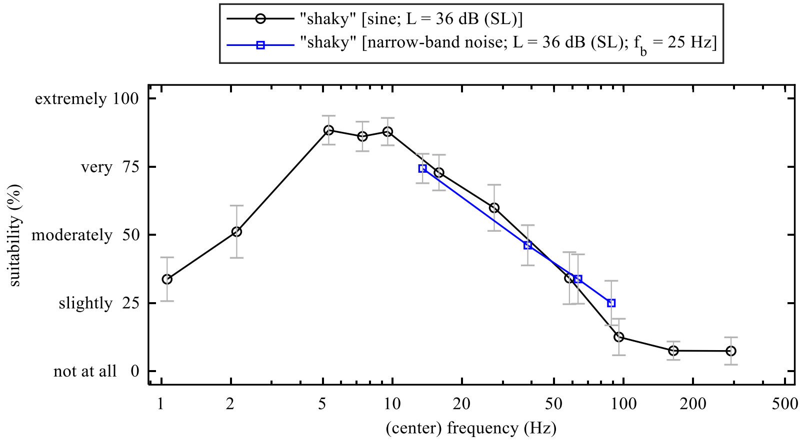

Figure 5.15.
Mean rating values and 95 % confidence intervals for the attribute “shaky” for sinusoidal vibration compared to bandlimited WGN vibrations (represented at their center frequencies) with varying center frequency

attributes in this group. Each attribute of this subgroup has a rating peak at frequency below 35 Hz that is depending on the attribute. The difference between peak and off-peak ratings is much more pronounced for the high SL (solid line) than for the low SL (dashed line). Thus, it can be concluded for sinusoidal vibration that in contrast to the other subgroup, this group is also utilized to communicate about stimulus frequency.

As with the previous subgroup, applying modulation to the sinusoidal vibration doesn't change the attribute rating of “up and down”. Similarly, impulse-like vibration elicit a rating comparable to sinusoidal vibration, due to the compensatory effect of level difference and decaying characteristic. Compared to the first subgroup, the stochastic excitation stimuli with small bandwidth show similar ratings but stimuli with high bandwidth show rating differences. The similarity of 25 Hz bandwidth noise stimuli compared sinusoidal stimuli is more easily observable for the attribute “shaky” displayed in Figure 5.15 However, for all attributes a general trend can be seen. If a noise stimulus spans the frequency of a sinusoidal stimulus both will have similar attribute ratings.

However, for all attributes, a general trend can be seen. If a WGN stimulus spans the frequency of a sinusoidal stimulus both will have similar attribute

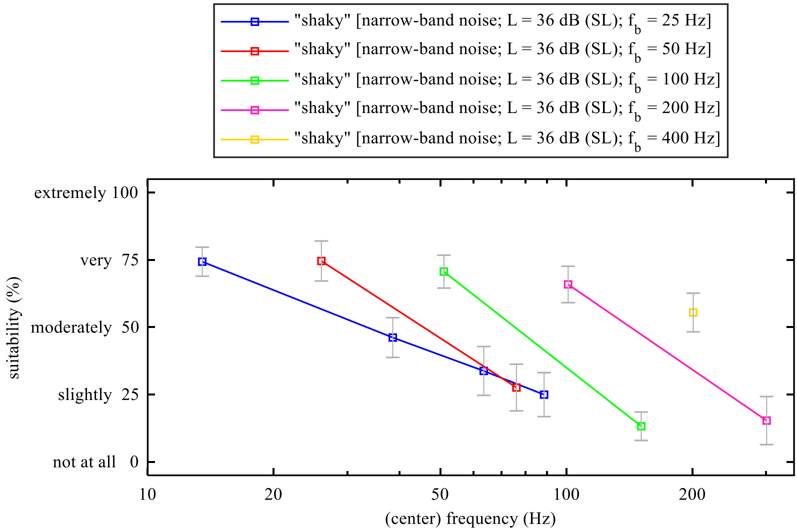

Figure 5.16.
Mean values and 95 % confidence intervals of attribute “shaky” for bandlimited WGN vibrations (represented at their center frequencies) with varying bandwidth

ratings. However, increasing the bandwidth leads to increasingly lower suitability ratings as evident from Figure 5.16 The explanation for this observation is likely to be found in the definition of the stimuli parameters (see section 5.2.2). As explained for sinusoidal stimuli, attributes in this group have a rating peak at a certain frequency. For sinusoidal stimuli, the total signal energy is concentrated into an infinitely small frequency range. Due to the constant RMS level of the noise stimuli remaining constant with bandwidth, the signal energy is distributed over an increasingly higher bandwidth. Therefore, the energy is shifted away from the frequency range of the rating peak towards a frequency range with a rating low. The remaining energy would be equivalent to the total energy of sinusoidal stimulus with a much lower total RMS level. This behavior is confirmed by the decreasing attribute rating when the bandwidth is increased. There is a high correlation between the vibration parameter (carrier-, center-, resonance-) frequency and the attribute “up and down” (Spearman’s $\rho = -0.732$, $p<0.01$). In summary, these attributes are used for describing low-frequency vibration in addition to describing vibration level. Their ratings are at their peaks if excitation is concentrated in the frequency range below 35 Hz.

5.4.3.3 Attributes Describing High Vibration Frequency

The perceptual attributes “tingling”, “humming”, “buzzing” and “grinding” show highly positive loadings onto the second component. In Figure 5.17 the attribute “tingling” shows the typical rating patterns demonstrated by attributes in this group. The ratting pattern suggests that these attributes are similar to low frequency describing attributes loading onto the first component e. g. Figure 5.14. However, in contrast to the rating peak being below 35 Hz for the former, the latter have a rating peak above 35 Hz varying from attribute to attribute. This similarity can not only be observed for the sinusoidal excitation pattern, but also the AM-sinusoidal excitation pattern, the WGN excitation pattern, and the impulse-like excitation pattern. There is a high correlation between the vibration parameter (carrier-, center-, resonance-) frequency and the attribute “tingling” (Spearman’s $\rho = 0.568$, $p<0.01$). In sum-

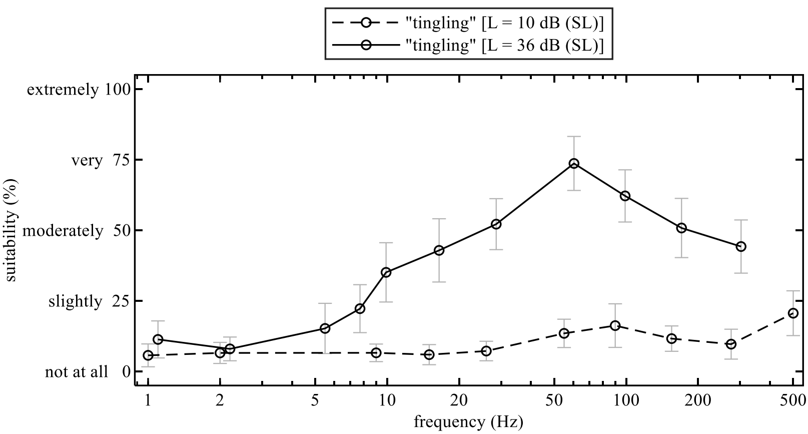

Figure 5.17.
Mean ratings and 95 % confidence intervals for the attribute “tingling” for sinusoidal vibration with a level of 10 dB (SL) and 36 dB (SL).

mary, these attributes are used for describing high frequent vibration in addition to describing vibration level. Their ratings are at their peaks if excitation is concentrated in the frequency range above 35 Hz.

5.4.3.4 Attributes Describing Modulation

The perceptual attributes “pulsating”, “ticking”, “repetitive”, and “throb bing” show highly positive loadings onto the third component. The attribute “uniform” loads onto this component only slightly less than the previously mentioned attributes, but loading higher onto the fourth factor. The rating pattern of the attribute “repetitive” as depicted in Figure 5.18 suggests that it is used to distinguish modulated vibration from unmodulated vibration The rating difference is most pronounced for the modulation frequency of 2 Hz. With rising modulation frequency it's rating is converging towards the ratings of sinusoidal vibration.

For the single impulse-like stimuli the rating of “repetitive” is low. Due to the envelope of multiple successive impulse-like vibration approximating the envelope of an AM-sinusoidal vibration, multiple impulse-like vibration also

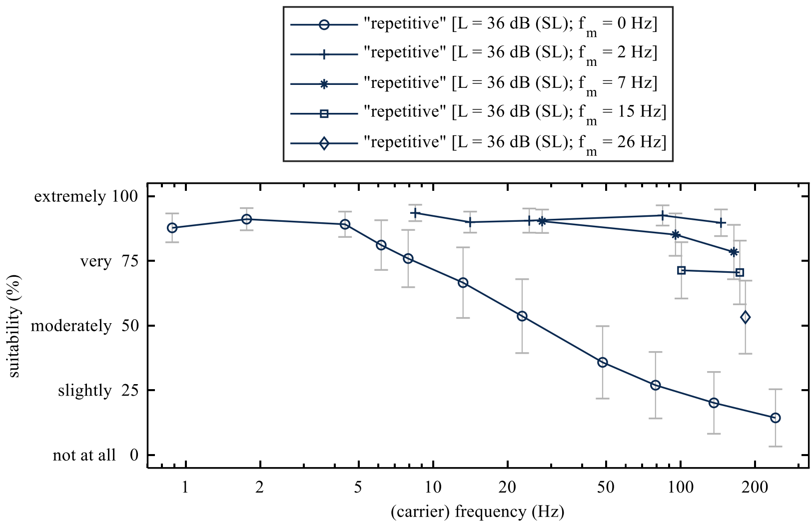

Figure 5.18.
Mean ratings and 95 % confidence intervals for the attribute “repetitive” for AM-sinusoidal vibration with varying modulation frequency and a level of 36 dB (SL).

produces a similar rating for that attribute (see section 7.3). The WGN excitation pattern and the sinusoidal excitation pattern both produce low ratings for attributes in this group above 20 Hz. However, the sinusoidal excitation pattern produces higher ratings in the frequency range of 20 Hz than the WGN excitation pattern. There is a high correlation between the vibration parameter modulation frequency and the attribute “repetitive” (Spearman’s $\rho = 0.583$, $p<0.01$). In summary, these attributes describe vibration with periodically rising and falling envelope, i.e. modulation.

5.4.3.5 Attributes Distinguishing Transient and Non-Transient Vibration

The fourth component also consists of two subgroups. The first group had high positive loadings while the second group had high negative loadings. The perceptual attributes “decaying”, “fading”, and “jolting” show highly positive loadings onto the fourth component. The rating pattern of the attribute “fading” is shown in Figure 5.19. These attributes suggest that they have

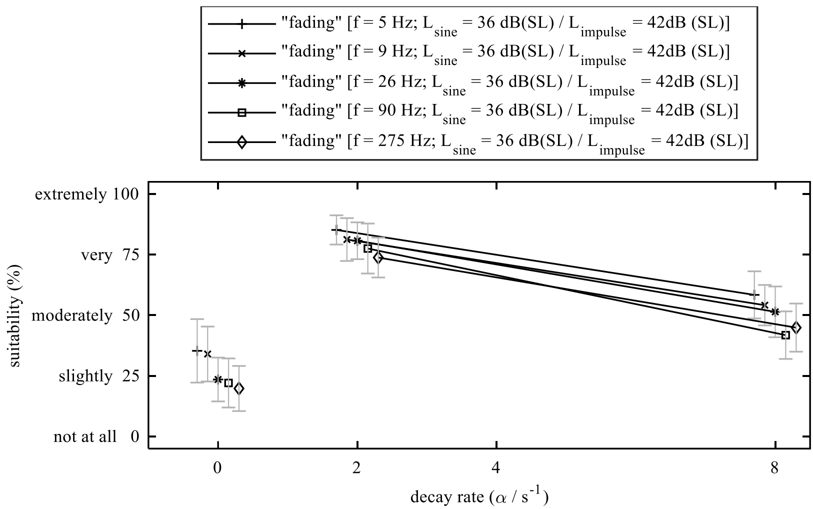

Figure 5.19.
Mean ratings and 95 % confidence intervals for the attribute “fading” for impulse-like vibration with varying decay rate frequency compared to sinusoidal vibration.

a high rating for slowly decaying vibration, i.e. impulse-like vibration and a low rating for non-decaying vibration i.e. sinusoidal vibration.
With increasing decay constants, the rating of “fading” decreases likely because the vibration is perceived only as a short burst with the decaying characteristic becoming less obvious. WGN vibration also showed ratings as low as sinusoidal vibration for this attribute. Repeating single impulse-like vibration to produce multiple successive impulses does not change the rating of fading if other parameters are kept constant. The similarity between multiple impulse-like vibration and AM-sinusoidal vibration due to their similar envelopes noted for “repetitive”, can also be found for “fading” (see section 7.3). There is a high correlation between the vibration parameter decay rate and the attribute “fading” (Spearman’s $\rho = 0.715$, $p<0.01$). In summary, this attribute is used for distinguishing between transient (single impulse and multi impulse) and non-transient vibration signals.

5.4.3.6 Attributes Distinguishing Periodic and Stochastic Vibration

The perceptual attribute “uniform” shows a highly negative loading onto the fourth component. For the attribute “uniform” ratings of bandlimited noise vibration produced low ratings while sinusoidal vibration produced much higher ratings as shown in Figure 5.20. In comparison to sinusoidal stimuli, AM-sinusoidal stimuli were rated as slightly more “uniform”. Also for “uniform” the similarity of the envelope of AM-sinusoidal vibration and multiple impulse-like vibration explained similar ratings (see section 7.3). In comparison to multiple impulses, single impulses were rated much lower regarding the attribute “uniform”. There is a medium correlation between the vibration parameter bandwidth and the attribute “uniform”, Spearman’s $\rho = -0.355$, $p<0.01$) In summary, this attribute was utilized to distinguish between stochastic and periodic vibration.

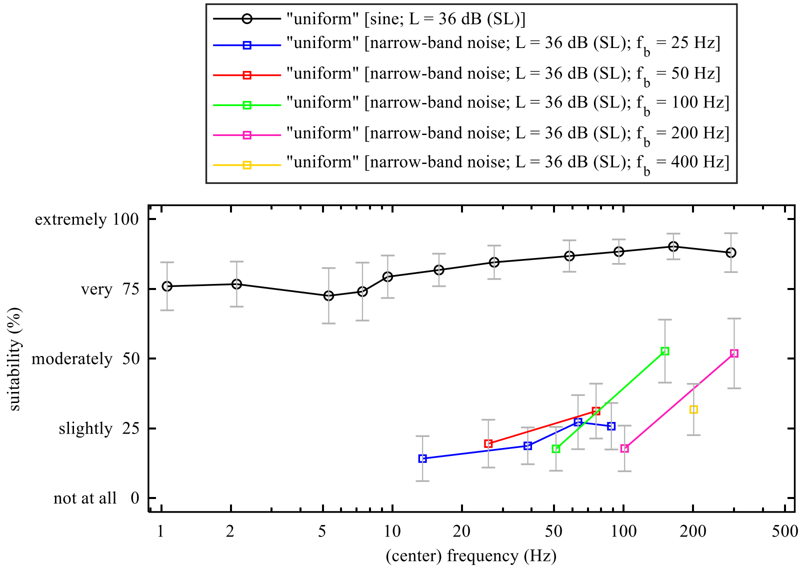

Figure 5.20.
Mean values and 95 % confidence intervals of attribute “uniform” for bandlimited WGN vibrations (represented at their center frequencies) with varying bandwidth.

5.4.4 Selecting the Elements of the Sensory Tactile Design Language

The PCA aids in the selection of a necessary and sufficient number of sensory-perceptual attributes forming the sensory tactile design language. Therefore, a compromise between minimizing redundancy and maximizing explanatory power should be selected. One option would be to utilize the four perceptual dimensions suggested by the four principal components directly. This would imply the extraction of four features based on the 21 attributes that were entered in the PCA. However, profiling vibration according to the four features would require to always rate the 21 attributes to determine the feature scores since underlying implicit dimensions cannot be rated directly. As stated in the requirements on the tactile design language in section 5.1.1 it should enable communication with laypersons and contain as few as possible attributes. The extractions of four features based on 21 attributes suggests that redundancy remains in the attribute set.

The resulting components output by the PCA on attribute ratings depends on the selection of attributes and the selection of stimuli for which the ratings were assessed in a semantic differential. Therefore, it is more reasonable to utilize the PCA to guide the selection of features i.e. perceptual attributes rather than to derive the features. Furthermore, the presence of redundancy implies that multiple subsets can be selected from the 21 attributes that represent the sensory-perceptual properties of WBV. A minimal subset would facilitate the efficient profiling of WBV. Since the PCA extracts four features explaining over 91 % of variance it seems reasonable to select one attribute with a high loading onto each of the four components. It can also be valuable to consider the subgroups identified in the previous section. For example, the attribute group one representable by “weak” and the attribute group representable by “up and down” both load onto the first factor but with inverse loadings. However, there is no perfect anticorrelation (Pearson correlation $-.83 < r < -.45$) between attributes of the two groups, and thus they are not completely redundant. These differences are obvious when examining the different rating patterns associated with the subgroups (see section 5.4.3), especially for some of the signal classes, such as sinusoidal vibration in the case of attributes loading onto the first factor. For the first component, “weak” is rather utilized to describe the overall level of vibration, while “up and down” is rather utilized to describe low-frequency vibration. For the fourth component, “fading” is utilized to distinguish transient from non-transient vibration, while “uniform” is utilized to distinguish periodic from stochastic vibration. Including one attribute for each of the two subgroups of a component is a reasonable compromise between explanatory power and compactness of the attribute set.

Previously suggested sensory dimensions of tactile perception of WBV and other locations of introduction can also aid with the selection of a set of sensory-perceptual attributes. One of the most investigated sensory-perceptual dimension is perceived intensity. The perceived intensity of WBV was investigated by [32]. However, due to the psychophysical focus, an expert unit “VIP” is defined for the perceived intensity instead of an attribute intuitively understandable by laypersons. The findings of [25] suggest that intensity is

an important perceptual dimensions for distinguishing vibration. Therefore, one attribute should be selected reflecting this dimension e. g. “weak”.
A low-frequency perceptual dimension and a high-frequency perceptual dimension were suggested by [20], [21], [58]. Thus, one attribute should be selected to represent the low-frequency dimension e.g. “up and down” and one attribute to represent the high-frequency dimension e. g. “tingling”.
A temporal irregularity dimension (“even” – “uneven”) was found by [25] when presenting stimuli with simple and complex rhythms. Therefore, also one attribute should be selected reflecting this dimension e. g. “uniform”.
The authors of [21] investigate AM-sinusoidal vibration with a carrier frequency of 150 Hz and a modulation frequency from 2 Hz to 80 Hz. They hypothesize that modulation is not a separate perceptual dimension but similar to the low-frequency perceptual dimension without presenting unmodulated, low-frequency vibration. In contrast to this study, the findings of [58] suggest a “pulsing” flowing dimension. Since a component was identified also in this study that the attributes “pulsating” and “repetitive” loaded on, but also a component for low-frequency vibration, one attribute should reflect the modulation property e. g. “repetitive”.
No previous study suggested a perceptual dimension reflecting transient impulse-like vibration. However, the attributes of the sixth subgroup had a very high suitability rating for impulse-like vibration compared to stationary vibration. Thus, it is reasonable to also include an attribute representing the transient vs. non-transient dimension as e.g. represented by “fading”.
Summarizing the findings, it is reasonable to represent the six perceptual dimensions by six sensory-perceptual attributes. The attribute “weak” was amongst the most frequently mentioned attributes and is thus selected to represent the level related perceptual dimension. The high frequency related perceptual dimension is represented by “tingling” since it has the highest loading on its component. Since the frequency-dependent attributes are characterized by having a rating peak at a specific frequency, it is reasonable to choose the attribute whose rating peak has the greatest frequency difference. Thus “up and down” is selected to represent the low frequency related perceptual dimension. The attribute “uniform” was the only attribute loading onto its component and is thus selected to represent the temporal irregularity dimension. The modulation related perceptual dimension is represented by “repetitive” since it has the highest loading on its component, too. The attribute “fading”

has a stronger tactile connotation in German language (“nachschwingend”) than “decaying” (“abklingend”) which has a more auditory connotation. Thus, the attribute “fading” was selected for the transient vs. non-transient dimension. The six sensory tactile perceptual attributes “weak”, “tingling”, “up and down”, “uniform”, “repetitive”, and “fading” are chosen as the elements of the sensory tactile design language. They represent a compromise between minimal redundancy and maximal explanatory power.

5.4.5 Summary

Starting with the thesaurus aggregated and occurrence prioritized set of 21 attributes, ratings were obtained for all attributes for all stimuli in a semantic differential. An ANOVA for each attribute, revealed stimulus as a significant factor influencing all attributes and thus their general suitability for describing perceptual differences of WBV. Based on the mean attribute ratings, a PCA was conducted which revealed six groups of attributes loading onto four components. In agreement with the perceptual dimensions suggested by previous studies, these groups of attributes represent a level related, low frequency related, high frequency related, modulation related, temporal irregularities related dimension. In addition, a transient vs. non-transient dimension is suggested by the results of this study. The six attributes “weak”, “up and down”, “tingling”, “repetitive”, “uniform”, and “fading” were selected to represent these dimensions for explicit communication, thus forming the sensory tactile design language.

5.5 Summary and Discussion

5.5.1 Summary

This chapter aimed to assess a sensory tactile design language for WBV. First, the requirements for such a design language were determined. It was suggested that it needs to be understandable by laypersons and enable explicit communication across individuals and situational contexts. Furthermore, it needs to enable efficient quantitative communication and the translation into physical vibration parameters. Inspired by previous studies a multistep procedure was adapted to this task.

According to the ecological approach to perception, in everyday life, WBV is a carrier of verbalizeable, information about the environment. Associating verbal labels to vibration is a categorization task, which groups perceptually non-identical vibration into a finite set of perceptually equivalent classes. Thus, instead of recording WBV of all potential situations, it is sufficient to construct a set of vibration that is similar to everyday life vibrations. Therefore, everyday life vibration was divided into four generalized excitation patterns. By systematically varying parameters of these excitation patterns, it was possible to cover all the perceivable range of vibration occurring in everyday life. This stimulus set enabled the elicitation of all relevant classes i.e. perceptual attributes.

In the first step, attributes mentioned by laypersons were collected in a free association task. They were instructed to come up with sensory tactile perceptual attributes relating directly to the vibration. In contrast to affective attributes, they are mostly independent from personal preference and situational context. Their direct relationship to vibration potentially enables the translation into physical vibration parameters. A subsequent thesaurus aggregation of the vast amount of non-frequent attributes shrunk the attribute set to the most relevant attributes by reducing redundancy on a language level. By prioritization of the most frequent and thus most understandable attributes, likely irrelevant attributes could be omitted. This made the subsequent semantic differential feasible. In the semantic differential, absolute ratings of all remaining 21 attributes were rated for each stimulus. A significant influence of stimulus on each attribute's ratings suggests their general suitability for describing perceivable differences of vibration. Based on the mean attribute ratings a PCA was conducted. It revealed six groups of attributes loading onto four underlying components explaining 91 % of the variance observed. An attribute was selected to represent each group thus forming the standardized design language.

5.5.2 Discussion

Compared to the previous studies discussed in section 2.1.3 that utilized stimulus sets with limited vibration parameter variation, e.g. only frequency variation, this study utilized a much larger stimulus set with variation representing everyday life vibration. However, the utilized stimulus set determines the identified perceptual dimensions. Thus previous studies identified only a subset of the suggests six perceptual dimensions each. A perceptual dimension related to vibration level was suggested by [25]. This perceptual dimension was confirmed in this study by the observed first attribute group represented by “weak” demonstrating a highly negative correlation to SL. A perceptual dimension related to high-frequency vibration and a perceptual dimension related to low-frequency vibration was observed by [20], [21], [88]. A low-frequency dimension was confirmed in this study by the emergence of the second attribute group represented by “up and down” showing a highly negative correlation to frequency. Also, a high-frequency dimension was confirmed in this study by the emergence of the third attribute group represented by “tingling” showing a highly positive correlation to frequency.
A perceptual dimension related to amplitude modulation was observed by [21] but due to a lack of low-frequency unmodulated vibration in their stimulus set, they hypothesized that no separate perceptual dimension exists for low-frequency vibration. In this study, a separate perceptual dimension for amplitude modulated vibration was observed in the form of the fourth attribute group represented by “repetitive” which is highly correlated to modulation frequency and which is distinct from the attribute group represented by “up and down”. Another perceptual dimension related to the temporal structure of vibration as described by “even - uneven” was suggested by [25] This dimensions was also confirmed by this study in the form of sixths attribute group represented by “uniform”, which has a medium correlation to the parameter bandwidth. This attribute can be used for distinguishing between broadband stochastic (noise) and narrowband periodic vibration.
Due to the inclusion of the majority of excitation patterns encountered in everyday life in the stimulus set of this study, the five previously found dimensions were observed simultaneously. This suggests that they are indeed non-

redundant. Other locations of vibration introduction were utilized in the previous studies, such as the finger [25], [88], the hand [20], [21], or the wrist [59]. while the current study presented vibration at the thighs. The emergence of these perceptual dimensions also for vibration introduced at the thighs suggests they are independent from the location of introduction and thus universal for vibrotactile perception. Since also impulse-like stimuli were included in this study in addition to the other excitation patterns, another perceptual dimension emerged in the form of the fifths attribute group represented by “fading”. By simultaneously including impulse-like stimuli. the existence of a sixth dimension related to transient changes represented by “fading” is suggested which is highly correlated to decay rate. These attributes can be used for distinguishing between instationary (impulse) and stationary vibration signals. To enable vibration synthesis for a wide range of situations, all relevant dimensions should be included. But also for the domain of product design, the inclusion of more perceptually distinguishable dimensions can facilitate the creation of tactons which are better distinguishable [88].

A disadvantage of most previous studies [20], [21], [25], [88] is that they are based on similarity judgments and thus could only show the existence of these underlying perceptual dimensions. These findings mostly enable the selection feedback vibration which maximizes perceivable differences. However, they did not show the explicit interpretation of these dimensions by laypersons, which is the prerequisite for explicit communication. Previously, explicit communication with laypersons about sensory tactile perceptual properties often resulted in different but often synonymous or antonymous attributes. Such qualitative user interview feedback is difficult to directly utilize for vibration design. The 21 investigated attributes are candidates for explicit, interindividually understandable communication about the six perceptual dimensions.

To enable efficient explicit communication, extracting four new features out of 21 attributes is not suitable. Instead, the PCA guided the selection of features i.e. attributes representing these four components. The results suggest six sensory-perceptual dimensions: a level related, low frequency related, high frequency related, modulation related, temporal irregularities related, and a transient vs. non-transient dimension. The six attributes “weak”, “up and down”, “tingling”, “repetitive”, “uniform”, and “fading” were selected to represent these dimensions for explicit communication. This set of attributes

forms the sensory tactile design language that is a compromise between compactness and explanatory power.
The absolute ratings of these attributes show small confidence intervals implying a reasonable degree of interindividual agreement. Thus, such ratings are not necessarily a subjective, but rather an interindividually shared perceptual representation of vibration. The design language containing these attributes enables standardized, quantitative communication about sensory tactile perceptual properties. Laypersons can efficiently profile sensory-perceptual properties of WBV by rating each element of the tactile design language providing the design engineer with a perceptual specification of vibration. The tactile design language will be utilized for this task in chapter 6. Each of the six sensory-perceptual attributes explains important aspects of the physical vibration properties as suggested by their rating vs parameter curves and the observed correlations. However, whether a translation of a rating profile consisting of the attributes contained in the design language into vibration is possible, cannot be confirmed or rejected based only on insights in this chapter. The attribute ratings profile vibration parameter set pairs obtained in this chapter will be the basis to attempt to build a model in chapter 7 that can translate a rating profile into vibration. If a recorded vibration can be perceptually profiled and translated back into a vibration that is perceptually equivalent to the recorded vibration, then the six attributes are complete i.e. sufficiently describing the sensory-perceptual properties of WBV.

6. Quantification of Expected Properties with the Sensory Tactile Design Language

In the previous chapter, everyday life WBV was generalized into abstract vibration patterns which cover the perceivable frequency range and the mostly encountered level range and temporal variations. Subsequently, a sensory tactile design language was extracted from the most frequently mentioned associations elicited by these vibration patterns. This tactile design language contains a set of six perceptual attributes explaining the majority of the variance observed for the most commonly mentioned associations. This chapter will build onto chapter 5 by utilizing the design language to quantify the sensory-perceptual properties of real, recorded WBV.
In the first step, a representative set of vehicle scenes will be selected. Multimodal recordings of all the scenes will be conducted to enable a simultaneous and reproducible presentation of optic, acoustic, and vibration stimuli. The first task is to confirm that the tactile design language can be meaningfully transferred from the abstract stimulus domain of the previous chapter to the complex vehicle stimuli domain for communicating about the sensory-perceptual properties. Therefore, a free association task is conducted to collect the most frequently mentioned associations elicited by this vehicle stimulus set. If communication with the sensory tactile perceptual design language is possible, then it's capability for quantifying the perceptual properties needs to be investigated. The sensory-perceptual attributes should enable the differentiation of the vibration in the representative stimulus set. Therefore, each multimodal scene is presented to participants in an experiment. They will rate each sensory-perceptual attribute of the tactile design language, thus quantifying the sensory-perceptual properties elicited by the vibration. Quantifying perceptual properties of presented vibration to translate them back into vibration would be of limited use. Therefore, it needs to be demonstrated that a

meaningful rating of the perceptual properties anticipated in a certain situational context, is possible without presenting vibration. [2]

6.1 Multimodal Stimuli

6.1.1 Selection of the Scenes

As argued in section 2.1.2, in our everyday lives we most frequently encounter WBV in vehicles. Therefore, vehicle scenes are optimal for selecting a representative set of scenes with WBV. Such a set is necessary to generalize the findings uncovered in this chapter and for the model validation in chapter 7. Inspired by the classification of vibration of [15], everyday life WBV can be categorized into seven classes based on the excitations processes in vehicles (see section 5.2.1). The focus of this work lies on vibration segments that would elicit quasi constant perceptual attribute ratings (see section 3.2.). Therefore, a subset of four categories relevant to this work were chosen: sinusoidal, AM-sinusoidal, stochastic, and impulse-like. The required set should contain scenes with vibrations of each of the four categories. An overview of the selected scenes can be seen in Table 6.1.
Twelve scenes were selected to cover non-transient vibration. Road condition is one major factor influencing the vibrations perceived by the driver. Two tarmac road types, concrete motorway, fine cobblestone, and rough cobblestone were included to cover a wide range of road roughness resulting in periodic as well as stochastic excitation. Vehicle speed was varied in the typical operating range of vehicles from 5 kph to 100 kph. Due to the focus on scenes that would elicit a constant perceptual profile (see section 3.2) mostly constant speed conditions were assessed. However, acceleration scenes were selected to include high engine load contribution to vibration, even though road contribution was changing with speed. Seven scenes were selected to cover transient vibration. Frequently encountered road irregularities such as extension joints, road surface changes, manhole covers, and tram tracks were in

[2] Some parts of this work were presented in [97] by Rosenkranz et al.

Table 6.1.
Overview of the selected scenes and their parameters.

No.	Category	Speed	Surface
1	non-transient	5 to 50	small cobblestone
2	non-transient	30	small cobblestone
3	non-transient	50	small cobblestone
4	non-transient	5 to 50	cobblestone
5	non-transient	30	cobblestone
6	non-transient	50	cobblestone
7	non-transient	50	tarmac (A-Road)
8	non-transient	70	tarmac (A-Road)
9	non-transient	5 to 50	Tarmac
10	non-transient	30	Tarmac
11	non-transient	50	Tarmac
12	non-transient	100	concrete motorway
13	transient	100	surface change
14	transient	40	tram tracks
15	transient	50	expansion joint
16	transient	100	expansion joint
17	transient	30	manhole cover cobblestone
18	transient	30	manhole cover tarmac
19	transient	50	manhole cover tarmac

cluded to cover transient, shock events. Middle-class vehicles have the highest market share and are thus the most frequently encountered car type. Therefore, a middle-class vehicle (second-generation Renault Scenic) was selected for the scenes. The vehicle has a 1.6-liter engine with 112 hp, a mass of 1400 kg, and standard 205 R15 tires. All scenes were recorded on public roads in and around the city of Dresden.

6.1.2 Recording of the Scenes

The selected scenes were recorded to enable a controlled and reproducible presentation to the participants. The goal was to capture the stimuli perceived by the driver in all scenarios. The recording setup is shown in Figure 6.1. Optical stimuli were recorded with a Canon EOS 600D Camera at a resolution of 1080p and a framerate of 30 frames per second. The camera was

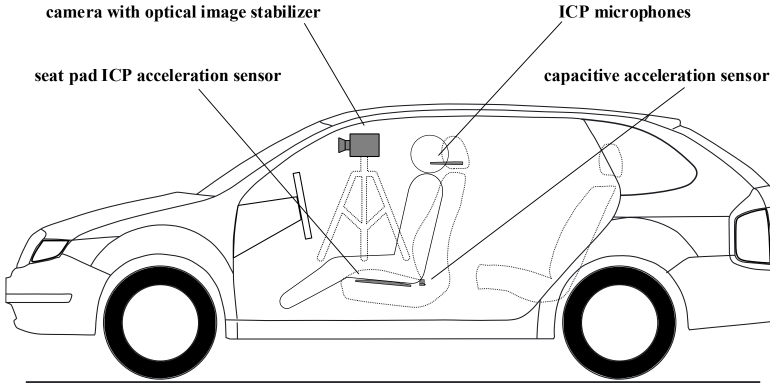

Figure 6.1.
Vehicle scenes were recorded with a seat pad accelerometer, a capacitive acceleration sensor, two microphones, and a camera.

mounted onto a tripod attached to the center console of the vehicle. The field of view was directed in the direction of driving with the front window frame excluded. The shoulder-neck section has a resonance frequency of about 4 to 5 Hz [30] and thus shows low pass characteristic for higher frequencies. Therefore, a mechanical motion stabilizer was used to eliminate motion in the video recordings that would not have been visually perceived by the driver in the vehicle.

Sound and vibration recordings were conducted with the software HEAD Recorder and a Zodiac DataRec 4 measurement frontend. Acoustical stimuli were recorded with two B&K 4188 microphones with B&K 2671 ICP amplifiers at 48 kHz sampling frequency. They were mounted at ear height on the headrest. Due to the utilization of a wave field synthesis reproduction system instead of headphones, in-ear recordings or dummy head recordings would have been unsuitable.

Vertical vibrotactile stimuli were recorded at the driver's seat with two triaxial sensors at a sampling frequency of 16 kHz. An ICP B&K 4515B seat pad acceleration sensor was placed between the driver and the seat's surface to record vibration above 4 Hz. A capacitive Kistler 8305B10 acceleration sensor was placed at the edge between the seat surface and the backrest to record vibration below 4 Hz.

6.1.3 Recorded Stimuli

The stimuli were created by choosing appropriate segments of the recordings that represented the scene types selected in section 6.1.1. The wide range of different scenes resulted in the duration varying between 4 s and 17 s providing enough time to get an impression of the scene while ensuring participants would focus on the same time segment of the scene. Stimuli were prepared for accurate reproduction according to the procedure described in chapter 4. The vertical recordings of the capacitive sensor were utilized for preparing the motion platform playback. The vertical recordings of the ICP sensor were utilized for preparing the electrodynamic shaker playback. For the seven impulse-like scenes, the beginning and the end of each impulse event was identified manually from the time signal. The beginning and the end were defined as the signal envelope rising and falling above the ambient envelope. However, these impulse like events lasted only from about 0.5 s to about 0.8 s. In a preliminary test, this short scene duration made it difficult for participants to immerse themselves in the multimodal scenes and it's auditory and visual context. Thus, the total duration was extended to at least 4 s, providing more time to gain an impression in each modality. A “rate now” subtitle was inserted into videos of the seven impulse-like scenes, to communicate to the participants the time span of the event to be rated.

Four of the 21 scenes are discussed in detail to provide a general idea about the range and properties of the vibrations contained in these scenes. The first scene was recorded on a cobblestone road at a constant speed of 30 kph. The acceleration signal obtained from the seat pad acceleration sensor is shown in Figure 6.2. The minimum frequency displayed here is 4 Hz, since vibration below 4 Hz were recorded with the capacitive sensor. There is broadband stochastic structural excitation as well as periodic excitation at about 60 Hz due to the cobblestone regularities. The total RMS acceleration level L_{acc} was at 130 dB, which is well above the perception threshold. The majority of the signal energy is in the frequency range of up to 200 Hz. Vibration components above 200 Hz are likely below the perception threshold, which is at about 90 dB for 200 Hz.

The second scene was recorded on a tarmac road at a constant speed of 70 kph. The acceleration signal of this scene is shown in Figure 6.3. The vibration of this scene consists of mostly broadband stochastic structural excitation. The total RMS acceleration level L_{acc} was at 111 dB, which is slightly above the perception threshold. The majority of the signal energy is in the frequency range up to 200 Hz for this scene as well.
The third scene was recorded on a tarmac road. Since the car was accelerating, the speed ranged from about 5 kph to about 50 kph. The vibration obtained from the accelerating vehicle is shown in Figure 6.4. As for the previous scene, the vibration of this scene consists of broadband stochastic structural excitation. However, there is also a periodic vibration component in the 70 Hz to 100 Hz range. This component is caused by the second motor order of the engine under high load. Since the JNDF of WBV is 30 % and higher [100] this component can still be considered as quasi non-transient from a perceptual perspective. The total RMS acceleration level L_{acc} was similar to the previously discussed scene at 110 dB. Again, the majority of the signal energy is in the frequency range up to 200 Hz.
The fourth scene was recorded on a tarmac road at a constant speed of 50 kph. The vehicle is passing over an extension joint of a bridge. The acceleration signal of the impulse-like event is shown in Figure 6.5. The vibration of this scene consists of broadband stochastic structural excitation with a shock event embedded in the middle of the scene. The peak acceleration level of the shock $\hat{L}_{acc}$ was at 135 dB, which is far above the perception threshold. The majority of the signal energy is in the frequency range up to 200 Hz also in this scene.

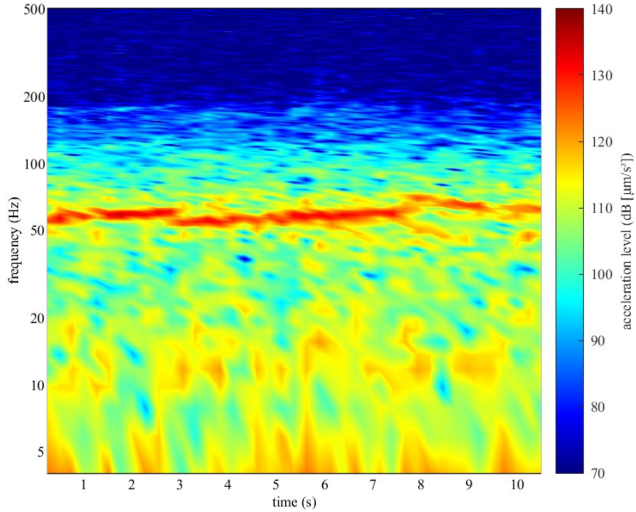

Figure 6.2.
Spectrogram (STFTs: 8,192 samples, 50 % overlapping Hann windows) of the vertical acceleration of the cobblestone road scene at 30 kph.

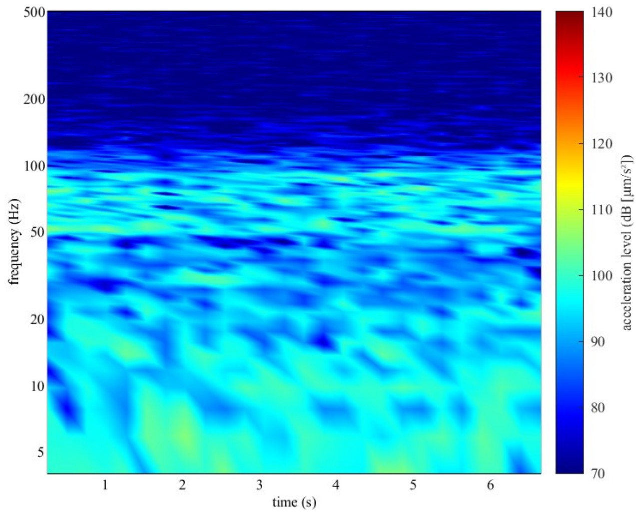

Figure 6.3.
Spectrogram (STFTs: 8,192 samples, 50 % overlapping Hann windows) of the vertical acceleration of the tarmac road scene at 70 kph.

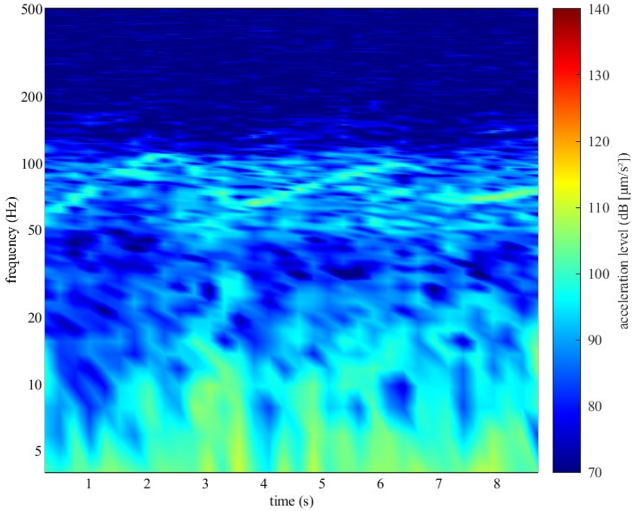

Figure 6.4.
Spectrogram (STFTs: 8,192 samples, 50 % overlapping Hann windows) of the vertical acceleration of the tarmac road scene from 5 kph to 50 kph.

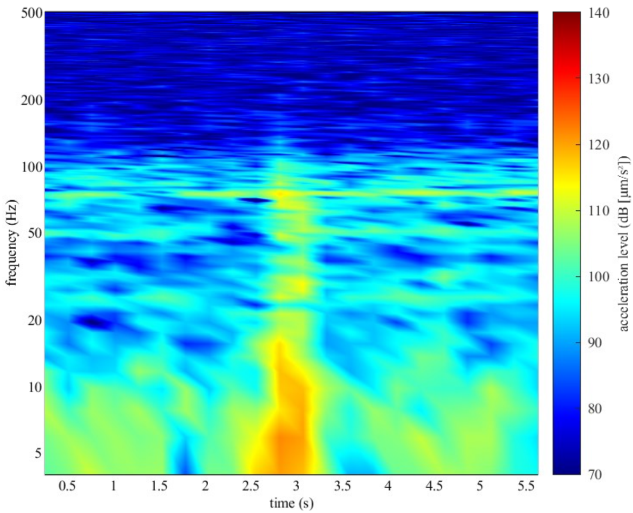

Figure 6.5.
Spectrogram (STFTs: 8,192 samples, 50 % overlapping Hann windows) of the vertical acceleration of driving over an extension joint on a tarmac road at 50 kph.

6.2 Qualitative Communication in the Presence of Vibration

In chapter 5 the sensory tactile design language was assessed with generalized vibration. The first investigation of this chapter is concerned with confirming that it can be meaningfully applied to complex vehicle stimuli. That implies that the tactile design language should be sufficient to capture the perceived variation of the representative stimulus set. However, the sufficiency could not be assessed directly at this stage of the overall investigation and will be assessed in chapter 7. Instead, it is possible to search for evidence contradicting this requirement. Such evidence might be observed in the form of additional attributes mentioned frequently for describing the representative stimulus set in comparison to the abstract stimulus set in chapter 5.

6.2.1 Experimental Design

Therefore, a free association task was conducted to determine the most frequently elicited sensory-perceptual attributes for describing the vibration of the 19 scenes. To enable comparison to the free association task conducted in section 5.3, the experiment design was identical, except for the stimuli. All multimodal stimuli described in section 6.1 were presented in the auditory-visual-tactile virtual environment described in chapter 4. All stimuli were randomized. For the non-transient scenes, participants were asked to come up with associations referring to the overall scene. For the transient scenes, participants were asked to come up with association referring only to the impulse-like event marked by the video subtitle “rate now”.

In section 5.3 an optional list with potential attributes, found by previous studies [50], [61], [105] was provided to participants to facilitate the association task. Since the occurrence of attributes in the previous study can be understood as evidence for their relevance in describing vibration, a potential bias towards attributes on the list was acceptable. Therefore, after coming up with associations themselves, subjects were handed a list that consisted of the attributes of the previous studies. As for section 5.3, the time to think about associations for one stimulus was not limited and varied by participant. Typically, it did not exceed one minute. Participants could play the stimulus again as often as necessary. Due to the comparably lower number of scenes, one session of one hour was more than sufficient to solve this task.

6.2.2 Participants

A total of 20 German native laypersons (15 male, 5 female) with an average age of 25 years (18 to 52 years) took part in the experiment. The study was conducted with the understanding and written consent of each participant.

6.2.3 Results

For the twelve non-transient scenes a total of 90 unique sensory-perceptual attributes were mentioned at least once. For the seven transient scenes, a total of 81 unique sensory-perceptual attributes were mentioned at least once. To identify a redundancy reduced set of most commonly used sensory-perceptual attributes, the same reduction steps were applied as in section 5.3.4 to the twelve non-transient and the seven transient scenes. First, low frequently mentioned attributes were merged to reduce redundancy. Attributes with the same word stem were merged, omitting the less frequent attribute and summing its occurrences to the more frequent attribute. This merge was also conducted for synonyms and antonyms according to the machine-generated *Openthesaurus* [106], but only the replacements utilized in section 5.3.4 were used. Second, low frequently mentioned attributes were omitted. Attributes not occurring more than 15 % for at least one stimulus were removed. Attributes mentioned in less than 2 % of the judgments (number of participants times the number of stimuli) were also removed. For the abstract stimulus set, the reduction steps resulted in 21 attributes. The list of the most frequently occurring attributes for the vehicle stimulus set is shown in Table 6.2.
When comparing the found attributes to the attributes identified in Table 5.2 of section 5.3.3, it is obvious that all attributes except the attribute “billowing” (“wabern”) are among the set of the 21 attributes found for the abstract stimulus set. However, when consulting *OpenThesaurus* [106], “billowing”

Table 6.2.
The most frequently mentioned attributes for the vehicle scenes and their number of occurrences across all subjects and scenes. Attributes were translated from German to English.

Attribute (Translation)	Attribute (German)	Number of Occurrences Non-Transient	Number of Occurrences Transient	Observed for generalized vibration
decaying	abklingend	0	0	yes
up and down	auf und ab	25	14	yes
humming	brummend	13	0	yes
uniform	gleichmäßig	19	0	yes
bumpy	holprig	114	59	yes
tingling	kribbelnd	27	8	yes
fading	nachschwingend	0	0	yes
pulsating	pulsierend	0	0	yes
rattling	ratternd	76	36	yes
grinding	rauschend	20	8	yes
calm	ruhig	67	15	yes
shaky	rüttelnd	45	19	yes
jolting	schlagend	24	43	yes
weak	schwach	43	27	yes
buzzing	summend	34	6	yes
ticking	tickend	0	0	yes
trembling	wackelnd	71	35	yes
soft	weich	24	22	yes
repetitive	wiederholend	0	0	yes
throbbing	wummernd	17	8	yes
shuddering	zittrig	29	8	yes
billowing	wabernd	14	0	no

(“wabernd”) is associated with waving (“wogend”), which is associated with “raise and sink” (“heben und senken”) it becomes obvious that this is simply a synonym for the percept associated with “up and down”. In line with the merging of synonyms and antonyms in section 5.3.4 simply merging the attribute “billowing” (“wabernd”) to “up and down” will remove the only additional attribute. It seems likely that participants of this experiment simply utilized a synonym to “up and down” to communicate about the similar percept.

6.2.4 Summary

The results of this study provide first tentative evidence that no other attributes than the 21 attributes identified in section 5.3 are required to describe the sensory properties of the representative stimulus set. Thus, all the sensory properties elicited by everyday life WBV are sufficiently represented by the 21 sensory attributes. The findings of section 5.4.4 suggest that the variation of these 21 attributes can be reasonably represented by only the six attributes forming the sensory tactile design language. Therefore, we can conclude that the sensory tactile design language can likely be used to communicate about the sensory-perceptual properties of complex everyday life stimuli.

6.3 Quantitative Communication in the Presence of Vibration

In the previous section, it was concluded that qualitative communication with the tactile design language about the sensory tactile perceptual properties is possible. However, this communication should not only happen qualitatively. It should also enable the quantification of the sensory tactile perceptual properties in rating scores. The scenes were selected to represent the range of everyday life encounters to WBV. It is obvious that there should be some perceivable differences. Thus, the six sensory-perceptual attributes of the tactile design language should reflect these differences by scenes influencing the rating score. Furthermore, the factor attribute should influence the rating score, to confirm that attributes are not completely redundant. However, the meaningfulness of these sensory-perceptual attribute ratings, i.e. how well they represent the perceptual properties, will be investigated in chapter 7. The assessment of ratings of the six sensory tactile perceptual attributes for each scene will be required for this step.

6.3.1 Experimental Design

The experimental design of this experiment was identical to section 5.4, except for the stimuli. All multimodal stimuli described in section 6.1 were presented in the auditory-visual-tactile virtual environment described in chapter 4. All stimuli were randomized. For the non-transient scenes, participants were asked to rate the overall scene according to the attributes shown. For the transient scenes, participants were asked to rate the impulse-like event marked by the video subtitle “rate now” according to the attributes shown. Each of the six selected German attributes “up and down” (“Auf und ab”), “tingling” (“kribbelnd”), “weak” (“schwach”), “repetitive” (“wiederholend”), “uniform” (“gleichmäßig”), and “fading” (“nachschwingend”) were rated for each scene. The attributes were grouped into two attribute triples. Attribute ratings were conducted on the quasi-continuous Rohrmann rating scale with verbal anchors [92] (see section 4.2). As for section 5.4, participants were allowed to repeat the stimulus, but rarely did so. The rating time of each attribute triple per stimulus was not limited, but typically took approximately 10 seconds. The experiment was conducted in one session of about 15 minutes.

6.3.2 Participants

A total of 31 German native laypersons (10 male, 21 female) with an average age of 26 years (18 to 61 years) took part in the experiment. The study was conducted with the understanding and written consent of each participant.

6.3.3 Results

The rating profiles for each of the 19 scenes consisting of judgments for each of the six perceptual attributes are shown in Figure 6.6. The range of the confidence intervals for the attribute ratings is 4 to 13 points, suggesting reasonably accurate quantifications. At first glance, the attribute ratings explain obvious differences between the scenes. The cobblestone road scenes contained vibration with a higher acceleration level compared to the tarmac scenes. For

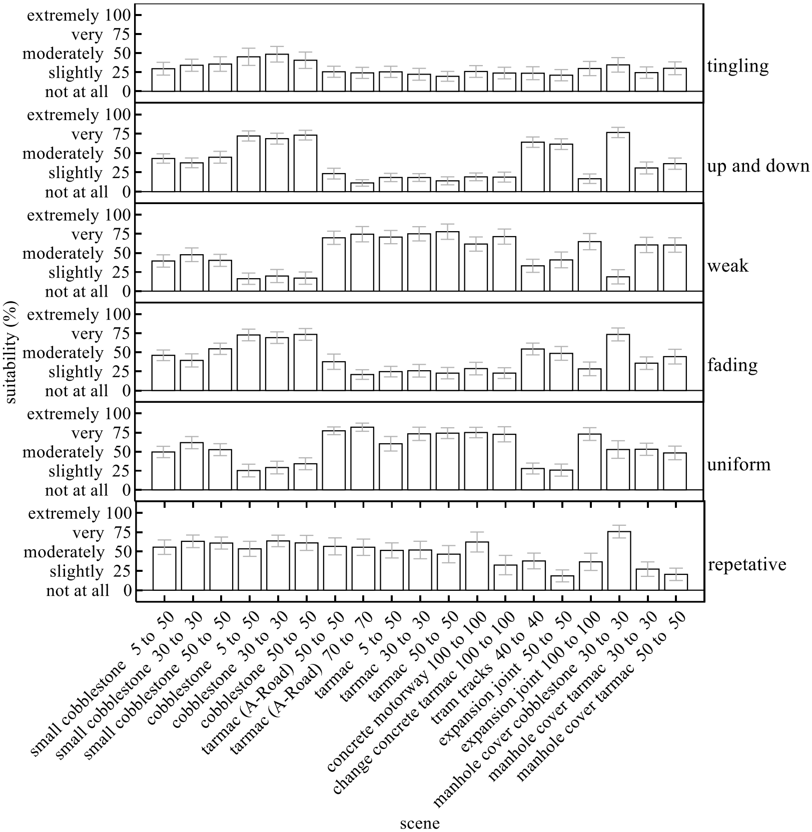

Figure 6.6.
Rating profiles for each of the six perceptual attributes consisting of mean ratings and 95 % confidence intervals for each of the 19 scenes.

the former, the attribute “weak” shows much lower ratings than for the latter. The attribute “up and down” shows higher ratings for the cobblestone scenes than for the tarmac scenes, while this difference is inverted for the attribute “tingling”. The cobblestone road is considered to be more uneven than the tarmac road. This difference was quantified by higher ratings of the attribute “uniform” for cobblestone scenes compared to tarmac scenes. Thus, there are obvious rating differences between different scenes and also between attributes.

To evaluate this finding from a statistical point of view, a two-way within-subjects repeated-measures ANOVA was conducted with the software IBM SPSS with factor scene and attribute. The factor scene shows a highly significant effect ($F(6.147,184.396) = 9.434$, $p < 0.001$), confirming that the sensory tactile attributes can explain perceptual differences of vibration. The factor attribute also shows a highly significant effect ($F(2.876,86.281) = 16.267$, $p < 0.001$) for the multimodal mode, confirming that not all attributes are redundant. The interaction of both factors is also highly significant ($F(16.215,486.458) = 35.041$, $p < 0.001$). The Greenhouse–Geisser adjustment was used to correct for violations of sphericity for both factors and the interaction.

6.3.4 Summary

It can be concluded that attribute ratings can likely reflect differences in perceptual properties of different scenes. Thus, there is a tentative evidence that the sensory tactile design language is not only sufficient for communicating qualitatively about the elicited percepts. It is also possible to utilize the language for communicating quantitatively about the sensory-perceptual properties of WBV.

6.4 Quantitative Communication in the Absence of Vibration

In the previous section, it was concluded that not only qualitative communication with the tactile design language about the sensory tactile perceptual properties is possible, but also quantitative communication. This scenario is useful for analyzing existing vibration, but not optimal for synthesizing new vibration based on expected properties. By rating one vibration out of many potentially occurring vibrations in a situational context, only elicited properties and not necessarily expected properties are assessed. This assumes that these properties are likely similar to the expected sensory-perceptual properties. However, it would be much more beneficial to quantify the expected

properties directly since it would enable refraining from presenting any vibration at all.
But why are expectations of such importance with regards to quality judgments and plausibility judgments alike? For both domains, it would be beneficial to assess the expected features of a situational context since they are the inner reference against which the elicited perceptual properties are compared. Such a comparison requires that the elicited perceptual object and the inner reference share the same features. According to [108] quality judgments can be defined as the comparison of the perceived composition of an entity to its desired composition. The perceived composition refers to the totality of the features of the entity, while the desired composition refers to the totality of expected features. In the domain of virtual environments [77] argues that plausibility judgments can be defined as a comparison between a perceptual object and inner reference, which is shaped by expectations and previous experiences. This comparison can be understood as a similarity judgment with respect to all relevant perceptual features.
Thus, it should be possible to assess the sensory tactile perceptual attributes of the inner reference, i.e. the expected sensory tactile perceptual properties. The findings of [109] suggest that it is possible to associate tactile stimuli to perceptual attributes conveyed in the auditory modality. The research question of this section is determined by this supposition. It will be investigated whether the anticipated sensory tactile perceptual attribute ratings of the inner reference of a situation context are reasonably similar to ratings of vibration experienced live in this situational context.

6.4.1 Experimental Design

The experimental design was similar to section 6.3, but the representative scene set was rated in two presentation modes. The first mode was auditory-visual-tactile scene presentation (multimodal), in which scenes were presented in the auditory-visual-tactile virtual environment identical to section 6.3. The multimodal scenes were preferred over tactile only scenes for two reasons. First, the inner references were likely formed by multimodal experiences of the situational contexts. Therefore, potential multimodal interaction or integration effects should also be included in the attribute ratings of vibration occurring in a situational context. Second, the vibration generated by the suggested approach will nearly always be embedded into an audio-visual situational context. The task of the multi-modal condition was to rate the tactile properties elicited by the vibration of the scene. For the non-transient scenes, participants were asked to rate the overall scene according to the attributes shown. For the non-transient scenes, participants were asked to rate the impulse-like event marked by the video subtitle “rate now” according to the attributes shown.

To enable the rating of the expected perceptual attributes for a situational context, the situational context needed to be communicated to the participants. Therefore, verbal descriptions of the scene content consisting of vehicle type, speed, operating condition (constant speed, accelerating), road surface as described in section 6.1.1 were created. The stimuli of the anticipated mode consisted only of these descriptions without any audio, visual or tactile reproduction. The task of this mode was to imagine the sensory tactile perceptual properties of the vibration from the written description of the scene, e.g. “Imagine you are driving with a middle-class vehicle over a rough cobblestone road at a constant speed of 30 kph. What would the vibration feel like?”.

The stimuli of the two presentation modes were presented in separate blocks. The anticipated mode was always presented as the first block to avoid the multimodal mode influencing the anticipated mode. All stimuli were presented in random order per block. Each of the six selected German attributes “up and down” (“Auf und ab”), “tingling” (“kribbelnd”), “weak” (“schwach”), “repetitive” (“wiederholend”), “uniform” (“gleichmäßig”), and “fading” (“nachschwingend”) was rated for each of the 19 scenes in each of the two presentation modes. Attribute ratings were conducted on the quasi-continuous Rohrmann rating scale with verbal anchors [92] (see section 4.2) again. A repeated-measures design was chosen i.e. each subject rated each scene and attribute in each presentation mode to facilitate the comparison between the ratings of the two modes. As for the ratings of the multimodal only experiment, participants were allowed to repeat the stimulus, but rarely did so. The experiment was conducted in two sessions of approximately 15 minutes each.

6.4.2 Participants

A total of 22 German native laypersons (13 male, 9 female) with an average age of 30 years (19 to 61 years) took part in the experiment. Participants were unfamiliar with the scenes of the multimodal presentation mode before the experiment to avoid influencing the ratings of the anticipated presentation mode. The study was conducted with the understanding and written consent of each participant.

6.4.3 Results

The rating profiles of the two presentation modes for each of the 19 scenes consisting of judgments for each of the six perceptual attributes are shown in Figure 6.7. Surprisingly, the confidence intervals for the sensory tactile perceptual attribute ratings of the anticipated mode exhibit a similar range from 4 to 16 scale points compared to the multimodal mode with a range from 4 to 15 points. This suggests that there is similar agreement across individual judgment of a presented scene (multimodal mode) and individual judgments of an imagined scene (anticipated mode) for a specific situational context. As for the previous experiment, the factor attribute and the factor scene clearly show an influence on the rating averages of both presentation modes. Furthermore, the factor presentation mode seems to have a much smaller influence. The rating difference based on estimated marginal means between anticipated mode and multimodal mode across all scenes and attributes is at only 2 points on the 100-point scale. The differences for the single attributes across all scenes is 2 for “up and down”, 4 for “uniform”, -1 for “tingling”, 3 for “fading”, 6 for “weak”, and -3 for “repetitive”.
These findings were assessed with a three-way within-subjects repeated measures ANOVA with factor scene, attribute, and presentation mode (multimodal vs. anticipated) to allow for statistical inferences. As for the multimodal mode only study (see section 6.2), the factor scene shows a highly significant effect ($F(6.213,130.467) = 20.594$, $p < 0.001$). Also the factor attribute shows a highly significant effect ($F(3.234,67.924) = 10.720$, $p < 0.001$). The Greenhouse–Geisser adjustment was used to correct for violations of

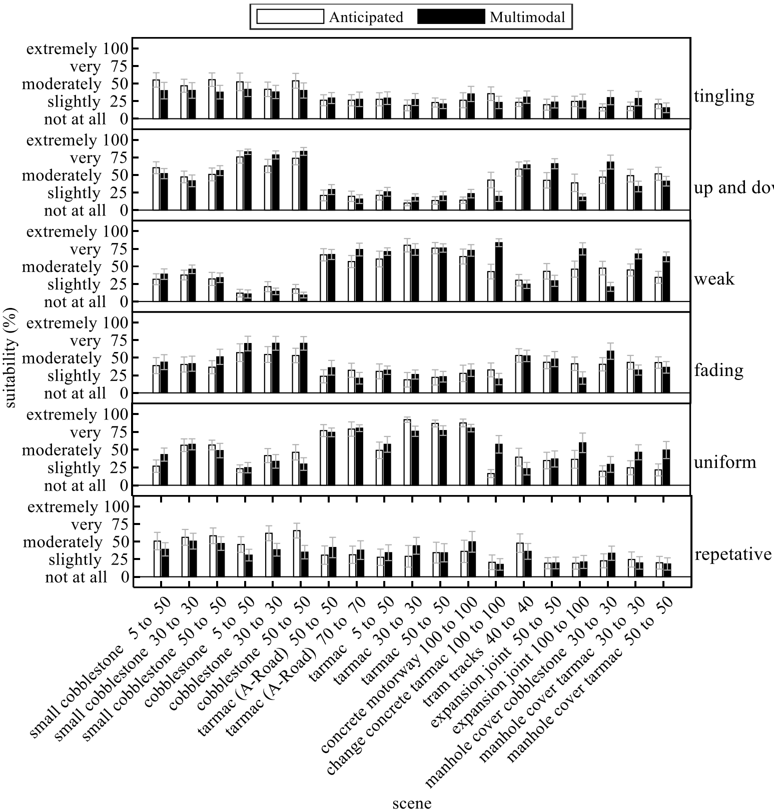

Figure 6.7.
Rating profiles of the anticipated presentation mode (white) and for the multimodal mode (black) for each of the six perceptual attributes consisting of mean ratings and 95 % confidence intervals for each of the 19 scenes.

sphericity for both factors. Contrary to these two factors, the factor presentation mode has no significant effect ($F(1,21) = 1.895$, $p = 0.183$) on mean attribute ratings overall attributes and all scenes. The interaction of factors scene and attribute is also highly significant (Greenhouse-Geisser corrected degrees of freedom $F(11.877,596.865) = 36.457$, $p < 0.001$). The interaction of factors scene and presentation mode is also significant (Greenhouse-Geisser corrected degrees of freedom $F(6.316,132.626) = 3.850$, $p < 0.01$). The interaction of factors attribute and presentation mode is not significant

(Huynh-Feldt correct degrees of freedom $F(4.878,102.428) = 4.878$, $p = 0.054$). The interaction of factors scene, attribute, and presentation mode is also highly significant (Greenhouse-Geisser corrected degrees of freedom $F(13.057,274.200) = 7.719$, $p < 0.001$). However, while this finding suggests that the presentation modes show similar attribute ratings, they are not necessarily identical. Instead, the effect size of the differences between multimodal and anticipated mode provides a basis for inferences about the degree of similarity between the two modes. A pairwise contrast between multimodal and anticipated mode shows a mean effect size of 2 points difference. The 95 % confidence interval of the difference ranges from -1 to 4 points. Therefore, the mean difference between presentation modes is unlikely to exceed a fifths of the 25-point scale tick interval or one twenty-fifth of the 100-point scale. Furthermore, the effect sizes of the factors and interactions on the dependent variable for within-subjects repeated measures designs can be estimated from the generalized η^2 as suggested by [110] and thus their influence compared. Almost all of the variance is explained by the factors attribute, scene, and their interaction, further confirming the negligible effect of presentation mode. Overall, this suggests that the difference caused by the different presentation modes is likely irrelevant for practical purposes.

Although the overall difference between the ratings of the anticipated and multimodal mode is not significant, differences of about a quarter of the rating scale could be observed at some of the scenes for some attributes. One potential explanation might be found in the very general descriptions that might not account for all of the variance potentially occurring for this scene category. When participants assess their expectations, they might think of a typical scene matching the description. If the multimodal scene was not very typical for the scene category conveyed by verbal description, then subjects might have e.g. compared the properties of a typical drive over cobblestone with an atypical. This underlines the necessity of providing accurate verbal characterizations of the scene to be rated.

6.4.4 Summary

It was demonstrated in this section that presenting the vibration of a situational context and assessing the expectations of a situational context directly result in extremely similar attribute ratings. This implies that the expected ratings of the attributes of the tactile design language are predictive of the attribute ratings that would have been elicited by experiencing live vibration of this situational context.

6.5 Summary and Discussion

First, a representative set of multimodal vehicle with WBV was selected and subsequently recorded. Subsequently, in a free association task, it was found that only attributes were elicited by these scenes, which were identical or synonymous to the attributes found in the process of assessing the sensory tactile design language in chapter 5. This suggests that the six sensory-perceptual attributes of the language are likely sufficient to qualitatively communicate about the perceptual properties of scenes with WBV. The following experiment demonstrated that absolute ratings of these attributes also reflect differences in the sensory-perceptual properties of the multimodal scenes. This suggests that the design language enables quantitative assessment of sensory-perceptual properties of WBV.

Since perceptual properties elicited by live vibration of a certain context are compared to the expected properties of that situational context in quality and plausibility judgments, the anticipated properties of vibration were assessed directly. Extremely similar sensory-perceptual attribute ratings of verbal descriptions of the scene compared to live vibration of a multimodal scene, suggests that anticipated ratings are predictive of the properties elicited by live vibration of the described situational context. Thus, the expected sensory-perceptual properties can likely be assessed directly from expectations without the need of presenting any vibration at all.

However, the question remains, if such quantifications sufficiently reflect the perceptual properties of vibrations to enable their translation into vibration matching user expectations. Therefore, attribute rating profiles assessed with the sensory tactile design language provided the necessary basis for the subsequent vibration synthesis for the scenes in chapter 7 to investigate this question.

7. Synthesis Models for the Translation of Sensory Tactile Properties into Vibration

In chapter 5 a sensory tactile design language was assessed which is suitable for the efficient profiling of sensory tactile perceptual properties of WBV. In the process of the assessment, the sensory tactile perceptual attributes eventually forming the design language were rated for a set of 99 stimuli. The stimulus set consists of four excitation patterns, whose parameters were systematically varied to cover the range of perceivable everyday life WBV. Thus, there is a database of vibration parameter tuples mapping onto rating profiles of the six sensory tactile perceptual attributes for each stimulus. This mapping is the basis to build models in this chapter for synthesizing vibration that will elicit the desired sensory tactile perceptual properties. The input of such a model should consist of profiles of the six sensory-perceptual attributes, too. The expected sensory tactile perceptual properties of a set of 19 scenes representing typical exposure to WBV in everyday life were quantified in the form of such rating profiles in chapter 6.

Building onto chapter 5 and chapter 6 this chapter will focus on creating a model that enables the systematic evocation of the plausibility illusion for arbitrary situational contexts. This model needs to translate the expected sensory tactile perceptual properties of a situational context into physical vibration parameters. From these parameters, WBVs need to be synthesized. These WBVs should be perceived as similarly plausible as the recorded vibrations in the situational context, since both should elicit the expected sensory tactile perceptual properties.

First, the tactile plausibility illusion needs to formalized to enable model building. Furthermore, while a linear relationship between vibration level and perceived magnitude has been reported [32] for a magnitude estimation experimental design, it is unclear if this finding can be transferred to absolute rating scales utilized in this work. Thus, the influence of vibration level on perceptual attribute ratings needs to be investigated as a prerequisite to model creation. Furthermore, in chapter 5 it was shown that AM-sinusoidal vibration and multiple successive impulse-like vibration elicit similar attribute ratings if their vibration signals have similar envelopes. If a vibration is to be generated from attribute ratings, such an ambiguity would be undesirable. Therefore, attribute ratings of AM-sinusoidal signals should be compared to successive impulse-like signals with identical RMS and modulation frequency. This would provide a basis for deciding whether to omit or keep the successive impulse-like vibration pattern.

The formalization of the plausibility illusion suggested a nearest neighbor model as the direct extension. Thus, such a model will be created in a first step to enable the translation of sensory tactile perceptual profiles into vibration. However, a k-nearest-neighbor model does not provide an explicit mathematical relationship, i.e. a function of vibration parameters depending on attribute ratings. To enable continuous predictions of vibration parameters from attribute ratings, also a regression model will be built. These regression models will also provide an inferential statistical point of view on the relationship between physical vibration parameters and perceptual attribute ratings in addition to the descriptive statistical point of view provided in section 5.4. Finally, both models need to be validated. Therefore, WBV will be generated from the attribute rating profiles obtained for the representative scene set of chapter 5 based on these models. The recorded vibrations of the 19 vehicle scenes will be replaced by the synthesized vibrations. A subsequent perceptual study will assess the perceived plausibility of the recoded vibration and synthesized vibration in the context of the audio-visual scenes. If the synthesized vibration is not perceived as less plausible than the recorded vibration, than the models are validated. Furtermore, this would also imply that the tactile design language sufficiently describes the sensory-perceptual properties of WBV. [3]

[3] Some parts of this work were presented in [81], [8], [97] and [98] by Rosenkranz et al.

7.1 Formalization of the Tactile Plausibility Illusion for Models

The basis for the development of a synthesis model that can produce plausible WBV is the formalization of the plausibility illusion. Furthermore, the synthesis model's inputs and outputs for the typical intended use case need to be defined.

7.1.1 Formalization of Plausibility

Mel Slater distinguishes place illusion and plausibility illusion as influencing factors for the perception and interaction with the virtual environment as with a real environment [6]. The place illusion is elicited when the user is immersed in the virtual environment. This illusion depends on the technical implementation of the stimulus presentation, e.g. display resolution or latency. The plausibility illusion is elicited when the content of the scene matches the user expectations. There are two approaches for generating content for virtual environments [74]. In many cases, the authentic approach is implicitly followed when attempting to elicit a perceptual object in the virtual environment, which is identical to the perceptual object elicited in the virtual environment. This implies that users can not perceive the difference between the real and the virtual environment in a paired comparison. However, such a comparison is impossible in the majority of usage scenarios of virtual environments. When a comparison to the real environment is not available, users need to rely on their expectations to judge the plausibility of the presented content. Thus, the plausible approach only attempts to elicit any perceptual object in the virtual which might have occurred in the corresponding real environment.

To judge whether the perceptual object might have occurred in this real environment, the users need to rely on their expectations on this real environment. Such expectations are necessarily dependent on previous experiences with such real environments. Therefore, the plausible approach to creating content relies on the user's ability to judge the conformity of the elicited perceptual object with the expected perceptual object. Thus, this ability is not only limited by the capabilities of the tactile receptors, but rather by the capabilities of the user to memorize the vibrotactile properties of real environments. Since the real environment is not available for a direct A-B comparison, the user cannot judge perceptual identity between virtual and real environment. Instead, he can only make judgments about perceptual equivalence since many similar WBVs might have occurred in the corresponding real environment. This implies that the plausibility illusion is a measure of agreement between the perceptual object elicited in the virtual environment and the expected perceptual object.

Indeed [77] argues that a plausibility judgment can be interpreted as a similarity judgment. Their study focused on the plausibility of auditory room simulations e.g. of reflections and reverberations. An MDS based on similarity judgments between different virtual rooms demonstrated two implicit dimensions underlying these judgments. Furthermore, plausibility judgments were obtained for each virtual room for a verbally communicated room type, e.g. “living room”. For each given room, they compared the Euclidean distance in this two dimensional to the perceived plausibility. Their results suggest that the plausibility decreases with increasing Euclidean distance in the two-dimensional perceptual space. This suggests that the perceptual object elicited in the virtual environment is compared to the recalled inner reference defined by expectations on a specific context i.e., a scene in an n-dimensional perceptual feature space.

However, to utilize this relationship, explicit perceptual properties are required that enable communication about the inner reference, i.e. the expectations for a specific situational context. Only then, it would be possible to quantitatively select vibrations by the distance between their elicited perceptual objects and the inner reference in the sensory tactile perceptual space, thus selecting a maximally plausible vibration. The tactile design language obtained in chapter 5 was assessed precisely to overcome this shortcoming. Thus, the six perceptual attributes “weak”, “up and down”, “tingling”, “repetitive”, “uniform”, and “fading” contained in the design language are initially assumed to sufficiently represent the sensory tactile perceptual space. These sensory tactile perceptual attributes can be quantified in a rating profile **r** which consists of ratings of six attribute $a_1 \ldots a_6$:

$$\boldsymbol{r} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{pmatrix} \qquad \textbf{7.1}$$

Such a rating profile **r** consists of r_w the attribute rating of “weak”, r_d the attribute rating of “up and down”, r_t the attribute rating of “tingling”, r_r the attribute rating of “repetitive”, r_u the attribute rating of “uniform”, and r_f the attribute rating of “fading”:

$$\boldsymbol{r} = \begin{pmatrix} r_w \\ r_d \\ r_t \\ r_r \\ r_u \\ r_f \end{pmatrix} \qquad \textbf{7.2}$$

The results from chapter 5 show that presenting multimodal recordings of a situational context will produce extremely similar sensory tactile perceptual attribute ratings as communicating the situational context with a verbal description. This suggests that the properties of the expected perceptual object i.e. the inner reference can be meaningfully assessed. The properties of the expected perceptual object **e** can be represented by the six-dimensional vector attribute rating profile $\mathbf{r}_{\mathbf{expected}}$:

$$\boldsymbol{e} = \boldsymbol{r}_{expected} \qquad \textbf{7.3}$$

When presenting a vibration stimulus, a perceptual object **s** is elicited that is characterized by its perceptual properties quantifiable in the rating profile $\mathbf{r}_{\mathbf{elicited}}$:

$$\boldsymbol{s} = \boldsymbol{r}_{elicited} \qquad \textbf{7.4}$$

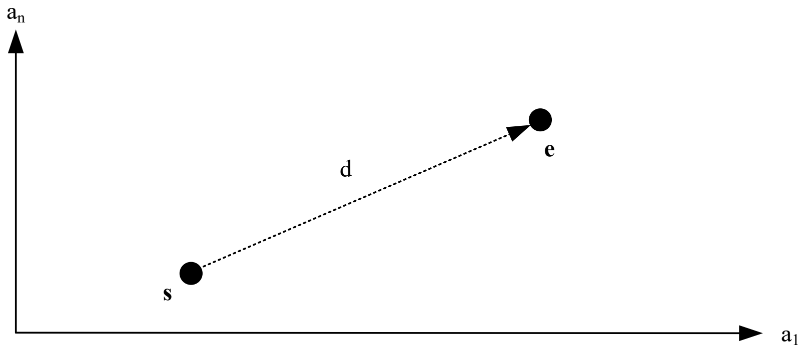

Figure 7.1.
The perceived plausibility of a vibration in a situational context can be interpreted as the distance between the perceptual object **s** elicited by this vibration and the perceptual object **e** expected for this context in the n-dimensional sensory tactile perceptual space.

The suggested relationship between plausibility p and the inverse Euclidean distance d between the expected perceptual object **e** and the elicited perceptual object **s** [77] is initially assumed to also apply to the n-dimensional sensory tactile perceptual space. The perceptual distance is depicted in Figure 7.1. The relationship between plausibility and the distance in the perceptual space is thus assumed to be formalizable according to the suggested proportional equation:

$$p \sim \frac{1}{d(\boldsymbol{s},\boldsymbol{e})} \qquad \textbf{7.5}$$

$$p \sim \frac{1}{\sqrt{\sum_{i=0}^{n}(e_i-s_i)^2}} \qquad \textbf{7.6}$$

Since the sensory tactile perceptual space is initially assumed to be representable by six perceptual attributes, the relationship can be reflected by:

$$p \sim \frac{1}{\sqrt{\sum_{i=0}^{6}(e_i-s_i)^2}} \qquad \textbf{7.7}$$

To translate the expected sensory tactile perceptual attribute profile **e** into plausible vibration, it is necessary to find a vibration, whose elicited attribute profile **s** has a minimal distance in the sensory tactile perceptual property space. Thus, the potential synthesis model needs to provide the vibration parameters that would ideally elicit an attribute rating profile, which is identical to the rating profile describing the inner reference, i.e. expected perceptual object.

7.1.2 Model Boundaries

Such a synthesis model can be integrated into a simulator generating the virtual environment as shown in Figure 7.2. A database is likely to contain the scene descriptions consisting of a duration of the scene, an expected sensory

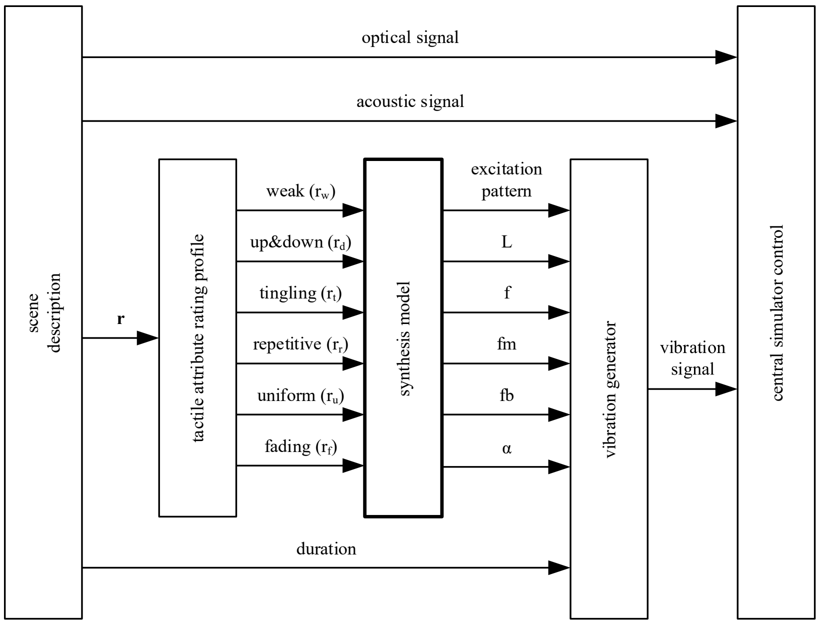

Figure 7.2.
Input and output of the synthesis model and its potential integration into a simulator for virtual environments.

tactile attribute rating profile **r**as well as signals of other modalities. These attribute ratings provide the input to the actual WBV synthesis model. The synthesis model needs to output the vibration parameters acceleration level L, vibration (carrier-, center- or resonance-) frequency f, bandwidth f_b, modulation frequency f_m, and decay constant α. These parameters, in turn, provide the input for a vibration signal generator. The resulting vibration signal as well as the signals of other modalities is relayed to the central simulator control, which is driving the simulator hardware.
As stated in section 3.2, this work focusses on vibration, which is considered to elicit a quasi-constant perceptual attribute rating profile for the selected duration. Thus, the synthesis in this work will investigate the general feasibility of the synthesis approach and will only focus on such scene segments that are represented by one attribute rating profile. However, it is very likely possible to provide a successive set of rating profiles and durations as an input, if some form of crossfading is added to the vibration signal generator for the transition between segments.

7.2 Investigation of the Influence of Vibration Level on Attribute Ratings

For the assessment of the tactile design language in section 5.2, the stimulus set was selected to cover the range of everyday life situations with WBV. However, only two acceleration levels were utilized, since small level differences would unlikely influence whether a perceptual attribute was elicited at all. For the subsequent absolute attribute ratings in the semantic differential, the same stimulus set was presented to participants. The two acceleration levels represent the lower boundary of everyday life exposure near the perception threshold and upper boundary of everyday life exposure near the exposure limits. These two levels would be sufficient to specify a linear model of the relationship between acceleration level and attribute ratings in the following sections. However, two levels are insufficient to decide whether it is necessary to increase model complexity beyond a simple linear model and thus whether ratings of more than two acceleration levels need to be obtained for each combination of the remaining vibration signal parameters.

The relationship between vibration level and perceived intensity has been studied before utilizing a magnitude estimation approach [32]. These findings were presented in Figure 2.4 of section 2.1.2. They suggest a linear relationship between acceleration level and perceived intensity. However, it is unclear if this finding can be transferred from relative magnitude estimation data to absolute rating scale data utilized in this work. Therefore, an experiment needs to be conducted to provide evidence that acceleration has a linear influence on the attribute ratings.

7.2.1 Stimuli

To enable comparisons to the study of [32] sinusoidal excitation was selected. The frequencies of the vertical WBV of 1 Hz, 15 Hz, and 150 Hz were in the range of this study as well. Furthermore, the SL was set to four steps in the range of this study at 10 dB (SL), 20 dB (SL), 30 dB (SL), and 36 dB (SL) to increase the resolution in the level range in comparison to section 5.2. The combination of the three frequency steps and the four level steps resulted in a total of twelve sinusoidal vertical WBV stimuli.

7.2.2 Experimental Design

The experimental design was identical to section 5.4, except for the stimuli. The twelve stimuli were presented with the tactile reproduction system described in chapter 4. Participants were allowed to repeat the stimulus, but rarely did so. Each of the six German attributes of the sensory tactile design language “up and down” (“auf und ab”), “tingling” (“kribbelnd”), “weak” (“schwach”), “repetitive” (“wiederholend”), “uniform” (“gleichmässig”), and “fading” (“nachschwingend”) was rated for each stimulus. The attributes were grouped into two attribute triples. Attribute ratings were conducted on the quasi-continuous Rohrmann rating scale with verbal anchors [92] (see section 4.2.2). The rating time of each of attribute triple per stimulus was not limited, but typically took approximately 10 seconds resulting in a total experiment duration of 15 minutes.

7.2.3 Participants

A total of 31 German native laypersons (10 male, 21 female) with an average age of 26 years (18 to 61 years) took part in the experiment. The study was conducted with the understanding and written consent of each participant.

7.2.4 Results

The attributes ratings and 95 % confidence intervals of the six sensory tactile perceptual attributes are shown in Figure 7.3. The attributes “up and down”, “tingling”, “fading”, and “weak” clearly demonstrate a linear relationship between acceleration level and attribute rating for a frequency of 1 Hz, 15 Hz, and 155 Hz. For the attribute “uniform” there is a slight deviation from a perfect linear relationship at 155 Hz towards the perception threshold at 10 dB (SL) of less than one rating scale tick. The attribute repetitive demonstrates a similar deviation from linearity towards the perception threshold for 15 Hz and 155 Hz. Therefore, it should be considered to introduce a more complex model for high frequency, near-threshold vibration. Over all attributes, there is no strong evidence against the assumption of a linear relationship between acceleration level and attribute rating for the six sensory tactile perceptual attributes forming the tactile design language especially in the range from 20 dB (SL) to 36 dB (SL). Assuming linearity would introduce a small error only for high frequency, near-threshold vibration for the two attributes “uniform” and “repetitive”. Thus, a linear relationship is a reasonable overall approximation for modeling the influence of the acceleration level.
It is also interesting to compare the intercept of the linear relationship of this data with the results from [32]. The attribute “weak” has a mean rating of 77 points at 10 dB (SL) and a mean rating of 22 points at 36 dB (SL) at 15 Hz. This would correspond to a 3.5-fold difference in perceived rating of “weak”. The results from [32] suggest that a level jump from 12 dB (SL) to 34 dB (SL) corresponds to an 8-fold difference in perceived intensity at 20 Hz. An explanation for the difference might be found in the susceptibility of ratio judgments to logarithmic response bias, as argued e.g. by [111] i.e. to an overestimation of the increase in perceived intensity with increasing acceleration level.

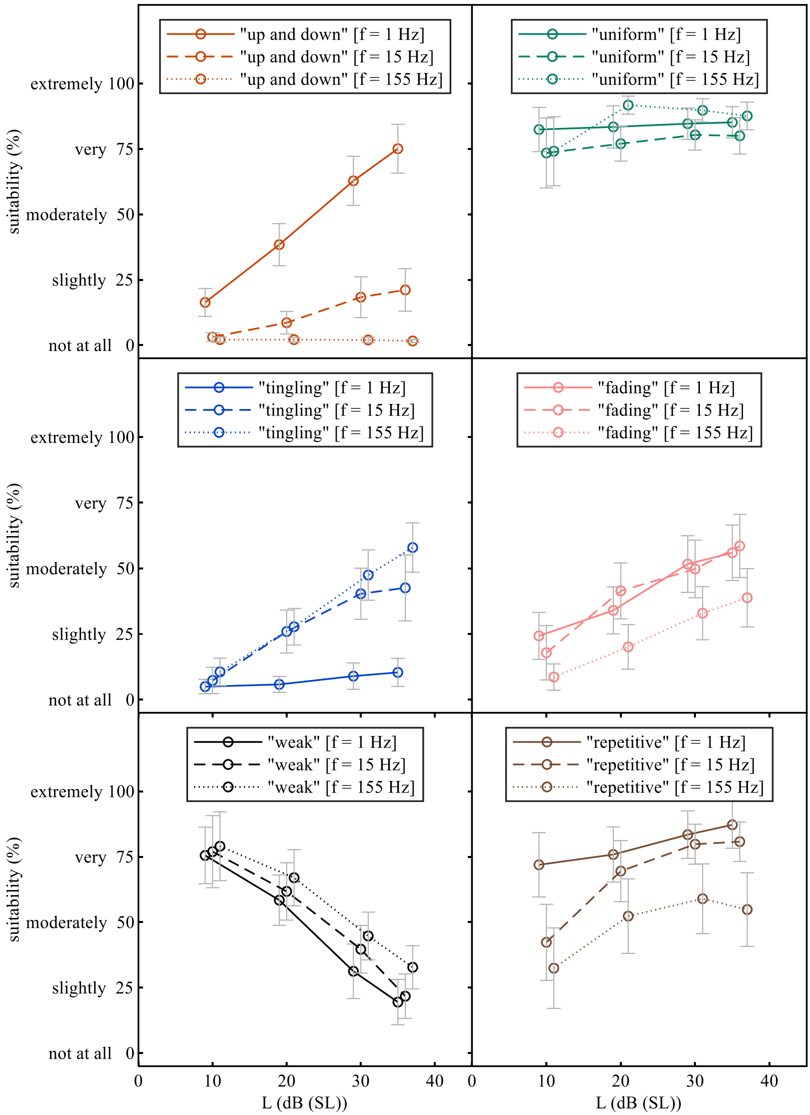

Figure 7.3
Mean attribute ratings and 95 % confidence intervals vs. acceleration level for the six sensory tactile perceptual attributes of the tactile design language.

7.2.5 Summary

Overall these findings suggest that there is an approximately linear relationship between acceleration level and attribute ratings in the level range of 10 dB (SL) to 36 dB (SL). Therefore, the subsequent modeling can be conducted with acceleration level as a linear term. This implies that only ratings of two acceleration levels are required for each combination of the remaining vibration signal parameters.

7.3 Comparison of Modulated Vibration to Successive Impulse-like Vibration

During the assessment of the tactile design language, both AM stimuli and multiple successive impulse-like stimuli were selected in 5.2 as excitation patterns occurring in everyday life situations. Subsequently, the 21 most frequently mentioned attributes were rated for these stimuli on a quasi-continuous scale in section 5.4. The results suggested very similar attribute ratings of AM stimuli and multiple successive impulse stimuli, for similar acceleration parameter values, i.e. acceleration level, carrier frequency or resonance frequency, and modulation frequency or effective repetition frequency. As discussed in section 5.2, the ecological approach to perception suggests that vibration is a carrier of information about the environment. Associating sensory-perceptual properties to vibration is related to the perceptual process of categorization in which many perceptually non-identical stimuli map onto one category and thus are perceptually equivalent. Therefore, different vibration signals mapping onto the same sensory tactile perceptual profile is to be expected.

While such a property is not an issue for an analysis model, it would be problematic for a potential synthesis model, if two vibration signals map onto the same sensory tactile perceptual profile. Thus, this ambiguous mapping would not enable the selection of one excitation pattern from a sensory tactile perceptual profile without introducing additional criteria for a synthesis model. Furthermore, it would needlessly complicate the model, since only one vibration needs to be output for each sensory tactile perceptual profile. However, the parameters RMS acceleration level and modulation frequency or effective repetition frequency were chosen independently. This complicates a direct assessment of the degree of similarity. Therefore, in this section, the similarity of AM stimuli and multiple successive impulse-like stimuli with identical RMS level and modulation frequency or effective repetition frequency will be compared regarding their ratings of the sensory tactile perceptual attributes of the tactile design language.

7.3.1 Stimuli

The characteristic parameters of the investigated excitation patterns were defined in section 5.2. Multiple successive impulse-like stimuli are characterized by their RMS SL, resonance frequency, decay constant, and repetition frequency. AM stimuli are characterized by their RMS SL, carrier frequency, and modulation frequency. To compare both excitation patterns, the parameter values of the AM stimuli need to be defined in such a way that the resulting vibration signal resembles the multiple successive impulses as closely as possible. However, the investigation should focus on the subset of the parameter range encountered in everyday life (see section 5.2), where the highest degree of dissimilarity between both excitation patterns is to be expected. Only if no major perceived difference is found in this range, then the multiple successive impulse excitation pattern could be substituted completely by the AM-sinusoidal excitation pattern.
The AM-sinusoidal excitation pattern is characterized by the slow rise and fall of the vibration signal envelope. The most dissimilar impulse-like excitation pattern should have a fast rise and fall in the vibration signal envelope shape as produced by a high decay constant. Therefore, the decay constant was set to the maximum value of 8 s^{-1} as used in section 5.2. Similarly, for low modulation frequencies or repetition frequencies, the difference in signal envelope shape should be maximal. However, when repetition frequency is very low, i.e. a fraction of one Hz, the resulting signal would increasingly be perceived as a sequence of independent single impulse events. Thus, the modulation frequency or repetition frequency was selected to be approximately 1.5 Hz, depending on the resonance frequency of the impulse-like vibration stimulus. Carrier frequency was selected to be 9 Hz, 15 Hz, and 90 Hz to

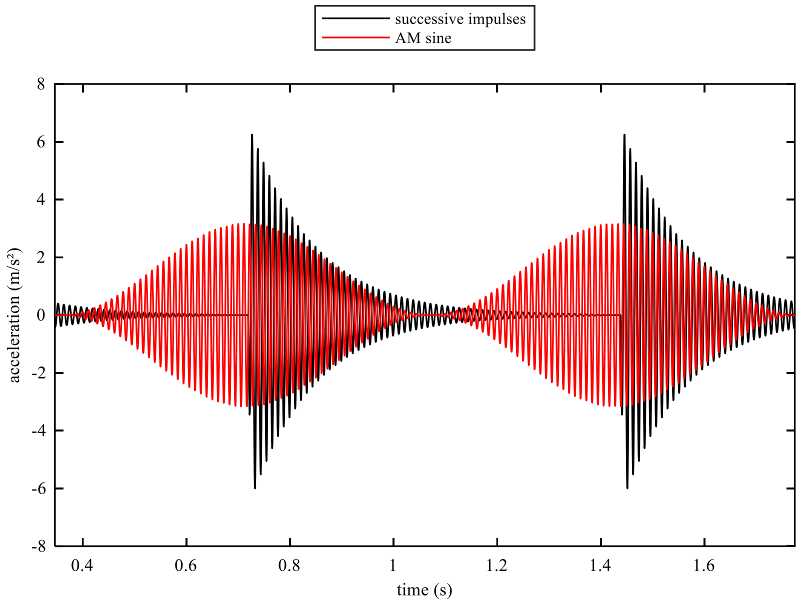

Figure 7.4.
AM sinusoidal signal and multiple successive impulses signal.

cover the range of attribute rating maxima of the frequency-dependent attributes (see section 5.4). A higher impulse acceleration level also produces larger differences in vibration signal envelope shape than a low level. Therefore, 42 dB (SL) level was selected for the multiple successive impulse-like vibration. Since most attributes are positively correlated to the acceleration level, the 42 dB (SL) level was selected for the multiple successive impulse-like vibration to produce high absolute ratings and thus potentially larger difference between the two excitation patterns. Since the SL for impulse-like vibration was defined as if the decay constant was zero, the actual RMS SL of the three multiple successive impulse-like vibration signals was calculated. The RMS SL of three AM vibration was scaled to these levels: 33 dB (SL) for 9 Hz, 31 dB (SL) for 15 Hz, and 30 dB (SL) at 90 Hz. This resulted in a total of three multiple impulse-like vibrations and three AM vibrations. One example of the time signals of two paired stimuli is shown in Figure 7.4.

7.3.2 Experimental Design

The experimental design was identical to section 5.4, except for the stimuli. Participants were allowed to repeat the stimulus, but rarely did so. The six stimuli described in the previous section were presented with the tactile reproduction system described in chapter 4. Five out of six German attributes of the sensory tactile design language “up and down” (“auf und ab”), “weak” (“schwach”), “repetitive” (“wiederholend”), “uniform” (“gleichmäßig”), and “fading” (“nachschwingend”) were rated for each stimulus. Since this experiment was conducted before the selection of the final attribute set, the attribute “buzzing” (“summend”) highly correlating (Pearson’s r > 0.9) with “tingling” (“kribbelnd”) was investigated instead. The attributes were grouped into two attribute triples. Attribute ratings were conducted on the quasi-continuous Rohrmann rating scale with verbal anchors [92] (see section 4.2.2). The rating time of each of attribute triple per stimulus was not limited, but typically took approximately 10 seconds resulting in a total experiment duration of 15 minutes.

7.3.3 Participants

A total of 14 German native speakers (10 male, 2 female) with an average age of 34 years (21 to 61 years) took part in the experiment. The study was conducted with the understanding and written consent of each participant.

7.3.4 Results

The mean attributes ratings and 95 % confidence intervals of these six sensory tactile perceptual attributes are shown in Figure 7.5. The ratings of AM-sinusoidal vibrations are clearly very similar to the ratings of multiple successive impulse-like vibrations. The mean difference is 2 scale points for “up and down”, 4 scale points for “weak”, 1 scale point for “repetitive”, 5 scale points for “uniform”, 2 scale points for “fading” and 9 points for “buzzing”. Only for buzzing the maximum difference exceeds half a tick of the 25-point scale tick interval.

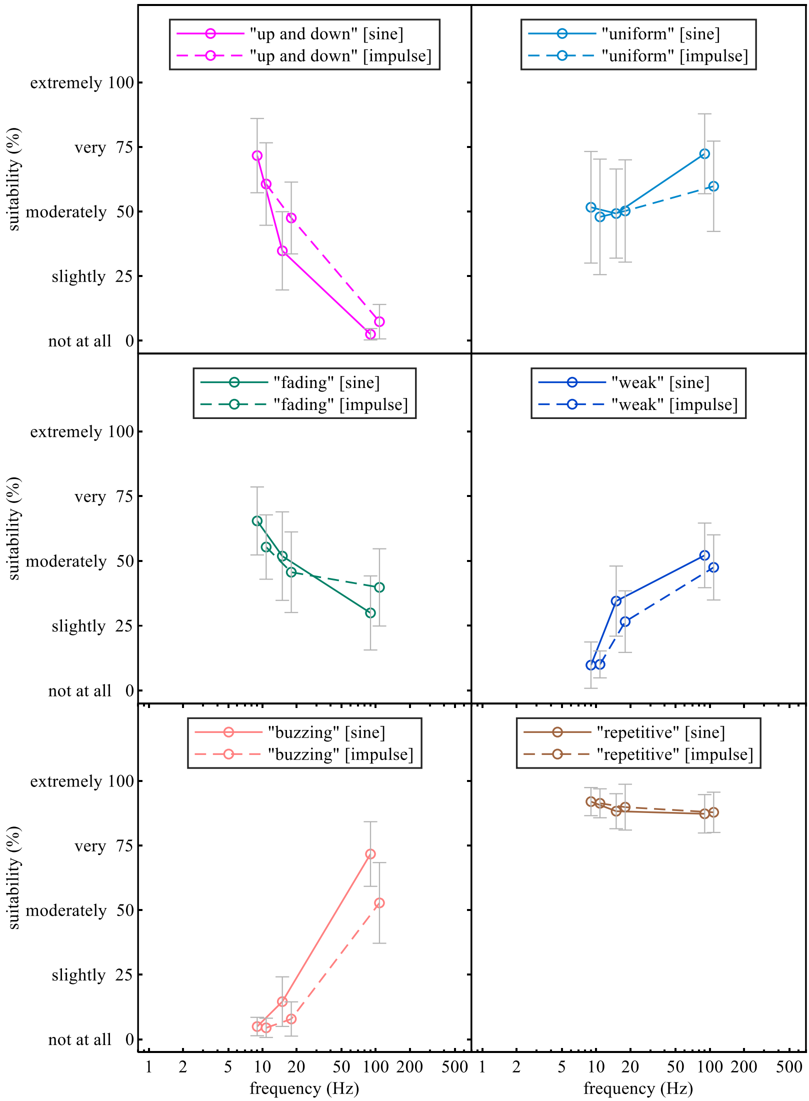

Figure 7.5.
Mean attribute ratings and 95 % confidence intervals vs. frequency for five sensory tactile perceptual attributes of the tactile design language and "buzzing" ("summend") instead of "tingling" ("kribbelnd") for AM-sinusoidal vibration in comparison to multiple successive impulse-like vibration.

The influence of the signal pattern (AM sinusoidal vs. multiple successive impulses) was analyzed from a statistical standpoint with a three-way within-subjects repeated-measures ANOVA with factors stimulus type, carrier frequency, or resonance frequency and attribute. The factor excitation pattern shows a significant effect ($F(1,13) = 9.466$, $p < 0.05$). However, a pairwise contrast between multiple successive impulse-like vibration and AM sinusoidal vibration shows that the mean effect size of the factor signal pattern is only approximately 3 scale points. The 95 % confidence interval of the effect size ranges from approximately 1 to 6 points. This measure of the overall similarity between the two signal patterns, suggests that the mean difference is unlikely to exceed a fourth of the 25-point scale tick interval and is thus likely irrelevant for practical purposes.

7.3.5 Summary

The comparison between multiple successive impulse-like vibration and AM sinusoidal vibration shows that these signal patterns produce very similar attribute ratings if their parameter values are identical. There is a small deviation for the attribute “buzzing” only for high frequencies, which is likely to be observable also for “tingling” due to their high correlation. Overall these findings suggest the multiple successive impulse-like excitation pattern can be approximately substituted by the AM sinusoidal excitation pattern without introducing major gaps in the sensory tactile perceptual profiles elicitable by the potential synthesis model. Therefore, the 8 multiple successive impulse-like vibration were eliminated from 99 vibration-rating profile pairs database of chapter 5, thus reducing it to 91 elements.

7.4 Synthesis Based on the Discrete Estimates of a k-Nearest-Neighbor Classifier

In chapter 5 a sensory tactile design language was assessed, which enables the efficient profiling of sensory tactile perceptual properties of WBV in the

form of rating profiles. During this assessment, a database of vibration parameter tuples mapping onto rating profiles **r** of the six sensory tactile perceptual attributes was obtained for 99 generalized WBV stimuli. Since the 8 multiple successive impulse-like vibrations were eliminated from the database in section 7.3, there are 91 vibration attribute rating profile pairs remaining. Thus, there is a discrete mapping of these vibrations onto sensory-perceptual attributes. Subsequently, sensory tactile perceptual property profiles were obtained to quantify expectations on a representative set of multimodal everyday life vehicle scenes with WBV in chapter 6. It was demonstrated that not only the properties elicited by vibration of a situational context can be profiled but also the expectations on that situational context.
The plausibility judgment was formalized as a similarity judgment between the expected perceptual object and the perceptual object elicited for the situational context in the sensory tactile perceptual feature space in section 7.1. Thus, the degree of similarity can be assessed by the Euclidean distance between the expected perceptual object and the elicited perceptual object in the six-dimensional sensory tactile perceptual space. The sensory properties of the expected perceptual object can be quantified in the rating profile **e** and of the perceptual objected elicited by a vibration stimulus in the rating profile **s**. To translate **e** into plausible vibration, it is necessary to find a vibration, whose **s** has a minimal Euclidean distance in the six-dimensional sensory tactile perceptual space.
Such a formalization of the plausibility illusion bears close resemblance to a k-nearest-neighbor (k-NN) classifier. Such a classifier attempts to classify an object into a set of classes according to a distance measure in the n-dimensional feature space. Therefore, the goal in this section is the implementation of a k-NN classifier based on the discrete mapping of WBV to the sensory tactile perceptual space, to translate a sensory tactile rating profile into a vibration parameter profile.

7.4.1 Definition of the K-Nearest-Neighbor Classifier

The k-NN is described e.g. by [112]. A k-NN classifier can be utilized if there is paired training data, consisting of m pairs of objects $\boldsymbol{x}$ and classes y:

$$(\boldsymbol{x}_1, \boldsymbol{y}_1), (\boldsymbol{x}_2, \boldsymbol{y}_2), \dots, (\boldsymbol{x}_m, \boldsymbol{y}_m) \quad \textbf{7.8.}$$

Each object $\boldsymbol{x}$ is defined by its n-dimensional feature vector, so that $\boldsymbol{x} \in \mathbb{R}^n$. For a norm $\| \cdot \|$ that assigns a distance measure number in $\mathbb{R}^n$, the objects of the training data $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ can be sorted by their distance to an object $\boldsymbol{x}$:

$$\|\boldsymbol{x}_{(1)} - \boldsymbol{x}\| \leq \cdots \leq \|\boldsymbol{x}_{(m)} - \boldsymbol{x}\| \quad \textbf{7.9.}$$

with $\boldsymbol{x}_{(1)}$ having the closest distance to $\boldsymbol{x}$. This produces the reordered training data:

$$(\boldsymbol{x}_{(1)}, \boldsymbol{y}_{(1)}), \dots, (\boldsymbol{x}_{(m)}, \boldsymbol{y}_{(m)}) \quad \textbf{7.10.}$$

The object $\boldsymbol{x}$ is classified by determining the most frequent class in amongst the classes $\boldsymbol{y}_{(1)}, \dots, \boldsymbol{y}_{(m)}$ of k-nearest neighbors $\boldsymbol{x}_{(1)}, \dots, \boldsymbol{x}_{(m)}$ of object $\boldsymbol{x}$.

7.4.2 Analysis Model

This definition is applied to the dataset of 99 vibration attribute rating profile pairs obtained in chapter 5 and reduced in section 7.3 to 91 pairs. The database consists of m generalized WBV stimuli **v** mapping onto rating profiles of the elicited perceptual objects **s** of the six sensory tactile perceptual attributes. Applying the k-NN approach, an analytic classifier can be easily created for this dataset.

The input objects **x** are vibrations **v**. Each of the 91 generalized vibration **v** of the database can be described by its features. These features are the vibration parameters level L, (carrier-, center- or resonance-) frequency f, bandwidth f_b, modulation frequency f_m, and decay constant α. However, it should be noted that not all vibration parameters are defined for each of the four excitation patterns (sinusoidal vibration, AM sinusoidal vibration, bandlimited WGN vibration, and impulse-like vibration).

The classes **y** are rating profiles elicited by each vibration stimulus **s**. Each elicited perceptual object **s** is defined by its 6-dimensional feature vector, so that $\boldsymbol{s} \in \mathbb{R}^6$. These **s** are characterized by their 91 rating profiles consisting of the attribute rating of “weak” r_w, the attribute rating of “up and down” r_d, the attribute rating of “tingling” r_t, the attribute rating of “repetitive” r_r, the attribute rating of “uniform” r_u, and the attribute rating of “fading” r_f.

If two vibrations are perceptually equivalent, i.e. eliciting the same rating profile, they would map onto the same class. Since vibration eliciting quasi identical attribute ratings were explicitly excluded in section 7.3, no two vibrations are eliciting identical rating profiles. Therefore, there would be 91 rating profile classes. Such a classifier would not provide a model of the continuous relationship between vibration parameters and the sensory tactile perceptual attribute ratings, but only a mapping of discrete vibration parameters onto discrete ratings. The prediction performance would thus be dependent on the intervals between vibration parameter steps.

7.4.3 Synthesis Model

However, the focus of this work is set on the creation of a vibration synthesis model. Therefore, object and class need to be swapped to create a synthesis classifier. Thus, the input objects **x** are perceptual objects elicited by the 91 vibration stimuli **s**. These perceptual objects are characterized by their features. Based on the PCA of chapter 5 six perceptual attributes were suggested as a compact set of features to represent the sensory tactile perceptual space. Thus, considering a further feature reduction is not necessary.

Each rating profile **s** is paired to an output class, which is a vibration **v**. Since vibration eliciting quasi identical attribute ratings were explicitly excluded in section 7.3, there are exactly 91 vibration classes, i.e. one for each of the 91 rating profiles. If there were multiple perceptually equivalent vibrations, i.e. multiple vibrations with the same rating profiles, it would be necessary to introduce additional criteria for vibration selection.

Rahloff suggests that the Euclidean distance can be a suitable norm for the perceptual space of auditory perceived room size [77]. Thus, a Euclidean distance is also assumed to be a suitable norm for the six-dimensional sensory tactile perceptual space. For a given rating profile of the expected perceptual object **e,** the Euclidean distance to the rating profile elicited by a vibration stimulus **s** of the database can be determined according to:

$$d(\boldsymbol{s}, \boldsymbol{e}) = \sqrt{\sum_{i=0}^{n}(e_i - s_i)^2} \quad \textbf{7.11.}$$

Based on the Euclidean distance to **e**, the rating profiles **s** can be sorted:

$$|\boldsymbol{s}_{(1)} - \boldsymbol{e}| \leq \cdots \leq |\boldsymbol{s}_{(m)} - \boldsymbol{e}| \quad \textbf{7.12.}$$

This produces the reordered paired training data:

$$(\boldsymbol{s}_{(1)}, \boldsymbol{v}_{(1)}), \ldots, (\boldsymbol{s}_{(n)}, \boldsymbol{v}_{(n)}) \quad \textbf{7.13.}$$

As suggested by [77], plausibility p is inversely proportional to Euclidean distance between in the perceptual space the expected perceptual object and the elicited perceptual object:

$$p \sim \frac{1}{|\boldsymbol{s} - \boldsymbol{e}|} \quad \textbf{7.14.}$$

Thus, it can be concluded that:

$$p(\boldsymbol{s}_{(1)}) \leq \cdots \leq p(\boldsymbol{s}_{(m)}) \quad \textbf{7.15.}$$

and that vibration $\mathbf{v}_{(1)}$ paired to $\mathbf{s}_{(1)}$ would be most plausible.
Finally, the parameter k needs to be chosen to determine the k-nearest neighbors, on which the classifier would base the selection of the vibration class **v**. Since there are as many classes of vibration **v** as there are rating profiles **s**, choosing a k > 1 is not meaningful. Furthermore, in the case of two rating profiles $\mathbf{s}_{(1)}$ and $\mathbf{s}_{(2)}$ having the same Euclidean distance to **e**, they would likely be equally plausible and thus it would be sufficient to base the classification exclusively on $\mathbf{s}_{(1)}$. Similarly, if the rating profile $\mathbf{s}_{(1)}$ has a smaller Euclidean distance to **e** then $\mathbf{s}_{(2)}$, the perceptual object described by $\mathbf{s}_{(1)}$ would be more plausible and it would thus be sufficient to base the classification exclusively on $\mathbf{s}_{(1)}$ also in that case.
Since the general synthesis approach will be validated in the form of a perceptual study in section 7.6, no separate cross-validation only for the model

stage is conducted. To enable a meaningful cross-validation of the classifier itself, a dataset with fewer classes than objects would be required. Since the investigated problem is structured in such a way that each rating profile **s** is associated with at least one class of vibration **v** a separate cross-validation step is not possible for a synthesis classifier. It would be possible for the previously described analysis classifier, since there are potentially multiple perceptually equivalent perceptual objects that have the same rating profile as suggested by section 7.3.

7.4.4 Interpolation of acceleration level for the vibration attribute profile pairs

A dataset of 99 vibration attribute rating profile pairs was obtained in chapter 5 and reduced in section 7.3 to 91 pairs. However, for the stimuli selected in section 5.2 only two SLs of 10 dB (SL) and 36 dB (SL) for non-transient und 30 dB (SL) and 42 dB (SL) for transient stimuli were utilized, representing the extremes of the range of everyday life WBV. The JNDL for WBV is approximately 1 dB [34]. This suggests that a much finer discrimination of vibration than 26 dB or 12 dB is possible and that instead level steps of a small multiple of the JNDL of about 3 dB are very likely perceivable. It has been shown by [80] that increasing the acceleration from the original acceleration level by 3 dB significantly decreased the perceived plausibility of the virtual environment of passing train. This suggests that the acceleration resolution of the database, on which the k-NN classifier would be built, is likely insufficient for eliciting a plausibility illusion that is comparable to the plausibility illusion elicited by the original vibration of the situational context.
To obtain a database with a finer acceleration level resolution of about 3 dB it would theoretically be necessary to repeat the attribute rating experiment of section 5.4 with an about four times larger stimulus set. However, it was demonstrated in section 7.2 that the acceleration level has an approximately linear influence on the ratings of the six attributes of the sensory tactile design language. Therefore, new vibration rating profile pairs can be obtained by linear interpolation.
Since reproduction in the limits defined in section 4.1.3 could not be achieved for the combinations 7 Hz and 9 Hz with 10 dB (SL) and 500 Hz at 36 dB

(SL) for the sinusoidal excitation pattern, the six attribute ratings needed to be interpolated between the ratings of 5 Hz at 10 dB (SL) and 15 Hz at 10 dB (SL) before the level interpolation. Thus, the ratings r(f) of an attribute for a sinusoidal vibration with the frequency f and 10 dB (SL) were linearly interpolated according to the formula:

$$r(f) = r_1 + \frac{r_2 - r_1}{f_2 - f_1}(f - f_1) \quad \textbf{7.16.}$$

Due to 500 Hz being the highest vibration frequency, a rating for 500 Hz at 36 dB could not be obtained. This interpolation resulted in a database with 92 vibration of pairings with a high acceleration level and a low acceleration level for each combination of remaining parameters. Similarly, new vibration rating profile pairs were obtained at the SL by linearly interpolating between the attribute rating r(SL) of the low SL's profile and the high SL's profile:

$$r(SL) = r_1 + \frac{r_2 - r_1}{SL_2 - SL_1}(SL - SL_1) \quad \textbf{7.17.}$$

For the non-transient vibration pairs, an SL increment if 3.25 dB was chosen for the interpolation resulting in 9 evenly spaced SLs between 10 dB (SL) and 36 dB (SL). For the transient vibration pairs, an SL increment of 3 dB was chosen for the interpolation resulting in 5 evenly spaced SLs between 30 dB (SL) and 42 dB (SL). These interpolation steps produced a database of 387 vibration rating profile pairs.

7.4.5 Implementation of the Synthesis

The algorithm for the 1-nearest neighbor classifier outlined in the previous section was implemented in MATLAB. The dataset on which it operates is the interpolated database of 387 vibration rating profile pairs. The MATLAB implementation is structured according to the model boundaries shown in Figure 7.2 of section 7.1. It consists of three functions for vibration parameter estimation, vibration signal generation, and multimodal scene construction. Besides these functions, a GUI was created to facilitate the usage of these functions.

7.4.5.1 Input

The input for the synthesis consists of a scene description. Such a description contains a rating profile of the expected sensory tactile properties as quantified in section 6.3 for multimodal scenes consisting of the attribute rating of “weak”, “up and down”, “tingling”, “repetitive”, “uniform”, and “fading”. Furthermore, it contains the multimodal scene's duration and the optical and acoustic signals associated with it. For the transient scenes, the beginning and end of the impulse event is an additional input. Another implicit input for the vibration parameter estimation is the interpolated database of 387 vibration rating profile pairs.

7.4.5.2 Translation

For the expected rating profile, the previously described 1-nearest neighbor classifier calculates the Euclidean distance to the 387 rating profiles of the vibration-rating profile pairs of the database. The vibration parameters (sensation level SL, vibration (carrier- or resonance-) frequency f, bandwidth f_b, modulation frequency f_m and decay constant α) associated to the rating profile with the smallest distance in the six-dimensional sensory tactile perceptual space are output by the classifier thus the vibration parameters are estimated. The vibration associated with this rating profile is perceptually most similar to the expected vibration and is thus likely the most plausible vibration of the database. An overview of the vibration parameter estimates output by the 1-nearest-neighbor (1-NN) synthesis model for the 19 vehicle scenes are shown in Table 7.1.

Based on the vibration parameter estimation the acceleration signal can be generated according to the formulas of the four excitation patterns in section 5.2. The information defining the specific excitation pattern is implicitly contained in the estimated parameter values. Impulse-like vibrations have a non-zero decay constant, bandlimited WGN vibrations have bandwidth greater or equal to 25 Hz, AM sinusoidal vibrations have a non-zero modula tion frequency and sinusoidal vibrations have none of the former. The expected rating profiles of two types of vehicle scenes were obtained in section 6.2: non-transient scenes and transient scenes.

Table 7.1.
Vibration parameter estimates of the 1-nearest neighbor classifier for the 19 vehicle scenes.

No.	Category	Speed	Surface	L (dB)	f (Hz)	f_b (Hz)	f_m (Hz)	α
1	non-transient	5 to 50	small cobblestone	112	9			
2	non-transient	30	small cobblestone	112	9			
3	non-transient	50	small cobblestone	112	9			
4	non-transient	5 to 50	cobblestone	124	26	50		
5	non-transient	30	cobblestone	121	26	50		
6	non-transient	50	cobblestone	124	26	50		
7	non-transient	50	tarmac (A-Road)	95	7		2	
8	non-transient	70	tarmac (A-Road)	109	155		15	
9	non-transient	5 to 50	Tarmac	103	9			
10	non-transient	30	Tarmac	97	15		5	
11	non-transient	50	Tarmac	105	155		15	
12	non-transient	100	concrete motorway	110	15			
13	transient	100	surface change	104	90			
14	transient	40	tram tracks	118	26	50		
15	transient	50	expansion joint	114	26	50		
16	transient	100	expansion joint	108	26			
17	transient	30	manhole cover cobblestone	122	9		2	
18	transient	30	manhole cover tarmac	109	201	400		
19	transient	50	manhole cover tarmac	112	201	400		

For the non-transient vehicle scenes, the rating profile relates to the vibration spanning the total scene duration. Therefore, vibration is generated for the total scene duration. A linear fade in and fade out of 300 ms was applied to the signal as in section 5.2.3.

For the non-transient vehicle scenes, the rating profile relates to the vibration occurring from the beginning to the end of the impulse event. Therefore, vibration is generated for the duration of the impulse event. If an impulse-like excitation pattern was estimated, a short linear fade in for half period of the resonance frequency was applied and no fade out due to the decaying characteristics of the excitation pattern as described in section 5.2.3. If another excitation pattern was estimated, it was necessary to apply a short fade-in of 100 ms at the beginning of the impulse event and a short fade out to the end of the impulse event to the signal to enable a correct playback on the reproduction system. For the remaining scene duration, the acceleration signal was padded with zeros. Based on the synthesized vibration a multimodal scene is constructed in the format specified in section 4.2.1 by replacing the recorded vibration signal with the synthesized vibration signal.

7.4.5.3 Output

For the subsequent model validation, 19 audio-visual scenes with the synthesized WBVs were generated based on these estimates. The spectra of these vibrations in comparison to the the recorded vibrations are shown in Figure 7.6.

7.4.5.4 Examples

The details of the synthesis algorithm are discussed for two examples. The sensory tactile perceptual attribute ratings of the scene with cobblestone at 30 kph are 20 for “weak”, 70 for “up and down”, 48 for “tingling”, 64 for “repetitive”, 29 for “uniform”, and 69 for “fading”. The algorithm determines the Euclidean distance to all the 387 rating profiles of the vibration-rating profile pairs of the database. Subsequently, the vibration-rating profile pairs of the database are sorted by their Euclidean distance to the rating profile of the cobblestone at 30 kph scene. The vibrations associated with the closest rating profiles are most similar to the expected vibration of the cobblestone at 30 kph scene from a sensory-perceptual standpoint and are thus anticipated to be most plausible in the context of this scene. The algorithm suggests that a bandlimited WGN vibration with an SL of 36 dB (SL), a center frequency of 26 Hz and a bandwidth of 50 Hz would elicit sensory-perceptual attribute ratings most similar to the expected ratings for this scene.
In Figure 7.6 (5) the spectrum of the recorded vibration is shown in comparison to this synthesized vibration. Both spectra are quite similar up to 50 Hz, while the recorded vibration spans a broader frequency range containing also components with relevant acceleration level above this frequency. The 9 profiles with the smallest distance are bandlimited noise WGN vibration. Their level is in the range from 29.5 dB (SL) to 36 dB (SL). Their center frequencies and bandwidths are combined in such a way that their spectra always cover the 1 Hz to 26 Hz frequency range at least partly. Thus, a bandlimited noise

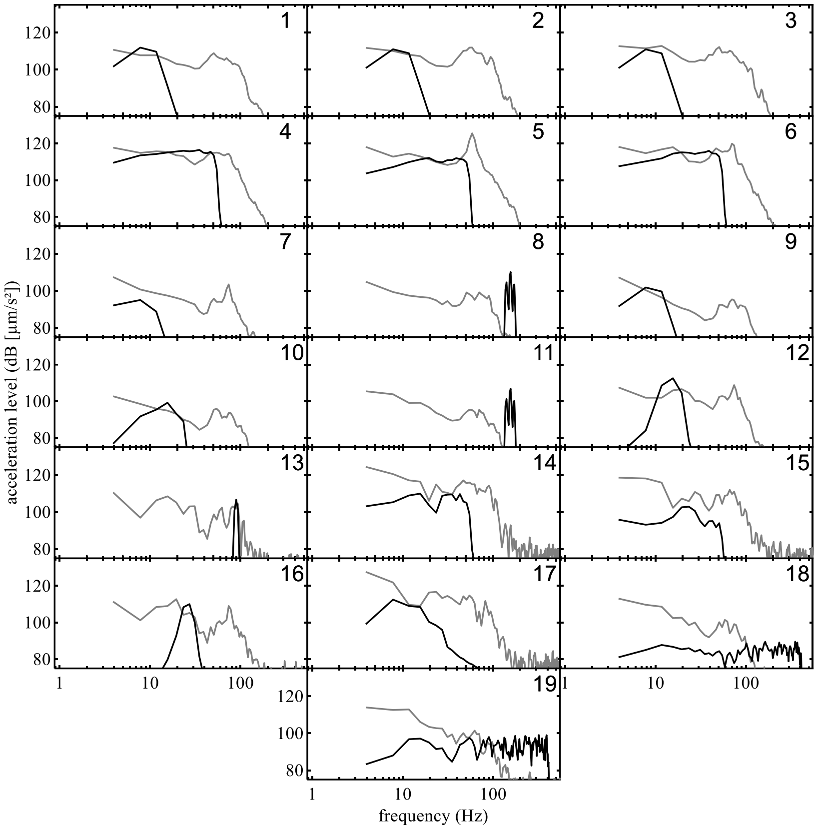

Figure 7.6.
Spectra (FFT, 4096 samples, 50 % overlapping Hann windows) of the recorded vibration (grey) and the vibration generated from the parameter estimates of the 1-nearest-neighbor model (black) for each of the 19 vehicle scenes.

vibration with a bandwidth of 400 Hz at a center frequency of 201 Hz is perceptually similar to a bandlimited noise vibration with a bandwidth of 25 Hz and a center frequency of 13.5 Hz. If perceptually similar vibration indeed elicited similar plausibility in the context of the scene as suggested by the proposed relationship between plausibility the distance in the sensory-perceptual space this would enable a substitution of broadband vibration by narrowband vibration.

The sensory tactile perceptual attribute ratings of the scene with tarmac at 30 kph are 78 for “weak”, 14 for “up and down”, 20 for “tingling”, 47 for “repetitive”, 74 for “uniform”, and 23 for “fading”. Again, the algorithm determines the Euclidean distance to all the 387 rating profiles of the vibration-rating profile pairs of the database. Subsequently, the vibration-rating profile pairs of the database are sorted by their Euclidean distance to the rating profile of the tarmac at 50 kph scene, too. Again, the vibrations associated with the closest rating profiles are most similar to the expected vibration of the tarmac at 30 kph scene from a sensory-perceptual standpoint and are thus anticipated to be most plausible in the context of this scene. The algorithm suggests that an AM-sinusoidal vibration with an SL of 10 dB (SL), a frequency of 15 Hz, and a modulation frequency of 5 Hz would elicit sensory-perceptual attribute ratings most similar to the expected ratings for this scene. In Figure 7.6 (10) the spectrum of the recorded vibration is shown in comparison to this synthesized vibration. In contrast to the cobblestone scene, the spectra for the tarmac scene are not very similar. While the recorded vibration has relevant acceleration levels in a broader frequency range from 1 Hz up to about 100 Hz, the synthetic vibration consists only narrowband AM-sinusoidal excitation at 15 Hz. Thus, similarly to the cobblestone scene, it might be possible to substitute a broadband signal for a narrowband signal also for the tarmac scene, if perceptually similar vibration indeed elicited similar plausibility in the context of this scene.

7.4.6 Advantages and Disadvantages

This simple approach has advantages as well as disadvantages that need to be discussed in detail. The main advantage of this approach is that it is the direct extension of plausibility as similarity judgment suggested by [77]. Since plausibility has been suggested as to be a similarity judgment in the sensory tactile perceptual space the 1-nearest neighbor classifier follows exactly this hypothesized relationship by determining the vibration which elicits sensory tactile properties which are similar to the expected sensory tactile properties. If the database of the classifier is populated by vibration-rating profile pairs that cover the range of vibration encountered in everyday life with sufficient granularity, no explicit modeling is required.

Thus, no generalization of the relationship between vibration parameters and sensory tactile attributes is conducted by any modeling procedure, potentially including deviations from the true underlying relationship and thus leading to less plausible vibration. Furthermore, a continuous mapping between vibration parameters and sensory tactile attributes is not required for this approach because the output is discrete, i.e. the vibration eliciting the one rating profile of the database with the smallest Euclidean distance to the input rating profile. The consequence of this approach is that the output will always be a valid set of vibration parameters, from which vibration can be generated.
Another advantage is that the classifier provides a ranking for each vibration in the database in the form of the Euclidean distance between the input rating profile and the rating profiles associated with each vibration. This could enable informed choices on the expected plausibility reduction depending on the vibration selection. As suggested by the cobblestone and tarmac example, narrowband vibration was perceptually similar to broadband vibration, i.e. their rating profile had a similar distance to the expected rating profile. This might enable the designer to consider trade-offs between reproduction system capabilities (e.g. cheap narrowband actuator vs. expensive broadband actuator) and plausibility of the vibration in the context of the scenes intended to be reproduced.
The main disadvantage lies in the discrete, i.e. non-continuous prediction of the vibration parameters by the 1-nearest neighbor classifier. The vibration output from the synthesis model depends on the vibrations and their manually assessed perceptual profiles contained in the database (see disadvantages of the database approach in 2.3). That implies that if there is insufficient granularity in the database of vibration-rating profile pairs then no close neighbor might be available impairing prediction performance and ultimately leading to less plausible vibration. However, this would depend on the actual distance between the expected rating profile and the elicited profile. [77] suggests that there is a correlation between plausibility and the inverse Euclidean distance. The exact relationship is still unknown. Thus, in the close range, in which this distance in the sensory-perceptual is low. Therefore, there might be multiple perceptually similar vibrations eliciting rating profiles with increasing distance to the expected rating profile only showing an influence above a certain threshold.

Another disadvantage inherent to the nearest neighbor approach is that it does not produce an explicitly interpretable model because any relationship is implicitly contained in the database with vibration-rating profile pairs. Such an explicit relationship might facilitate design choices for the tactile designer by providing e.g. an analytic Formula that can predict vibration parameter changes from sensory tactile perceptual attribute rating changes. However, such a relationship is likely more beneficial for an analysis model predicting the attribute ratings from vibration properties. Since the nearest neighbor classifier only does lazy learning, i.e. it stores the training data and does not create a classifier function [112] it needs to compare the input rating profile to all the rating profiles in the database. Thus, simply interpolating an increasing number of vibration-rating profile pairs would thus decrease prediction performance from a computational load standpoint. However, this is unlikely to impact prediction runtime performance problems because even the utilized interpolated database only contains 387 items.
For a continuous prediction, it seems a good idea to extend the nearest neighbor classifier into a nearest neighbor regression model. However, in section 5.2.1 four discrete excitation patterns (sinusoidal vibration, AM sinusoidal vibration, bandlimited WGN vibration, and impulse-like vibration) were defined to represent everyday life WBV which cannot be reflected by a single generation formula, since some parameters are only valid for some excitation patterns. Finally, since a validation of the nearest neighbor classifier performance separately from the plausibility of the vibrations produced by it were not possible, no cues about the model error can be obtained.

7.5 Synthesis Based on the Quasi-Continuous Estimates of Regression Models

In chapter 5 a stimulus set was created that contains 99 generalized vibrations. Since the 8 multiple successive impulse-like vibrations were eliminated from the database in section 7.3, there are 91 vibration attribute rating profile pairs remaining. Each vibration was profiled with the sensory tactile design language to obtain rating profiles. This resulted in a database of discrete mappings of vibration onto sensory-perceptual attributes. The plausibility judgment was formalized as the Euclidean distance between the expected perceptual object and the elicited perceptual object in the six-dimensional sensory tactile perceptual space in section 7.1. Based on this relationship a 1-NN classifier was constructed, which can determine the vibration in the database whose rating profile has a minimal distance to the expected rating profiles of vehicle scenes obtained in chapter 6. The vibration associated with the rating profile pair with the minimum distance is output by the classifier.

The main disadvantage of such an approach is the discrete prediction limited to the vibration items and their manually profiled sensory tactile perceptual properties. The performance of this synthesis model is thus limited by the range and resolution of rating profiles and their associated vibration. Since the vibration stimuli of chapter 5 were systematically constructed by varying the parameters of the four excitation patterns, not only discrete mappings between vibration and rating profiles could be observed but continuous mappings, i.e. systematic relationships in the form of physical parameter vs. rating curves. E.g. for the vibration level a linear influence could be observed on attribute ratings in section 7.2. The observed relationships enable inferences not possible from discrete mappings. To partially rectify the shortcoming of the 1-NN model, new vibration rating profile pairs were obtained at the intermediate SLs by linearly interpolating between the attribute rating of the low SL's profile and the high SL's profile. This potentially decreased the distance to the most similar attribute ratings profile and thus should enable a more plausible vibration synthesis.

However, instead of utilizing the observed relationships to interpolate more vibration-rating profile pairs, it would be more efficient to create models from these relationships to enable a continuous vibration prediction from attribute rating profiles. The modeling conducted in this section will incorporate aspects of explanatory modeling as well as predictive modeling (see [113] for comparison of these two modeling goals). For the aim of explanatory modeling, it is attempted to keep the bias of the model with respect to the true underlying relationship between physical parameters and attribute ratings minimal. This will provide a statistical assessment of the observed relationships (see section 5.4). For the aim of predictive modeling, it is attempted to minimize the prediction error of the models. This will provide a second synthesis

model, based on the vibration parameter vs. attribute ratings curves observed in section 5.4. The modeling procedure will consist of three steps.
In section 5.2.1 everyday life vibration was generalized into four discrete excitation patterns (sinusoidal vibration, AM sinusoidal vibration, bandlimited WGN vibration, and impulse-like vibration). However, not all characteristic vibration parameters of each excitation pattern are defined for the other excitation patterns. Therefore, the synthesis model needs to be split into four submodels. For each excitation pattern regression models will be built predicting ratings of the relevant attributes of tactile design language from vibration parameters. In this step, the relationships between vibration parameters and attribute ratings observed in section 5.4 will also be analyzed from an inferential statistical point of view. However, since there is no established theoretical framework for sensory tactile perceptual attributes yet, the modeling is not solely driven by explanatory aspects but also needs to include exploratory aspects. Subsequently, for each excitation pattern, an equation system formed by the regression model equations will be solved. This will produce a set of synthesis equations predicting the vibration parameters from attribute ratings.

7.5.1 Overall Model Structure

The model boundaries shown in Figure 7.2 of section 7.1 are identical to the 1-NN model to facilitate a switch between the two synthesis models. As for the 1-NN model, the model input is an expected perceptual object in the sensory tactile perceptual space characterized by its rating profile **e**. The rating profiles consist of the attribute rating of "weak" r_w, the attribute rating of "up and down" r_d, the attribute rating of "tingling" r_t, the attribute rating of "repetitive" r_r, the attribute rating of "uniform" r_u, the attribute rating of "fading" r_f. Again, the model output is a vibration **v** characterized by its parameters level L, (carrier- or resonance-) frequency f, bandwidth f_b, modulation frequency f_m, and decay constant α.
In contrast to the previous model, the translation from an attribute rating profile to vibration parameters is not a discrete prediction but will be a quasi-continuous prediction. The dataset on which the modeling will be conducted contains vibrations of four discrete generalized excitation patterns (see sec-

tion 5.2) that are each characterized by a different set of parameters. The sinusoidal excitation pattern is characterized by RMS sensation level SL and frequency f. The AM-sinusoidal excitation pattern is characterized by RMS sensation level SL, carrier frequency f, modulation frequency f_m, and modulation index m. The bandlimited WGN excitation pattern is characterized by the RMS sensation level SL, center frequency f_c, and bandwidth f_b. The impulse-like excitation pattern is characterized by RMS sensation level SL (defined as if the impulse was a non-decaying sinusoidal), resonance frequency f, and decay rate α.

It is obvious that all four excitation patterns only share SL as a common parameter. Since the bandlimited WGN excitation pattern is generated from a stochastic process and not from a periodic process as the other excitation patterns, it is not possible to define a single formula that can generate vibration of all four excitation patterns. Theoretically, it would be possible to create a single formula for three of the four excitation patterns (sinusoidal, AM-sinusoidal, impulse like) since they also share a periodic component characterized by (carrier- or resonance-) frequency. However, combining the models of these excitation patterns would likely require additional investigations of the interaction of the remaining parameters modulation frequency, modulation index, and decay rate in comparisons to the parameter variation investigated in section 5.2. Due to the enormous resources required to obtain such a large amount of ratings, it is indicated to investigate a simpler model first. Such a simple model could be created by dividing the overall model into four discrete submodels for each excitation pattern.

If the sensory tactile perceptual attribute profiles elicited by everyday life vibration cluster around the range of rating profiles elicited by the vibration of the four excitation patterns but not of the superimposition of the excitation patterns, such an abstraction would likely be meaningful. Furthermore, if the rating profiles elicited by vibrations of two excitation patterns overlap, a vibration could be generated according to either of the two excitation patterns. If the rating profile of two vibrations are identical they could be considered to be perceptually equivalent and would likely elicit identical plausibility ratings. Therefore, this synthesis model will be split into four submodels for the four excitation patterns. Whether the vibration synthesized with such an aggregate model can achieve satisfying performance regarding the perceived plausibility will be investigated in the subsequent model validation section.

Since the synthesis model will be split into four submodels for the four excitation patterns, it is necessary to determine the submodel for the synthesis based on the expected rating profile input into the model. Before the submodels can be utilized for a synthesis, the most suitable submodel needs to be selected and thus the general excitation pattern generating the vibration. The four excitation patterns can be considered as classes that need to be selected according to the input expected perceptual object characterized by its rating profile.

After building a classifier that determines the most suitable excitation pattern, the submodels for each excitation pattern need to be built. The dataset obtained in section 5.4 contains pairs of vibration parameters and attribute rating profiles. The physical parameters vs. rating curves suggest obvious systematic relationships to sensory tactile perceptual attributes to which regression models can be fit. The naive solution would be to model the physical vibration parameters depending on the observed sample mean perceptual attribute ratings of the stimuli. However, the physical vibration parameters were set in a controlled way and are thus practically without error. In contrast, the population mean perceptual attribute ratings could only be observed as the sample mean perceptual attribute ratings superimposed with measurement error caused by the variance in ratings of participants. Linear regression models estimated by ordinary least squares assume weak endogeneity of regressors, i.e. that errors can be present in the dependent variable but the independent variable should be uncorrelated to the error term [114].

There are two solutions to solve this problem: reversing dependent and independent variables or modeling the inverse relationship and subsequent inversion of the resulting equation [115]. The reverse approach would violate the assumption of weak endogenity of regressors would thus introduce attenuation bias or regression dilution into the estimated coefficients of the regression model [116]. Unsurprisingly, the inverse approach produces a greater bias than the reverse modeling approach [115]. Further complications can arise with this method as data transformations and the introduction of additional predictors magnify the bias [117]. While correction of such bias is sometimes possible, for the case of multivariate prediction, regression coefficients can be overestimated or underestimated, and thus such correction should be avoided in favor of improved study designs, if possible [118]. In

contrast to classical linear regression, this method is far less established in statistical literature [115].
The inversion approach suggests to treat the observed variable as the dependent variable as in the case of classical linear regression and subsequently invert the identified equation for prediction. Thus, the physical parameters should be considered as the independent variable i.e. the regressor and the sample mean perceptual attribute ratings as the dependent variable i.e. the regressand. In contrast to the previous approach, no problems with multiple predictors or data transformations are introduced. Only regression models built this way enable unbiased statistical inferences about the relationships between physical vibration parameters and population mean attribute ratings from the investigated sample. However, the goal of the synthesis model is to predict physical vibration parameters from perceptual attribute ratings. Thus, the equations of the analysis regression models need to be inverted. For each excitation pattern, one equation for each characteristic parameter needs to be included in an equation system. By solving the equation system by the physical vibration parameters the synthesis equations can be determined.

7.5.2 Classification of the Excitation Pattern with a Support Vector Machine

7.5.2.1 Selection of a Classification Algorithm

To determine the correct excitation pattern for the synthesis, a classifier needs to be built, which classifies a given expected attribute rating profile into one of the four excitation patterns. Important classification algorithms can be grouped into four categories: probabilistic, linear, prototype-based, and hierarchical [112]. The applicability of the presented classifiers belonging to each category to the problem at hand is discussed in the following.
The naive Bayes classifier falls in the probabilistic category. This method is preferable for discrete input features and is less suitable for continuous data such as the continuous attribute ratings contained in the acquired rating profiles.

The nearest neighbor classifier belongs to the prototype-based category. The disadvantage of this classifier is that no explicit classification function is designed. Since this was argued to be one of the major disadvantages of the 1-NN synthesis model, this classifier will not be utilized again to overcome this shortcoming.

Decision trees fall into the hierarchical category. They can be utilized if not all features need to be considered at a time. For example, feature A can be only relevant for classification on the top level of the decision tree and feature B only relevant for the sublevel. Indeed, observations gained in section 5.4 shows that the difference between excitation patterns is communicated by different attributes. The impulse-like excitation pattern in contrast to the non-transient excitation patterns is characterized by high ratings of the attribute “fading” as evident from Figure 5.19. In the subgroup of non-transient excitation patterns, the bandlimited WGN excitation pattern can be discriminated from periodic excitation patterns by low ratings of the attribute “uniform” as obvious from Figure 5.20. Furthermore, in the subgroup of periodic excitation patterns, the sinusoidal excitation pattern can be distinguished from the AM-sinusoidal excitation pattern by low ratings of “repetitive” as evident from Figure 5.18. However, similar to the naive Bayes classifier, the decision tree is mostly restricted to discrete input features. If the input features are treated as continuous, then the class borders are necessarily parallel to coordinate axes, preventing modeling of more complex class borders.

The support vector machine (SVM) classifier falls into the linear category. The advantage is that it is suitable for continuous features and produces an explicit classification function. The disadvantage lies in the fact that it is most suitable for the classification of two classes, but four classes are required for the problem at hand. However, this can be overcome by simply building a set of multiple binary classifiers.

Given the advantages and disadvantages of the different classification approaches, a combination of the decision tree approach and the SVM approach will be utilized. Therefore, three SVM classifiers will be utilized that distinguish between the excitation patterns in a hierarchical decision tree. The ratings of the perceptual attribute “fading” demonstrated the largest observed differences between excitation patterns (see Figure 5.19). Thus, the tree will first discriminate transient from non-transient excitation patterns with a first

SVM. For the ratings of the perceptual attribute “uniform” also large differences between excitation patterns could be observed (see Figure 5.20). Therefore, in the subgroup of non-transient excitation patterns, the bandlimited WGN excitation pattern will be discriminated from periodical excitation with a second SVM. Finally, a third SVM will discriminate between the sinusoidal excitation pattern and the AM-sinusoidal excitation pattern. The ratings of the perceptual attribute “repetitive” show differences between these two excitation patterns (see Figure 5.18).

7.5.2.2 Classifier for the Discrimination of the Transient Excitation Pattern from the Non-Transient Excitation Patterns

The dataset with 91 vibration-rating profile pairs was partitioned according to the excitation patterns. The transient excitation pattern contains the 18 impulse-like stimuli, while the non-transient excitation pattern group contains the 21 sinusoidal stimuli, 30 AM-sinusoidal stimuli, and 22 bandlimited WGN stimuli. As observed in Figure 5.19 the attribute “fading” is likely a useful feature for the classification of vibration into transient and non-transient excitation patterns. Unfortunately, the two excitation pattern groups are not linearly separable. Classification failure is an especially unfavorable property for these two excitation pattern groups. The duration of impulse-like vibration is constrained by the decay rate, while the duration of the other excitation pattern is externally constrained by the defined duration.
However, the PCA conducted in section 5.4 suggests that the attribute “uniform” is also loading highly onto the component. Impulse-like vibration is usually perceived as having low “uniform” ratings, but bandlimited WGN vibration is also perceived as having low “uniform” ratings in contrast to periodic vibration (see Figure 5.20). Thus, the attribute “uniform” alone is insufficient for separation of transient from non-transient vibration requiring the introduction of another attribute of the sensory tactile design language to the classifier. The impulse-like vibrations are rated “not at all repetitive” while bandlimited WGN is rated “slightly to moderately repetitive”. But repetitive alone would also be insufficient since high-frequency sinusoidal vibration is also rated as “not repetitive” (see Figure 5.18). Therefore, another SVM was built with the both attributes “uniform” and “repetitive”.

The binary linear SVM is built by finding the hyperplane that maximizes the margin between the two classes in the n-dimensional feature space [112]. It is possible to find non-linear class borders by introducing a kernel function. Many libraries have implemented algorithms for the construction of SVMs. MATLAB offers the *fitcsvm* function for low dimensional datasets, which was utilized for SVM building. A linear kernel function was utilized at first to examine whether the vibration a linearly separable into the two excitation pattern groups.
The resulting improved hyperplane and its support vectors for the classification of vibration into transient and non-transient excitation patterns by the attribute ratings of “uniform” r_u and “repetitive” r_r is shown in Figure 7.7. The hyperplane in the two-dimensional feature space enabling the discrimination of transient from non-transient vibration has the following equation:

$$\text{isTransient} = \begin{cases} 0, if\ r_u > -3.95 r_r + 100.7 \\ 1, if\ r_u \leq -3.95 r_r + 100.7 \end{cases} \tag{7.18}$$

The improved classifier has a much wider margin around the hyperplane and thus enables the classification of the training data with 0 % error.

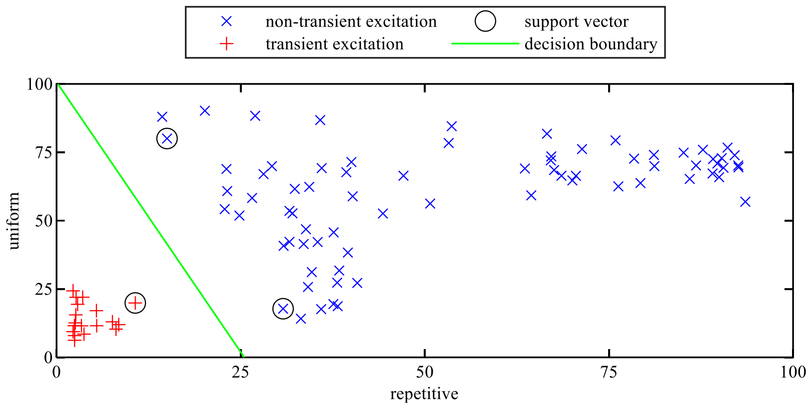

Figure 7.7.
Classification of vibration into transient and non-transient excitation patterns by the attributes “uniform” and “repetitive” according to the hyperplane of a linear SVM and its support vectors.

7.5.2.3 Classifier for the Discrimination of the Stochastic Excitation Pattern from the Periodic Excitation Patterns

The remaining dataset of non-transient vibration with 73 vibration-rating profile pairs was again partitioned into stochastic excitation containing 22 bandlimited WGN vibrations and periodic excitation containing 30 AM-sinusoidal vibrations and 21 sinusoidal vibrations. Also for this classifier, an obvious choice was suggested from the ratings of the attribute “uniform” observed in Figure 5.20. Since periodic vibration was observed to be rated less “uniform” with increasing ratings of the attribute “weak”, the attribute was also included as an input feature for the classifier. In Figure 7.8 the resulting hyperplane and its support vectors for the classification of vibration into stochastic and periodic excitation patterns by the attribute ratings of “uniform” r_u and “weak” r_w is shown.

The hyperplane in the two-dimensional feature space that can be utilized to distinguish periodic from stochastic excitation has the following equation:

$$\text{isPeriodic} = \begin{cases} 0, if\ r_u > -0.1r_w + 60.27 \\ 1, if\ r_u \leq -0.1r_w + 60.27 \end{cases} \qquad \textbf{7.19.}$$

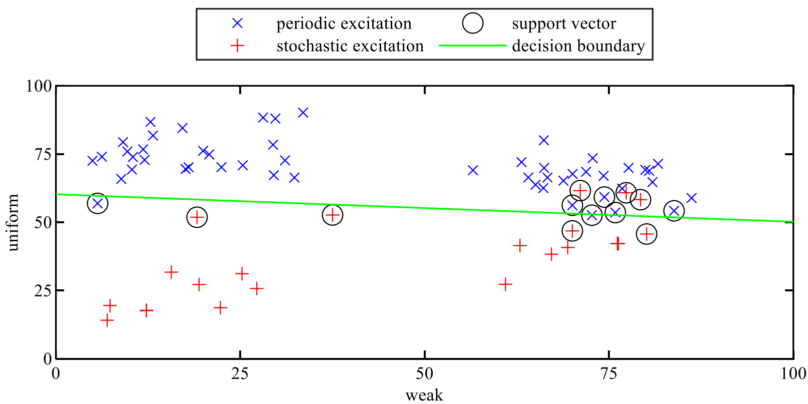

Figure 7.8.
Classification of vibration into stochastic and periodic excitation patterns by the attributes “uniform” and “weak” according to the hyperplane of a linear SVM and its support vectors.

Overall, the classifier has a good performance with a classification error of only 7 %. Such misclassification occurs mostly for bandlimited WGN vibration, which does not have low-frequency components. Such bandlimited noise is tending towards uniform ratings of sinusoidal vibration (see Figure 5.20). Thus, such misclassification is likely intrinsic to the problem at hand and not caused by a suboptimal classifier. If a misclassification results in a signal of the other excitation pattern being synthesized but which elicits similar attribute ratings as a vibration that is synthesized to the correctly classified excitation pattern, it will likely not negatively affect plausibility.

7.5.2.4 Classifier for the Discrimination of the AM-Sinusoidal Excitation Pattern from the Sinusoidal Excitation Pattern

Again the remaining dataset of periodic vibration with 51 vibration-rating profile pairs was partitioned into 30 AM-sinusoidal vibrations and into 21 sinusoidal vibrations. The attribute "repetitive" is a reasonable predictor for amplitude modulation as suggested by the high ratings observed for AM-sinusoidal vibration in contrast to sinusoidal vibration in Figure 5.18. However, it is problematic that low-frequency sinusoidal vibration can also be perceived as very "repetitive". The attribute "up and down" can be utilized to separate these cases as it is rated high for low-frequency sinusoidal vibration but low for amplitude modulated vibration. Therefore, the attributes "repetitive" and "up and down" were utilized as input features for the classifier. Again, the resulting hyperplane and its support vectors for the classification of vibration into AM-sinusoidal and sinusoidal excitation patterns by the ratings of attributes "repetitive" r_r and "up and down" r_d is shown in Figure 7.9. The hyperplane in the two-dimensional feature space enabling the discrimination of AM-sinusoidal from sinusoidal vibration has the following equation:

$$\text{isModulated} = \begin{cases} 0, if\ r_d < 1.23r_r - 46.75 \\ 1, if\ r_d \geq 1.23r_r - 46.75 \end{cases} \qquad \textbf{7.20.}$$

Overall, also this classifier has a good performance with a classification error of approximately 18 %. Introducing a polynomial kernel function or more attributes does not produce lower errors. As argued for the previous classifier,

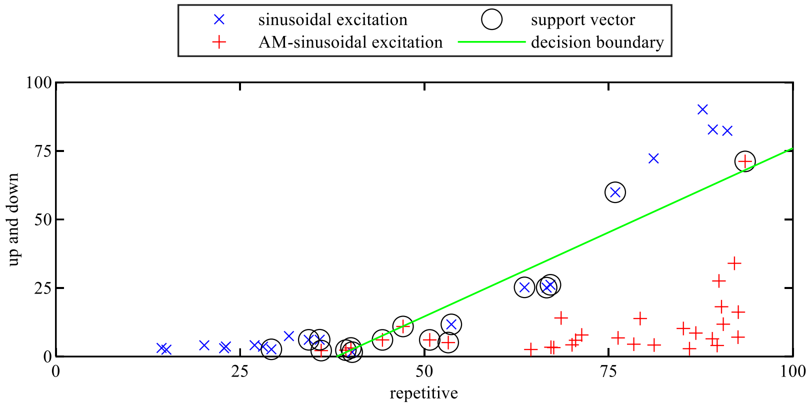

Figure 7.9.
Classification of vibration into AM-sinusoidal and sinusoidal excitation patterns by the attributes "repetitive" and "up & down" according to the hyperplane of a linear SVM and its support vectors.

some misclassification can likely be attributed to the vibration of two excitation patterns eliciting similar rating profiles and thus is not likely to lead to a decrease in plausibility ratings.

7.5.2.5 Decision Tree Based on the three Classifier

The three classifiers were joined to the decision tree shown in Figure 7.10. Often SVM classifiers are cross-validated with a separate dataset not utilized for the training to avoid overfitting. Theoretically, the recorded vehicle scenes might be utilized for such validation. Practically, the recorded vehicle scenes were not synthesized according to one of the four excitation patterns and thus the true excitation pattern is not defined. Therefore, the formal correctness of the classification cannot be assessed with this dataset. However, the overall synthesis model will be validated in the form of a perceptual study in section 7.6 with this dataset which will provide an overall measure of performance.

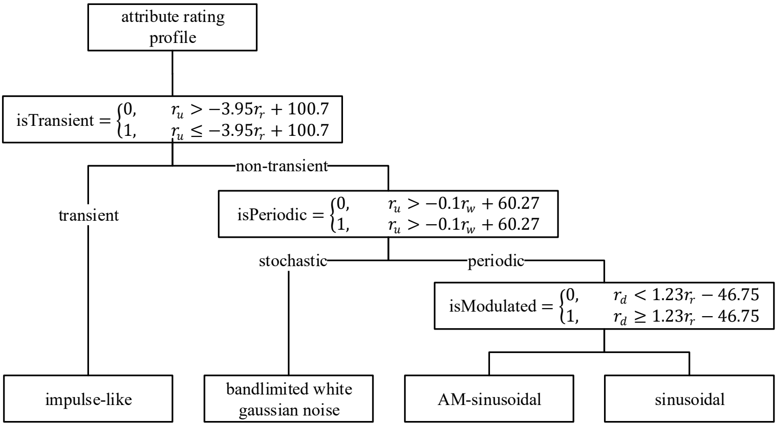

Figure 7.10.
Decision tree for classifying an attribute rating profile into one of the four excitation patterns.

7.5.3 General Approach to the Regression Models of each Excitation Pattern

After the excitation pattern for the synthesis is determined by the decision tree, it's physical vibration parameters need to be estimated for the synthesis. Thus, for each of the four excitation patterns synthesis equations need to be determined for their defining physical vibration parameters. The dataset with 91 vibration-rating profile pairs was partitioned into four excitation pattern groups of 18 impulse-like pairs, 22 bandlimited WGN pairs, 30 AM-sinusoidal pairs, and 21 sinusoidal pairs. Since the physical vibration parameters were set without error but the population sensory tactile perceptual attribute ratings were obtained with error, regression functions for the inverse problem need to be fitted.

The regressors are vibration parameters level L, (carrier-, center- or resonance-) frequency f, bandwidth f_b, modulation frequency f_m, and decay constant α. However, only a subset is relevant for each excitation pattern. The acceleration level was already log-transformed from acceleration. Such transformation is frequently applied to physical parameters to approximate linear relationships to perceptual quantities e.g. in psychoacoustics [28] and is also useful for acceleration as log-transformed acceleration shows an approximately linear relationship to perceived vibration intensity [32]. Similar to the JNDL being an approximately constant fraction of the reference level, the JNDF is an approximately constant fraction of reference frequency and not a constant offset. Therefore, a log-transform was also applied to (carrier-, center- or resonance-) frequency f, bandwidth f_b, and modulation frequency f_m. Since the modulation frequency might be zero for the unmodulated case, an offset of 1 was applied before the log-transform. The regressands are the attribute ratings of “weak” r_w, “up and down” r_d, “tingling” r_t, “repetitive” r_r, “uniform” r_u, and “fading” r_f of the expected rating profile.

In section 5.4 for each attribute of the tactile design language correlations were assessed. The attribute “weak” is highly negatively correlated with the vibration level. The attribute “up and down” is highly negatively correlated with (carrier-, center- or resonance-) frequency. The attribute “tingling” is highly positively correlated with (carrier-, center- or resonance-) frequency. The attribute “repetitive” is highly correlated to the modulation frequency. The attribute “uniform” has a moderate correlation to the bandwidth parameter. The attribute “fading” is highly correlated to the decay rate. For each physical vibration parameter, the most predictive regressand can be selected accordingly. However, the non-perfect correlations suggest that multiple regressors are likely necessary for sufficient explanatory power.

Subsequently, these regression functions need to be inverted. On the one hand, there is not necessarily an analytical solution to the inversion problem and thus the complexity of the regression models is constrained. On the other hand, the model fit is likely to be higher for increasing complexity. Therefore, a compromise between model complexity and resolvability needs to be found that still enables the identification of an analytical solution to the inversion problem. Thus, additional terms of physical vibration parameters for the prediction of the attribute ratings can only be added to the regression model, if an analytical solution to the inversion problem can be found. If only the prediction of attribute ratings from physical parameters were the goal of the regression models, more explanatory terms would potentially be included in the regressions.

The linear regression model cannot simply be applied to the pooled attribute ratings for all stimuli and test subjects. Since there is a hierarchical clustering

of ratings of vibration within test subjects due to the repeated measures this would violate the assumption of independence of observations and thus lead to inflated type I error in estimated coefficients. There are two solutions to this problem: aggregation of observations of each vibration or accounting of hierarchical clustering in a multilevel model i.e. linear mixed-effects model.
First, the repeatedly measured ratings can be aggregated across subjects by calculating the mean rating for each stimulus. Linear multiple regression models were fitted on the mean attribute ratings with the MATLAB function *fitlm*. Such a model is sufficient if the prediction of attribute ratings on the individual level is not of interest but rather the prediction on the population level. Since the goal of the synthesis is to provide vibration which will be perceived as maximally plausible on average, modeling the relationship between physical vibration parameters and attribute ratings on a population level is required. The R^2 of the population average models offers a measure of model performance directly relevant to the use case. The disadvantage of aggregating repeated measures ratings is the loss of statistical power due to the reduced number of observations.
Second, a linear mixed-effects multiple regression model can be extended from the linear model by adding subjects as random effects to the fixed effects of predictors of the simple linear model and thus enabling the fitting of the model directly to repeated measures of subjects for one vibration. For the goal of confirmatory hypothesis testing for data with one single observation per treatment level per unit, a by-unit intercept only random effect is recommended [119]. For the data obtained in section 5.4, there is indeed only one observation of a stimulus for each participant and thus a by-subject random intercept was included in the linear mixed-effects regression models. The subsequent models were estimated with the MATLAB function *fitlme*. Since continuous data was observed and the sample size is not very high, the linear mixed-effects model was estimated with restricted maximum likelihood [120]. The disadvantage of this type of model is that the conditional R^2 of the mixed-effects model is a measure of explained variance by fixed effects and random effects i.e. a measure of model fit on both the individual level and the population level and thus offers no measure of model performance directly relevant to the use case. However, for a balanced design, where each participant rated each vibration as in the case of the present data set, the coefficient estimates of the fixed effects are identical to the estimates of linear model.

Due to the higher number of observations the standard deviation of the estimates is potentially lower enabling more accurate inferences about the estimates.

7.5.4 Synthesis for the Impulse-like Excitation Pattern

For the impulse-like excitation pattern, there are 18 vibration-attribute rating profile pairs. The physical vibration parameters level L, resonance frequency f, and decay constant α need to be estimated for the synthesis of vibrations of this excitation pattern. The attributes “fading”, “weak”, and “up and down” or “tingling” are thus relevant predictors for the synthesis. For each of the four attributes, a multiple regression needs to be fitted to the dataset.

7.5.4.1 Regression for “Fading”

Obviously, the decay rate is an important predictor of the perceptual attribute “fading”. Indeed, a highly positive correlation with the decay rate was found in section 5.4 for all 99 vibration-attribute rating pairs. Furthermore, vibration level and resonance frequency seem to influence the rating of “fading” as suggested by Figure 7.3 in section 7.2. First, a simple linear regression model was fitted to the dataset of by-stimulus averages with the previously selected features. The model is shown in Table 7.2. R^2 was 0.954 suggesting a very good model fit. The terms decay rate and resonance frequency are highly significant, and the term level and intercept are also significant. Therefore, the null hypothesis can be rejected for the reported model. Second, a linear mixed-effects model was fitted to the dataset of 28 subjects with the selected features as fixed effects and by-subject random intercepts. The fixed effects of the model are shown in Table 7.3. The linear mixed-effects model demonstrated a similar standard deviation as the linear model.

7.5.4.2 Regression for “Weak”

The semantics of the attribute “weak” suggests level as a highly relevant predictor. Furthermore, some frequency dependence was observed in Figure 5.12 of section 5.4 with high-frequency vibration being perceived as slightly more

Table 7.2.
Effects in the linear regression model of the attribute "fading" as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Adjusted R^2 of the model: $R^2 = 0.954$.

	Estimate	Std. error	t-value	p-value	
Intercept	43.731	14.638	2.987	0.010	*
L	0.448	0.124	3.613	0.003	**
$\log_{10}(f)$	-9.539	1.560	-6.117	0.000	***
α	-4.968	0.299	-16.602	0.000	***

Table 7.3.
Fixed effects in the linear mixed effects regression model of the attribute "fading" as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Conditional adjusted R^2 of the model: $R^2 = 0.515$. SD of the residual is 19.174. SD for the random effect of participant is 12.359.

	Estimate	Std. error	t-value	p-value	
Intercept	43.731	15.328	2.853	0.005	**
L	0.448	0.128	3.491	0.001	**
$\log_{10}(f)$	-9.539	1.614	-5.911	0.000	***
α	-4.968	0.310	-16.042	0.000	***

"weak" than low-frequency vibration. The decay rate determines how fast the envelope approaches the perception threshold and thus is also likely to affect perceived intensity as communicated with the attribute "weak". Again, a simple linear regression model was fitted to the dataset of by-stimulus averages with the selected regressors. In Table 7.4 the model is reported. R^2 was 0.949 also suggesting a very good model fit. The terms resonance frequency, level, and intercept are highly significant, and the decay rate is also significant. Thus, the null hypothesis for the model can be rejected. Furthermore, a linear mixed-effects model was also fitted to this dataset of 29 subjects with the selected features as fixed effects and by-subject random intercepts. The fixed effects of the model are shown in Table 7.5. Due to the higher number of observations, the standard errors for the fixed effect terms intercept, decay rate, level, and resonance frequency decreased by almost half relative to the linear model. All the reported terms are highly significant for the linear mixed-effects regression.

Table 7.4.
Effects in the linear regression model of the attribute “weak” as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Adjusted R^2 of the model: $R^2 = 0.949$.

	Estimate	Std. error	t-value	p-value	
Intercept	354.494	20.942	16.927	0.000	***
L	-2.879	0.177	-16.242	0.000	***
$\log_{10}(f)$	31.429	2.231	14.087	0.000	***
α	1.610	0.428	3.761	0.002	**

Table 7.5.
Fixed effects in the linear mixed effects regression model of the attribute “weak” as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Conditional adjusted R^2 of the model: $R^2 = 0.680$. SD of the residual is 15.930. SD for the random effect of participant is 9.304.

	Estimate	Std. error	t-value	p-value	
Intercept	354.494	12.487	28.390	0.000	***
L	-2.879	0.105	-27.504	0.000	***
$\log_{10}(f)$	31.429	1.317	23.855	0.000	***
α	1.610	0.253	6.369	0.000	***

7.5.4.3 Regression for “Up and down”

Intuitively, frequency would be associated with the attribute “up and down” which is confirmed by the highly negative correlation with frequency obtained in section 5.4. However, also vibration level affects the rating of “up and down” as implied by Figure 5.13. Also for this attribute, a simple linear regression model was fitted to the dataset of by-stimulus averages with the selected regressors first. This model is reported in Table 7.6. The R^2 was 0.886, which suggests a very good model fit. The term resonance frequency is highly significant, and the level is significant. However, the estimated intercept is not significant implying that when all the predictors equal zero, the estimate of the intercept is not significantly different from zero. Since the model will only be utilized above the perception threshold, i.e. in a level range far away from zero the intercept can remain in the model despite being non-significant [121] thus enabling predictions. To narrow down the standard deviation of the intercept, a linear mixed-effects model was fitted to this dataset of 29 subjects, which included by-subject random intercepts in addition to the fixed effects of the linear model. The resulting fixed effects of this

Table 7.6.
Effects in the linear regression model of the attribute "up and down" as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Adjusted R^2 of the model: $R^2 = 0.886$.

	Estimate	Std. error	t-value	p-value	
Intercept	-53.574	36.901	-1.452	0.167	
L	1.098	0.313	3.511	0.003	**
$\log_{10}(f)$	-42.635	3.893	-10.953	0.000	***

Table 7.7.
Fixed effects in the linear mixed effects regression model of the attribute "up and down" as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Conditional adjusted R^2 of the model: $R^2 = 0.718$. SD of the residual is 16.572. SD for the random effect of participant is 10.015.

	Estimate	Std. error	t-value	p-value	
Intercept	-53.574	12.979	-4.128	0.000	***
L	1.098	0.109	10.086	0.000	***
$\log_{10}(f)$	-42.635	1.355	-31.465	0.000	***

model are shown in Table 7.7. The higher number of observations leads to a decrease in the standard errors of the fixed effects in comparison to the linear model. All the reported terms are highly significant for the linear mixed-effects regression. Thus, the null hypothesis for all the fixed effects terms of the model can be rejected.

7.5.4.4 Regression for "Tingling"

Intuitively, frequency would also be associated with the attribute "tingling" which is confirmed by the highly positive correlation with frequency obtained in section 5.4. As argued for the attribute "weak" the decay rate determines how fast the envelope approaches the perception threshold. For longer decays, there is more time available for the energy integration of the high-frequency oscillations associated with "tingling". Therefore, the decay rate also needs to be included as a regressor. Again, for "tingling" a simple linear regression model was fitted to the dataset of by-stimulus averages with the selected regressors. This model is shown in Table 7.8. Again, a very good model fit is implied by the R^2 of 0.870. The term resonance frequency is highly significant, and the decay rate is significant. As for "up and down" the intercept

Table 7.8.
Effects in the linear regression model of the attribute "tingling" as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Adjusted R^2 of the model: $R^2 = 0.870$.

	Estimate	Std. error	t-value	p-value	
Intercept	-2.894	4.259	-0.679	0.507	
$\log_{10}(f)$	27.332	2.566	10.653	0.000	***
α	-2.190	0.609	-3.595	0.003	**

Table 7.9.
Fixed effects in the linear mixed effects regression model of the attribute "tingling" as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Conditional adjusted R^2 of the model: $R^2 = 0.550$. SD of the residual is 18.671. SD for the random effect of participant is 9.949.

	Estimate	Std. error	t-value	p-value	
Intercept	-2.894	2.825	-1.024	0.306	
$\log_{10}(f)$	27.332	1.270	21.520	0.000	***
α	-2.190	0.302	-7.262	0.000	***

is not significantly different from zero for this model. To also narrow down the standard deviation of the intercept for this attribute, a linear mixed-effects model was fitted to this dataset of 28 subjects, which included by-subject random intercepts besides the fixed effects of the linear model. The resulting fixed effects of the model are shown in Table 7.9. The higher number of observations leads to a large decrease in the standard errors of the fixed effects. The terms are resonance frequency and decay rate are highly significant for this linear mixed-effects regression. However, the intercept is still not significantly different from zero. Therefore, the null hypothesis can only be rejected for the decay rate and the resonance frequency. Since the model will not be utilized for a decay rate close to zero, the intercept can remain in the model despite being non-significant [121] enabling more accurate predictions.

7.5.4.5 Synthesis Equations

The linear models provide the regression equations for the prediction of the attribute ratings of "fading" r_f, "weak" r_w, "up and down" r_d, and tingling r_t.

Since the frequency was log-transformed before the regression, the logarithm of frequency instead of frequency is present in the equations. These equations are forming the following equation system:

$$\begin{array}{lllllll} r_f & = & -4.968\alpha & +0.448L & -9.539\log_{10}(f) & +43.731 & (I) \\ r_w & = & 1.61\alpha & -2.879L & +31.429\log_{10}(f) & +354.494 & (II) \\ r_d & = & & 1.098L & -42.635\log_{10}(f) & -53.574 & (III) \\ r_t & = & -2.19\alpha & & +27.332\log_{10}(f) & -2.894 & (IV) \end{array} \quad \textbf{7.21.}$$

However, there are four equations but only the three unknowns level L, resonance frequency f, and decay rate α, thus producing an overdetermined equation system. The decay rate has the largest influence on the attribute “fading”, the level on the attribute “weak”, and the frequency on the attributes “up and down” or “tingling”. Since both “up and down” and “tingling” are highly correlated to frequency, they are most likely redundant for this excitation pattern. Therefore, it is reasonable to omit one equation from the overdetermined equation system to find an analytical solution and to avoid numeric approximation. Therefore, two subsets can be selected from the four equations that both include equation the equations I and II for “fading” and “weak” but either the equation III or IV for “up and down” or “tingling”. These two equation systems can be solved analytically in MATLAB by the physical vibration parameters using the *solve* command.

Two solutions of synthesis equations were obtained from the equation systems. The physical parameters of the impulse-like excitation pattern can be estimated from the ratings of “fading”, “weak”, and “up and down” with the following equations for the first equation system:

$$\begin{array}{cccccc} L & = & -0.162r_f & -0.499r_w & -0.332r_d & +166.0 \\ \log_{10}(f) & = & -0.00417r_f & -0.0129r_w & -0.032r_d & +3.02 \\ \alpha & = & -0.208r_f & -0.0203r_w & +0.0316r_d & +18.0 \end{array} \quad \textbf{7.22.}$$

The R^2 of the synthesis equations for the 18 vibration-attribute rating profile pairs of the impulse-like excitation pattern can be calculated, by comparing the defined vibration parameters to the vibration parameters estimated from the rating profiles by the synthesis equations. The vibration level can be predicted with an R^2 of 0.84, the frequency with an R^2 of 0.87, and the decay rate with an R^2 of 0.92.

Similarly, the physical parameters of the impulse like-excitation pattern can be estimated from the ratings of “fading”, “weak”, and “tingling” with the following equations for the second equation system:

$$\begin{array}{ccccccc} L & = & -0.282r_f & -0.391r_w & +0.351r_t & +152.0 \\ \log_{10}(f) & = & -0.0157r_f & -0.00245r_w & +0.0339r_t & +1.65 \\ \alpha & = & -0.196r_f & -0.0305r_w & -0.0334r_t & +19.3 \end{array} \qquad \textbf{7.23.}$$

The vibration level can be predicted with an R^2 of 0.9, the resonance frequency with an R^2 of 0.9, and the decay rate with an R^2 of 0.94. Since the R^2 of the synthesis equations was higher for this solution, they were chosen for the synthesis model.

7.5.5 Synthesis for the Bandlimited White Gaussian Noise Excitation Pattern

For the bandlimited WGN excitation pattern, there are 22 vibration-attribute rating profile pairs. The physical vibration parameters level L, carrier frequency f, and bandwidth f_b need to be estimated for the synthesis of vibration of this excitation pattern. The attributes “uniform”, “weak”, and “up and down” or “tingling” are thus relevant predictors for the synthesis. For each of the four attributes, a multiple regression needs to be fitted to the dataset.

7.5.5.1 Regression for “Uniform”

For the attribute “uniform” a negative correlation to bandwidth was found in section 5.4 for all 99 vibration-attribute rating pairs. While this attribute is suitable to distinguish stochastic from periodic vibration, it demonstrates little variation depending on the bandwidth for the subdomain of bandlimited white noise vibration. In contrast, the center frequency seems to be positively correlated to “uniform” as observed in Figure 5.20. Apart from this predictor, vibration level seems to slightly influence it's rating as suggested by Figure

Table 7.10.
Effects in the linear regression model of the attribute "uniform" as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Adjusted R^2 of the model: $R^2 = 0.743$.

	Estimate	Std. error	t-value	p-value	
Intercept	50.472	14.213	3.551	0.002	**
L	-0.629	0.120	-5.227	0.000	***
$\log_{10}(f)$	31.402	4.432	7.086	0.000	***

Table 7.11.
Fixed effects in the linear mixed effects regression model of the attribute "uniform" as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Conditional adjusted R^2 of the model: $R^2 = 0.315$. SD of the residual is 24.697. SD for the random effect of participant is 11.707.

	Estimate	Std. error	t-value	p-value	
Intercept	50.472	9.083	5.557	0.000	***
L	-0.629	0.075	-8.433	0.000	***
$\log_{10}(f)$	31.402	2.747	11.432	0.000	***

7.3 in section 7.2. First, a simple linear regression model was again fitted to the dataset of by-stimulus averages with the previously selected features. The model is shown in Table 7.10. The R^2 of 0.743 suggested a good model fit. The terms level and center frequency are highly significant, and the term intercept is also significant. The null hypothesis can thus be rejected for the reported model. Second, also a linear mixed-effects model was fitted to the dataset of 28 subjects with the selected features as fixed effects and by-subject random intercepts. The model's fixed effects are shown in Table 7.11. The linear mixed-effects model narrowed down the standard deviation of the coefficients in comparison to the linear model, too. All the reported terms are highly significant for the linear mixed-effects model.

7.5.5.2 Regression for "Weak"

Again, the semantics of the attribute "weak" suggest level as a highly relevant predictor. Similar to impulse-like vibration, frequency might be an additional predictor. Again, a simple linear regression model was fitted to the dataset of by-stimulus averages with the selected regressors. The model consisting of

Table 7.12.
Effects in the linear regression model of the attribute "weak" as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Adjusted R^2 of the model: $R^2 = 0.761$.

	Estimate	Std. error	t-value	p-value	
Intercept	246.671	24.597	10.028	0.000	***
L	-1.764	0.214	-8.247	0.000	***

Table 7.13.
Fixed effects in the linear mixed effects regression model of the attribute "weak" as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Conditional adjusted R^2 of the model: $R^2 = 0.570$. SD of the residual is 23.136. SD for the random effect of participant is 10.783.

	Estimate	Std. error	t-value	p-value	
Intercept	246.671	7.898	31.234	0.000	***
L	-1.764	0.066	-26.553	0.000	***

intercept, level, and center frequency had an R^2 of 0.908. However, the inclusion of center frequency as a predictor ultimately led to worse performance for the bandwidth parameter for the subsequent synthesis equation. Therefore, the term center frequency was dropped from the model. In Table 7.12 the reduced model is reported. R^2 was 0.761, also suggesting still a good reduced model fit. The terms level and intercept are highly significant. Thus, the null hypothesis for the model can be rejected. Furthermore, a linear mixed-effects model was also fitted to this dataset of 29 subjects with the selected features as fixed effects and by-subject random intercepts. The fixed effects of the model are shown in Table 7.13. Due to the higher number of observations, the standard errors for both fixed-effect terms intercept and level decreased to about a third of standard errors of the linear model. Both terms are highly significant for the linear mixed-effects regression.

7.5.5.3 Regression for "Up and down"

In section 5.4 it was argued that the ratings of low frequency describing attributes such as "up and down" have a rating peak at a specific frequency below 35 Hz and falling ratings above that frequency for sinusoidal vibration. Similarly, for bandlimited noise, the rating is high, if energy is falling into the frequency band where the maximum rating can be observed, as suggested by

Table 7.14.
Effects in the linear regression model of the attribute “up and down” as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Adjusted R^2 of the model: $R^2 = 0.628$.

	Estimate	Std. error	t-value	p-value	
Intercept	-43.187	28.860	-1.496	0.152	
L	0.888	0.242	3.675	0.002	**
$\log_{10}(f)$	-67.379	12.419	-5.425	0.000	***
$\log_{10}(f_b)$	47.794	11.060	4.321	0.000	***

Table 7.15.
Fixed effects in the linear mixed effects regression model of the attribute “up and down” as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Conditional adjusted R^2 of the model: $R^2 = 0.595$. SD of the residual is 17.494. SD for the random effect of participant is 6.974.

	Estimate	Std. error	t-value	p-value	
Intercept	-43.187	6.229	-6.933	0.000	***
L	0.888	0.051	17.393	0.000	***
$\log_{10}(f)$	-67.379	2.624	-25.678	0.000	***
$\log_{10}(f_b)$	47.794	2.337	20.452	0.000	***

Figure 5.16. Whether energy does fall into the range of maximum suitability depends on the center frequency and bandwidth of the bandlimited noise. Furthermore, also vibration level affects the rating of “up and down” as implied by Figure 5.14. Also for this attribute, a simple linear regression model was fitted to the dataset of by-stimulus averages with the selected regressors first. This model is reported in Table 7.14. The R^2 was 0.628, which suggests a reasonable model fit. The terms frequency and bandwidth are highly significant, and the level is significant. Similarly, to the “up and down” model for the transient excitation pattern, the estimated intercept is not significant.
To narrow down the standard deviation of the intercept, a linear mixed-effects model was fitted to this dataset of 30 subjects, which included by-subject random intercepts in addition to the fixed effects of the linear model. The resulting fixed effects of this model are shown in Table 7.15. The higher number of observations lead to a decrease in the standard errors of the fixed effects in comparison to the linear model. All the reported terms are highly significant for the linear mixed-effects regression model. Thus, the null hypothesis for all the fixed effects terms of this model can be rejected.

7.5.5.4 Regression for "Tingling"

The argument for predictors for "tingling" is identical to "up and down". Therefore, frequency, bandwidth, and level were included as regressors initially. For "tingling" a simple linear regression model was fitted to the dataset of by-stimulus averages with the selected regressors. However, in a model with the predictors intercept, center frequency, level, and bandwidth, the estimate for bandwidth had a very large p-value and was thus not significant. Therefore, the term bandwidth was dropped from final the model. This model is shown in Table 7.16. Again, a good model fit is implied by the R^2 of 0.773. The terms level and intercept are highly significant, but the center frequency is above the significance threshold for the linear model. To also narrow down the standard deviation of the center frequency for this attribute, a linear mixed-effects model was fitted to this dataset of 28 subjects, which included by-subject random intercepts besides the fixed effects of the linear model. The resulting fixed effects of the model are shown in Table 7.17. The higher number of observations lead to a large decrease in the standard errors of the fixed effects. The terms level and intercept are highly significant, and the term center frequency is significant for this linear mixed-effects regression model.

Table 7.16.
Effects in the linear regression model of the attribute "tingling" as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Adjusted R^2 of the model: $R^2 = 0.773$.

	Estimate	Std. error	t-value	p-value	
Intercept	-111,863	16,363	-6,836	0,000	***
L	1,098	0,139	7,923	0,000	***
$\log_{10}(f)$	6,065	5,102	1,189	0,249	

Table 7.17.
Fixed effects in the linear mixed effects regression model of the attribute "tingling" as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Conditional adjusted R^2 of the model: $R^2 = 0.509$. SD of the residual is 17.756. SD for the random effect of participant is 8.822.

	Estimate	Std. error	t-value	p-value	
Intercept	-111.863	6.549	-17.080	0.000	***
L	1.098	0.054	20.469	0.000	***
$\log_{10}(f)$	6.065	1.975	3.071	0.002	**

7.5.5.5 Synthesis Equations

Again, the linear models provide the regression equations for the prediction of the attribute ratings of "uniform" r_u, "weak" r_w, "up and down" r_d, and tingling r_t. Since the frequency and the bandwidth were log-transformed before the regression, the logarithm of frequency and the logarithm of bandwidth instead of frequency and bandwidth are present in the equations. These equations are forming the following equation system:

$$
\begin{array}{lclllll}
r_u & = & & -0.629L & +31.402\log_{10}(f) & +50.472 & (I) \\
r_w & = & & -1.764L & & +246.671 & (II) \\
r_d & = & +47.794\log_{10}(f_b) & 0.888L & -67.379\log_{10}(f) & -43.187 & (III) \\
r_t & = & & 1.098L & +6.065\log_{10}(f) & -111.863 & (IV)
\end{array}
\tag{7.24.}
$$

Again, there are four equations but only the three unknowns Level L, center frequency f, and bandwidth f_b, thus producing an overdetermined equation system. Since the bandwidth is only contained in the equation of the attribute "up and down", it needs to be included in the equation system. Furthermore, level has the largest influence on the attribute "weak" and is thus also included. However, "up and down" and "tingling" are both correlated to frequency, but "up and down" is already included in the equation system. Therefore, it is reasonable to omit the equation of "tingling" from the overdetermined equation system to find an analytical solution for the equation system. This equation system was again solved analytically in MATLAB by the physical vibration parameters. One solution of synthesis equations was obtained from the equation system. The physical parameters of the bandlimited WGN excitation pattern can be estimated from the ratings of "uniform", "weak", and "up and down" with the following equations:

$$
\begin{array}{lcllll}
L & = & & & -0.567r_w & +140.0 \\
\log_{10}(f) & = & 0.0318r_u & & -0.0114r_w & +1.19 \\
\log_{10}(f_b) & = & 0.0449r_u & +0.0209r_d & -0.00548r_w & -0.0109
\end{array}
\tag{7.25.}
$$

However, the bandwidth is not independent from the center frequency. The bandlimited WGN excitation pattern is only well defined for a center frequency which is larger than double the bandwidth. Furthermore, the resulting signal should not contain components below one Hz, to not introduce a static component into the acceleration signal. Therefore, the following constraint was introduced to the synthesis equation of the bandwidth:

$$f_b = min(f_b, 2(f - 1\ Hz)) \quad \textbf{7.26.}$$

Again, the R^2 of the synthesis equations for the 22 vibration-attribute rating profile pairs of the bandlimited WGN excitation pattern can be calculated, by comparing the defined vibration parameters to the vibration parameters estimated from the rating profiles by the synthesis equations. The vibration level can be predicted with an R^2 of 0.71, the center frequency with an R^2 of 0.68, and the constraint bandwidth with an R^2 of 0.5.

7.5.6 Synthesis for the Amplitude Modulated Sinusoidal Excitation Pattern

For the AM-sinusoidal excitation pattern, there are 30 vibration-attribute rating profile pairs. The physical vibration parameters level L and frequency f need to be estimated for the synthesis of vibration of this excitation pattern. The attributes “weak” and “up and down” or “tingling” are thus relevant predictors for the synthesis. For each of the three attributes, a multiple regression needs to be fitted to the dataset.

7.5.6.1 Regression for “Repetitive”

Modulation frequency is an obvious predictor of the perceptual attribute “repetitive”, which is evident from the highly positive correlation with modulation frequency found in section 5.4. Furthermore, the vibration level seems to influence the rating of “repetitive” as suggested by Figure 7.3 of section 7.2. First, a simple linear regression model was again fitted to the dataset of by-stimulus averages with the previously selected features. The model is shown in Table 7.18. The R^2 of 0.848 suggested a good model fit. The terms

Table 7.18.
Effects in the linear regression model of the attribute “repetitive " as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Adjusted R^2 of the model: $R^2 = 0.848$.

	Estimate	Std. error	t-value	p-value	
Intercept	6.061	11.609	0.522	0.606	
L	0.902	0.100	9.051	0.000	***
$\log_{10}(f_m+1)$	-43.631	4.494	-9.709	0.000	***

Table 7.19.
Fixed effects in the linear mixed effects regression model of the attribute “repetitive " as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Conditional adjusted R^2 of the model: $R^2 = 0.512$. SD of the residual is 21.720. SD for the random effect of participant is 14.946.

	Estimate	Std. error	t-value	p-value	
Intercept	6.061	7.097	0.854	0.393	
L	0.902	0.056	16.086	0.000	***
$\log_{10}(f_m+1)$	-43.631	2.529	-17.255	0.000	***

level and modulation frequency are highly significant. Similarly, to the “up and down” model for the transient and stochastic excitation pattern, the estimated intercept is not significant. To narrow down the standard deviation of the intercept, a linear mixed-effects model was fitted to the dataset of “repetitive” of 29 subjects, which included by-subject random intercepts in addition to the fixed effects of the linear model. The model's fixed effects are shown in Table 7.19. The linear mixed-effects model narrowed down the standard deviation of the coefficients in comparison to the linear model, too. The reported terms level and modulation frequency are highly significant for the linear mixed-effects model. Similar to the “tingling” model for the transient excitation pattern, the intercept is still not significantly different from zero. Therefore, the null hypothesis can only be rejected for the modulation frequency and the level. Again, since the model will not be utilized for a level close to zero, the intercept can remain in the model despite being non-significant [121] enabling more accurate predictions.

7.5.6.2 Regression for "Weak"

As for the transient and the stochastic excitation pattern, the semantics of the attribute suggest vibration level as a highly relevant predictor. Similarly, carrier frequency might be an additional predictor as observed in Figure 5.12 of section 5.4. Again, a simple linear regression model was fitted to the dataset of by-stimulus averages with the selected regressors. In Table 7.20 the model is reported. R^2 was 0.976 also suggesting a very good model fit. All the terms are highly significant for the linear model. Thus, the null hypothesis for the model can be rejected. Furthermore, a linear mixed-effects model was also fitted to this dataset of 29 subjects with the selected features as fixed effects and by-subject random intercepts. The fixed effects of the model are shown in Table 7.21. All the reported terms are also highly significant for the linear mixed-effects regression. The linear mixed-effects model demonstrated similar standard deviation as the linear model.

Table 7.20.
Effects in the linear regression model of the attribute "weak" as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Adjusted R^2 of the model: $R^2 = 0.976$.

	Estimate	Std. error	t-value	p-value	
Intercept	228.866	6.950	32.931	0.000	***
L	-1.994	0.059	-33.730	0.000	***
$\log_{10}(f)$	23.757	1.814	13.093	0.000	***

Table 7.21.
Fixed effects in the linear mixed effects regression model of the attribute "weak" as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Conditional adjusted R^2 of the model: $R^2 = 0.698$. SD of the residual is 19.287. SD for the random effect of participant is 11.933.

	Estimate	Std. error	t-value	p-value	
Intercept	228.866	6.315	36.243	0.000	***
L	-1.994	0.050	-39.643	0.000	***
$\log_{10}(f)$	23.757	1.544	15.388	0.000	***

Table 7.22.
Effects in the linear regression model of the attribute “up and down” as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Adjusted R^2 of the model: $R^2 = 0.667$.

	Estimate	Std. error	t-value	p-value	
Intercept	101.302	31.855	3.180	0.004	**
L	0.393	0.110	3.565	0.001	**
$\log_{10}(f)$	-146.993	38.099	-3.858	0.001	**
$\log_{10}(f)^2$	37.734	11.492	3.284	0.003	**

Table 7.23.
Fixed effects in the linear mixed effects regression model of the attribute “up and down” as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Conditional adjusted R^2 of the model: $R^2 = 0.553$. SD of the residual is 12.549. SD for the random effect of participant is 8.439.

	Estimate	Std. error	t-value	p-value	
Intercept	101.302	31.855	3.180	0.004	**
L	0.393	0.110	3.565	0.001	**
$\log_{10}(f)$	-146.993	38.099	-3.858	0.001	**
$\log_{10}(f)^2$	37.734	11.492	3.284	0.003	**

7.5.6.3 Regression for “Up and down”

As for the transient excitation pattern, the frequency would be intuitively associated with the attribute “up and down” which is confirmed by the highly negative correlation with frequency obtained in section 5.4. Similarly, also vibration level affects the rating of “up and down” as implied by Figure 5.14. A simple linear regression model was fitted to the dataset of by-stimulus averages with the selected regressors first. However, the R^2 of only 0.543 for the linear model with the terms intercept, frequency, and level suggested only a mediocre fit. For sinusoidal vibration, the rating curve of “up and down” (see Figure 5.14) shows a parabola shape for frequencies above 9 Hz, as contained in the AM-sinusoidal dataset. Therefore, a quadratic term for frequency was introduced to the model to increase the model fit. The improved model is reported in Table 7.22. The R^2 was 0.667, which suggests a reasonable model fit. All the terms of the model are significant. Thus, the null hypothesis for all the terms of the model can be rejected. To narrow down the standard deviation of the estimates, also a linear mixed-effects model was fitted to this dataset of 30 subjects, which included by-subject random intercepts in addition to the fixed effects of the linear model. The resulting fixed effects of this model are shown in Table 7.23. The higher number of observations lead to a decrease in the standard errors of the fixed effects in comparison to the linear model. All the reported terms are highly significant for the linear mixed-effects regression.

7.5.6.4 Regression for "Tingling"

As for the stochastic excitation pattern, the frequency would be intuitively associated with the attribute "tingling" which is confirmed by the highly negative correlation with frequency obtained in section 5.4. Again, also vibration level affects the rating of "tingling" as implied by Figure 5.17. Again, for "tingling" a simple linear regression model was fitted to the dataset of by-stimulus averages with the selected regressors. This model is shown in Table 7.24. Again, a good model fit is implied by the R^2 of 0.775. The terms level and intercept are highly significant, and the carrier frequency is significant. Thus, the null hypothesis for all the three terms of the model can be rejected. To also narrow down the standard deviation of the estimates, a linear mixed-effects model was fitted to this dataset of 28 subjects, which included by-subject random intercepts besides the fixed effects of the linear model. The resulting fixed effects of the model are shown in Table 7.25. The higher number of observations lead to a large decrease in the standard errors of the fixed effects. Thus, the all terms are highly significant for this linear mixed-effects regression model.

7.5.6.5 Synthesis Equations

The linear models provide the regression equations for the prediction of the attribute ratings of "repetitive" r_r, "weak" r_w, "up and down" r_d, and tingling r_t. Since the frequency and the modulation frequency were log-transformed before the regression, the logarithm of frequency and the logarithm of modulation frequency instead of frequency and modulation frequency are present in the equations. These equations are forming the following equation system:

Table 7.24.
Effects in the linear regression model of the attribute “tingling” as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Adjusted R^2 of the model: $R^2 = 0.775$.

	Estimate	Std. error	t-value	p-value	
Intercept	-132.244	15.941	-8.296	0.000	***
L	1.210	0.136	8.925	0.000	***
$\log_{10}(f)$	13.227	4.162	3.178	0.004	**

Table 7.25.
Fixed effects in the linear mixed effects regression model of the attribute “tingling” as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Conditional adjusted R^2 of the model: $R^2 = 0.536$. SD of the residual is 19.021. SD for the random effect of participant is 10.239.

	Estimate	Std. error	t-value	p-value	
Intercept	-132.244	6.243	-21.184	0.000	***
L	1.210	0.050	23.971	0.000	***
$\log_{10}(f)$	13.227	1.550	8.536	0.000	***

$$
\begin{aligned}
r_r &= +0.902L - 43.631\log_{10}(f_m + 1) + 6.061 && (I)\\
r_w &= -1.994L + 23.757\log_{10}(f) + 228.866 && (II)\\
r_d &= 0.393L + 37.734\log_{10}(f)^2 - 146.993\log_{10}(f) + 101.302 && (III)\\
r_t &= 1.21L + 13.227\log_{10}(f) - 132.244 && (IV)
\end{aligned}
$$

7.27.

Again, there are four equations but only the three unknowns Level L, carrier frequency f, and modulation frequency f_m, thus producing an overdetermined equation system. The modulation frequency has the largest influence on the attribute “repetitive” and the level on the attribute “weak”. As mentioned for the previous excitation patterns, both “up and down” and “tingling” are highly correlated to frequency and are thus most likely redundant also for this excitation pattern. For the AM-sinusoidal excitation pattern, the carrier frequency is usually higher than the modulation frequency. For higher carrier frequencies the attribute “tingling” is utilized for describing vibration perception. Furthermore, the adjusted R^2 of the tingling model is 0.775 while the adjusted R^2 of the “up and down” model is smaller with only 0.667 for the AM-sinusoidal excitation pattern. Therefore, it is reasonable to omit the “up

and down" equation from the overdetermined equation system to find an analytical solution. MATLAB was again utilized to solve the resulting equation system.
One solution of synthesis equations was obtained from the equation system for the AM-sinusoidal excitation pattern. The physical parameters of this excitation pattern can be estimated from the ratings of "repetitive", "weak" and "tingling" with the following equations:

$$\begin{array}{rcrrrr} L & = & & 0.431r_t & -0.24r_w & +112.0 \\ \log_{10}(f) & = & & 0.0362r_t & +0.022r_w & -0.241 \\ \log_{10}(f_m+1) & = & -0.0229r_r & +0.00891r_t & -0.00496r_w & +2.45 \end{array} \tag{7.28.}$$

The vibration level can be predicted with an R^2 of 0.9, the carrier frequency with an R^2 of 0.39, and the modulation frequency with an R^2 of 0.43.

7.5.7 Synthesis for the Sinusoidal Excitation Pattern

For the sinusoidal excitation pattern, there are 21 vibration-attribute rating profile pairs. The physical vibration parameters level L and resonance frequency f need to be estimated for the synthesis of vibration of this excitation pattern. The attributes "weak" and "up and down" or "tingling" are thus relevant predictors for the synthesis. For each of the four attributes, a multiple regression needs to be fitted to the dataset.

7.5.7.1 Regression for "Weak"

As for the transient pattern, the stochastic and the AM-sinusoidal excitation pattern, the semantics of the attribute "weak" suggests vibration level as a highly relevant predictor. Again, the frequency might be an additional predictor as observed in Figure 5.12 of section 5.4. A simple linear regression model was fitted to the dataset of by-stimulus averages with the selected regressors. In Table 7.26 the model is reported. R^2 was 0.892 also suggesting a very good model fit. All the terms are highly significant also for this linear model. Thus,

Table 7.26.
Effects in the linear regression model of the attribute “weak” as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Adjusted R^2 of the model: $R^2 = 0.892$.

	Estimate	Std. error	t-value	p-value	
Intercept	243.078	18.583	13.081	0.000	***
L	-2.070	0.172	-12.044	0.000	***
$\log_{10}(f)$	26.601	2.990	8.896	0.000	***

Table 7.27.
Fixed effects in the linear mixed effects regression model of the attribute “weak” as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Conditional adjusted R^2 of the model: $R^2 = 0.663$. SD of the residual is 21.807. SD for the random effect of participant is 9.275.

	Estimate	Std. error	t-value	p-value	
Intercept	243.078	7.558	32.160	0.000	***
L	-2.070	0.068	-30.411	0.000	***
$\log_{10}(f)$	26.601	1.184	22.462	0.000	***

the null hypothesis for the model can be rejected. Furthermore, the fixed effects of a linear mixed-effects model are shown in shown in
Table 7.27. It was also fitted to this dataset of 29 subjects with the selected features as fixed effects and by-subject random intercepts. All the reported terms are also highly significant for the linear mixed-effects regression. This linear mixed-effects model demonstrated a similar standard deviation as the linear model.

7.5.7.2 Regression for “Up and down”

As for the transient and the AM-sinusoidal excitation pattern, frequency is intuitively associated with the attribute “up and down” as evident from the highly negative correlation with frequency obtained in section 5.4. Again, also vibration level affects the rating of “up and down” as implied by Figure 5.14. Also for this attribute, a simple linear regression model was fitted to the dataset of by-stimulus averages with the selected regressors first. This model is reported in Table 7.28. The R^2 was 0.718, which suggests a reasonable model fit. The term frequency is highly significant while the term level is

Table 7.28.
Effects in the linear regression model of the attribute "up and down" as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Adjusted R^2 of the model: $R^2 = 0.718$.

	Estimate	Std. error	t-value	p-value	
Intercept	-55.276	30.372	-1.820	0.085	
L	1.122	0.281	3.993	0.001	**
$\log_{10}(f)$	-34.849	4.887	-7.131	0.000	***

Table 7.29.
Fixed effects in the linear mixed effects regression model of the attribute "up and down" as dependent variable (* : $p < 0.05$; ** : $p < 0.01$; *** : $p < 0.001$). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Conditional adjusted R^2 of the model: $R^2 = 0.625$. SD of the residual is 22.583. SD for the random effect of participant is 13.451.

	Estimate	Std. error	t-value	p-value	
Intercept	-55.276	6.817	-8.326	0.000	***
L	1.122	0.061	18.494	0.000	***
$\log_{10}(f)$	-34.849	1.055	-33.028	0.000	***

significant for this model. Similarly, to some of the previous models, the estimated intercept is not significant for the linear model. To narrow down the standard deviation of the estimates, also a linear mixed-effects model was fitted to this dataset of 30 subjects, which included by-subject random intercepts in addition to the fixed effects of the linear model. The resulting fixed effects of this model are shown in Table 7.29. The higher number of observations lead to a decrease in the standard errors of the fixed effects in comparison to the linear model. All the reported terms are highly significant for the linear mixed-effects regression. Thus, the null hypothesis for all the terms of this model can be rejected.

7.5.7.3 Regression for "Tingling"

As for the stochastic and the AM-sinusoidal excitation pattern, frequency is intuitively associated with the attribute "tingling" as evident from the highly negative correlation with frequency obtained in section 5.4. Similarly, also vibration level affects the rating of "tingling" as implied by Figure 5.17. Again, for "tingling" a simple linear regression model was fitted to the dataset of by-stimulus averages with the selected regressors. However, the R^2 of only

Table 7.30.
Effects in the linear regression model of the attribute “tingling” as dependent variable (* : p < 0.05; ** : p < 0.01; *** : p < 0.001). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Adjusted R^2 of the model: $R^2 = 0.588$.

	Estimate	Std. error	t-value	p-value	
Intercept	-111.453	25.998	-4.287	0.000	***
L	1.050	0.230	4.569	0.000	***
$\log_{10}(f)$	33.930	13.087	2.593	0.019	*
$\log_{10}(f)^2$	-11.767	4.842	-2.430	0.026	*

Table 7.31.
Fixed effects in the linear mixed effects regression model of the attribute “tingling” as dependent variable (* : p < 0.05; ** : p < 0.01; *** : p < 0.001). Reported are the regression coefficient estimates, standard errors, t-values, and p-values. Conditional adjusted R^2 of the model: $R^2 = 0.439$. SD of the residual is 20.877. SD for the random effect of participant is 8.600.

	Estimate	Std. error	t-value	p-value	
Intercept	-111.453	7.746	-14.389	0.000	***
L	1.050	0.067	15.686	0.000	***
$\log_{10}(f)$	33.930	3.812	8.900	0.000	***
$\log_{10}(f)^2$	-11.767	1.411	-8.342	0.000	***

0.475 for the linear model with the terms intercept, level, and frequency, which had a very high p-value, suggested only a mediocre fit. In section 5.4 it was argued that ratings of high frequency describing attributes such as “tingling” have a rating peak at a specific frequency above 35 Hz and falling ratings below and above that frequency for sinusoidal vibration (see Figure 5.17). Therefore, also a quadratic term for frequency was introduced to the model to increase the model fit. The improved model is reported in Table 7.30. A reasonable model fit is implied by the R^2 of 0.588. The terms level and intercept are highly significant, and the linear and quadratic frequency terms are significant. Thus, the null hypothesis for all the four terms of the model can be rejected. To also narrow down the standard deviation of the estimates, a linear mixed-effects model was fitted to this dataset of 28 subjects, which included by-subject random intercepts besides the fixed effects of the linear model. The resulting fixed effects of the model are shown in Table 7.31. The higher number of observations lead to a large decrease in the standard errors of the fixed effects. Thus, the all terms are highly significant for this linear mixed-effects regression model.

7.5.7.4 Synthesis Equations

The linear models provide the regression equations for the prediction of the attribute ratings of "weak" r_w, "up and down" r_d, and tingling r_t for the sinusoidal excitation pattern. Since the frequency was log-transformed before the regression, the logarithm of frequency instead of frequency are present in the equations. These equations are forming the following equation system:

$$\begin{array}{lcllll} r_w & = & -2.07L & & +26.601\log_{10}(f) & +243.078 \quad (I) \\ r_d & = & 1.122L & & -34.849\log_{10}(f) & -55.276 \quad (II) \\ r_t & = & 1.05L & -11.767\log_{10}(f)^2 & +33.93\log_{10}(f) & -111.453 \quad (III) \end{array} \tag{7.29.}$$

For this excitation pattern, there are three equations but only the two unknowns Level L and frequency f, thus producing an overdetermined equation system, too. Three equation systems of two equations each can be formed to find an analytical solution also for this excitation pattern. MATLAB was again utilized to solve each resulting equation system.
Four solutions of synthesis equations were obtained from the equation system for the sinusoidal excitation pattern. The first equation system produces the following synthesis equations for the physical parameters of the sinusoidal excitation pattern depending on the ratings of "weak" and "up and down":

$$\begin{array}{lcccc} L & = & -0.824r_w & -0.629r_d & +165.0 \\ \log_{10}(f) & = & -0.0265r_w & -0.0489r_d & +3.74 \end{array} \tag{7.30.}$$

This synthesis equations predict the vibration level with only an R^2 of 0.21 and the frequency with only an R^2 of 0.02 and are thus discarded.
The second equation system produces synthesis equations depending on the ratings of "weak" and "tingling". There are two solutions to this equation system:

$$
\begin{aligned}
L_1 &= 6.42\sqrt{20.3 - 0.172r_w - 0.34r_t} && -0.483r_w && +143.0 \\
\log_{10}(f_1) &= 0.5\sqrt{20.3 - 0.172r_w - 0.34r_t} && && +2.01 \\
L_2 &= -6.42\sqrt{20.3 - 0.172r_w - 0.34r_t} && -0.483r_w && +143.0 \\
\log_{10}(f_2) &= -0.5\sqrt{20.3 - 0.172r_w - 0.34r_t} && && +2.01
\end{aligned}
$$

7.31.

However, these synthesis equations are only valid under the following condition:

$$20.3 - 0.172r_w - 0.34r_t > 0$$ **7.32.**

Therefore, the domain of these synthesis equations is limited to approximately 34 % of the two-dimensional perceptual subspace of the attributes “weak” and “tingling”. The calculation of R^2 was adapted, to include only the 17 of 21 vibration-rating profile pairs which fulfill the condition of the synthesis equations. For the first solution, the synthesis equations predict level as well as the frequency with an R^2 below zero, i.e. the synthesis is less accurate than simply utilizing the average level or frequency. Therefore, also these synthesis equations need to be discarded. For the second solution, the synthesis equations predict level with an R^2 of 0.502 and frequency with an R^2 of 0.469.

The third equation system produces synthesis equations depending on the ratings of “up and down” and “tingling”:

$$
\begin{aligned}
L &= -15.5\sqrt{0.318r_d - 0.34r_t + 11.7} && +0.892r_d && +137.0 \\
\log_{10}(f) &= -0.5\sqrt{0.318r_d - 0.34r_t + 11.7} && && +2.83
\end{aligned}
$$

7.33.

However, again these synthesis equations are only valid under the following condition:

$$0.318r_d - 0.34r_t + 11.7 > 0$$ **7.34.**

Therefore, the domain of these synthesis equations is limited to approximately 88 % of the two-dimensional perceptual subspace of the attributes “up and down” and “tingling”. Again, the calculation of R^2 was adapted, to include only the 16 of 21 vibration-rating profile pairs which fulfill the condition of the synthesis equations. The five omitted vibrations are at the upper end of the frequency range of the upper end of the level range and represent extrema of vibration exposure that are unlikely to be encountered. For the second solution, the synthesis equations predict level with an R^2 of 0.406 and frequency with an R^2 of 0.473. The domain of the synthesis equations depending on “up and down” and “tingling” is far larger than the domain of the synthesis equations depending on “weak” and “tingling”. For the former synthesis equations, almost the complete range of perceptual attribute ratings can be utilized as a valid input, which makes them preferable.

7.5.8 Implementation of the Synthesis

The algorithm for the regression model based synthesis outlined in this section was also implemented in MATLAB. As for the 1-NN model, the MATLAB implementation is structured according to the model boundaries shown in Figure 7.2 of section 7.1. It consists of three functions for vibration parameter estimation, vibration signal generation, and multimodal scene construction. Only the vibration parameter estimation was changed in comparison to the 1-NN model, while the vibration signal generation and multimodal scene construction are identical. The GUI for the synthesis was extended to allow the switching of the vibration parameter estimation method.

7.5.8.1 Input

The input format of the synthesis is identical to the input format of the synthesis by the 1-NN model described in section 7.4.

7.5.8.2 Translation

After the determination of the excitation pattern, the vibration parameters are estimated with the syntheses equations from the expected rating profile. For

Table 7.32.
Vibration parameter estimates of the regression model for the 19 vehicle scenes.

No.	Category	Speed (kph)	Surface	L (dB)	f (Hz)	f_b (Hz)	f_m (Hz)	α
1	non-transient	5 to 50	small cobblestone	117	211	420		
2	non-transient	30	small cobblestone	117	12			
3	non-transient	50	small cobblestone	117	258	513		
4	non-transient	5 to 50	cobblestone	131	65	128		
5	non-transient	30	cobblestone	129	79	156		
6	non-transient	50	cobblestone	130	122	242		
7	non-transient	50	tarmac (A-Road)	108	17		2	
8	non-transient	70	tarmac (A-Road)	105	179		9	
9	non-transient	5 to 50	Tarmac	107	22			
10	non-transient	30	Tarmac	106	19			
11	non-transient	50	Tarmac	103	21			
12	non-transient	100	concrete motorway	110	145		8	
13	transient	100	surface change	108	22			
14	transient	40	tram tracks	122	50	97		
15	transient	50	expansion joint	119	3			9
16	transient	100	expansion joint	113	37			
17	transient	30	manhole cover cobble-stone	130	379	756		
18	transient	30	manhole cover tarmac	109	11			
19	transient	50	manhole cover tarmac	106	119	235		

the impulse-like excitation pattern, the synthesis equations are assessed in 7.5.4, for the bandlimited WGN excitation pattern in 7.5.5, for the AM-sinusoidal excitation pattern in 7.5.6 and the sinusoidal excitation pattern in 7.5.7. The attribute ratings elicited by the vibration generated from these equations will maximize the similarity to the expected attribute ratings and are thus most likely plausible. An overview of the vibration parameter estimates the output of the regression synthesis model for the 19 vehicle scenes are shown in Table 7.32. The subsequent vibration signal generation from the parameter estimates and the multimodal scene construction are identical to the 1-NN implementation.

7.5.8.3 Output

For the subsequent model validation, 19 audio-visual scenes with the synthesized WBV were generated also based on these estimates. The spectra of these vibrations in comparison to the the recorded vibrations are shown in Figure 7.11.

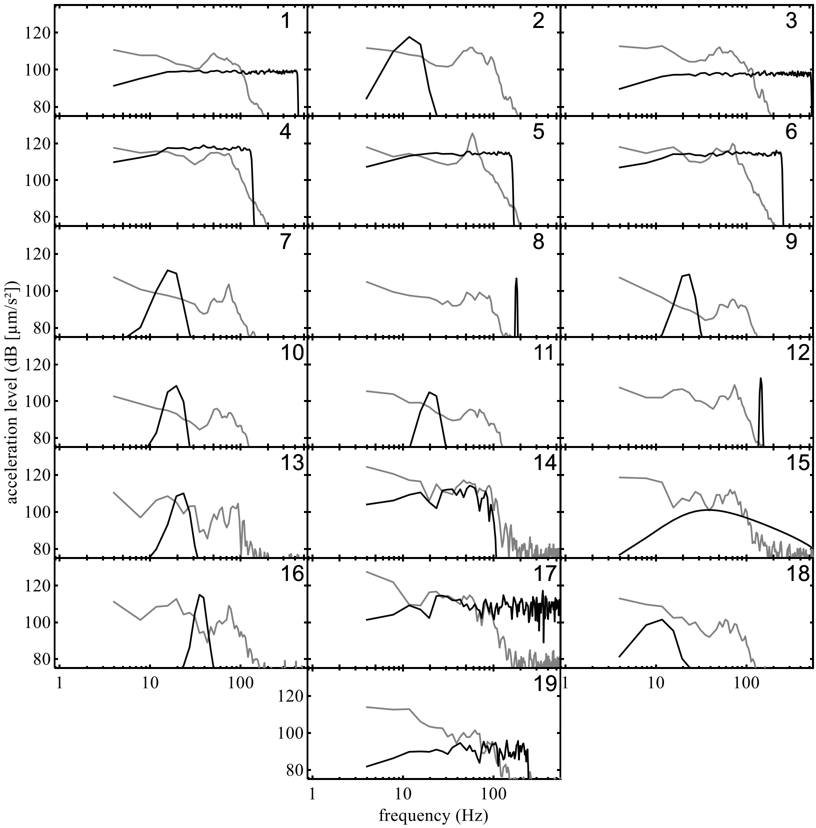

Figure 7.11.
Spectra (FFT, 4096 samples, 50 % overlapping Hann windows) of the recorded vibration (grey) and the vibration generated from the parameter estimates of the regression model (black) for each of the 19 vehicle scenes.

7.5.9 Advantages and Disadvantages of the Approach

The regression model approach has advantages as well as disadvantages, e. g. compared to the 1-NN approach. One of the main goals of this section was to improve on the shortcomings of the previously presented 1-NN model. The 1-NN approach only allows for discrete predictions, since it only selects the perceptually most similar vibration-attribute rating pair amongst a finite set of vibration items of a previously assembled database. Each new vibration items need to be manually profiled, to extend the vibration items potentially output by the synthesis. This is also the shortcoming of the previously studied, similar approach by [60]. Tuning the vibration items in the database can only help to partially overcome this shortcoming, since the physical parameters affect the perceptual attributes differently depending on the starting point vibration for the tuning [12] and thus do not represent a universal mapping between physical and perceptual domain. In contrast to these approaches, the presented regression model approach enables a continuous, universal mapping between the physical domain and the tactile sensory-perceptual domain. The model provides quasi-continuous predictions of physical parameters from perceptual attribute ratings, i.e. continuous predictions for each of the four generalized excitation patterns. In contrast to database approaches, the prediction performance is not limited by the range and resolution of rating profiles and their associated vibrations and thus supports much more flexible vibration generation.

Another disadvantage of the 1-NN prediction is its reliance on an implicit relationship between physical vibration parameters and their associated attribute rating profiles that are contained in the vibration-rating profile pairs in the database. Since the vibration stimuli of chapter 5 were systematically constructed by varying the characteristic parameters of the four excitation patterns, it was possible to observe physical parameter vs. rating curves. Based on these observations the regression models could be created that generalize the implicit relationship into an explicit relationship. These models provide insights on the explicit relationship between physical vibration parameters and sensory tactile perceptual attribute ratings in the form of synthesis equations. Such relationships enable the designer to draw inferences about e.g. the required changes in vibration level to match the higher expected "weak" sensation.

Furthermore, the sample size of subjects providing perceptual attribute ratings is always limited by time and costs, the true population average of attribute ratings elicited by a vibration can only be assessed with limited accuracy. The regression models take the measurement error of attribute ratings into account and thus minimize the bias of predictions. Furthermore, regression models provided the opportunity for statistical tests of the predictors of the models. Statistically significant predictors suggest that the coefficient estimates of the predictors based on the subjects' ratings of this sample are unlikely to have arisen by pure chance. Therefore, the utilization of spurious relationships for the prediction can be avoided.

Finally, the goodness of fit statistics of the regression models provide cues about the model error and thus would enable comparisons of the performance of different models. However, the level of acceptable error needs to be judged in a separate validation study, which assesses the plausibility elicited by the vibration produced by the regression synthesis model.

One of the main disadvantages of the inverted regression approach is that the model complexity and thus potentially the model fit is constrained by the analytic resolvability of the inversion problem. Therefore, the fit of the synthesis equations of the model cannot easily be improved while following the current approach. However, any $R^2 > 0$ shows an improvement compared to the zero model that simply predicts the mean physical parameter value. Whether the error is acceptable should be evaluated from the plausibility judgments elicited by the predicted vibration.

Another disadvantage compared to the 1-NN approach is that the domain of the function might be limited. For the 1-NN synthesis model, any rating profile will lead to the estimation of a valid physical parameter set. Since the ratings of some attributes are at least partially correlated, contradictory attribute ratings would be possible from a theoretical standpoint. Thus, extreme combinations of attribute ratings might fall out of the domain of the synthesis equation. From a practical standpoint, laypersons should be aware that a vibration perceived as "slightly weak" (or as very intensive) is implicitly usually either "very up and down" or "very tingling" at the same time

Similarly, the synthesis equations might output parameter estimates, which do not enable the synthesis of vibrations according to the equations of an

excitation pattern because of an inherent dependency, e.g. center frequency and bandwidth for the bandlimited WGN excitation pattern. However, such a case is likely caused by prediction error and thus it should be possible to correct the estimates into a valid combination without shifting elicited attribute ratings away from the expected ratings.
A more general disadvantage lies in abstraction of everyday life vibration into four separate excitation pattern that are each characterized by a different set of vibration parameters. A unified definition of excitation patterns might enable not only a quasi-continuous vibration parameter prediction but a truly continuous prediction. However, if vibrations of different excitation patterns elicit overlapping rating profiles for specific parameter combinations, then a gradual shift from one excitation pattern to the other excitation pattern would likely not be clearly perceivable.

7.6 Validation of the Synthesis Models

In section 7.4 and section 7.5 two synthesis models were created. These models require an input in the form of an expected rating profile consisting of ratings of the six sensory tactile perceptual attributes representing the sensory tactile perceptual space (see chapter 5). The expected rating profiles for a set of scenes representing the range of everyday life exposure to WBV was obtained in chapter 6. Both models can predict a set of vibration parameters based on which WBV can be synthesized.
The goal of this section is to validate the two synthesis models and the vibration which is generated from their estimates based on the ratings of the six sensory-perceptual attributes forming the design language. Since a completeness of the tactile design language could not be determined in chapter 5 directly, it will be determined here indirectly. If the attribute set is complete then it should be possible to obtain ratings only of these attributes for vibrations of any situational context and translate the ratings into vibration. Such synthesized vibration and recorded vibration should both elicit very similar perceptual attribute ratings. If they are perceived as similarly plausible, then the attribute set is complete with regards to describing the sensory-perceptual

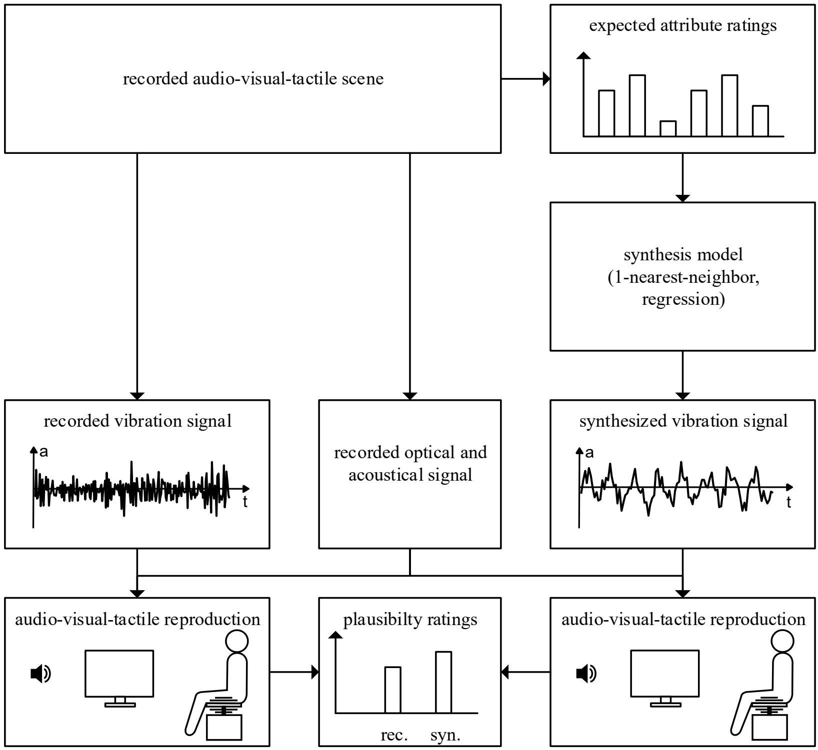

Figure 7.12.
Overview of the model validation procedure.

properties of vibration. Together this would allow inferences on whether the synthesis approach of this work in suitable in general.
To conduct this validation, the general method shown in Figure 7.12 was followed. Obviously, the perceived plausibility of the synthesized vibrations in the situational context of their audio-visual scenes needs to be obtained in a perceptual study. To support the interpretation of the plausibility judgments of the synthesized vibration also plausibility judgments of the recorded vibration in the situational context of their audio-visual scenes should be obtained. By comparing the perceived plausibility of the synthesized vibrations to the perceived plausibility of the recorded vibrations inferences about the success or failure of the approach can be derived.

7.6.1 Stimuli

The stimuli of the validation study were multimodal vehicle scenes according to section 6.1 consisting of recorded video, recorded audio, and recorded or synthesized WBV. Multimodal scenes with synthesized vibration were generated according to the synthesis models described in the previous sections. For the 1-NN model, the 19 multimodal vehicle scenes were generated as described in section 7.4. Similarly, for the regression model, the 19 multimodal vehicle scenes were generated as described in section 7.5.
For comparison, 19 multimodal vehicle scenes with recorded vibration were recorded in a vehicle as described in section 6.1. The recorded vibration of the 12 non-transient scenes were left unchanged. For the 7 transient scenes, a rating profile was obtained in section 6.3 for the duration of the single event contained in these scenes and not before or after the event. Subsequently, vibration was synthesized for the duration of the single event, not for the total duration of the scene. To make the transient scenes with recorded vibration comparable to the scenes with synthesized vibration, the recorded vibration should only be present for the duration of the single event. If recorded vibration was present for the total duration of the scene but not for the synthesized vibration, it might bias the results towards the scenes with recorded vibration. Therefore, the recorded vibration was set to zero and only faded in 300 ms before the impulse-like event and faded out 300 ms after it for each of the 7 transient scenes. All audio-visual-tactile vehicle scenes were presented with the multimodal reproduction system described in chapter 4.

7.6.2 Experimental Design

To assess the elicitation of the plausibility illusion, a suitable plausibility measure needed to be selected. One possibility would be a procedure involving a paired comparison, i.e. presenting the scene with recorded vibration and subsequently presenting the scene with synthesized vibration. If the authenticity was to be assessed, this would be a preferred measure, since the perceived similarity between the synthesized vibration and the vibration originally occurring in the context of the scene is compared [74]. However, in the majority of use cases for the virtual environment, the original vibration is not available for such a comparison. Instead, the user is typically forced to rely on his expectations on the situational context to judge the plausibility of the virtual environment [77]. To obtain results that are representative for this use case, recorded and synthesized vibration should not be presented in short succession to suppress the possibility of a direct comparison between recorded and synthesized vibration. Instead, the plausibility of the synthesized and recorded vibrations should be assessed separately requiring absolute plausibility judgments.

Therefore, the multimodal scenes were split into three separate blocks to impede A/B comparison between recorded and synthesized vibration. The first block contained the 19 scenes with vibration synthesized from the 1-NN classifier estimates. The second block contained the 19 scenes with vibration synthesized from the regression model estimates. The 19 scenes with recorded vibration were presented in the last block to prevent subjects from comparing them to scenes with synthesized vibration and instead forcing them to rely on their expectations as a reference. Participants were oblivious to the type of vibration presented. All stimuli of a block were presented in random order. Participants were allowed to repeat the stimulus, but rarely did so. The rating time of the plausibility of each stimulus was not limited, but typically took only a few seconds resulting in a total block duration of 10 minutes.

In section 2.2.2 it was argued that plausibility can be assessed with questionnaires inquiring the perceived plausibility explicitly. Therefore, the perceived plausibility of the presented vibration in the context of the audio-visual-tactile scene was rated on a quasi-continuous Rohrmann scale as in section 5.4 implemented as a MATLAB graphical user interface (see Figure 5.11).

For the non-transient scenes, participants were instructed to rate the plausibility of the presented vibrations in the context of the audio-visual scene. For the transient scenes, participants were presented the complete audio-visual scene containing only vibration of the single event as described in the stimuli section. They were instructed to rate the plausibility of the presented vibrations of the single events. A “rate now” subtitle was shown for the duration of the single event to prevent misunderstandings about the segment to be rated.

7.6.3 Participants

A total of 22 German native laypersons (13 male, 9 female) with an average age of 30 years (19 to 61 years) took part in the experiment. To avoid interference to the synthesized vibrations participants were unfamiliar with multimodal scenes before the experiment. The study was conducted with the understanding and written consent of each participant.

7.6.4 Results

In each of the three blocks, 19 plausibility ratings were obtained for each of the 22 participants. The ratings of vibrations of the 1-NN model and the regression model are compared to the ratings of the recorded vibration separately.

7.6.4.1 1-Nearest-Neighbor Model

The mean plausibility ratings and 95 % confidence intervals of the vibrations synthesized by the 1-NN model in comparison to the ratings of recorded vibrations are shown in Figure 7.13. The plausibility ratings of different scenes span the range from moderately plausible (44 points) to between very and extremely plausible (86 points). Scenes with cobblestone roads are rated slightly higher than scenes with tarmac or concrete road or transient scenes. It is obvious that if the rating pairs of the scenes are compared that recorded and synthesized vibrations are perceived as similarly plausible in the context of their respective audio-visual vehicle scenes. The rating difference based on estimated marginal means between recorded vibration and synthesized vibration overall scenes is at only 5 points on the 100-point scale. The plausibility difference between recorded and synthesized vibration ranges from -16 points to 18 points on the 100-point plausibility scale, suggesting that some synthesized vibrations are perceived as more plausible than their recorded counterparts in the context of the audio-visual scene. Six of the 19 scenes with synthesized vibrations are rated as slightly more plausible than their corresponding counterparts with recorded vibrations. For the driving on coarse cobblestone at 50 kph scene, the driving on tarmac at 50 kph scene and the

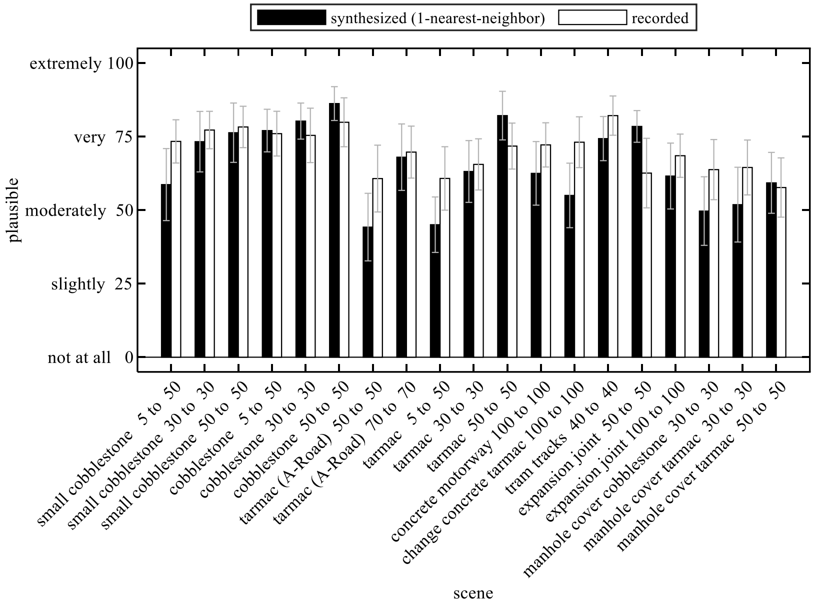

Figure 7.13.
Mean plausibility ratings and 95 % confidence intervals of recorded vibration vs. vibration synthesized by the 1-NN model in the context of their respective audio-visual vehicle scenes.

crossing of an extension joint at 50 kph differences of up to 16 points can be found. For 13 scenes the synthesized vibration is less plausible than the recorded vibration. For two of the three scenes, in which the speed is increasing, this difference is above average.

These findings were assessed with a two-way repeated-measures ANOVA with factor scene and model (recorded vs. synthesized) to allow for statistical inferences. The factor scene showed a highly significant effect ($F(7.432,156.070) = 9.818$, $p < 0.001$). The Greenhouse–Geisser adjustment was used to correct for violations of sphericity. In contrast to the factor scene, the factor model did not show a significant effect ($F(1,21) = 2.413$, $p = 0.135$) on plausibility rating. The interaction of factor model and factor scene showed a highly significant effect ($F(8.213,172.480) = 3.758$, $p < 0.001$).

From the lack of a significant effect of the factor model, an absence of any effect of model for very large sample sizes cannot be inferred. Thus, besides the significance of the factor model, the unstandardized effect size of the difference between recorded and synthesized vibration was also assessed as a measure of similarity between the plausibility ratings of synthesized and recorded vibration. An unstandardized effect size in the units of measurement i.e. plausibility rating points was preferred since it provides a direct measure of difference. A pairwise contrast between recorded vibration and synthesized vibration revealed a mean difference of 5 points. The 95 % confidence interval of the difference ranges from -2 to 11 points difference. Therefore, the mean difference between presentation modes is unlikely to exceed half of the 25-point scale tick interval or about 10 % of the 100-point scale. These findings suggest that this plausibility difference is likely irrelevant for practical purposes.

7.6.4.2 Regression Model

The mean plausibility ratings and 95 % confidence intervals of the vibrations synthesized by the regression model in comparison to the ratings of recorded vibrations are shown in Figure 7.14. The plausibility ratings of different scenes span the range from moderately plausible (37 points) to between very and extremely plausible (86 points). Again, it is obvious that if the rating pairs of the scenes are compared that recorded and synthesized vibrations are perceived as similarly plausible in the context of their respective audio-visual vehicle scenes. The rating difference based on estimated marginal means between recorded vibrations and synthesized vibrations overall scenes is at 4 points on the 100-point scale, which is slightly lower than for the 1-NN model.

Also for this model, some synthesized vibrations are perceived as up to 10 points more plausible than their recorded counterparts in the context of the audio-visual scene. Seven of the 19 scenes with synthesized vibration are rated as more plausible than their corresponding counterparts with recorded vibration. Similarly, to the 1-NN model, the synthesized vibrations of the driving on coarse cobblestone at 50 kph scene and the crossing of an extension joint at 50 kph differences are perceived as slightly more plausible. For the remaining 12 scenes, the synthesized vibrations are slightly less plausible than the recorded vibrations. The driving over a manhole embedded in a cobblestone road is one outlier with a 37 points difference between recorded and

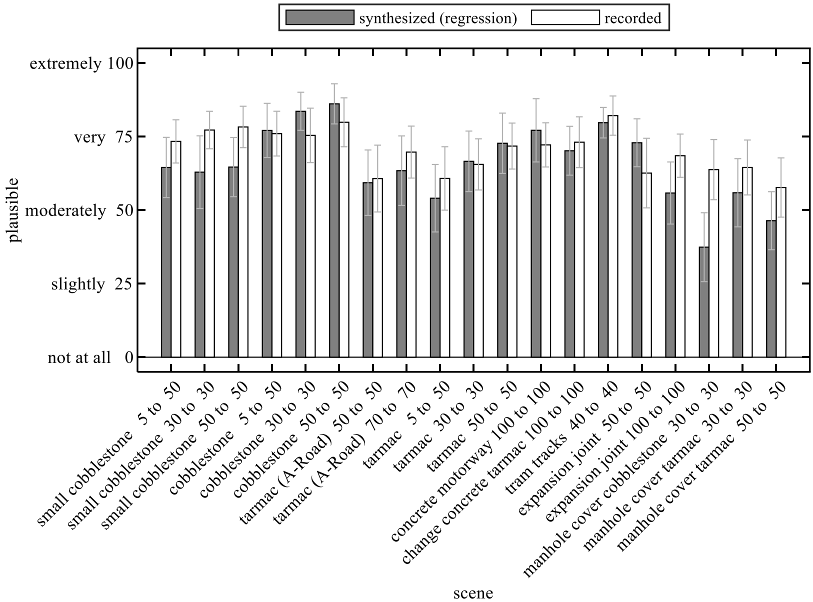

Figure 7.14.
Mean plausibility ratings and 95 % confidence intervals of recorded vibration vs. vibration synthesized by the regression model in the context of their respective audio-visual vehicle scenes.

synthesized vibration, with the next largest difference being 14 points. Also for this comparison, a two-way repeated-measures ANOVA with factor scene and model (recorded vs. synthesized) was conducted to enable statistical inferences. The Greenhouse–Geisser adjustment was used to correct for violations of sphericity. Again, the factor scene showed a highly significant effect ($F(7.729,162.303) = 11.216$, $p < 0.001$). The factor model did not show a significant effect ($F(1,21) = 3.832$, $p = 0.064$) on plausibility rating for the regression model vs. recorded, too. The interaction of factor model and factor scene showed a significant effect ($F(7.851,164.876) = 3.151$, $p < 0.05$).

A pairwise contrast between recorded vibration and synthesized vibration revealed a mean difference of 4 points. The 95 % confidence interval of the difference ranges from 0 to 9 points difference. As for the nearest neighbor model, the mean difference between presentation modes is thus unlikely to exceed half of the 25-point scale tick interval or about 10 % of the 100-point scale. These findings confirm again that this plausibility difference is likely irrelevant for practical purposes.

Based on the expected sensory tactile perceptual properties of a situational context quantified with the suggested sensory tactile design language, vibration parameters can be estimated with the 1-NN model as well as with the regression model. The results of the study validate that vibrations generated based on these estimates is perceived as quasi equally plausible as vibrations originally occurring in this situation in a virtual environment, in which optical and acoustic recordings of the situation are simultaneously presented. A non-significant influence of the factor model (recorded vs. synthesized) in combination with a significant interaction effect between model and scene suggests a disordinal or crossover interaction. This type of interaction means that the effect of model (recorded vs. synthesized) on plausibility in one scene is different from the effect of model on plausibility in another scene. Thus, while there is no significant effect of model overall, for some scenes synthesized vibration tends to be more plausible but for other scenes recorded vibration tends to be more plausible.

As argued in section 2.2.2 the perception of contradictory stimuli in different modalities is dominated by the most convincing cue for the context [9]. This suggests that the expected sensory tactile perceptual properties of a scene might be different from the elicited properties Since vibration is synthesized directly from the expected sensory tactile perceptual properties, the properties elicited by the synthesized vibration are identical to the expected properties for an ideal synthesis model. If the perceived plausibility is proportional to the inverse Euclidean distance between expected sensory tactile perceptual properties and the elicited properties, it would explain the surprising observation of synthesized vibration being perceived as more plausible than recorded vibration. E.g. for the case of cobblestone road at 30 kph loud noise and visually perceived rough road might have influenced the expected sensory tactile perceptual properties from which vibration was synthesized that was perceived as more plausible than the recorded vibration. This would also explain why the vibrations in a basketball game were perceived as more plausible [7] despite them not being noticeable from the audience in the real environment.

For most of the acceleration scenes, the plausibility of the synthesized vibration was lower than of the recorded vibration. As shown in section 6.1 the physical vibration of such scenes is changing slightly. Since only an overall rating profile was obtained for each scene, it might not be completely representative of the expected profile at the beginning and the end of the acceleration scene. Therefore, a vibration is synthesized that does not elicit the expected sensory tactile perceptual properties which is thus perceived as less plausible. Such scenes with slow changes of the sensory tactile perceptual properties, should be split in a set of successive segments with quasi constant perceptual properties, as discussed in section 3.2. However, the focus of this study was set on scene segments that can be represented by quasi constant perceptual properties. The large decrease in the plausibility of the synthesized vibrations of driving over a manhole embedded in a cobblestone road might be explainable by the model estimation falling out of the range ($f = 379$ Hz, $f_b = 756$, see section 5.2.2) for which vibration-rating profile pairs were assessed. Since the focus of the validation study was to produce vibrations which are maximally plausible, it might be argued that any vibration might have been as plausible as the presented vibration. However, the findings of [80] contradict such a hypothesis, since the increase or decrease of the acceleration level significantly decreased perceived plausibility.

7.6.5 Summary

It can be concluded that the 1-NN model and the regression model produce vibration which are overall perceived as almost equally plausible as recorded vibration. Therefore, both models provide satisfactory performance in producing plausible WBV for virtual environments for the representative scene set and are thus validated.

7.7 Summary and Discussion

7.7.1 Summary

First, the plausibility judgment was formalized as a similarity judgment in a perceptual feature space. The six sensory-perceptual attributes of the design language (“weak”, “up and down”, “tingling”, “repetitive”, “uniform”, and “fading”) were assumed to represent the sensory tactile perceptual space. It was hypothesized that the perceived plausibility of vibration in a situational context is proportional to the inverse Euclidean distance between the expected sensory tactile perceptual properties of this context and the sensory tactile perceptual properties elicited by these vibrations. For example, the better the elicited properties (e.g. “up and down”) match the expected properties of WBV in a scene (e.g. “driving on a cobblestone road”) the more plausible it will be perceived. Since the expected properties are quantified in the form of a rating profile consisting of ratings of the six sensory-perceptual attributes of the design language, the synthesis model should estimate parameters of vibration that would elicit a rating profile which has a minimal Euclidean distance to the expected rating profile and is thus perceived as maximally plausible in the corresponding situational context. This formalization was the basis for the subsequent models, which estimate vibration parameters from the expected sensory tactile perceptual properties thus translating the user expectations into vibration.

The data basis for the models were the vibration-rating profile pairs assessed in section 5.4. Only two acceleration levels were utilized in this dataset, which can be problematic as small deviations in vibration level significantly affect plausibility. Previous findings with a magnitude estimation experimental design suggested a linear relationship between vibration level and perceived intensity for WBV. An approximately linear influence of vibration level on the absolute ratings of the six sensory tactile perceptual attributes was confirmed in the perceptual experiment with four vibration levels. Thus acceleration level was assumed to have a linear influence on attribute ratings for the subsequent modeling.

The results of section 5.4 suggested that the AM-sinusoidal excitation pattern and the multiple successive impulse-like excitation pattern can produce very similar attribute ratings. Since the utilized RMS acceleration values, modulation frequencies, or repetition rates were not identical for both the excitation patterns in this previous experiment, AM-sinusoidal vibrations were generated to match multiple successive impulse-like vibrations according to these parameters. A perceptual study, in which the six sensory tactile perceptual attributes were rated, showed that these excitation patterns produce very similar attribute ratings if these parameter values are identical. If one rating profile maps onto multiple excitation patterns, additional criteria for selecting one excitation pattern would need to be defined for a synthesis model. However, if two vibrations elicit the same rating profile, they would be perceptually equivalent under the assumptions of the six sensory-perceptual attributes sufficiently representing the sensory tactile perceptual space. Thus, the multiple successive impulse-like excitation pattern was omitted from the subsequent modeling since it can be substituted by the AM sinusoidal excitation pattern.

Building onto the formalization of plausibility as being proportional to the inverse Euclidean distance in the sensory tactile perceptual space, a 1-NN classifier was created. For a given rating profile consisting of ratings of the six sensory-perceptual attributes, the classifier determines the rating profile with the smallest Euclidean distance in the database of vibration-rating profile pairs obtained in section 5.4 with generalized vibration of four excitation patterns. This database was extended from 91 pairs to 387 pairs by interpolating additional rating profiles for vibration with levels between the low and high SL utilized in section 5.4. The parameters (sensation level SL, vibration (carrier- or resonance-) frequency f, bandwidth f_b, modulation frequency f_m, and decay constant α) of the vibration associated with this perceptually most similar rating profile represent the estimates of the model. The relevant excitation pattern for the vibration synthesis is implicitly contained in the vibration parameters.

One of the disadvantages of the 1-NN models is the vibration parameter estimates being discrete. Therefore, also a regression-based model was created, which produces quasi-continuous estimates. However, due to the abstraction of everyday life WBV into four distinct excitation patterns (sinusoidal, AM-sinusoidal, bandlimited WGN, and impulse like, see section 0) without unified parameters, the creation of one single regression model covering all excitation patterns is not possible. Thus, a set of three SVMs was created based on the vibration-rating profile pairs obtained in section 5.4 to enable the selection of one excitation pattern based on the input rating profile. For each excitation pattern the two to three sensory tactile perceptual attributes, which ratings were most affected by the defining vibration parameters of this excitation pattern, were selected. Subsequently, two to three regression equations

of the relationships between the defining vibration parameters and each selected attribute was modeled also based on the vibration-rating profile pairs obtained in section 5.4. The equation system formed by the two to three regression equations was solved by the defining vibration parameters of the excitation pattern. This produced two to three synthesis equations for estimating the defining vibration parameters of each excitation pattern. Vibration can be synthesized based on the parameter estimates of both models.
Subsequently, the synthesis models were validated in a user study. In section 6.1 a representative set of 19 audio-visual-tactile vehicle scenes with WBV was selected and recorded. These scenes were presented in a virtual environment to obtain the expected rating profiles (see section 6.3), which formed the input for the synthesis models. Based on the parameter estimates, output by the two synthesis models, vibration was synthesized. The recorded vibration of the 19 audio-visual-tactile scenes was replaced by the recorded vibration of both models, forming three sets of 19 scenes. In a perceptual study, each set of scenes was presented separately and the perceived plausibility of the WBVs of each scene was rated on a quasi-continuous rating scale. The results suggested that vibration generated based on these estimates is perceived as quasi equally plausible as vibration originally occurring in this situation in a virtual environment, in which optical and acoustic recordings of the situation are simultaneously presented. Therefore, the 1-NN model and the regression model provide satisfactory performance in producing plausible WBV for virtual environments for the representative scene set and are thus validated.

7.7.2 Discussion

In the process of identifying the sensory tactile design language in section 5.4, the PCA suggested that six groups of attributes are necessary to explain the majority of variance in attribute ratings observed. It was left open whether the six selected sensory tactile perceptual attributes (“weak”, “up and down”, “tingling”, “repetitive”, “uniform”, and “fading”) selected to represent these six groups sufficiently represent the sensory tactile perceptual space. The validation demonstrated that vibration recorded in a situational context can be perceptually profiled with these six attributes and translated back into a vibration that is perceptually quasi equivalent to the recorded vibration in this situational context. Since the validation scenes were selected to represent typical everyday life WBV exposure, this likely implies that these six attributes are completely i.e. sufficiently describing the sensory-perceptual properties of WBV.

For the auditory perception of room properties [77] showed that perceived plausibility of room reflections is proportional to the inverse distance between expected and elicited properties in the underlying perceptual space. The distance of two stimuli in the perceptual space was determined by rating their similarity. However, no explicit attributes were assessed to represent the auditory perceptual space of room properties. Therefore, the expected perceptual features could not be quantified explicitly and no synthesis of sounds could be conducted based on the proposed relationship between perceived plausibility the inverse Euclidean distance in the perceptual space. For the tactile perception, six sensory tactile perceptual attributes were assessed to represent the sensory tactile perceptual space. This enabled the explicit quantification of the expected sensory tactile properties of WBVs for a situational context in the form of rating profiles consisting of the six sensory-perceptual attributes. The synthesis models synthesize vibration which elicits a rating profile with minimal Euclidean distance to the rating profile expected in the situational context. Since the vibration recorded in the situational context and the synthesized vibration are perceived as quasi equally plausible, the proposed relationship between perceived plausibility and the inverse Euclidean distance in the perceptual space is likely applicable to the tactile domain. Furthermore, it suggests that this relationship is not only of theoretical value for the understanding of the plausibility illusion. Instead, it can also be practically utilized for the synthesis of plausible WBVs for virtual environments. Overall, this confirms the general method for synthesizing plausible WBV for virtual environments proposed in this thesis.

Previous approaches to user perception centered tactile design relied on libraries of discrete vibration effects [55], [59], [60]. Assembling one vibration for each potential situational context [55] requires extreme effort due to each small change in situational context requiring a new vibration effect, resulting in the necessity to collect a practically infinite number of effects.

However, if two vibrations of two different situational contexts elicit the same sensory-perceptual properties, then it is only necessary to include one of these vibrations and its elicited sensory tactile perceptual properties into the effects library. This enables the selection of a vibration according to perceptual properties instead of the selection by situational context, as demonstrated by [59], [60]. This approach is still inflexible because it only allows the selection of the perceptually most similar effect amongst a finite set of discrete vibration effects of a previously assembled into the library. Unfortunately, it requires a lot of effort because it necessitates manual perceptual profiling of each new vibration effect. For the design of scenes with continuously changing perceptual profiles, it will potentially result in artifacts at the transitions from a previous effect to a successive effect.

One solution to overcoming this limitation is the tuning of vibration effects in the library [12]. However, since each vibration effect demonstrated different influences of physical vibration tuning parameters onto perceptual attributes elicited by the vibration effect, a specific tuning curve would need to be assessed for each vibration effect. Furthermore, if the transfer function of the reproduction system is not compensated as in [59], [60], the observed mappings are difficult to generalize to other reproduction systems.

Instead of investigating the mapping between physical vibration parameters and changes in perceptual attribute ratings of each vibration effect in a library as necessary for the tuning approach, it is more efficient to investigate the relationship between physical vibration parameters and perceptual attribute ratings. The generalization of vibration into excitation patterns enabled the controlled variation of physical vibration parameters and thus provided the basis for obtaining universal and continuous mappings, i.e. systematic relationships. The created regression models approximate these relationships. The coefficients of the physical parameters of these models are all significant and thus the null hypothesis that no relationship exists between physical parameters and attribute ratings can be rejected. The presented models enable quasi-continuous predictions, i.e. continuous predictions for each excitation pattern. Thus, the synthesized vibration is not limited to a pool of discrete vibration effects contained in a library constrained in range or resolution which provides a much more flexible vibration generation. This flexibility is beneficial for a future extension from a synthesis for short scenes of quasi-constant sensory tactile attribute rating profiles towards a synthesis for longer scenes containing multiple successive segments with different rating profiles, since the continuous predictions would minimize artifacts at the transitions between segments. By compensating the transfer function of the utilized reproduction system, the presented synthesis equations can be easily utilized with other calibrated reproduction systems. Thus, design engineers can use these models to translate expected sensory tactile perceptual attribute ratings into vibrations that ensure the elicitation of these attribute ratings in users if exposed to these vibrations.

A shortcoming of linear regression is that it relationships between independent and dependent variable are approximated as linear predictor functions. In contrast, artificial neural networks, such as multilayer perceptrons, enable universal function approximation and could potentially produce a model with a higher goodness of fit and thus better prediction performance [112]. Multilayer perceptrons usually consist of an input layer consisting of n input neurons for each model input and an output layer consisting of m output neurons for each model output. Between the input layer and the output layer, there are zero to multiple hidden layers. Neurons from adjacent layers are connected via directed edges, which have associated weights each.

Multilayer perceptrons require a sufficient sample size for training to produce a model that properly generalizes the relationship of interest [122]. Generalization refers to the proper input-output mapping of the trained network being approximately correct also for new data not used for training. Based on [123], a recommendation for sufficient sample sizes N is provided by [122] as $N > \frac{W}{e}$, where W are the total number of input weights and e is the error rate. Assuming six input neurons for six perceptual attributes each linked to each of the five output neurons for five physical vibration parameters, the estimation of at least 30 weights for the connections to the output neurons would be required. This approximation ignores potential additional hidden layers. For an error rate of 10 %, 300 training cases would be recommended to avoid overfitting and to enable sufficient generalization of the relationship of interest in the model. However, only 91 training cases were obtained in section 5.4. If the training cases are too small, the chance of overfitting the data is high. In this case, the artificial neural network would memorize the noisy training data but fail to generalize the relationship of interest [122]. As a consequence, the prediction error would be inflated for new cases that were not used for training. Unfortunately, acquiring many vibration-rating-profile pairs is costly, because it relies on a system capable of controlled vibration

reproduction. Thus, unlike many other problems where neural networks are utilized, crowdsourcing ratings e.g. with Mechanical Turk is difficult. Another disadvantage of artificial neural networks is that they usually do not offer an interpretable model.
In the course of the synthesis, it was observed that there are major temporal spectral differences between the synthesized and recorded vibration. Despite these differences, the vibrations were perceived to be quasi equally plausible. However, since the perceived plausibility is likely inversely proportional to the inverse Euclidean distance in the sensory tactile perceptual space and both vibration elicit the same sensory tactile perceptual attribute ratings this is to be expected. If two vibration elicit the same sensory tactile perceptual attribute ratings, they could be considered to be perceptually equivalent. Such perceptual equivalences can guide the simplification of reproduction systems. If broadband signals are perceptually equivalent to narrowband signals e.g. for the tarmac road, broadband signals can be substituted by narrowband signals.

8. General Discussion and Outlook

The goal of this thesis was to assess a tactile design language enabling the quantification of user expectations and subsequently the translation into plausible WBV (i.e. vibration on seats) for virtual environments. To achieve this goal three main problems were solved. First, a compact set of sensory tactile perceptual attributes, which sufficiently characterize the perceivable variation of WBV and their direct relationship to the physical vibration parameters was identified to form a sensory tactile design language. Second, it was demonstrated that the sensory tactile design language can be utilized to quantify perceptual attributes expected in a situational context, even when this context is communicated by a verbal description instead of a multimodal presentation. Third, after the tactile plausibility illusion was formalized as a Euclidean distance measure in the sensory tactile perceptual space, synthesis models were created. These models can produce vibration from the expected sensory tactile perceptual attribute ratings, which are perceived as quasi equally plausible in a virtual environment as vibration originally occurring in the situational context.

To enable this study an experimental setup was developed, which enabled the presentation of WBV covering the frequency- and level range of everyday life of whole-body exposure. Thus, a reproduction system was created by combining a hydraulic motion platform suitable for low-frequency reproduction with an electrodynamic shaker for high-frequency reproduction. To create an auditory-visual-tactile virtual environment for validating the perceived plausibility of the synthesized vibrations acoustic stimuli were presented with a wave field synthesis system and optical stimuli were presented with a projector. In the following overview, the main findings of each sub-step will be summarized.

1) *Generalization of WBV encountered in everyday life into of excitation patterns*

   According to the ecological approach to perception, everyday life vibrations are a carrier of information about the environment. To assess a tactile design language all tactile perceptual attributes that might be elicited WBV needed to be assessed. Since recording each vibration potentially occurring in everyday life is infeasible, real vibration was generalized into four excitation patterns (sinusoidal vibration, AM-sinusoidal vibration, bandlimited WGN vibration, and impulse-like vibration). This abstraction enabled the variation of the defining parameters of these excitation patterns and thus to produce a vibration stimulus set covering the level and frequency range resolvable by tactile receptors.

2) *Assessment of the most relevant sensory-perceptual attributes elicited by generalized WBV*

   The generalized vibration stimulus representing everyday life exposure to WBV was the prerequisite to identify all relevant tactile perceptual attributes. The focus was set on sensory tactile perceptual attributes (e.g. "tingling"), since they have a direct relationship to the vibration signal and are thus most suitable for a translation into physical vibration parameters. In a free association task participants came up with a very large set of sensory-perceptual attributes. A design language requires efficient profiling making this set unsuitable for the direct utilization. However, natural language is full of synonyms and antonyms and each attribute was not mentioned with equal frequency. By utilizing a machine-generated thesaurus, it was possible to merge low frequently mentioned attributes with their more frequently mentioned synonyms or antonyms heuristically reducing redundancy. The most frequently occurring attributes are most familiar for laypersons and thus are most relevant as candidates for inclusion in the tactile design language.

3) *Identification of a compact sensory tactile design language*

   The main requirement of a tactile design language is that it enables the efficient and effective profiling of sensory tactile properties. This implies that the number of sensory tactile perceptual attributes included should be as small as possible while ensuring explanatory power. Therefore, the 21 most relevant sensory-perceptual attributes were rated absolutely on a quasi-continuous scale. Subsequently, a

PCA was conducted, revealing six groups of sensory-perceptual attributes. By selecting one attribute for each group (“weak”, “up & down”, “tingling”, “repetitive”, “uniform” and “fading”) redundancy is further reduced and the majority of the encountered variance in attribute ratings can be explained. Thus, the tactile sensory-perceptual space can be represented by these six attributes forming the sensory tactile design language.

4) *Confirmation of the sufficiency of the sensory tactile design language for qualitative communication about sensory tactile properties of real WBV*

However, the design language was assessed using generalized excitation patterns and not real WBV. Thus, it needed to be confirmed that no other attributes are elicited by such vibration. In everyday life WBV exposure most often occurs in vehicles. Therefore, a set of 19 auditory-visual-tactile scenes varying in road surface, vehicle speed, and operating condition was selected to represent typical exposure. The vibration segments contained in the scenes with a duration of 4 s to 17 s were either non-transient or contained single transient events to avoid having to account for temporal changes throughout the scene. Again, a free association task was conducted for these multimodal scenes to assess sensory tactile perceptual attributes, and mentions were aggregated and prioritized just like for the generalized vibration. The results confirm that the sensory tactile design language can be applied also for qualitative communication about typical real WBV.

5) *Demonstration of quantitative communication about WBV with the tactile design language.*

However, to enable translation into vibration, also quantitative communication about the elicited sensory tactile perceptual properties of WBV is required. Therefore, the sensory tactile design language consisting of the six attributes (“weak”, “up & down”, “tingling”, “repetitive”, “uniform”, and “fading”) was utilized to obtain rating profiles of the 19 auditory-visual-tactile scenes in a perceptual study.

The selection of a representative scene set suggested that there should be some obvious, perceivable differences which should be reflected in differences in ratings of the attributes of the design language. Indeed, the attributes showed significant rating differences between the scenes. This suggests that the design language can be utilized for efficient profiling of the sensory tactile perceptual properties.

6) *Demonstration of quantitative communication about expectations in a situational context with the tactile design language.*

   The ultimate goal is to utilize profiles consisting of sensory tactile perceptual attribute ratings for vibration synthesis. Thus, instead of recording vibration of a situational context and presenting them to quantify the elicited sensory tactile perceptual properties, it would be much more efficient to directly quantify sensory tactile perceptual properties expected of vibration in this situational context. If the expected sensory tactile perceptual properties of a situational context were representative of the elicited properties, a vibration synthesis directly from expectations would be possible. To investigate this hypothesis, verbal scene descriptions of the 19 scenes were created, e.g. driving on a cobblestone road at 50 kph constant speed. In the first block of a perceptual study, participants rated the six sensory tactile perceptual attributes of the verbal description and in a second block the multimodal scenes. A pairwise contrast between multimodal and anticipated mode shows a non-significant mean difference in attribute ratings of 2 points difference with a 95 % confidence interval of the difference ranging from -1 to 4 points. This suggests that the expected properties of a situational context are reasonably predictive of the elicited properties. This enables the quantification of sensory tactile perceptual properties for scenes where recording vibration is difficult or impossible.

7) *Formalization of the tactile plausibility illusion enabling model creation*

The previously defined rating profiles consisting of sensory tactile perceptual attribute ratings were to be translated into vibration eliciting the tactile plausibility illusion. The plausibility illusion is elicited when the content of the scene matches the user expectations. A previous study for the auditory domain suggested that the perceived plausibility is proportional to the inverse Euclidean distance between the expected and elicited properties in the underlying perceptual space. Applying this relationship to the tactile domain, it was assumed that the six perceptual attributes (“weak”, “up & down”, “tingling”, “repetitive”, “uniform”, and “fading”) forming design language sufficiently represent the sensory tactile perceptual space. Thus, it can be concluded that the perceived plausibility of vibration in a situational context is proportional to the inverse Euclidean distance between the expected sensory tactile perceptual properties of this context and the sensory tactile perceptual properties elicited by these vibrations. For example, the better the elicited properties (e.g. “up and down”) match the expected properties of WBV in a scene (e.g. “driving on a cobblestone road”) the more plausible it will be perceived. Since the expected properties are quantified in the form of a rating profile consisting of ratings of the six sensory-perceptual attributes of the design language, the synthesis model should estimate parameters of vibration that would elicit a rating profile which is has a minimal Euclidean distance to the expected rating profile and is thus perceived as maximally plausible in the corresponding situational context. This formalization enables the creation of models, which estimate vibration parameters from the expected sensory tactile perceptual properties thus translating the user expectations into vibration.

8) *Plausible vibration synthesis based on discrete estimates of a 1-NN classifier*

Building onto the formalization of plausibility as being proportional to the inverse Euclidean distance in the sensory tactile perceptual space, a 1-NN classifier was created. The classifier operates on the database of vibration-rating profile pairs collected during investigation of the relationship between attribute ratings and vibration generalized into four excitation patterns (sinusoidal vibration, AM-sinusoidal vibration, bandlimited WGN vibration and impulse-like vibration). For a given rating profile consisting of ratings of the six sensory-perceptual attributes (“weak”, “up & down”, “tingling”, “repetitive”, “uniform”, and “fading”) the classifier determines the rating profile with smallest Euclidean distance in the database. The parameters (sensation level SL, vibration (carrier- or resonance-) frequency f, bandwidth f_b, modulation frequency f_m and decay constant α) of the vibration associated with this perceptually most similar rating profile represent the discrete estimates of the model. The relevant excitation pattern for the vibration synthesis is implicitly contained in the vibration parameters.

9) *Plausible vibration synthesis based on quasi-continuous estimates of a regression model*

One of the disadvantages of the 1-NN models is the vibration parameter estimates being discrete. Therefore, also a regression-based model was created, which produces quasi-continuous estimates. The modeling was again based on the investigation of the relationship between attribute ratings and vibration generalized into four excitation patterns (sinusoidal vibration, AM-sinusoidal vibration, bandlimited WGN vibration, and impulse-like vibration). However, due to the abstraction of everyday life WBV into four distinct excitation patterns without unified parameters, the creation of one single regression model covering all excitation patterns is not possible. Thus, a set of three SVMs was created based on the vibration-rating profile pairs obtained in to enable the selection of one excitation pattern based on the input rating profile. For each excitation pattern the two to three sensory tactile perceptual attributes, which ratings were most affected by the defining vibration parameters of this excitation pattern, were selected. Subsequently, two to three regression equations of the relationships between the defining vibration parameters and each selected attribute was modeled. The equation system formed by the two to three regression equations was solved by the defining vibration parameters of the excitation pattern. This produced two to three synthesis equations for estimating the defining vibration parameters of each excitation pattern.

10) *Validation of the synthesis models and the design language*

The expected rating profiles obtained by presenting 19 audio-visual-tactile vehicle scenes were input into both models to estimate vibration parameters and to synthesize vibration for both scenes. The recorded vibration of the 19 audio-visual-tactile scenes was replaced by the recorded vibration of both models, forming three sets of 19 scenes. In a perceptual study, each set of scenes was presented separately and the perceived plausibility of the WBV of each scene was rated on a quasi-continuous rating scale. The results of suggested that vibration generated based on these estimates is perceived as quasi equally plausible as vibration originally occurring in this situation in a virtual environment, in which optical and acoustic recordings of the situation are simultaneously presented. Therefore, the 1-NN model and the regression model provide satisfactory performance in producing plausible WBV for virtual environments for the representative scene set and are thus validated. The validation demonstrated that vibration recorded in a situational context can be perceptually profiled with the sensory tactile design language and translated back into a vibration that is perceptually quasi-equivalent to the recorded vibration in this situational context. Since the validation scenes were selected to represent typical everyday life WBV exposure, this likely implies that the sensory tactile design language is suitable for completely i.e. sufficiently describing the sensory-perceptual properties of WBV.

In summary, the sensory tactile design language sufficiently describes the sensory-perceptual properties of WBV and also enables efficient, standardized, and quantitative communication about these properties. Since expected sensory properties can be quantified without the presence of vibration it is possible to assess these properties in advance in user surveys (e.g. web-based) by providing a verbal description of the situation. In contrast to physical vibration parameters, the design language is understandable by laypersons. The regression model suggests that there is a continuous relationship between physical vibration parameters and sensory-perceptual attributes, which shrinks the semantic gap. Therefore, a model-driven, i.e. automatic translation of from a profile quantified with the standardized sensory tactile design language into WBV is possible. The translation of user expectations into WBV with controlled sensory-perceptual properties enables the systematic elicitation of the expected sensory tactile attributes and thus of the tactile plausibility illusion in virtual environments where an authentic reproduction is not required and the user needs to rely on their expectations. This reduces or even eliminates the need for trial and error, i.e. iterative user studies with live vibration presentation. Furthermore, the proposed method offers an alternative to recording vibration, where this is either impossible or difficult, e. g. because vibrations need to be added to audio-visual recordings in retrospect. In some cases, the vibrations synthesized from user expectations tend to be even more plausible than recorded vibrations. Overall, the demonstrated method can be a new tool for designers authoring vibration for virtual environments and enable further automatization and thus potential time and cost reductions.

In the course of the synthesis, it was observed that there are major temporal spectral differences between the synthesized and recorded vibration. Despite these differences the vibrations were perceived to be quasi equally plausible, suggesting it is not necessary synthesized vibration to be indistinguishable from real vibration but that it is rather sufficient if they elicit the same sensory-perceptual properties. If two vibration elicit the same sensory tactile perceptual attribute ratings, they could be considered to be perceptually equivalent. If broadband signals are perceptually equivalent to narrowband signals e.g. for the tarmac road, broadband signals can be substituted by narrowband signals. Such perceptual equivalences can guide the simplification of reproduction systems, i.e. replacing an expensive system with a flat frequency response spanning a wide frequency range with a cheaper narrow-band exciter. Overall, the success of the proposed approach shows that sensory tactile perceptual attributes are not only arbitrary labels for perceivable vibration properties. If a set of sensory tactile perceptual attributes can be chosen to represent the sensory tactile perceptual space sufficiently, it hints at a sensory layer of abstraction in tactile perception. The results of the validation study demonstrated that vibration of a situational context can be perceptually profiled with the set of six sensory tactile perceptual attributes and translated back into vibration with different temporal and spectral properties but similar sensory tactile attribute ratings. Since both vibrations are quasi equally plausible, the information contained in the sensory layer is likely sufficiently characterizing the perceivable variation in vibration under the condition that no A/B paired comparison is utilized. Arguably such A/B comparisons are most often utilized in laboratory settings but are less representative of real use cases, where users mostly have to rely on comparisons to their expectations. Furthermore, it was demonstrated that the information available on this layer for a situational context can explicitly assessed and quantified in the form of the sensory tactile perceptual attributes, possibly suggesting a relationship to the perceptual process of categorization. By combining an analysis model that estimates the sensory-perceptual attribute ratings with a synthesis model, it might be possible to produce a sensory tactile compression algorithm for the plausible transmission of vibration.

There are multiple ways to extend this study. The approach should also be tested for longer scenes with changing perceptual attribute ratings over time. Such scenes could be treated as a set of successive segments with constant perceptual attribute ratings. Artifacts at the transitions of segments could be eliminated by e. g. crossfading. The models could be improved by incorporating more vibration-rating profile pairs. It would be beneficial to merge the utilized four excitation patterns with partly separate vibration parameters into one more general excitation pattern with one unified set of vibration parameters. Future studies should confirm the finding for more scenarios with vibration e.g. with scenarios unfamiliar to users. It can be interesting to investigate whether the general approach can be applied to other locations or directions of vibration introduction. The results from [96] suggest that hand-arm vibration and WBV are described with the same sensory tactile perceptual attributes and elicit similar ratings for identical vibration signals.

Instead of an iterative design process were user expectations play an implicit role when a design decision is approximated incrementally by repeated user validations, the tactile design language could enable an explicit, standardized assessment of expected sensory tactile perceptual properties by future users without prior training even before any prototype is available. The suggested

models can automatically translate these quantifications into vibration parameters suitable for engineers, which ensure the desired perceptual properties will be elicited and thus will likely match tactile user expectations. This might reduce the reliance on the experience of experts as well as time and cost requirements in the product development process. By acknowledging that specific vibrations are inherently eliciting specific sensory-perceptual attributes, a set of intended feedback messages could be purposefully assigned to such vibration that makes them more intuitively understandable, thus reducing training. E.g. a confirmation message might be rather linked to a "uniform" vibration than a warning message.
For many products, vibration is produced by e.g. actuators of limited capabilities or by machines whose excitation cannot be varied arbitrarily by dampening or resonance frequency shifts. By building an analysis model from the observed relationship between vibration and sensory-perceptual attributes, the attribute rating range elicited by the constrained vibration could predict. The degree of discrepancy between sensory tactile perceptual properties of expected and elicited vibration might provide a metric for tactile quality prediction. Since pleasantness or annoyance in particular and quality, in general, are dependent on user expectations, which differ from situational context to context, each context would require a different model. By interpreting the quantified expected sensory tactile perceptual properties as an additional input, more universally applicable quality models might be possible.

Acknowledgment

Research funded by German Research Foundation (DFG, Deutsche Forschungsgemeinschaft) AL1473/5-1 and as part of Germany's Excellence Strategy – EXC 2050/1 – Project ID 390696704 – Cluster of Excellence "Centre for Tactile Internet with Human-in-the-Loop" (CeTI) of TU Dresden.

References

[1] U. Jekosch, "Basic Concepts and Terms of 'Quality', Reconsidered in the Context of Product-Sound Quality," *acta acustica united with Acustica*, vol. 90, no. 6, pp. 999–1006, 2004.

[2] M. Slater and M. V. Sanchez-Vives, "Enhancing Our Lives with Immersive Virtual Reality," *Frontiers in Robotics and AI*, vol. 3, p. 74, 2016.

[3] A. El Saddik, "The potential of haptics technologies," *IEEE Instrumentation & Measurement Magazine*, vol. 10, no. 1, pp. 10–17, 2007.

[4] M. M. North, S. M. North, and J. R. Coble, *Virtual Reality Therapy*. IPI Press, 1996.

[5] K. Walker and W. L. Martens, "Perception of audio-generated and custom motion programs in multimedia display of action-oriented DVD films," in *International Workshop on Haptic and Audio Interaction Design*, 2006, pp. 1–11.

[6] M. Slater, "Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1535, pp. 3549–3557, 2009.

[7] H. Hayakawa, Y. Kamiyama, S. Matsuzono, M. Xu, M. Tanaka, H. Motoyama, N. Suzuki, and K. Minamizawa, "An Empirical

Practice of Live Feeling of a Professional Basketball Game with Haptic Transmission," in *Euro Haptics 2018*, 2018.

[8] R. Rosenkranz, M. E. Altinsoy, M. Stamm, and S. Merchel, "Nutzung von taktilen semantischen Informationen zur Synthese von plausiblen Ganzkörperschwingungen," in *DAGA 2013 - 39th German Annual Conference on Acoustics*, 2013.

[9] G. Calvert, C. Spence, and B. E. Stein, *The Handbook of Multisensory Processes*. The MIT Press, 2004.

[10] O. Schneider, K. MacLean, C. Swindells, and K. Booth, "Haptic experience design: What hapticians do and where they need help," *International Journal of Human-Computer Studies*, vol. 107, pp. 5–21, 2017.

[11] O. Franzén, R. Johansson, and L. Terenius, "Somesthesis and the Neurobiology of the Somatosensory Cortex," O. Franzén, R. Johansson, and L. Terenius, Eds. Birkhauser Verlag, 1996, pp. 49–58.

[12] H. Seifi, M. Chun, and K. E. Maclean, "Toward Affective Handles for Tuning Vibrations," *ACM Transactions on Applied Perception (TAP)*, vol. 15, no. 3, p. 22, 2018.

[13] H. Nagano, S. Okamoto, and Y. Yamada, "Modeling Semantically Multilayered Affective and Psychophysical Responses Toward Tactile Textures," *IEEE Transactions on Haptics*, vol. 11, no. 4, pp. 568–578, Oct. 2018.

[14] H. Dupuis and G. Zerlett, *The effects of whole-body vibration*. Springer Science & Business Media, 2012.

[15] M. J. Griffin, *Handbook of Human Vibration*. Academic press, 1990.

[16] M. Morioka and M. J. Griffin, "Absolute Thresholds for the Perception of Fore-And-Aft, Lateral, and Vertical Vibration at the Hand, the Seat, and the Foot.," *Journal of Sound and Vibration.*, vol. 314, pp. 357–370, 2008.

[17] G. Robles-De-La-Torre, "The importance of the sense of touch in virtual and real environments," *Ieee Multimedia*, vol. 13, no. 3, pp. 24–30, 2006.

[18] S. J. Bolanowski Jr, G. A. Gescheider, R. T. Verrillo, and C. M. Checkosky, "Four channels mediate the mechanical aspects of

touch," *The Journal of the Acoustical society of America*, vol. 84, no. 5, pp. 1680–1694, 1988.

[19] S. Okamoto, H. Nagano, and Y. Yamada, "Psychophysical Dimensions of Tactile Perception of Textures," *IEEE Transactions on Haptics*, vol. 6, no. 1, pp. 81–93, 2013.

[20] I. Hwang and S. Choi, "Perceptual space and adjective rating of sinusoidal vibrations perceived via mobile device," in *Haptics Symposium, 2010 IEEE*, 2010, pp. 1–8.

[21] G. Park and S. Choi, "Perceptual space of amplitude-modulated vibrotactile stimuli," in *World Haptics Conference (WHC), 2011 IEEE*, 2011, pp. 59–64.

[22] I. Hwang, J. Seo, and S. Choi, "Perceptual Space of Superimposed Dual-Frequency Vibrations in the Hands," *PloS one*, vol. 12, no. 1, p. e0169570, 2017.

[23] K. MacLean and M. Enriquez, "Perceptual design of haptic icons," in *Proc. of EuroHaptics*, 2003, pp. 351–363.

[24] L. A. Jones and A. Singhal, "Perceptual dimensions of vibrotactile actuators," in *2018 IEEE Haptics Symposium (HAPTICS)*, 2018, pp. 307–312.

[25] D. Ternes and K. E. Maclean, "Designing large sets of haptic icons with rhythm," in *International Conference on Human Haptic Sensing and Touch Enabled Computer Applications*, 2008, pp. 199–208.

[26] S. Choi and K. J. Kuchenbecker, "Vibrotactile display: Perception, technology, and applications," *Proceedings of the IEEE*, vol. 101, no. 9, pp. 2093–2104, 2013.

[27] L. A. Jones and H. Z. Tan, "Application of psychophysical techniques to haptic research," *IEEE Transactions on Haptics*, vol. 6, no. 3, pp. 268–284, 2013.

[28] H. Fastl and E. Zwicker, *Psycho Acoustics - Facts and Models*, 3rd ed. Springer Verlag, Berlin, 2010.

[29] T. Miwa, Y. Yonekawa, and K. Kanada, "Thresholds of perception of vibration in recumbent men," *The Journal of the Acoustical Society of America*, vol. 75, no. 3, pp. 849–854, 1984.

[30] M. A. Bellmann, "Perception of Whole-Body Vibrations - From Basic Experiments to Effects of Seat and Steering Wheel

Vibrations on the Passenger's Comfort Inside Vehicles," Universität Oldenburg, 2002.
[31] A. Schwendicke, M. E. Altinsoy, and S. Merchel, "Die Grenzen der Wahrnehmung - Fühlschwellen für Ganzkörperschwingungen bei hohen Frequenzen," in *DAGA 2015 - 41st German Annual Conference on Acoustics*, 2015.
[32] S. Merchel, *Auditory-Tactile Music Perception*. Shaker Verlag, 2014.
[33] M. J. Griffin, "Subjective Equivalence of Sinusoidal and Random Whole-Body Vibration," *The Journal of the Acoustical Society of America*, vol. 60, no. 5, pp. 1140–1145, 1976.
[34] M. Morioka and M. J. Griffin, "Difference Thresholds for Intensity Perception of Whole-Body Vertical Vibration: Effect of Frequency and Magnitude," *The Journal of the Acoustical Society of America*, vol. 107, no. 1, pp. 620–624, 2000.
[35] A. Schwendicke, J. Dou, and M. E. Altinsoy, "Temporal Masking Characteristics of Whole Body Vibration Perception," in *IEEE World Haptics 2017*, 2017.
[36] I. R. Summers, P. G. Cooper, P. Wright, D. A. Gratton, P. Milnes, and B. H. Brown, "Information from time-varying vibrotactile stimuli," *The Journal of the Acoustical Society of America*, vol. 102, no. 6, pp. 3686–3696, 1997.
[37] S. Brewster and L. M. Brown, "Tactons: structured tactile messages for non-visual information display," in *Proceedings of the fifth conference on Australasian user interface-Volume 28*, 2004, pp. 15–23.
[38] J. Ryu, J. Chun, G. Park, S. Choi, and S. H. Han, "Vibrotactile feedback for information delivery in the vehicle," *IEEE Transactions on Haptics*, vol. 3, no. 2, pp. 138–149, 2010.
[39] H. Z. Tan, N. I. Durlach, C. M. Reed, and W. M. Rabinowitz, "Information transmission with a multifinger tactual display," *Perception & Psychophysics*, vol. 61, no. 6, pp. 993–1008, 1999.
[40] S. Wang, X. Zhong, M. F. Dorman, W. A. Yost, and J. M. Liss, "Using tactile aids to provide low frequency information for cochlear implant users," *The Journal of the Acoustical Society of America*, vol. 134, no. 5, pp. 4235–4235, 2013.

[41] E. Steinbach, M. Strese, M. Eid, X. Liu, A. Bhardwaj, Q. Liu, M. Al-Ja'afreh, T. Mahmoodi, R. Hassen, A. El Saddik, and others, "Haptic codecs for the tactile internet," *Proceedings of the IEEE*, vol. 107, no. 2, pp. 447–470, 2018.

[42] B. Winter, *Sensory linguistics: Language, perception and metaphor*, vol. 20. John Benjamins Publishing Company, 2019.

[43] J. J. Gibson, *The ecological approach to visual perception: classic edition*. Psychology Press, 2014.

[44] W. W. Gaver, "How Do We Hear in the World? Explorations in Ecological Acoustics," *Ecological psychology*, vol. 5, no. 4, pp. 285–313, 1993.

[45] W. W. Gaver, "What in the World Do We Hear? An Ecological Approach to Auditory Event Perception," *Ecological psychology*, vol. 5, no. 1, pp. 1–29, 1993.

[46] H. Cohen and C. Lefebvre, *Handbook of categorization in cognitive science*. Elsevier, 2005.

[47] E. Rosch, C. B. Mervis, W. D. Gray, D. M. Johnson, and P. Boyes-Braem, "Basic objects in natural categories," *Cognitive psychology*, vol. 8, no. 3, pp. 382–439, 1976.

[48] S. Harnad, "Psychophysical and cognitive aspects of categorical perception: A critical overview," in *Categorical perception: The groundwork of cognition*, Cambridge University Press, 1987, pp. 1–52.

[49] D. Dubois, C. Guastavino, and M. Raimbault, "A Cognitive Approach to Urban Soundscapes: Using Verbal Data to Access Everyday Life Auditory Categories," *Acta acustica united with acustica*, vol. 92, no. 6, pp. 865–874, 2006.

[50] A. Hoffmann, J. Verhey, and R. Weber, "Charakteristische Empfindungen bei sinusförmigen Ganzkörpervibrationen," in *DAGA 2010 proceedings*, 2010.

[51] J. W. Tanaka and M. Taylor, "Object categories and expertise: Is the basic level in the eye of the beholder?," *Cognitive psychology*, vol. 23, no. 3, pp. 457–482, 1991.

[52] A. Giboreau, C. Dacremont, C. Egoroff, S. Guerrand, I. Urdapilleta, D. Candel, and D. Dubois, "Defining sensory

descriptors: Towards writing guidelines based on terminology," *Food quality and preference*, vol. 18, no. 2, pp. 265–274, 2007.

[53] S. D. Deyne, W. Voorspoels, S. Verheyen, D. J. Navarro, and G. Storms, "Accounting for graded structure in adjective categories with valence-based opposition relationships," *Language, Cognition and Neuroscience*, vol. 29, no. 5, pp. 568–583, 2014.

[54] C. V. Jansson-Boyd, "Touch matters: exploring the relationship between consumption and tactile interaction," *Social Semiotics*, vol. 21, no. 4, pp. 531–546, 2011.

[55] A. Israr, S. Zhao, K. Schwalje, R. Klatzky, and J. Lehman, "Feel Effects: Enriching Storytelling with Haptic Feedback," *ACM Trans. Appl. Percept.*, vol. 11, no. 3, pp. 11:1–11:17, 2014.

[56] C. O'Sullivan and A. Chang, "An activity classification for vibrotactile phenomena," in *International Workshop on Haptic and Audio Interaction Design*, 2006, pp. 145–156.

[57] S. Guest, J. M. Dessirier, A. Mehrabyan, F. McGlone, G. Essick, G. Gescheider, A. Fontana, R. Xiong, R. Ackerley, and K. Blot, "The development and validation of sensory and emotional scales of touch perception," *Attention, Perception, & Psychophysics*, vol. 73, no. 2, pp. 531–550, 2011.

[58] M. Obrist, S. A. Seah, and S. Subramanian, "Talking about tactile experiences," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 1659–1668.

[59] H. Seifi, K. Zhang, and K. E. MacLean, "VibViz: Organizing, visualizing and navigating vibration libraries," in *World Haptics Conference (WHC), 2015 IEEE*, 2015, pp. 254–259.

[60] H. Seifi and K. E. MacLean, "Exploiting haptic facets: Users' sensemaking schemas as a path to design and personalization of experience," *International Journal of Human-Computer Studies*, vol. 107, pp. 38–61, 2017.

[61] M. E. Altinsoy, "What Can We Learn from Psychoacoustics Regarding the Perception of Whole-Body Vibrations? Tactile Descriptors for Whole-Body Vibrations," in *Proceedings of 2nd Polish-German structured conference on acoustics. The 58th open seminar on acoustics. Jurata, Poland*, 2011.

[62] M. E. Altinsoy, "Identification of Quality Attributes of Automotive Idle Sounds and Whole-Body Vibrations," *International Journal of Vehicle Noise and Vibration*, vol. 9, no. 1/2, pp. 4–27, 2013.

[63] Y. Yoo, T. Yoo, J. Kong, and S. Choi, "Emotional responses of tactile icons: Effects of amplitude, frequency, duration, and envelope," in *2015 IEEE World Haptics Conference (WHC)*, 2015, pp. 235–240.

[64] J. Blauert and U. Jekosch, "Sound-quality evaluation–a multi-layered problem," *Acta acustica united with acustica*, vol. 83, no. 5, pp. 747–753, 1997.

[65] R. M. Held and N. I. Durlach, "Telepresence," *Presence: Teleoperators & Virtual Environments*, vol. 1, no. 1, pp. 109–112, 1992.

[66] D. C. Dennett, "Quining qualia," in *Consciousness in modern science*, Oxford University Press, 1988.

[67] A. Noë, A. Noë, and others, *Action in perception*. MIT press, 2004.

[68] M. Slater, "Measuring presence: A response to the Witmer and Singer presence questionnaire," *Presence*, vol. 8, no. 5, pp. 560–565, 1999.

[69] M. Slater, "A Note on Presence Terminology," *Presence connect*, vol. 3, no. 3, pp. 1–5, 2003.

[70] M. Garau, D. Friedman, H. R. Widenfeld, A. Antley, A. Brogni, and M. Slater, "Temporal and spatial variations in presence: Qualitative analysis of interviews from an experiment on breaks in presence," *Presence: Teleoperators and Virtual Environments*, vol. 17, no. 3, pp. 293–309, 2008.

[71] B. G. Witmer and M. J. Singer, "Measuring Presence in Virtual Environments: A Presence Questionnaire," *Presence: Teleoperators and virtual environments*, vol. 7, no. 3, pp. 225–240, 1998.

[72] M. Usoh, E. Catena, S. Arman, and M. Slater, "Using presence questionnaires in reality," *Presence: Teleoperators & Virtual Environments*, vol. 9, no. 5, pp. 497–503, 2000.

[73] M. Slater, "How Colorful Was Your Day? Why Questionnaires Cannot Assess Presence in Virtual Environments," *Presence*, vol. 13, no. 4, pp. 484–493, 2004.

[74] P. Novo, "Auditory Virtual Environments," in *Communication Acoustics*, J. Blauert, Ed. Springer, 2005, p. 278.

[75] H. Fastl, B. Lichtinger, and S. Kerber, "Psychoakustische Experimente zur Knallhaftigkeit," in *DAGA 2012 - 38th German Annual Conference on Acoustics*, 2012.

[76] R. Skarbez, "Plausibility illusion in virtual environments," University of North Carolina at Chapel Hill Graduate School, 2016.

[77] K. Kuhn-Rahloff, "Prozesse der Plausibilitätsbeurteilung am Beispiel ausgewählter elektroakustischer Wiedergabesituationen," Technische Universität Berlin, 2011.

[78] A. Lindau and S. Weinzierl, "Assessing the Plausibility of Virtual Acoustic Environments," *Acta Acustica united with Acustica*, vol. 98, no. 5, pp. 804–810, 2012.

[79] Z. Wang, J. Lu, A. Peer, and M. Buss, "Influence of Vision and Haptics on Plausibility of Social Interaction in Virtual Reality Scenarios," in *Haptics: Generating and Perceiving Tangible Sensations*, 2010, pp. 172–177.

[80] S. Sakamoto, T. Ohtani, Y. Suzuki, and J. Gyoba, "Effects of Vibration Information on the Senses of Presence and Verisimilitude of Audio-visual Scenes," in *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, 2016, vol. 253, no. 3, pp. 4890–4895.

[81] R. Rosenkranz, M. E. Altinsoy, M. Stamm, and S. Merchel, "Taktile Wahrnehmung im Kontext multimodaler virtueller Umgebungen," in *DAGA 2012 - 38th German Annual Conference on Acoustics*, 2012.

[82] D. Gongora, H. Nagano, M. Konyo, and S. Tadokoro, "Vibrotactile rendering of camera motion for bimanual experience of first-person view videos," in *2017 IEEE World Haptics Conference (WHC)*, 2017, pp. 454–459.

[83] J. Lee, B. Han, and S. Choi, "Motion Effects Synthesis for 4D Films," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 10, pp. 2300–2314, 2016.

[84] M. Kim, S. Lee, and S. Choi, “Saliency-Driven Real-Time Video-to-Tactile Translation,” *IEEE Transactions on Haptics*, vol. 7, no. 3, pp. 394–404, 2014.

[85] J. Lee and S. Choi, “Real-time perception-level translation from audio signals to vibrotactile effects,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2013, pp. 2567–2576.

[86] ISO2631, “Mechanical Vibration and Shock – Evaluation of Human Exposure to Whole-Body Vibration,” International Organization for Standardization, Geneva, Switzerland, 1978.

[87] M. E. Altinsoy, U. Jekosch, J. Landgraf, and S. Merchel, “Progress in Auditory Perception Research Laboratories - Multimodal Measurement Laboratory of Dresden University of Technology,” in *Audio Engineering Society Convention 129*, 2010.

[88] M. Azadi and L. A. Jones, “Evaluating vibrotactile dimensions for the design of tactons,” *IEEE Transactions on Haptics*, vol. 7, no. 1, pp. 14–23, 2014.

[89] M. E. Altinsoy and S. Merchel, “BRTF (Body-Related Transfer Function) and Whole-Body Vibration Reproduction Systems,” in *Audio Engineering Society Convention 130*, 2011.

[90] M. E. Altinsoy, *Auditory-Tactile Interaction in Virtual Environments*. Shaker Verlag, 2006.

[91] N. F. Dixon and L. Spitz, “The detection of auditory visual desynchrony,” *Perception*, vol. 9, no. 6, pp. 719–721, 1980.

[92] B. Rohrmann, “Verbal qualifiers for rating scales: Sociolinguistic considerations and psychometric data,” *University of Melbourne, Melbourne2007*, 2007.

[93] R. Rosenkranz, S. Gruschwitz, M. E. Altinsoy, and S. Merchel, “Untersuchung der Pegelabhängigkeit von Wahrnehmungsmerkmalen von Ganzkörperschwingungen,” in *DAGA 2015 - 41th German Annual Conference on Acoustics*, 2015.

[94] R. Rosenkranz, S. Gruschwitz, M. E. Altinsoy, and S. Merchel, “Wahrnehmungsmerkmale von breitbandigen Ganzkörperschwingungen im Vergleich zu sinusoidalen Ganzkörperschwingungen,” in *DAGA 2016 - 42th German Annual Conference on Acoustics*, 2016.

[95] R. Rosenkranz, S. Gruschwitz, M. E. Altinsoy, and S. Merchel, “Nutzung von Wahrnehmungsmerkmalen zur Charakterisierung von amplitudenmodulierten sinusoidalen Ganzkörperschwingungen,” in *DAGA 2017 - 43nd German Annual Conference on Acoustics*, 2017.

[96] R. Rosenkranz, S. Gruschwitz, M. Wilberg, M. E. Altinsoy, and S. Merchel, “Identification and Evaluation of Perceptual Attributes for Periodic Whole-Body and Hand-Arm Vibration,” in *Euro Haptics 2018*, 2018.

[97] R. Rosenkranz and M. E. Altinsoy, “Tactile Design: Translating User Expectations into Vibration for Plausible Virtual Environments,” in *IEEE World Haptics 2019*, 2019.

[98] R. Rosenkranz and M. E. Altinsoy, “Mapping the Sensory-Perceptual Space of Vibration for User-Centered Intuitive Tactile Design,” *To appear in IEEE Transactions on Haptics*, 2020.

[99] M. E. Altinsoy and U. Jekosch, “The Semantic Space of Vehicle Sounds: Developing a Semantic Differential with Regard to Customer Perception,” *Journal of the Audio Engineering Society*, vol. 60, no. 1/2, pp. 13–20, 2012.

[100] S. Merchel, M. E. Altinsoy, and M. Stamm, “Just-Noticeable Frequency Differences for Whole-Body Vibrations.,” in *Proceedings of Internoise 2011*, 2011.

[101] A. Israr, H. Z. Tan, and C. M. Reed, “Frequency and amplitude discrimination along the kinesthetic-cutaneous continuum in the presence of masking stimuli,” *The Journal of the Acoustical society of America*, vol. 120, no. 5, pp. 2789–2800, 2006.

[102] S. Merchel, M. E. Altinsoy, and M. Stamm, “Equal Intensity Contours for Whole-Body Vibrations Compared with Vibrations Cross-Modally Matched to Isophones.,” in *HAID*, 2011, pp. 71–80.

[103] C. Lalanne, *Mechanical Vibration & Shock Volume II: Mechanical Shock*. Hermes Penton Ltd., New York, 2002, p. 160.

[104] S. Webb, “RECEPTIVE AND PRODUCTIVE VOCABULARY SIZES OF L2 LEARNERS,” *Studies in Second Language Acquisition*, vol. 30, no. 1, pp. 79–95, 2008.

[105] M. E. Altinsoy, “Wahrnehmungsmerkmale von Ganzkörperschwingungen,” *Lärmbekämpfung*, vol. 11, no. 1, 2016.

[106] C. Biemann, G. Heyer, U. Quasthoff, and M. Richter, “The Leipzig corpora collection-monolingual corpora of standard size,” *Proceedings of Corpus Linguistic*, vol. 2007, 2007.

[107] B. Rohrmann, “Empirische Studien zur Entwicklung von Antwortskalen für die sozialwissenschaftliche Forschung,” *Zeitschrift für Sozialpsychologie*, vol. 107, no. 5, pp. 222–245, 1978.

[108] U. Jekosch, *Voice and speech quality perception: assessment and evaluation*. Springer Science & Business Media, 2006.

[109] E. Hoggan and S. Brewster, “Designing Audio and Tactile Crossmodal Icons for Mobile Devices,” in *Proceedings of the 9th International Conference on Multimodal Interfaces*, 2007, pp. 162–169.

[110] R. Bakeman, “Recommended effect size statistics for repeated measures designs,” *Behavior research methods*, vol. 37, no. 3, pp. 379–384, 2005.

[111] E. C. Poulton and S. Poulton, *Bias in quantifying judgments*. Taylor & Francis, 1989.

[112] T. A. Runkler, *Data Analytics*, 2nd ed. Springer, 2012.

[113] G. Shmueli, “To Explain or to Predict?,” *Statist. Sci.*, vol. 25, no. 3, pp. 289–310, 2010.

[114] F. Hayashi, “Econometrics,” Princeton, 2000, p. 187.

[115] P. A. Parker, G. G. Vining, S. R. Wilson, J. L. Szarka III, and N. G. Johnson, “The prediction properties of classical and inverse regression for the simple linear calibration problem,” *Journal of quality technology*, vol. 42, no. 4, pp. 332–347, 2010.

[116] N. R. Draper and H. Smith, “Applied regression analysis,” 3rd ed., vol. 326, John Wiley & Sons, 1998, p. 90.

[117] S. Pischke, “Lecture notes on measurement error,” *London School of Economics, London*, 2007.

[118] A. N. Phillips and G. D. Smith, “How independent are independent? effects? Relative risk estimation when correlated exposures are measured imprecisely,” *Journal of clinical epidemiology*, vol. 44, no. 11, pp. 1223–1231, 1991.

[119] D. J. Barr, R. Levy, C. Scheepers, and H. J. Tily, “Random effects structure for confirmatory hypothesis testing: Keep it maximal,”

Journal of memory and language, vol. 68, no. 3, pp. 255–278, 2013.

[120] D. M. McNeish and L. M. Stapleton, “The effect of small sample size on two-level model estimates: A review and illustration,” *Educational Psychology Review*, vol. 28, no. 2, pp. 295–314, 2016.

[121] V. Bewick, L. Cheek, and J. Ball, “Statistics review 7: Correlation and regression,” *Critical care*, vol. 7, no. 6, p. 451, 2003.

[122] S. Haykin, *Neural networks: a comprehensive foundation*. Prentice Hall PTR, 1994.

[123] E. B. Baum and D. Haussler, “What size net gives valid generalization?,” in *Advances in neural information processing systems*, 1989, pp. 81–90.