AUDITORY-TACTILE INTERACTION IN VIRTUAL ENVIRONMENTS

AUDIOTAKTILE INTERAKTION IN VIRTUELLEN UMGEBUNGEN

DISSERTATION

zur Erlangung des akademischen Grades eines
Doktor-Ingenieurs
der Fakultät für Elektrotechnik und Informationstechnik
der Ruhr-Universität Bochum

vorgelegt von

M. Ercan Altınsoy, M.Sc.
aus Eskişehir, Türkei

Aachen 2005

Tag der mündlichen Prüfung	: 08.02.2006
Erstreferent	: Prof. Dr.-Ing. techn. h.c. J. Blauert
Korreferent	: Prof. Dr. rer. nat. Armin Kohlrausch

Acknowledgements

The following work was accomplished at the Institute of Communication Acoustics of the Ruhr-University Bochum, Germany. The completion of this thesis could not have been realised without the support and presence of several people.

First of all, I am very grateful to Prof. Dr.-Ing. Dr. techn. h.c. Jens Blauert, the head of the institute, for all his support and the valuable discussions during the supervision of my thesis. He has played big role in the attempt to make me a researcher.

Furthermore, I am so grateful to Prof. Dr. Armin Kohlrausch for giving me valuable support and guidance on the multimodal interaction research and electrifying my imagination, and ultimately my research.

I would like to thank Ass. Prof. Dr.-Ing. Jonas Braasch for his excellent support and spending so much time for our discussions on research and life over the last 5 years.

I would like to express my appreciation to Prof. Dr. Temel Belek who initiated my interest in acoustics and vibration and encouraged my path to higher education.

Many thanks also to Dr. Thomas Hempel, Dr.-Ing. Sebastian Möller and Prof. Dr. Ute Jekosch for numerous helpful discussions and for forcing me to think about new and exciting research problems.

I warmly thank to Mr. Peter Salzsieder for his outstanding technical support.

I would like to thank my friends and colleagues at the Institute of Communication Acoustics for all their support, the helpful discussions and the good working relationship we had through all the years. Thanks also due to those who participated so patiently in my psychophysical tests.

My sincere thanks go to all the organizations, which gave me their valuable generous sponsoring, including the DAAD (Deutscher Akademischer Austausch Dienst), the IGSN (International Graduate School of Neuroscience, Ruhr-University Bochum) and the European Union project MUVII (Multi User Virtual Interactive Interface, IST-2000-28463). I am especially grateful to the supervisors and fellows of the IGSN for all the valuable discussions and activities we had during the course of the graduate school.

I would like to express my gratitude to my wife Hanife for all her support and her love during the years. Life is always beautiful whenever I am with her. I would also like to thank my parents Burçin and Zeki Altınsoy. They always let me know that they are proud of me, which motivates me work harder and do my best.

"Hearing is a form of touch. Something that's so hard to describe, something that comes, sound that comes to you... You feel it through your body, and, sometimes, it almost hits your face."

(Evelyn Glennie," Touch The Sound")

Contents

Chapter 1

Introduction

In our daily life, sound is usually produced by the vibrations of a body. These vibrations in our natural environment lead to both auditory and tactile perceptions. In many situations, such as driving a car, drilling a hole, playing a guitar etc., we are exposed to sound and vibration simultaneously.

Perception is a multisensory phenomenon and one major capability of our perceptual system is the integration of the multisensory stimuli which are generated by a multimodal event. Dependencies in multisensory stimuli, which result from the physical processes that generate the stimuli are important hints for our brain (Kohlrausch and van de Par, 1999). The multimodal inputs which have a single origin (same event or object) are correlated in time and other properties such as frequency and intensity etc., according to physical laws.

Haptic feedback brings the sense of touch (tactile sense) and force-feedback to multi-media applications in addition to the mostly utilised modalities i.e., the auditory and visual one. In many new applications like virtual reality, flight simulators, and medical surgery, the user receives simultaneous auditory and tactile information. To benefit from multimodal displays, the users must be able to experience a coherent perception within the (virtual) environment by integrating the inputs from multiple modalities (Vogels, 2001). Computer processing time (which can cause delayed feedback reproduction), difficulties to generate large feedback forces for simulating hard contact, and limitations on the mechanical force-feedback bandwidth are some of the well-known problems of multi-modal interfaces. Multi-modal interface designers should optimise all these physical and technical constraints. From the point of view of a multi-modal interface designer, an understanding of the factors contributing to the integration of auditory and tactile information is necessary to obtain more realistic and compelling products.

The aim of this thesis is to gain a better understanding of the auditory-tactile integration and interaction. In the first part of this work, it is investigated which physical factors and which physical conditions can cause a perceptual segregation of auditory and tactile events. Most investigations dealing with multi-modal segregation (or integration) are related to auditory and visual modalities, while only very few investigations are addressing auditory and tactile interaction. A considerable amount of the auditory–visual integration studies has focused on

simultaneity, which is the most powerful cue available for determining whether two inputs have been generated by a single multimodal event or multiple unimodal events. In these studies artificial conflict situations were used in the laboratory and the sensitivity to delays between auditory and visual stimuli was measured (Dixon and Spitz, 1980; van de Par and Kohlrausch, 1999). The general findings of these studies are:

- Humans have great tolerance to intersensory asynchrony, and
- They are more sensitive to visual delays than to auditory delays.

Another factor that plays a role in multimodal integration is the ***spatial origin*** of events. If multisensory inputs belong to a single event, their perceived locations must coincide. Previous studies showed the importance of the spatial origin congruency for the auditory–visual integration by measuring the tolerance of humans to the difference of the spatial origins of auditory and visual stimuli (Bertelson and Radeau, 1981; Blauert, 1970). In this study, auditory-tactile event segregation based on spatial location will be investigated.

Two other factors, ***frequency*** and ***intensity,*** which contribute to the perceptual binding of multisensory stimuli, will be introduced, and their influence on the segregation of auditory-tactile events will be investigated. Although some auditory-tactile interaction studies have focused on the frequency of the auditory-tactile stimuli, they did not provide any theoretical or experimental information related to the influence of the auditory and tactile stimuli frequency to the *segregation or binding* of auditory-tactile events (Lederman et al., 1999; McGee, 2002; Jousmäki, and Hari, 1998). There have been relatively few attempts to examine the influence of stimulus intensity on the segregation of multimodal events. Early studies have concentrated upon the cross-modal intensity matching, by trying to understand whether subjects are able to make a direct comparison between the intensities of two different sensory modalities. However, the focus of these studies was not the segregation of the sensory modalities.

When our brain receives sensory inputs which are provided by different sensory channels, these multiple sensory inputs can be combined and result in a unified percept, but it is also possible to observe that the percept can segregate into isolated percepts in each modality. If multisensory inputs are combined, and result in a unified percept, it is important to investigate the laws behind the integration process and, consequently, interaction issues between these sensory inputs.

Multisensory perception can be defined as a “construction” of the brain that is derived from a weighted combination of multiple sensory inputs. During the evaluation of multi-modal events, tactile and auditory information interact and possess a substantial influence. Two modalities can be combined and the resulting multi-modal percept may be a weaker, stronger, or altogether different percept. Research questions dealing with intersensory interaction are: how does the brain weigh the inputs it receives from the different senses to produce a final percept? In other words, what are the relative contributions of the different sensory modalities to the multimodal percept? Can a perception of an event in one sensory modality change due to the presence of a stimulus in another sensory modality? Multisensory interaction gives us an opportunity to develop multisensory displays which may be used to overcome the present limitations of the multisensory user interfaces.

In the second part of this work, two fundamental auditory-tactile interaction experiments were performed to investigate the rules governing multisensory perception regarding auditory-tactile stimuli. The aim of the first one is the investigation of the effect of loudness on haptic force-feedback perception. For two reasons, hitting an object by a human subject was selected as stimulus condition. On the one hand, hitting is a very common multisensory event in our daily life (e.g. knocking on the door, playing a drum) and several possible applications in the realm of user-interfaces come to the mind. On the other hand, the relationship between physical attributes of the auditory and tactile stimuli and the human perception of these attributes are clearly observable.

In our daily life, another common multisensory event is the exploration of surfaces (textures) using bare fingers. By exploring a surface with our finger (e.g. scraping), we get information simultaneously from the auditory, tactile, and visual sensory channels. Roughness is the most important physical and perceptual property for surfaces. The aim of the second example is to determine the relative contributions of the auditory and tactile systems on multimodal roughness perception, and to investigate the influence of the incongruent auditory-tactile roughness information on the multimodal roughness perception. In this context, it is note Worthing that people are able to judge the roughness of different surfaces using tactile feedback alone, using the sounds produced by touching the surfaces alone (Lederman, 1979), or using visual information of the surfaces alone (Lederman and Abbott, 1981).

We obtain information from different sensory modalities when using different industrial products, e.g. driving a car, using a vacuum cleaner or a hand mixer etc. Consequently, the cross-modal information has a substantial influence on the product quality evaluation of the user. So, when designing an industrial product that is capable of addressing several modalities, it is important to provide the appropriate stimuli for the respective modalities at the right time for the purpose of perceptual integration. However, traditionally the perceived quality of one mode has been studied in isolation from one another. The importance of the multi-modal aspects of product perception was emphasized by Bednarzyk (1996), Blauert and Jekosch (1997), Kohlrausch, and van de Par (1999), and Quehl (2001). With an increased interest in multimedia applications, the influence of the visual information on the product sound quality or the influence of the auditory information on the video quality have been studied in recent years (Patsouras, 2003; Fastl, 2004; Kohlrausch and van de Par, 1999). However a systematic study which investigates the multi-modal aspects of product perception is missing. In this study, the combined influence of auditory and tactile stimuli in the overall product quality assessment was investigated by presenting electrical drill noise and vibrations to the subjects simultaneously.

Nowadays, virtual reality systems play a pronounced role in scientific research and, thus, become more and more important as tools for psychophysical and acoustical research. Therefore, investigations in the auditory-tactile interaction were carried out using an auditory-tactile virtual environment. The tactile subsystem of this environment was developed in this study. It consists a tactile glove system (which applies to the user's hand vibrotactile and force feedback information) and a whole-body vibrations system. While in some experiments recorded sounds were used, in other experiments physically modelled synthesized sounds were employed. Due to increasing usage of multimodal interfaces, interactive sounds which are the result of the user's haptic contact with virtual objects (hitting, rubbing, and stroking etc.) are becoming more and more important. In this work, touch-induced scraping sounds were physically modelled and synthesized.

Chapter 2 of this thesis gives a short overview on the psychophysical, physiological, and neurophysiological aspects of the auditory and tactile senses. The physiological aspects is described in a more detailed manner, because the understanding of physiological processes is fundamental for developing suited applications for the tactile channel—a field cordially inviting today's designers.

Chapter 3 explains present auditory and tactile virtual displays, and introduces the concept and the characteristics of the experimental system which was developed for this study and will be used to conduct further experiments.

In **Chapter 4**, the physical factors that play a role in the perceptual integration of auditory and tactile events are presented, and the psychophysical experiments which were conducted to determine the physical conditions which can cause the segregation of auditory and tactile events (e.g. incongruence between auditory and tactile stimuli) are described. In Section 4.2, the influence of the simultaneity on the segregation of auditory and tactile events is investigated. Section 4.3 investigates the influence of stimulus level on the segregation of auditory and tactile events. The influence of the stimulus frequency on the segregation of auditory and tactile events is presented in section 4.4. Section 4.5 reports the experiments in which spatial origins of the auditory and tactile stimulus are varied, and the influence of the spatial origin of the stimulus on the segregation of the auditory and tactile events is studied.

Chapter 5 is concerned with auditory-tactile interaction. To gain a better understanding of the interaction of auditory and tactile information, some investigations that are necessary to be able to specify the important criteria, are described. The investigations are concerned with:

- the influence of the loudness on the haptic force-feedback perception (Section 5.1),
- the influence of the incongruent auditory-tactile roughness information on the multimodal roughness perception (Section 5.2).

In **Chapter 6,** the combined influence of auditory and tactile information on product quality is investigated.

The thesis concludes with a general discussion in **Chapter 7**. Prospects for the further research are also outlined in this chapter.

Chapter 2

Auditory and Tactile Systems/Perception

2.1 Introduction

An understanding of physiological and psychophysical processes is fundamental to developing suited applications for the auditory and the tactile channels. Therefore a short overview about the physiological and the psychophysical aspects of the auditory and the tactile sense is given in this chapter.

Compared to the visual and auditory modalities, our understanding of the tactile modality is very limited. The tactile modality was seen as little more than an inferior version of vision, and it has been relatively neglected (Klatzky and Lederman, 2000). Therefore information processing in the tactile system is described in a more detailed manner here. This chapter starts by presenting properties of the auditory system.

2.2 The Auditory system

2.2.1 Physiology and information processing in the auditory system

Hair cells in our ear are responsible for transducing mechanical energy to electrical signals and transmitting these signals to the brain. The two most important attributes of the sound which should be encoded by the cochlea are the stimulus frequency and intensity. The basilar membrane in the inner ear acts like a frequency analyzer by distributing stimulus energy to the hair cells arrayed along its length according to the spectral components that make up the stimulus (place theory) (Hudspeth, 1999). It means that frequency and intensity processing begins already in the cochlea. Another process, which occurs in the cochlea, is the amplification of the sound energy. This amplification causes also noise similar to technical amplifiers. Nature solves this problem by reducing the sensitivity for the low frequency range (Bekesy, 1974).

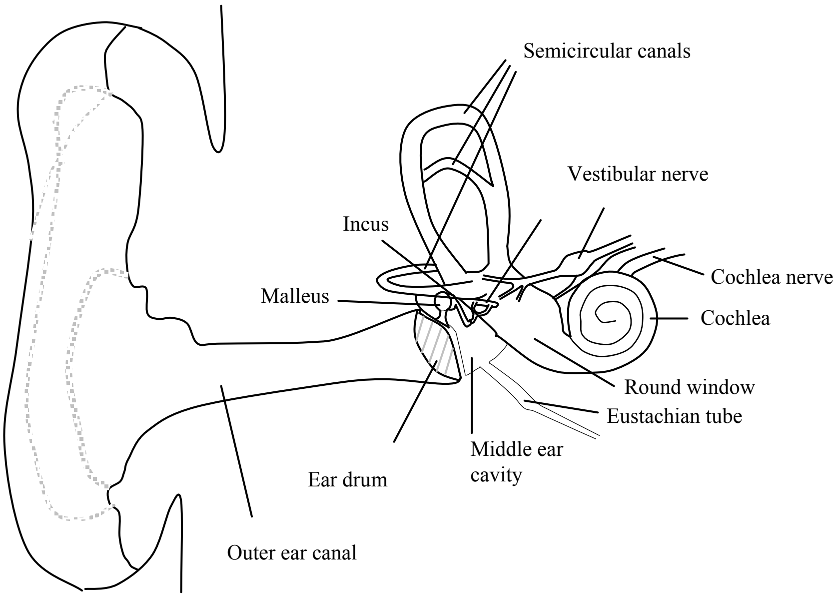


Figure 2.1: A cross-section of the ear.

The hair cells are frequency selective and this selectivity has a V-shaped tuning curve characteristic. It means that each hair cell is strongly selective for one frequency and this selectivity decreases with the increasing distance from this specific frequency. Ashmore (1995) reported that two tones that differ by 1% or less will excite distinct populations of hair cells. At low sound levels, a stimulus containing several frequencies stimulates several separated small groups of nerve fibres. At higher levels, a single tone stimulates many fibres at different location (Zwicker and Fastl, 1999).

Our hearing system decomposes a broad spectrum into parts that correspond to critical bands. The audible frequency range between 20 Hz to 16 kHz can be subdivided into 24 critical bands. The unit of the critical band rate is “Bark”.

Increasing loudness causes an increase on the vibration amplitude of the cochlea, and so hair cells are activated more strongly. Therefore there is a relationship between action potentials of the hair cells and the sound pressure level.

The next step in auditory information pathway after hair cells in cochlea is the information transfer to the axons in the cochlear nerve. Each axon is sensitive to a specific frequency similar to the hair cells. The number of action potentials depends on the loudness of the stimuli. The relation between sound-pressure level and firing rate in each fiber of the cochlear nerve is approximately linear (Hudspeth, 1999).

Acoustical information is processed in parallel pathways, each particular feature of the auditory stimulus is transmitted by different pathways (Fig. 2.10).

In higher levels of the brain, the cells are more complex and specialized. They are sensitive to level differences, interaural time delays, frequency modulated tones, and amplitude changes

(Zwicker and Fastl, 1999). An important capability of our brain is the localization of the sound sources. The brain uses two different binaural cues, the interaural time delay which is detected by neurons in the medial superior olivary nuclei, and the interaural level difference which is detected by neurons in lateral superior olivary nuclei, for detecting the location of the sound. A detailed overview in the field of spatial hearing is given by Blauert (1997).

2.2.2 Psychoacoustics

Attributes of the auditory sensation can be grouped into two categories (Blauert (2003)):

1) Feature attributes: Loudness, pitch, timbre, roughness, etc.
2) Spatial attributes: Distance, direction, spaciousness, etc.

Loudness is the attribute of sound that allows us to organize sounds on a scale from soft to loud. Loudness is a function of sound intensity and frequency. Equal-loudness contours which are to determine at what intensities tones of different frequencies appear equal in loudness as compared to a standard tone (1 kHz) at various intensities are shown in Figure 2.2.

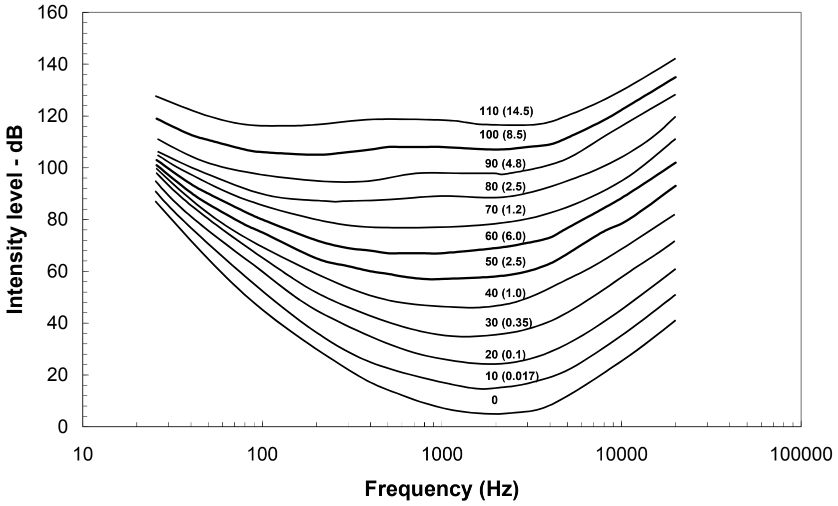


Figure 2.2: Equal-loudness contours (according to Stevens, 1975).

The specific loudness (N') is a function of the critical band rate (z) and the loudness is the integral of the specific loudness over critical-band rate (Zwicker and Fastl, 1999).

$$N = \int_0^{24\,Bark} N'\,dz \tag{2.1}$$

The pitch is that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high. Pitch depends mainly on the frequency content of the sound stimulus, but it also depends on the sound pressure and the waveform of the stimulus." ANSI

(1994). For periodic or nearly periodic signals, the pitch is related to the fundamental frequency of the tone, Stevens called that as tonal pitch. Real-world sounds have many harmonics above the fundamental frequency. The perception of pitch changes with this harmonic content as well. A richer spectrum seems to reinforce the sensation of pitch, making the octave seem more "in-tune" (Gerhard, 2003).

Modulated sounds elicit two different kinds of auditory sensations: at low modulation frequencies up to a modulation frequency of about 20 Hz, the auditory sensation of "fluctuation strength" is produced. At higher modulation frequencies, the auditory sensation of "roughness" is observed. At about 15 Hz, roughness starts to increase, and reaches it's maximum near a modulation frequency of 70 Hz and decreases at higher modulation frequencies. Roughness is influenced by the speed of change, i.e. it is proportional to the frequency of modulation and modulation depth (Zwicker and Fastl,1990).

2.3 The Tactile sense

This thesis includes the investigations regarding to both hand and whole-body stimulations. While the sensations on the hand and the sensations on the whole body do differ in several aspects such as physiology, neurophysiology and psychophysics, the physiological and the psychophysical aspects of the tactile sense are described for both hand and whole-body sensations.

2.3.1 Physiology

2.3.1.1 Touch

The skin is the largest sensory organ in the body. In the average adult, it covers close to $2m^2$ and weighs about 3-5 kg (Quilliam, 1978; Klatzky and Lederman, 2002). Hairless (glabrous) skin, which covers the palmar and fingertip regions of the body, plays the most important role in tactile explorations. There are two types of sensory receptors in the glabrous skin to be regarded: mechanoreceptors and thermoreceptors. Both types of cells are located near the surface of the skin.

Mechanoreceptor cells are responsible for the sensation of vibration and object surface parameters e.g. roughness, shape and orientation of an object. They transduce mechanical energy into neural responses and can be grouped into two categories according to the rate of adaptation: *rapidly adapting* (RA) and *slowly adapting* (SA) mechanoreceptors.

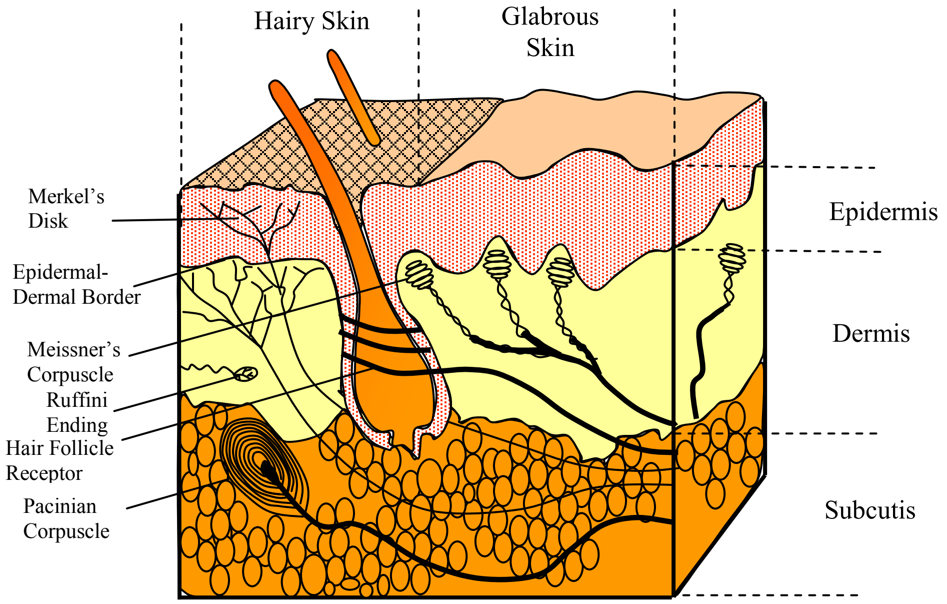


Figure 2.3: Cross-section of the skin.

RA mechanoreceptors, that are Pacinian corpuscles and Meissner corpuscles, only respond when the skin is moving (Barlow and Mollon, 1982). SA mechanoreceptors are Merkel cells and Ruffini endings. Pacinian corpuscle (PC) fibers are found in the deep subcutaneous tissue (Figure 2.3). They sense vibrations also when the skin is compressed and by frictional displacement of the skin, but they are not sensitive to fine spatial discrimination and steady pressure.

Table 2.1: Mechanoreceptor types

	Mechanoreceptor Cells			
	Rapidly adapting		**Slowly adapting**	
	Pacinian corpuscle (PC)	**Meissner corpuscle (RA)**	**Merkel disks (SA-I)**	**Ruffini ending (SA-II)**
Location	deep subcutaneous tissue	dermal papillae	base of the epidermis	dermis and deep subcutaneous tissue
frequency range	50-1000 Hz	10-60 Hz	5-15 Hz	0.4 - 100 Hz
spatial resolution	very poor	poor	good	fair
sensitive to	• vibrations, also when skin is compressed • frictional displacement of the skin	• low-frequency vibrations • detection and localization of small bumps and ridges	• compressing strain • does not have the capabilities of spatial summation	• directional stretching • local force

The PC channel functions as a linear spatial integrator of stimulus energy, similarly to a critical band filter in audition (Verrillo and Gescheider, 1975; Marks, 1979; Gescheider et al., 1994; Makous et al., 1995; Bensmaia and Hollins, 2000).

The sensation of roughness is the principal dimension of texture perception. Some physiological studies have shown that RA mechanoreceptors are responsible for the sensation of roughness (Blake, Hsiao and Johnson, 1997; Connor, Hsiao, Philips and Johnson, 1990; Connor and Johnson 1992). The RA response plays a role in roughness perception of surfaces such as raised dots of varying spacing and diameter. SA-I afferents are mainly responsible for information about form and texture whereas RA afferents are mainly responsible for information about flutter, slip, and motion across the skin's surface.

It is important to note that the thresholds of different receptors overlap, and it is believed that the perceptual qualities of touch are determined by the combined inputs from different types of receptors (Bolanowski et al, 1988; Youngblut et al., 1996). The transmission delay of these receptors ranges from about 50 to 500 msec.

The sensation of force is detected by different receptors in the skin, joints and muscles. Some receptors, which are found in deeper layers of the skin, e.g. merkel disks and ruffini endings were already introduced. The other receptors which are situated in joints are Golgi endings and Ruffini type endings. They respond to joint torque and capsule stretch respectively. Muscle spindle primary, muscle spindle secondary and Golgi tendon organs are other receptors that are located in muscles. Muscle spindle primary is sensitive to muscle length and changes with the length of the muscle. Muscle spindle secondary is sensitive to muscle stretch, and Golgi tendon organ is sensitive to changes in the muscle tension (Pearson and Gordon, 2000).

2.3.1.2 Whole-body vibration

The ***whole-body vibration*** sensation is a complex phenomenon. Several combinations of effects of whole-body vibration occur simultaneously. Therefore different physiological mechanisms such as body proprioceptors, muscles, tendons, joints, touch, pressure, and the vestibular system, can play a role in the sensation of the whole-body vibrations.

Skin on the different parts of the body such as the hands, feet, legs, etc., can contact with the vibrating surface by whole-body vibration exploration and mechanoreceptors. Meisner's and pacinian corpuscles, as explained above, are responsible to obtain the information. Particularly at intermediate and high frequencies, the somatosensory (a combination of the cutaneous[1], kinesthetic[2] and visceral[3] sensory systems) information received from surface end organs may be most important (Griffin, 1990). At intermediate frequencies the forces and movements within the body may yield a kinesthetic sense of motion (Griffin, 1990). Muscle spindle primary, muscle spindle secondary and Golgi tendon organ are responsible to receive this type of information. At low frequencies, the vestibular system provides the information about movement of the head and the position of the head with respect to gravity and any other acting inertial forces. The vestibular system consists of the otolith organs, saccule and utricle, and three semicircular channels Figure 2.4. The human utricle contains about 30,000 hair cells, while the saccule contains some 16,000. Utricle and saccule allow us to sense the direction and speed of linear acceleration (horizontal acceleration: like riding a car; vertical acceleration: like riding in an elevator or plane) and the tilt of the head (Fig. 2.4) and the semicircular channels allow us to sense angular acceleration (Fig. 2.4) by using hair cells which transduce mechanical stimuli into receptor potentials like other mechanoreceptors.

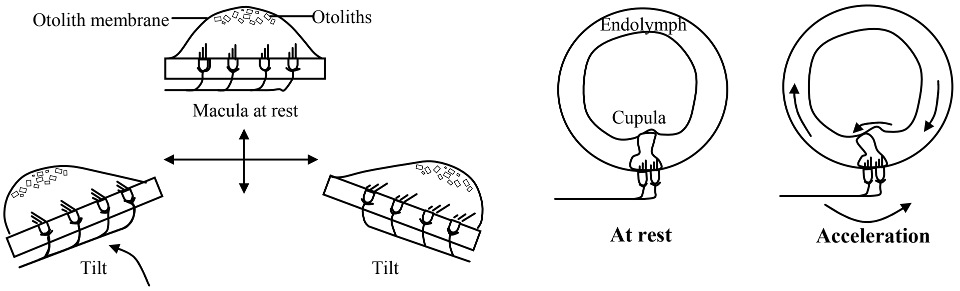


Figure 2.4: Vestibular system.

Temperature is one of the important tactile features, but in this study it will be concentrated on somatosensory, kinaesthetic and whole-body vibration feedbacks. General overviews on the thermal sensations can be found in Jones (1997), Jones and Berris (2002).

2.3.2 Psychophysics

2.3.2.1 Touch

One of the most important spatial features of the human tactile sense is the two-point discrimination threshold, which is the smallest separation at which the subject can consistently discriminate between one and two points (Corkin et al., 1970). Two-point discrimination threshold ranges from 2-5 mm on the fingertips (Louis et al., 1984). Threshold values vary throughout the body surface, Figure 2.5 shows the thresholds of different body parts.

---

[1] Cutaneous sensory system is responsible for the sense of touch, pressure, warmth, cold and pain. Its receptors are in or near the skin.

[2] Kinesthetic is the sense of force and motion. Kinesthetic sensory information is derived from receptors in the muscles, joints and tendons.

[3] Sensations from the internal organs of the body (heart, digestive and endocrine systems etc.) (Griffin, 1990)

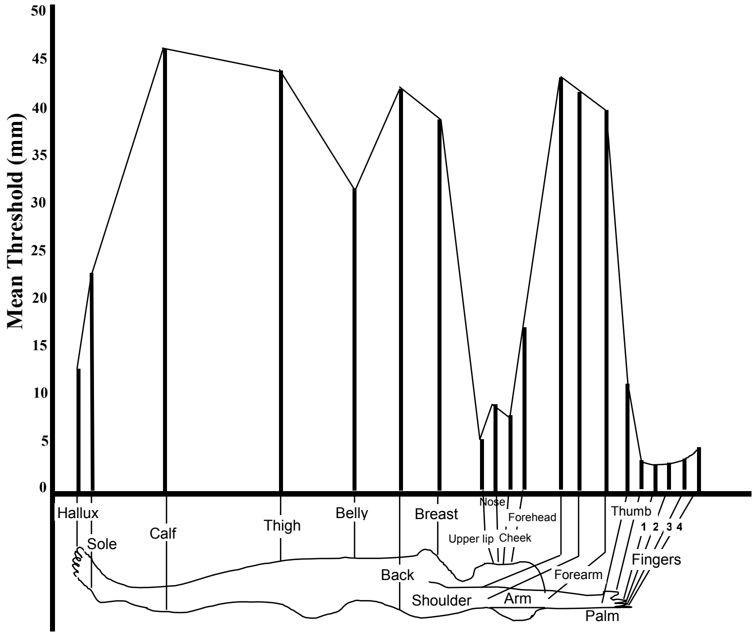


Figure 2.5: Two point discrimination threshold values are shown for various body sites. The data represent the mean value of left and right sides (adapted from S. Weinstein, 1968).

Spatial frequency resolving capacity of the fingertip was measured using grating surfaces and has been found to be 40-50 µm in a spatial period[4] of 0.7-1.0 mm (Morley et al., 1983). The threshold for detecting an indentation of the skin on the fingertip is 11.2 µm (Johansson and Vallbo, 1979b). When the finger scans a surface, a person can detect a raised element such as a dot or a bar on that surface: The threshold height for detecting such as element is 0.1 µm. The other important tactile feature is the shape of an object. Tactile discrimination of a straight edge was studied by Philips and Johnson (1981). Wheat and Goodwin (2001) have also conducted some experiments to quantify the human scaling and discriminating capacity of the curved edges of a flat stimulus. The smallest difference in curvature that could be discriminated by subjects was about 20 m^{-1}.

The temporal resolution of the fingertip can be defined by the successiveness limen (SL), which is the minimum time for which subjects are able to detect two consecutive stimuli. The minimum separation time between two 1 ms pulse stimuli has been found to be 5.5 ms (Klatzky and Lederman 2002). However adult subjects are able to detect vibrations up to about 700 Hz, which suggests that they can resolve temporal intervals as small as about 1.4 ms (Verrillo, 1963).

Similar to the human auditory system, the human tactile system is not equally sensitive to all frequencies. Our skin is sensitive to the frequency range from 8 Hz to 1000 Hz and the

---

[4] Spatial period is the distance between two adjacent ridge onsets.

highest sensitivity is reached in the range of 200-300 Hz. Thresholds for the detection of vibrotactile stimuli measured as a function of sinusoidal frequency at the thenar eminence of the right hand by Verrillo (1963) are shown in Figure 2.6. This figure also shows equal sensation curves which show the vibration magnitudes necessary to produce equal subjective intensity with each vibration frequency.

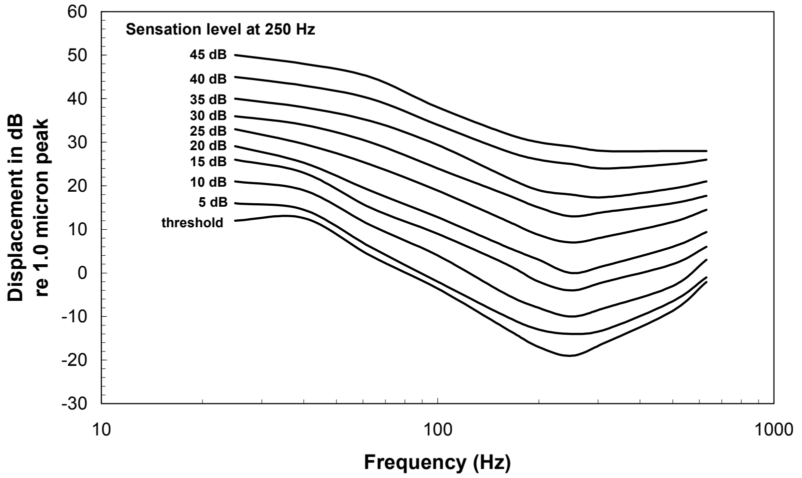


Figure 2.6: Curves of equal subjective intensity plotted as a function of frequency (adapted from Verrillo and Gescheider, 1992).

Subjective scaling of apparent force was studied by Stevens and Mack (1959). They were able to show that the subjective force of handgrip grows with the 1.7 power of the physical force exerted. The Just-noticeable-difference for human force sensing was measured by Jones and found to be 7 % (Jones, 1989).

Burdea and Coiffet (1994) measured for the average person that the index finger can exert 7 N, the middle finger 6 N, and the ring fingers 4.5 N without experiencing discomfort or fatigue. The maximum exertable force from a finger is approximately 30 to 50 N (Salisbury and Srinavasan, 1997).

Tan, Srinivasan, Eberman and Cheng (1994) conducted some psychophysical experiments to define human factors for the design of force-reflecting haptic interfaces. They measured the average pressure JNDs as a function of the contact area (see Table 2.2). It seems that JND was independent of tested body part. Pressure JND decreased by a factor of roughly 4 (from 15.6% to 3.7%) when the contact area was increased by a factor of 16 (from 1.3 cm^2 to 20.4 cm^2). The JND was roughly 0.06-0.09 N/cm regardless of the contact area.

The joint angle JNDs for the wrist, elbow, and shoulder were also measured by them. The JND values are 2,0°, 2,0°, 0.8° for the wrist, elbow and shoulder respectively. The other parameter which was also measured by Tan et al. was the minimum stiffness required to simulate a rigid object (e.g., a wall) without visual information. A rectangular aluminium beam was clamped at a 90° angle to a wall. The subjects closed their eyes and pressed on the wider surface of the beam. The measured average stiffness required to simulate a rigid object was 242 N/cm. They reported the interesting observation as to which, when a subject had reached the "threshold point" where the beam still felt rigid, the displacement caused by the probing was visually detectable. Since the subjects eyes were closed, they were obviously unable to detect this displacement with purely haptic perception.

Table 2.2: The average pressure JNDs (adapted from Tan, Srinivasan, Eberman and Cheng, 1994)

Body site	**Contact Area (cm^2)**		
	1.27	5.06	20.27
Elbow (Volar)	16.7 %	6.2 %	4.0 %
Elbow (Dorsal)	11.3 %	5.2 %	3.3 %
Wrist (Dorsal)	18.8 %	4.4 %	-
Overall Average JND	15.6 %	5.3 %	3.7 %

2.3.2.2 Whole-body vibration

Whole-body vibration usually occurs when the whole environment is undergoing motion and the effect of interest is not local to any particular point of contact (Griffin, 1990). Interest in human responses to whole-body vibration has grown, particularly due to the increasing usage of vehicles, e.g. cars, trucks, and helicopters etc. International Standard 2631 defines methods of quantifying whole-body vibration in relation to the evaluation of vibration perception, human health and comfort, and the incidence of motion sickness. According to ISO 2631, the principal relevant basicentric coordinate systems for seated person are shown in Fig. 2.7.

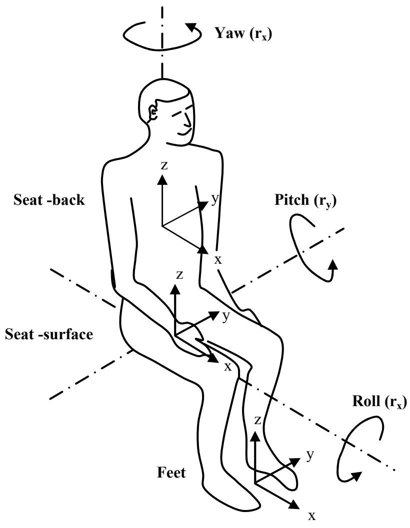


Figure 2.7: A 12-axis basicentric coordinate system for seated person (According to ISO 2631).

Using the coordinate system of ISO 2631, the human body has the same sensitivity for vibrations in x and y directions. The range of frequencies of whole-body vibrations is approximately 0.5 to 100 Hz. Perception thresholds for z-axis whole-body vibration of seated persons can be seen in Figure 2.8. Growth in subjective magnitude (exponent of Stevens' power law) of whole-body vibration was measured by Howarth and Griffin (1988). In the z-direction, the exponent of growth starts with 1.21 at 4 Hz and it increases up to 1.29 at 60 Hz in the y-direction, it is 0.68 at 4 Hz and it increases up to 1.69 at 60 Hz.

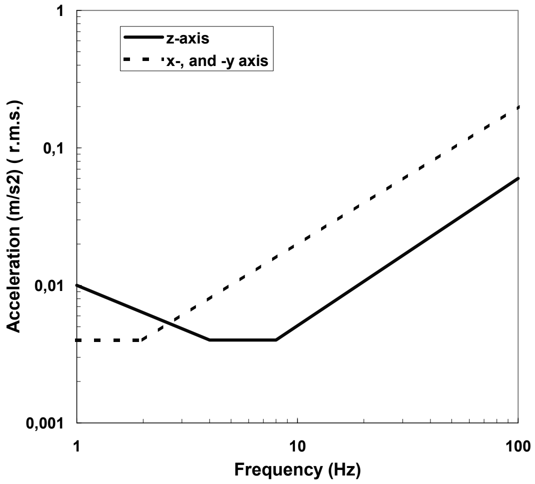


Figure 2.8: Perception thresholds for vertical whole-body vibration of a seated person.

Using an engineering approach, the dynamical response of the body to whole-body vibration can be modeled by using masses, dampers, and springs. Such models are useful in understanding the nature of the body movements and to provide information necessary for the optimization of isolation systems and the dynamics of other systems coupled to the body (Griffin, 1990).

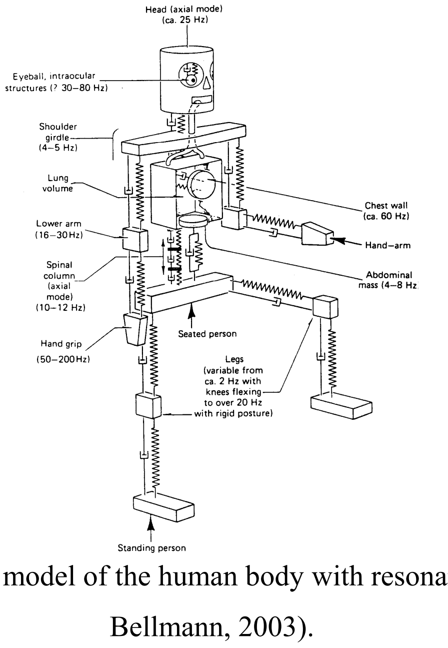


Figure 2.9: A simple theoretical model of the human body with resonance frequencies. (adapted from Bellmann, 2003).

The neural firing in the vestibular nerve is proportional to head velocity over the range of frequencies in which the head commonly moves, that is, 0.5 to 7 Hz. However, the semi-circular canals provide the best response in the first second or so, and output decays exponentially with a time constant of about 7 s (Youngblut et al., 1996). There is an interesting phenomena in the perception of angular acceleration. If sustained acceleration (10 – 20 seconds) takes place in one direction, the fluid in the appropriate canal also remains continually displaced and after a brief moment the hair follicles will return to the vertical position, therefore the brain will perceive that the acceleration has stopped. However, if the angular movement has a constant acceleration of under 2 degrees per second, humans can not sense any rotation at all. The perception of angular motion varies with frequency, falling at around 0.2 log unit/decade between 0.1 and 1.0 Hz and falling at -1 log unit/decade below 0.1 Hz. For stimuli shorter than 15 seconds, this perception of angular motion is related to the time, t, taken to detect angular acceleration, and the product has an average constant value of 3.7°/sec. For sustained rotational stimulation with prolonged acceleration (such as what occurs in an aircraft), the sensory threshold for angular rotation is determined by the magnitude of angular acceleration rather than the velocity change and the mean threshold for angular accelerations of the head about the z axis has been demonstrated as 0.32°/s with a range of 0.05 to 2.2°/s. With respect to the perception of linear acceleration, for a linear oscillation at approximately 0.3 Hz in the horizontal plane, the mean threshold was around 0.03 m/s^2 for oscillations in the x, y axes and around 0.06 m/s^2 for oscillations in the z body axis. The common peak angular velocity for passive nodding of the head, such as occurs during walking or running, is ±10°/s. Volitional head movements usually exhibit a peak angular velocity of at least 100°/s but may be as high as 500°/s (Benson, 1990; Youngblut et al., 1996).

2.3.3 Information processing in the somatosensory system

When perceiving an object, our brain integrates different characteristics of the object which are obtained and transmitted to the brain by different receptors and forms a coherent image of a single object. The somatosensory system integrates the object properties, such as shape, texture, mass, and temperature. In neuroscience this ability is known as “stereognis.” Various somatosensory information is sensed by different receptors and conveyed by anatomically separate pathways (Gardner and Kandel, 1999).

Figure 2.10 shows the neural information transfer pathways of tactile sensations. Sensory receptors and primary sensory neurons responsive to pressure or vibration are connected to clusters of cells in the dorsal column nuclei and thalamus (Gardner and Kandel, 1999).

In the somatosensory cortex, Broadman's area 3a receives inputs from muscle and stretch receptors in the deep tissue. Information from slowly and rapidly adapting receptors in the skin is transmitted to the different columns of area 3b regarding the adaptation type of the receptors (Figure 2.10). Neurons in area 3b respond to a particular form and amount of energy at a specific location of space and together reproduce its shape (Philips et al., 1988; Gardner and Kandel, 1999). Area 1 receives information from the rapidly adapting receptors, which have larger receptive fields than the cells in area 3b. Broadman area 2 receives pressure and joint position inputs from the mechanoreceptors which are in the underlying muscles and joints (Sur et al., 1984). Areas 1 and 2 are sensitive to the orientation of edges, the direction of motion across the skin, the surface curvature of objects, or the spatial arrangement of repeated patterns that form textures (Gardner and Kandel, 1999).

The posterior parietal cortex (Broadman's areas 5 and 7) is responsible for information integration. Integration of tactile information from mechanoreceptors in the skin with the pressure and joint position inputs from the muscles and joints takes place in area 5. Area 5 is also responsible for the integration of the inputs from two hands. Multimodal integration of the somatesensory and visual inputs takes place in area 7, which is important for the eye-hand coordination and movements.

Gardner and Kandel (1999) reported the factors which are involved in the integration tactile features:

- The size of the receptive field becomes larger at each level of processing, so that eventually the entire object rather than a single edge is sensed by a neuron.
- The profile of activity in the active population of neurons changes through the action of inhibitory networks. Inhibitory activity serves to sharpen the peak of activity within the brain. In this manner, when the skin is touched at two or more points simultaneously, inhibition makes the identification of different points easier
- At successive levels of sensory processing in the cortex, individual neurons respond to more complex inputs
- The sub modalities converge on individual neurons in association cortical areas

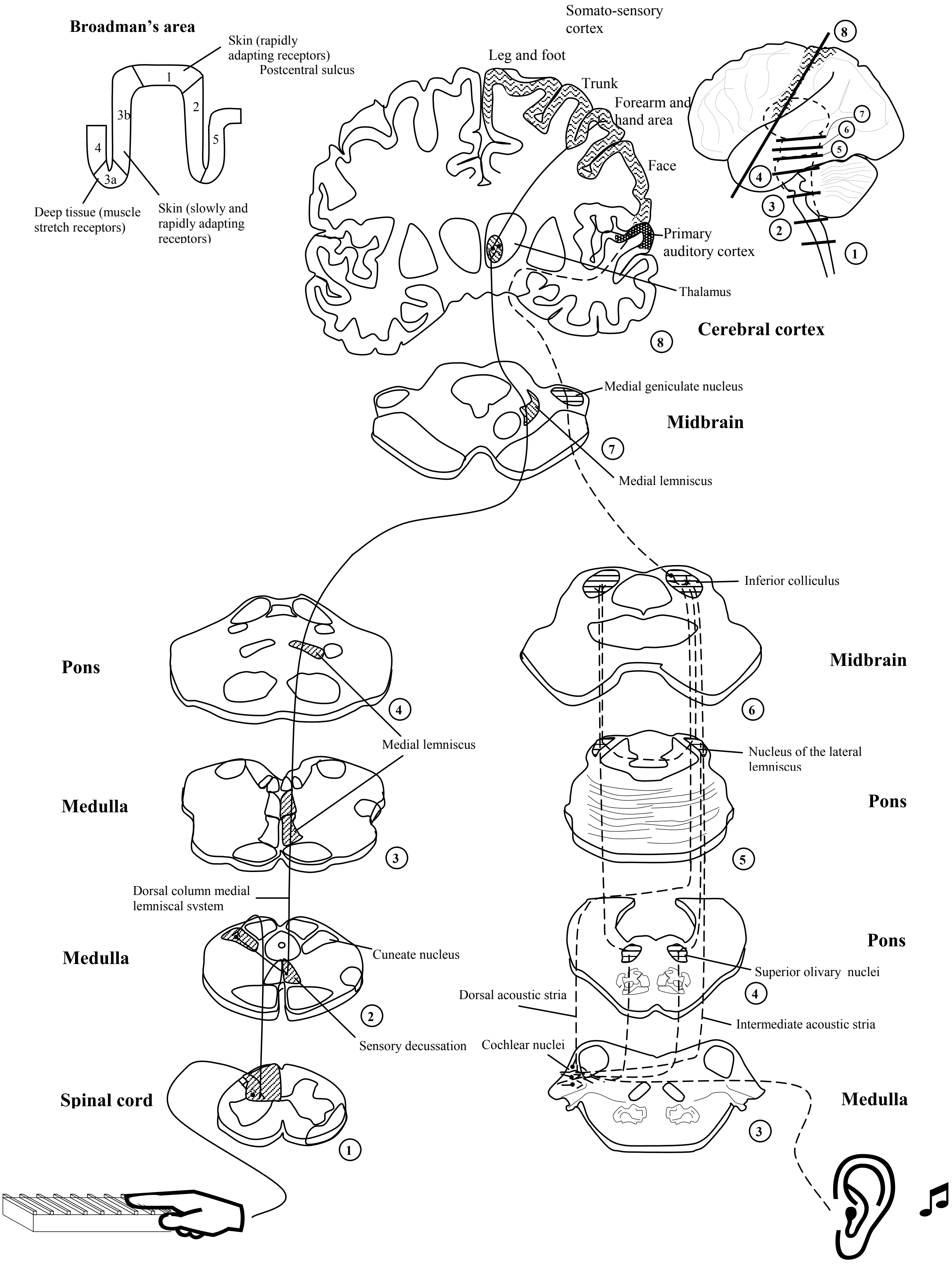


Figure 2.10: Afferent auditory and tactile pathways (adapted from Hudspeth, 1999; Gardner and Kandell, 1999).

The somatosensory information is processed in parallel by distinct areas of the cortex. Tactile sensory information must be compared with more recent information being processed at the early stages. Responses in areas 3a and 3b occur 20 ms after touch or movement, and the more posterior cortical areas receive sensory information at longer latencies, and process stimuli which were presented 30-100 ms earlier. The knowledge on the visual binding process suggests that the various stimulus features may have been bound by synchronising firings in different cortical areas (Singer and Gray, 1995).

2.4 Multisensory responses in the brain

Although a lot of areas of the brain receive converging multisensory inputs, most of the studies of multisensory processing and integration have focused on the Superior Colliculus (SC), which is situated in the midbrain.

Multisensory stimuli evoke responses on the neurons in the deeper layers of SC that vary with the relative timing, location, and the intensity of the stimuli. Higher discharge rates are observed when stimuli of different modalities are delivered in close temporal and spatial proximity (King, 2004). When stimuli from different modalities are not matched in space and/or time, multisensory information can produce a suppression of neural responses to non-meaningful, or distracting, signals (Stein et al., 2004).

The various receptive fields of SC neurons are organized into overlapping auditory, visual, and somatosensory maps, in effect creating a multisensory map of space (Stein and Meredith, 1993). The multisensory map of space is used to discriminate single-origin co-occurring inputs from invalid ones and to detect event singleness (one multimodal event or several unimodal events) (King, 2004).

In a study reported by Stein et al. (2004), the processes which were performed by the SC multisensory neurons were investigated. The results showed that different crossmodal stimulus combinations resulted in 15% superadditive (multisensory response is greater than the linear summation of the unimodal responses), 21% subadditive (multisensory response is less than the sum of the unimodal responses), and 63% additive (multisensory response is equal to the linear summation of the unimodal responses) multisensory response. For the approximately 80% of the multisensory neurons in the SC, the number of impulses evoked by a combination of stimuli is significantly different from that evoked by the most effective of these stimuli alone.

The delays for transduction from the receptors to the cells in the SC were about 16 ms for auditory stimulus and 23 ms for tactile stimulus (adult cat) (Wallace and Stein, 1997).

The development of multisensory neurons and the multisensory integration was examined in the deep layers of the SC of kittens (Wallace and Stein, 1997; Stein et al., 1973). The results showed that somatosensory-responsive neurons are present at birth (and before), auditory-responsive neurons appear late in the first postnatal week, visual-responsive neurons appear in the third postnatal week, and multisensory neurons first appear toward the end of the second postnatal week (auditory- somatosensory neurons appear at 12 day of post natal life (dpn), followed by visual multisensory neurons at 20 dpn).

Studies in infants have shown that infants can start perceiving the intersensory relations based on common synchrony, duration, presentation rate, and rhythm during the first few months (Spelke, 1979; Lewkowicz, 1986, 1996). They can put together the visual and auditory components of speech by 4 months of age (Kuhl and Meltzoff, 1982).

Chapter 3

Development of an Experimental System for Investigations on Auditory-Tactile Interactions

3.1 Introduction

Virtual reality provides the user with real-time multi-sensory interaction with the computer-generated environment. Virtual-reality generators have proved to be potent tools for research and development. They allow for flexible and economic presentation of complex experimental scenarios which can be modified without any physical effort (Blauert, 2000). These advantages make also virtual reality applications very attractive for the commercial market (training, education, medical applications, and entertainment).

Although the first virtual environment concept was introduced in 1965 (Sutherland, 1965), until 1990's the human computer interaction was rather limited. Many of the virtual reality applications concentrated on visual feedback. Representation of an interactive auditory virtual world through headphones or loudspeakers was possible just at the end of 1980's. Tactile feedback was almost an input modality using a keyboard, a mouse or a joystick as an input device. One reason for the lack of virtual tactile information could be that the visual and auditory modalities are more dominant than the sense of touch in humans. The other reason might be the computational and technical complexity required to generate convincing touch sensations in a computerised environment.

This chapter begins with a description of auditory virtual and haptic virtual displays and auditory-tactile virtual environments. Subsequently an experimental system for the investigation of auditory-tactile interaction which was developed and used for the experiments in this study will be introduced.

3.2 Auditory virtual displays

The purpose of an Auditory Virtual Environment (AVE) is to create situations in which humans have auditory perceptions that do not correspond to their real environment, but to a virtual one (Blauert, 1990). AVE's are based on a technique called "Binaural Room Simulation". This method facilitates listening into spaces that only exist in the form of computer models. Therefore temporal and spectral properties of the virtual sound field and spatial distribution of sound sources are reproduced.

Early interactive real-time binaural auditory display systems had been presented by Boerger, Laws, and Blauert (1977); Boerger, and Kaps (1978). In these systems interaction was already provided by monitoring the position of the listeners heads and modifying the acoustic input signals to their ears accordingly. The first interactive real-time auditory-tactile virtual environment was developed by Institute of Communication Acoustics in the framework of an EU-funded project "SCATIS" (SCATIS Project Report 1996, Blauert, Lehnert, Sahrhage and Strauss, 2000). The real-time sound field model of SCATIS was running on two dedicated Silicon Graphics R4000-Indigo Workstations and 40 24-bit fixed point Motorola DSP56002 (40 MHz) were used for the required signal processing tasks.

Fundamental structures of an AVE were reported by Silzle, Novo and Strauss (2004). It consists a sound generator, a binaural room simulator module (environment model), and an auralization unit. Sound source signals can be pre-recorded or synthesized utilising different synthesizing methods (more explanation can be found in Section 3.5.1). Binaural room simulator module consists:

- a model of the geometrical and acoustical properties of the environmental boundaries,
- a model of the effects of the air on the sound propagation,
- a model of the complex directivity characteristics of the sound sources and
- a model of the wall-reflection characteristics.

Two different reproduction formats can be used in AVEs: headphones and loudspeaker formats. If a Head-Related Transfer Function (HRTF) based reproduction format will be used, auralization unit contains HRTF filters.

The dynamic behaviour of the virtual reality systems can be evaluated using two parameters: "system-latency" and "frame rate (update rate)" (Blauert et. al., 2000). System latency can be defined as the time elapsing between the application of a stimulus and the first indication of a response. Wenzel (1997) gave an example for the calculation of the AVE system latency as the time elapsed from the transduction of an event or action, such as movement of the head, until the consequences of that action are available to the listener, e.g., as a change in the relative location of a sound source. The system latency of the SCATIS's auditory pathway is 80 ms on average (Blauert, 2000).

The frame rate can be defined as the frequency at which input changes are processed (update rate of a display) (Pellegrini, 2002). SCATIS has achieved update rates of about 60 Hz.

3.3 Haptic virtual displays

Haptics comes from a greek word “haptesthai” meaning “the science of touch” (Webster, 1985). In recent years, its meaning extended to the scientific study for applying tactile and force feedback sensations of humans into the computer-generated world. Haptic devices can be grouped in two categories: input and output devices.

The required input information for haptic interfaces (virtual environments) are the position and the orientation of the hand and fingers. Most input devices, such as keyboard, mouse, and joystick, are very well known. However these devices can not transmit information related to the fingers and mostly they are applicable only for 2D applications. A data glove is a type of input device that can be used to give spatial position or movement information of the hand and fingers to the computer. Different tracking sensors, e.g. mechanical, optical, acoustic (ultrasonic), magnetic etc. are used in data gloves.

Haptic output devices can be categorized based on the feedback type. There are mainly two feedback types, tactile and force feedback. *Tactile feedback* is defined as “the sensation applied to the skin, typically in response to contact or other actions in a virtual world” and *force-feedback* is defined as “the sensation of weight or resistance in a virtual world” (Burdea, 1996).

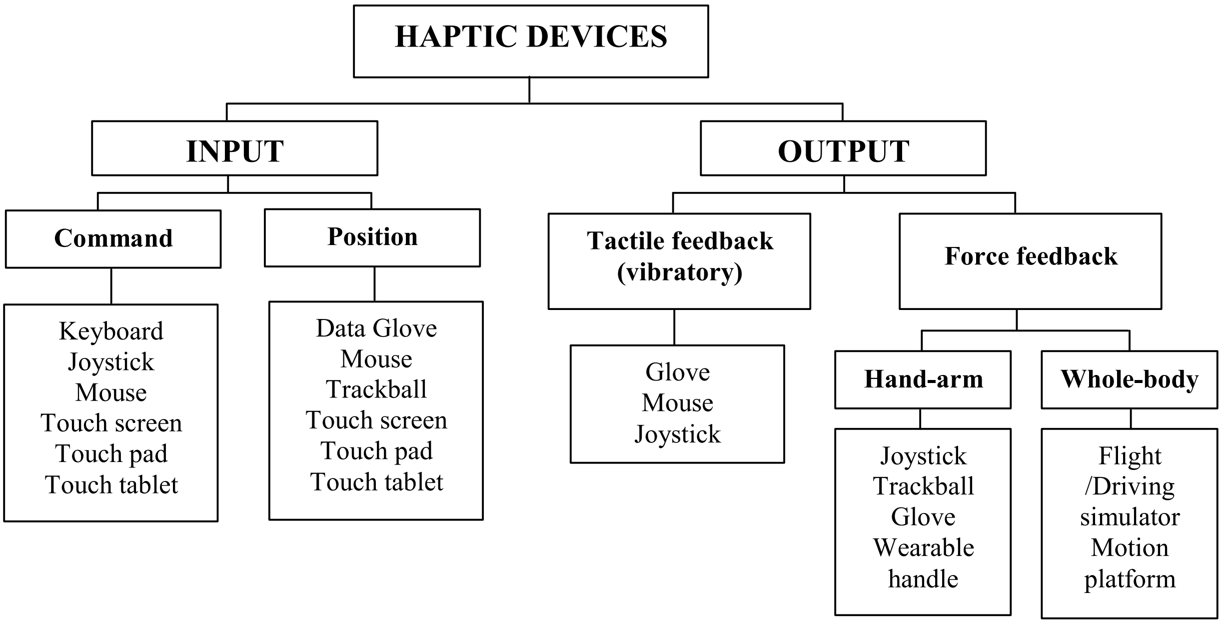


Figure 3.1: Categorization of types of haptic devices

3.3.1 Tactile feedback

In our daily life, if we touch an object, we receive information from the surface of the object. Texture perception is a critical part of the human tactile exploration. The texture of a virtual surface can both increase the sense of realism of an object as well as convey information about object identity, type, location, function, and so on (McGee, 2002). Different tactile feedback interfaces are used to simulate the texture information and mechanical vibrations in virtual environments. The most widely used method is the vibrotactile stimulation technique. Through this technique, voice coils or micro-pin arrays convey texture information on the user's fingertip applying vibrations. The advantages of the vibrotactile devices are the small size and the low weight. The disadvantages are the high complexity and cost. The pneumatic stimulation technique sends compressed air to the user's finger through small arrays, or a ring which consists of a rubber balloon and where the rubber balloon makes vibrations according to air pressure fluctuations. The advantages of the pneumatic stimulation technique are its simplicity and lower cost compared to vibrotactile devices. Less suitability for portable haptic interfaces and high weight are the disadvantages. Another technique which is used to present texture information is the electrotactile stimulation. In this technique, texture information is presented to the fingertip of the user by transmitting small currents through the electrodes. Nonreactive metallic electrodes of titanium, gold, platinum, silver, or stainless steel were used for the electrotactile stimulation (Kaczmarek and Bach-y-Rita, 1995). The advantages and disadvantages of these generation techniques are summarized in Table 3.1.

Table 3.1: The major strengths and weaknesses of different tactile stimulation techniques according to Burdea (1996)

Display Type	Advantages	Disadvantages
Vibrotactile (e.g. voice coil)	• Small size • Low weight • High temporal resolution	• High complexity • Poor spatial resolution • Noise
Pneumatic	• Low cost • Simplicity	• Less suitable for portable applications • Poor spatial and temporal resolution • Limited bandwidth
Electrotactile	• Lower power consumption • No moving parts • Low weight • No noise	• User discomfort and even pain when using improper electrodes or driving current

3.3.2 Force feedback

Force feedback devices apply physical forces and torques on the user's hand or finger. Different types of force feedback generation methods are used in haptic feedback devices (e.g., electromagnetic motors, hydraulics, pneumatics).

Before comparing the force-feedback actuators, which are used today, it is important to know the requirements for the ideal actuators. Jex (1998) described four criteria for an ideal haptic interface:

- ability to stimulate a piece of light balsa wood, with negligible inertia, friction, or vibrations being perceived by the operator,
- it should be able to simulate a crisp hard stop,
- and also simulate Coulomb friction without sponginess or jitter,
- simulate a mechanical centering detent with crisp transitions and no lag.

These criteria are stringent and today's technology is not so far for developing such interfaces. Burdea (1996) reported the physical parameters which are important for the haptic interface design. Haptic actuators should be light, powerful, simple, cheap, safe, and compact, therefore designers should maximize the power-to-mass ratio and power-to-volume ratio. Another important aspect is that the interface should not apply forces to the user's hand, when no physical interaction exists. This means that the interface should have minimal static friction and low actuator inertia. System bandwidth (sampling rate) is an important parameter for the stiffness. High stiffness requires high sampling rate. A comparison list is given in Table 3.2 according to Burdea (1996). DC Motors have low power-to-mass ratio, therefore they are not proper for portable haptic interfaces. Hydraulic actuators have high power-to-mass ratio and good bandwidth, however, they are dirty and expensive.

Table 3.2: Comparison of force feedback interfaces according to power-to-mass ratio and mechanical bandwidth

Actuator Type	Power-to-mass ratio	Mechanical Bandwidth
DC Motor	Low (it is inadequate to be used on a portable haptic interface)	Good
Brushless Servomotor	Low	Good
Pneumatic	Middle range	Low
Hydraulic	Very high	Good
Shape memory metals	High	

3.3.3 Whole-body vibrations

Motion platforms distribute whole-body vibration (tactile information) and feedback forces over the user's body (Burdea, 1996). They are driven by hydraulic, pneumatic or electromagnetic power. People are usually exposed to whole-body vibration while traveling. Motion platforms are used for different purposes, e.g. driving simulator, flight simulator. The performance of the motion platforms can be expressed by frequency and displacement (or force) limits. The frequency ranges of the hydraulic and electrodynamic type exciters are given in Fig. 3.2 (Booth, 1958). While hydraulic exciters are more convenient for low frequency and high load applications, electrodynamic exciters are suitable for high frequency applications. Pneumatic exciters have smaller force capability than hydraulic exciters. But they have lighter construction and are cleaner. The main disadvantage of the motion platforms is the high level noise which occurs when they generate signals. (Hydraulic exciters are louder than electrodynamic exciters.).

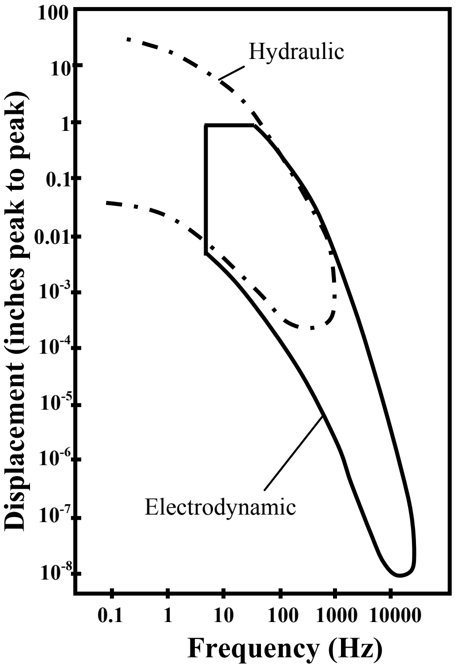


Figure 3.2: The frequency range and displacement limits of motion platforms.

3.3.4 Tactile virtual environments

Computational aspects of the haptic virtual environments show similarity with auditory virtual environments. There are different applications which are based on workstations or PCs (multicomputers). The PC or workstation should receive the input data (position or velocity) from the sensors (e.g. data glove), then calculate the interaction forces or vibration and finally send to actuators force-feedback or vibration information. Physical models are used to calculate interaction forces or vibration. These physical models are based on Newtonian physical laws. The first step of the physical modeling is the collision detection (Burdea, 1996). Required information related to collision are when and where the collision occurs, and determination of the interpenetration between the two objects. Different methods are used to detect the collision, e.g. Cyrus-Beck algorithm (Moore and Wilhelms, 1998), Bounding boxes (Foley et al., 1990), multi-body collision detection algorithm (Cohen et al., 1995). According to collision type, different interactions can occur, e.g. grasping, fall to the ground, hitting, etc. and the physical parameters which should be calculated changes according to the interaction type. For example, by grasping the virtual hand applies forces on the grasped object, and required physical information are tangential and normal components of the grasping forces which can be calculated using interpenetration volumes and stiffness of the object, and grasp stability. Rutgers Master I is an example of the single-user network-distributed architecture, which consists of four workstations (Burdea, 1996). One of the workstations reads information from a Data Glove, maintains virtual-object state information, performs collision detection, collision response, and physical modeling (surface deformation and contact forces) and sends force-feedback information to the control interface.

The system latency of the tactile VE-subsystems are estimated to be in the range of 40 ms to 60 ms (MUVII Project).

3.4 Dynamic behaviour of the multimodal virtual environments

In multimodal VEs, each unimodal information can be delayed with respect to the action of the user. For example, in auditory-tactile VEs, both auditory and tactile feedback can be delayed with respect to the action, when both information are delayed by the same amount of time, auditory and tactile events still are synchronous. A multimodal VE system latency can be defined as the time elapsing between the unimodal feedback occurrences (e.g. auditory-visual, auditory – tactile, visual – tactile). If a user hits an object with his/her hand, the central controller should receive information related (e.g. applied velocity, location of the event, location of the listener head, etc.) to the hitting event, and transmit this information to the

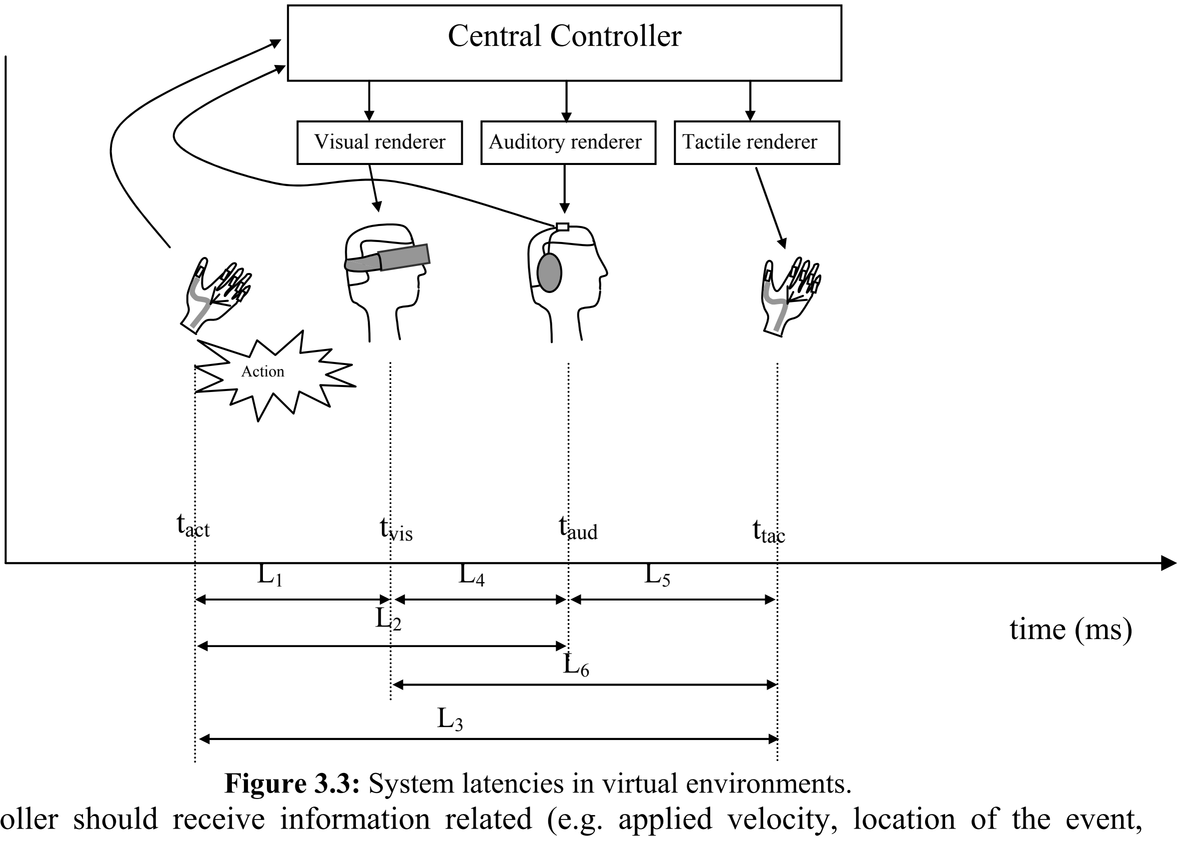


Figure 3.3: System latencies in virtual environments.

auditory, tactile, and visual renderers. Each renderer makes the required calculations and then the tactile renderer transmits the force-feedback information to the tactile actuator, the auditory renderer sends sound data to the loudspeakers or the headphones, and the visual renderer transmits the data to a head-mounted display or a projection screen.

The latencies which are important for the VE designer in the design of VE generators are shown in Figure 3.3. L_1, L_2, and L_3 are the latencies of the each unimodal subsystems; visual, auditory, tactile, respectively. L_4, L_5, and L_6 are the latencies between modalities; visual-auditory, auditory-tactile, visual-tactile, respectively. t_{act} time which the action occurs, t_{vis} arriving time of the visual information which is related to action, to the user, t_{aud} arriving time of the auditory information to the user, t_{tac} arriving time of the tactile information to the user.

An approximate latency for an auditory-tactile virtual environment can be estimated to be in the range of 20 ms to 40 ms.

Blauert et al. (2000) reported that although the frame rate of the auditory subsystem is about 60 Hz, if the tactile/thermal subsystem was integrated into the VE generator, the frame rate would decrease drastically to about 12 Hz at maximum utility.

3.5 An experimental system for investigations on the auditory-tactile interaction

This section of the chapter describes the concept and the characteristics of an experimental system which has been developed for this study and will be used for conducting further experiments (Figure 3.13).

3.5.1 Auditory subsystem

For the *auditory subsystem* of this experimental system, matlab-written algorithms and some parts of the software tool, IKA-SIM, which was developed by Strauss and published by Silzle, Novo and Straus (2004) from the Institute of Communication Acoustics, was used. IKA-SIM runs on Microsoft’s Windows NT/2000/XP platforms and basically does not require any special hardware. The structure of IKA-SIM consists of a controller which collects information on the listener status (e.g. position, orientation), and decides about the appropriate system reaction, a sound field model, which calculates the direct and reflected sound paths using geometric and acoustic data, a signal processing module, which adds a delay according to the propagation time, makes spectral weighting, and spatial filtering corresponding to the sound direction of incidence. At the final stage auditory information can be presented to the listener either through headphones or loudspeakers (in this study only headphones presentation will be used). The dynamic behaviour characteristics of IKA-SIM were reported by Silzle, Novo and Strauss (2004). The end-to-end system latency of IKA-SIM (including an optical renderer) is about 60 ms, and the frame rate is higher than 30 Hz. IKA-SIM can calculate, for a headphone reproduction on a standard computer with an Intel P4 processor with 2.6 GHz, a total of 40 to 60 sources or mirror sources in real-time (for further details about IKA-SIM see Silzle, Novo and Strauss, 2004).

As digital sound source signal, both recorded and "interactive" sounds were used. Due to the development of a haptic feedback as an input and output display in a computer environment, sounds which are generated in real time according to the haptic interaction of the user with sound-producing virtual objects are required. To synthesize this type of "interactive" sounds, different methods are suggested.

Boundary Element (BEM), Finite Element (FEM) and Digital Waveguide Modeling Methods are frequently used to generate sound (Smith 2000). These numerical methods are available for sound design problems but they are not capable of real time feedback in computer-human interaction applications, e.g. virtual reality. Most sounds currently used in VE are sampled from real sound sources or synthesized by MIDI devices (Klatzky, Pai and Krotkov 2000). However, huge memory capacities are required to achieve a continuous interactive sound synthesis.

Synthesizing sounds using physical models is another generation method and recently used in virtual reality applications (Cook, 2002). The advantages of this method are its flexibility to generate almost all possible situations by using input parameters, e.g. physical attributes of the objects and spatial properties and the significant reduction of the required memory capacities. The disadvantage of this method is that it requires a high level of computational power and long calculation time. However, recently developed faster processors and algorithmic advances in digital signal processing allow to use physical models in real time applications.

The underlying algorithms can be classified as additive and subtractive synthese. The approach for additive synthesis is based on the summation of sinusoids (Fig. 3.4). The origin of this approach is the Fourier transform. According to this approach, any complex sound can be created by combining multiple sine waves at different frequencies, phase angles, and amplitudes. In Figure 3.4, the rectangular blocks indicate the amplitude (Amp_n) and frequency ($Freq_n$) functions of the filters, the half-ellipse blocks indicate the sine signal generators.

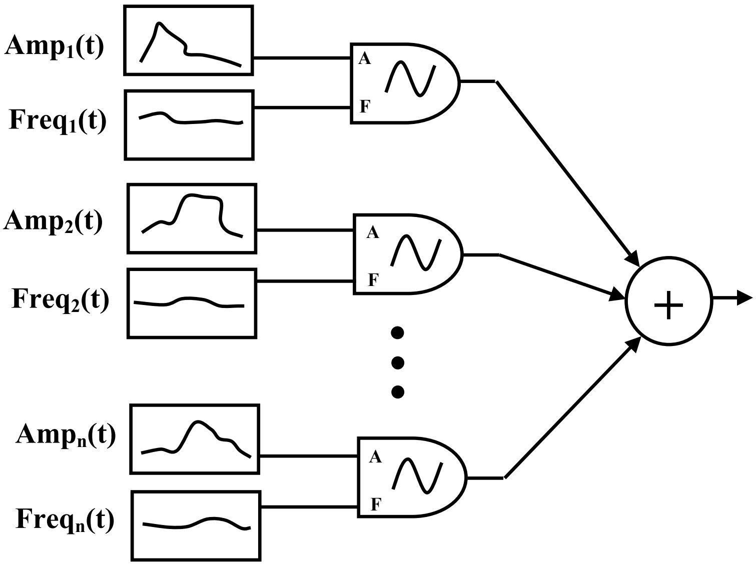


Figure 3.4: Additive Synthesis.

The subtractive synthesis starts with a broadband input signal. Subsequently some frequency components are filtered out from the input signal to achieve the target sound (Figure 3.5). White noise and a chain of periodic impulses are the mostly used source signals with this method. The advantage of subtractive synthesis method is the easiness to work and its suitability to model noisy and percussive signals.

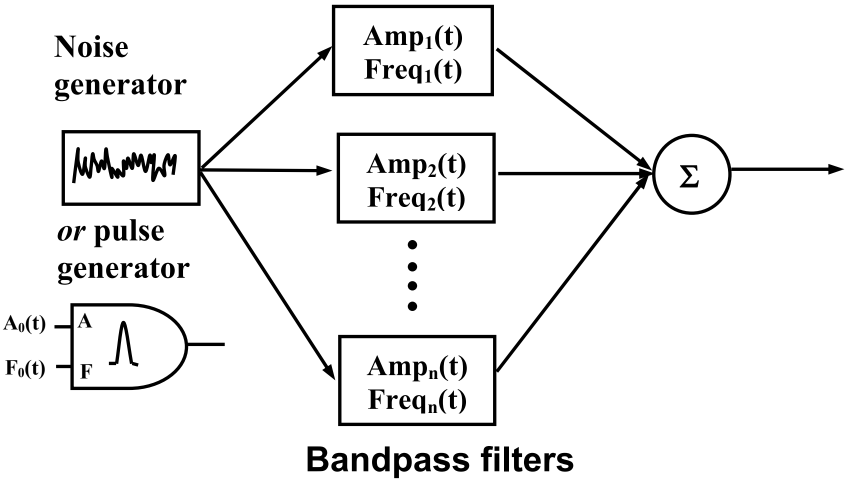


Figure 3.5: Subtractive Synthesis.

3.5.1.1 Touch-induced scraping sound synthesis

Taking into account the further investigations which will be conducted in the framework of this thesis, one of the interactive multimodal events, namely “scraping” was selected and two different situations were modelled: fingertip scraping across the surface of a grooved block and fingertip scraping across the surface of a sandpaper.

3.5.1.2 Physical modelling of touch-induced scraping sounds

Grooved wood

During the fingertip scraping across the surface of the grooved wood, the ridges experience the force which is applied by the fingertip and their movement is transmitted to the block (Fig. 3.6a). The vibrations of the wooden block and the ridges are the predominant sources of the noise. The fundamental frequency component is proportional to the ridge number and the scraping velocity

$$f_{r1} = \nu\, r / L \qquad (3.1)$$

where r is the total ridge number, L is the length of the block and ν is the scraping velocity.

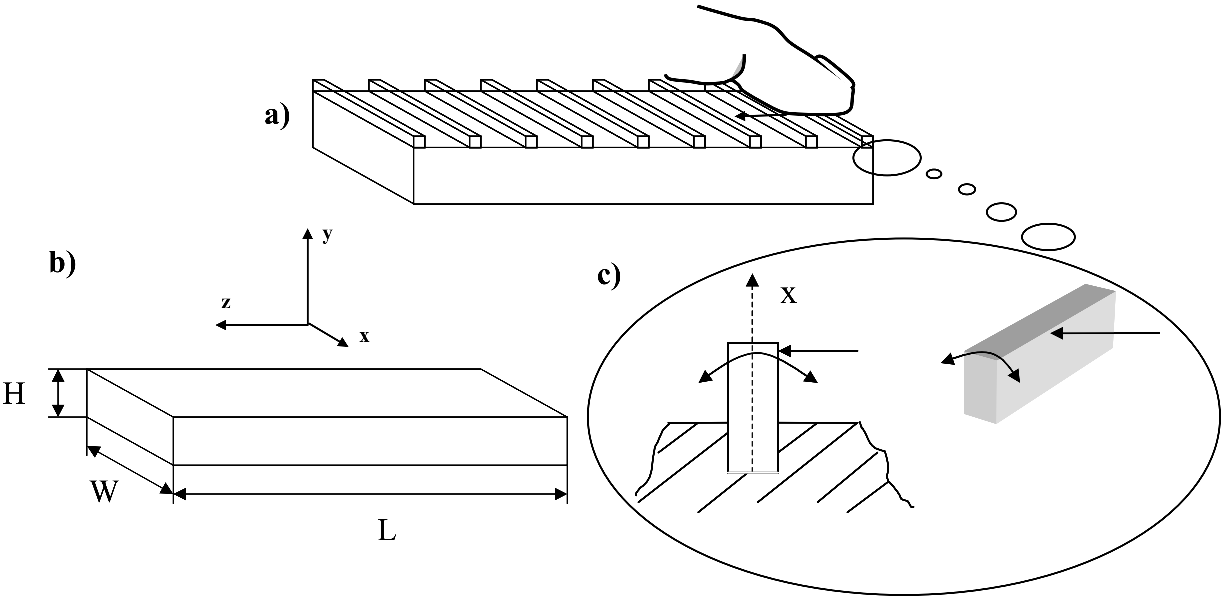


Figure 3.6: a) Fingertip scraping across the surface of the grooved block, **b)** block as a rectangular resonator, **c)** ridge as a bar

If we assume the ridge as a bar with clamped ends at its bottom ($y=0, \partial y / \partial x = 0$) and free ends at its top (Fig. 3.6c), the second frequency component is related to the ridge oscillations (Fletcher and Rossing, 1991);

$$f_{r2}(n) = \frac{\pi K}{8L^2}\sqrt{\frac{E}{\rho}}, \quad n = 1.194^2, 2.988^2, 5^2, \ldots, (2n-1)^2 \qquad (3.2)$$

If we assume the block as a rectangular resonator (Fig. 3.6b), the physical description of the behaviour of the resonators oscillations can be defined with four principal modes (Fletcher and Rossing, 1991, Lakatos, McAdams and Caussé, 1997):

- Longitudinal modes (along the z axis);

$$f_L(n) = \frac{n}{2L}\sqrt{\frac{E}{\rho}}, \quad n = 1,2,3,\ldots.., \tag{3.3}$$

where E is Young’s modulus, ρ is the density of the material.

- Torsional modes;

$$f_T(n) = \frac{n\alpha}{2L}\sqrt{\frac{E}{2\rho(1+\nu)}}, \quad n=1,2,3,\ldots.., \tag{3.4}$$

where υ is Poisson ratio, α is the ratio of the block width to its height.

- Transverse bending modes in the y-z plane;

$$f_H(n) = \frac{n\pi H}{8L^2\sqrt{12}}\sqrt{\frac{E}{\rho}}, \quad n=3.011^2, 5^2, 7^2, \ldots.., \tag{3.5}$$

where H is the height of the block.

- Transverse bending modes in the x-z plane;

$$f_W(n) = \frac{n\pi W}{8L^2\sqrt{12}}\sqrt{\frac{E}{\rho}}, \quad n=3.011^2, 5^2, 7^2, \ldots.., \tag{3.6}$$

where W is the width of the block.

Sandpaper

During the fingertip scraping across the surface of the sandpaper, the grits experience the force which is applied by the fingertip and then their movement is transmitted to the paper (Figure 3.7). The first frequency component is proportional to the grit number (g) and the scraping velocity;

$$fg = g\,\nu / L_x \tag{3.7}$$

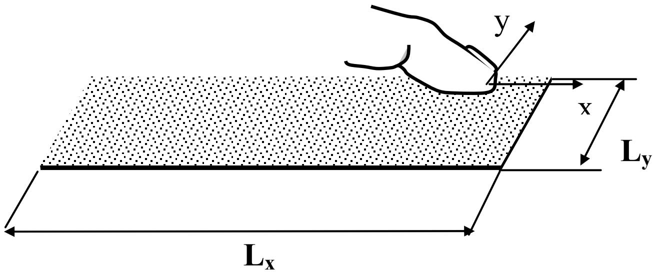


Figure 3.7: Fingertip scraping across the surface of the sandpaper

If we assume the sandpaper is a rectangular membrane, principal vibration modes can be derived as:

$$f_{mn} = \frac{c}{2}\sqrt{\left(\frac{n_x}{L_x}\right)^2 + \left(\frac{n_y}{L_y}\right)^2} \tag{3.8}$$

where c is the sound velocity.

3.5.1.3 Synthesis of the touch induced sounds using physical models

To synthesize touch-induced scraping sounds, the subtractive synthesis method was used. The reason for this selection was that touch-induced scraping sounds remind either of a pulse train or of white noise and the subtractive synthesis approach models the mechanisms of sound generation advantageously.

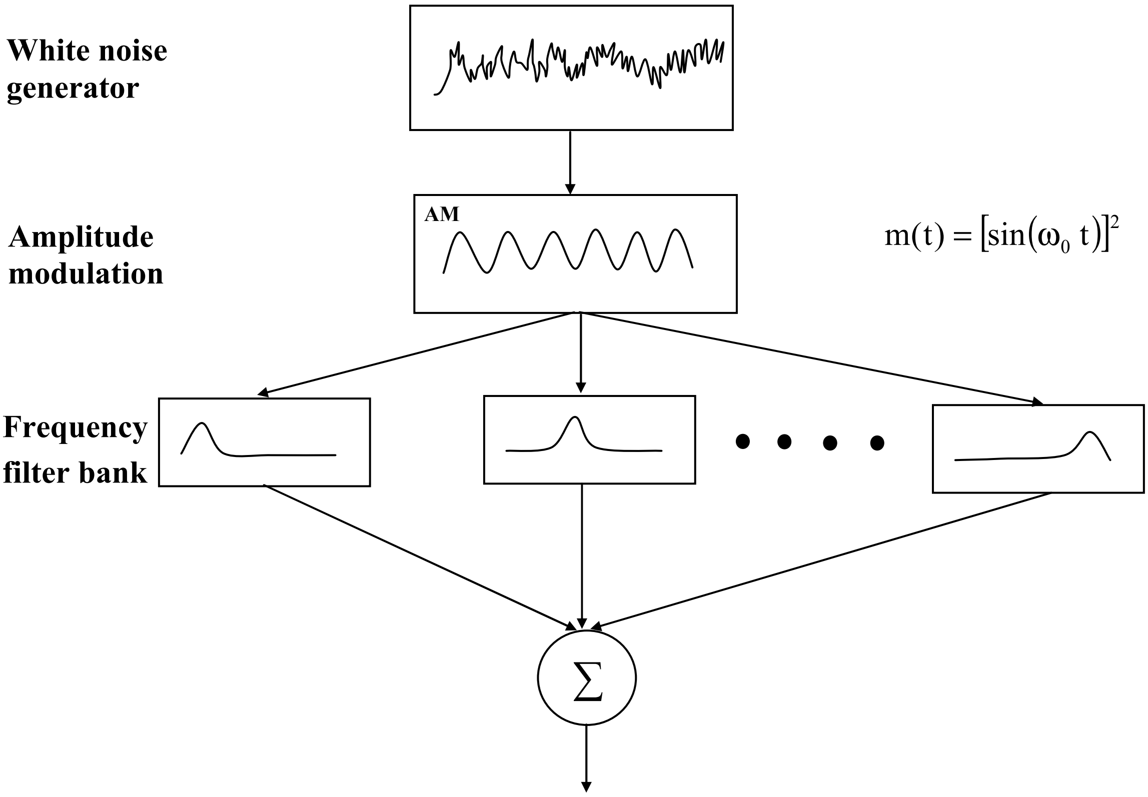


Figure 3.8: Schematic representation of the touch-induced scraping sound synthesis

White noise was selected as input signal (Fig. 3.8). It was amplitude modulated using the sinusoidal signal $m(t) = [\sin(\omega_0 t)]^2$, $\omega_0 = 2\pi f_0$, at a frequency of $f_0 = f_{r1}$ or f_g to simulate the effects of the grooves (grooved block) or grits (sandpaper). The resulting signal was filtered by a bandpass filterbank. Their centre frequencies were calculated using Equations which are originated from Eq. 3.3, 3.4, 3.5, 3.6, or 3.8. The amplification gains of the bandpass filters were adjusted using spectral data of pre-recorded signals.

3.5.2 Tactile subsystem

The tactile subsystem consists of a tactile glove system, and a whole-body vibrations system. **The Tactile Glove System** consists of an electro-tactile feedback unit, a force-feedback unit, and a data glove.

By representing the texture information, electro-tactile stimulation technique was selected according to its advantages. Its operation doesn't cause any noise (vibrotactile and pneumatic devices function relatively noisy), this makes it especially suitable for auditory-tactile virtual environment applications. Its low weight, simplicity, lower power consumption, and low cost are other reasons for the selection. Self-adhesive electrodes were used to excite the user's fingertip. The active electrodes (+) are attached to the fingertips of the user, and return electrodes (-) are attached to the upper side of the hand (Figure 3.9. b,c). The soundcard of the PC was used to generate current signals. The signals are fed through an amplifier-transformer box to the electrodes and transmitted to the user's fingertip (Figure 3.9 a,d). The frequency range of the electrotactile system is 10 to 1000 Hz. The results of experiments on the number of discriminable levels led to an estimate of 8 bits/s for the information capacity of an electrotactile channel (Brown and Stevens, 1992).

A simple non-portable one-degree-of-freedom mechanism was designed for the virtual force-feedback information (Figure 3.13). Two handles which transmit forces to the index finger and thumb of the user, function dependently of each other (one handle is attached to the index finger of the user, and other handle is attached to the thumb of the user). A DC motor provides up to 7-N force feedback.

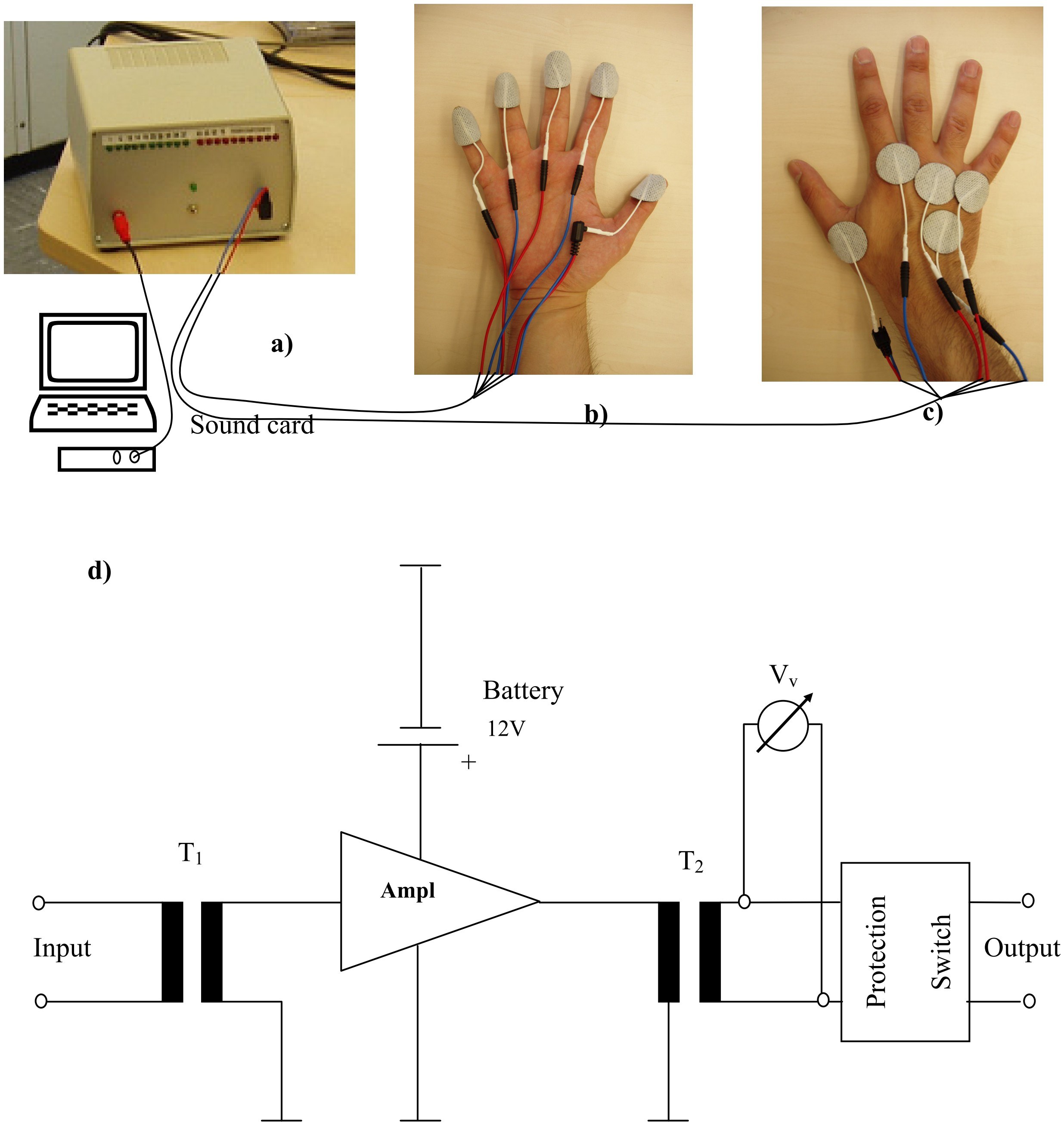


Figure 3.9: Electrotactile system: a) amplifier – transformer box, b) electrodes – palm of the hand c) electrodes – back of the hand, and d) circuit diagram of the amplifier – transformer box

Position and orientation of the hand and fingers are required input information for tactile virtual environments. To track this information, a data glove is used in the tactile system. The data glove is a low cost P5 type device from the company "Essential Reality" (Figure 3.10). It is based upon proprietary bend sensor and optical remote tracking technology and enables 6 degrees of tracking (X, Y, Z, Yaw, Pitch and Roll). It has a resolution of about 0.5 degrees, its update rate is 60 Hz, and it weighs 125 gr. Infrared light which is emitted by the glove and information from bend sensors (fingers) are received by a receiver unit which should be placed to the right of the PC monitor. The receiver unit is plugged to the USB port of the PC.

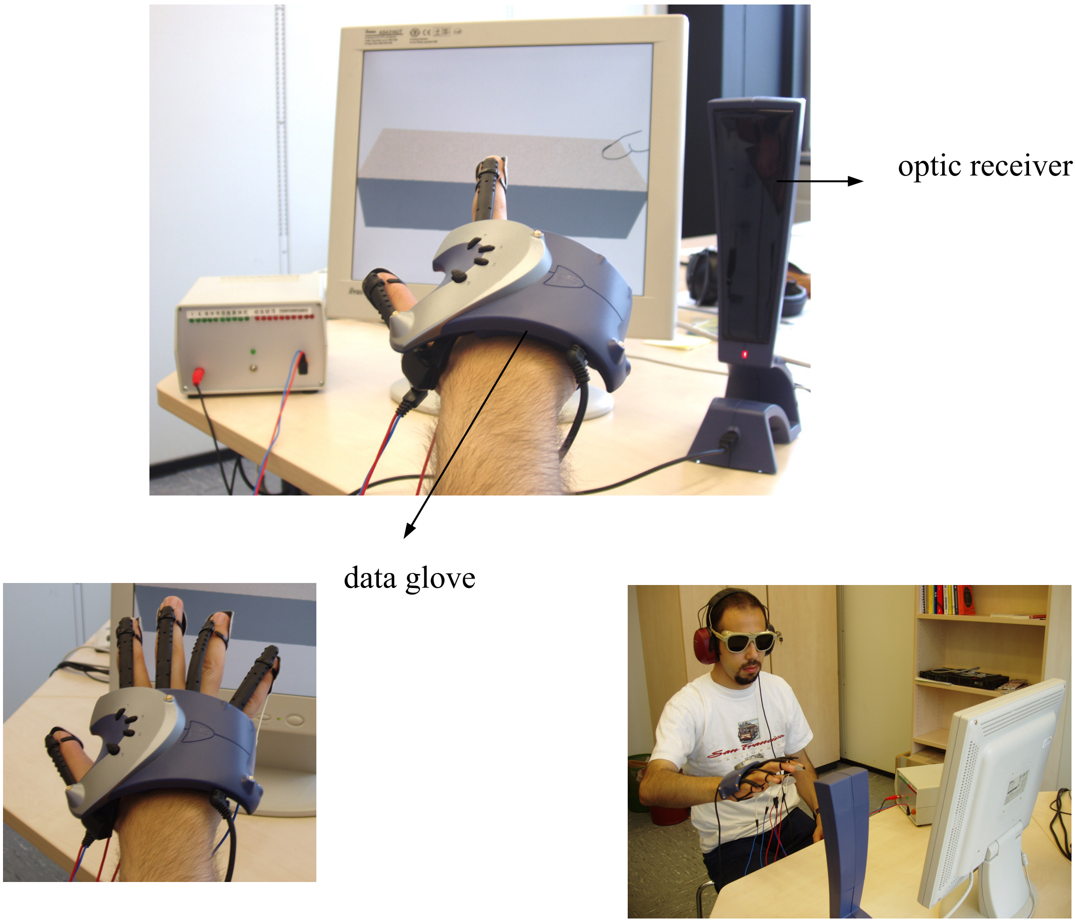


Figure 3.10: Data glove and optic receiver.

To simulate the whole-body vibrations, a system which produces only vertical vibrations was used. The whole-body vibration system contains a shaker and a chair (Figure 3.11). The chair consists of a steel-tube construction (Radius of the tube: d = 9 mm) which is made up of a leg consul and an additional tube. The leg construction is connected in three points to the wooden seat plate (thickness: t = 8 mm). Each leg of the chair has an elastic band which disconnects the chair from the ground (passive isolation). An additional tube, which supports the wooden back of the chair, is connected to the cross-point of the legs. The height of the seat is about 50 cm. The shaker is connected to the underside of the chair at two points. It weighs 1.5 kg and the maximum exciting power is 50 W.

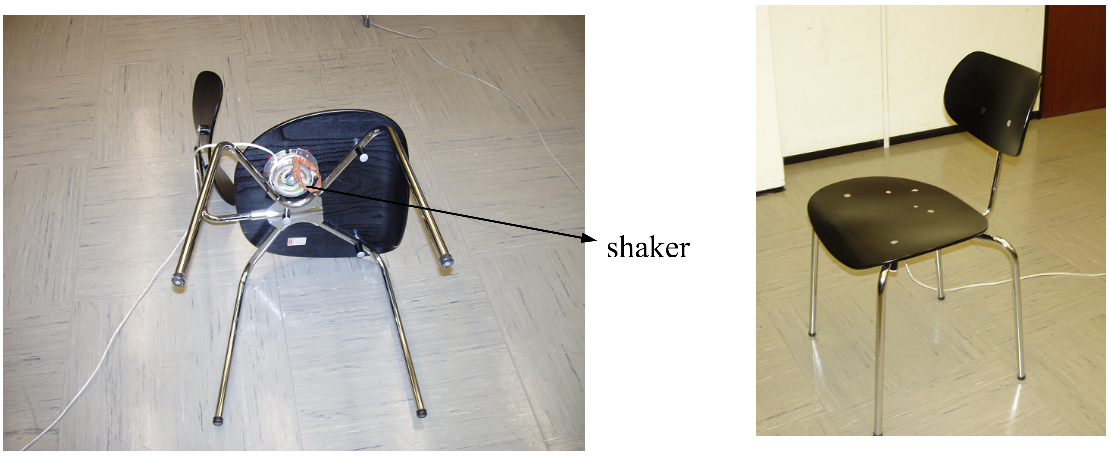


Figure 3.11: Whole-body vibrations chair.

The system is capable of producing vibrations in a frequency range from 10 to 1000 Hz. The transfer function of the shaker–chair system was measured with white noise as input vibration signal in vertical direction. The generated signals (PC) are transmitted digitally from the sound card to an amplifier (Kenwood KA-1100 SD) and from the amplifier to the shaker. The subject’s weight may influence the transfer function. Therefore a subject was instructed to sit on the rigid wooden seat, and the transfer function was measured while she/he was seating. The subjects weight was 75 kg. The measurement of vibration at a body-seat interface requires that special transducer is located between body and the seat (Griffin, 1990). A semi-rigid pad containing accelerometer (B&K 4322) was used in this measurement and placed between the seat and the subject. The transfer function of the system can be seen in Figure 3.12 (input signal: amplifier input, mV; output signal: vibration at the body-seat interface, m/s^2).

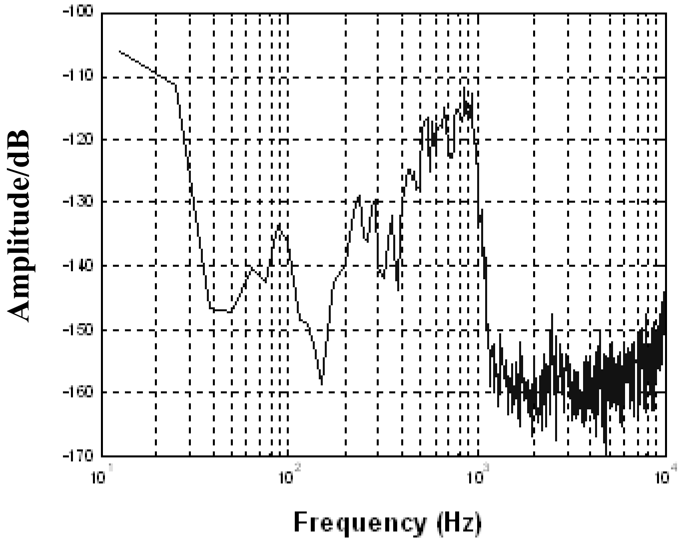


Figure 3.12: Transfer function of the shaker-chair system.

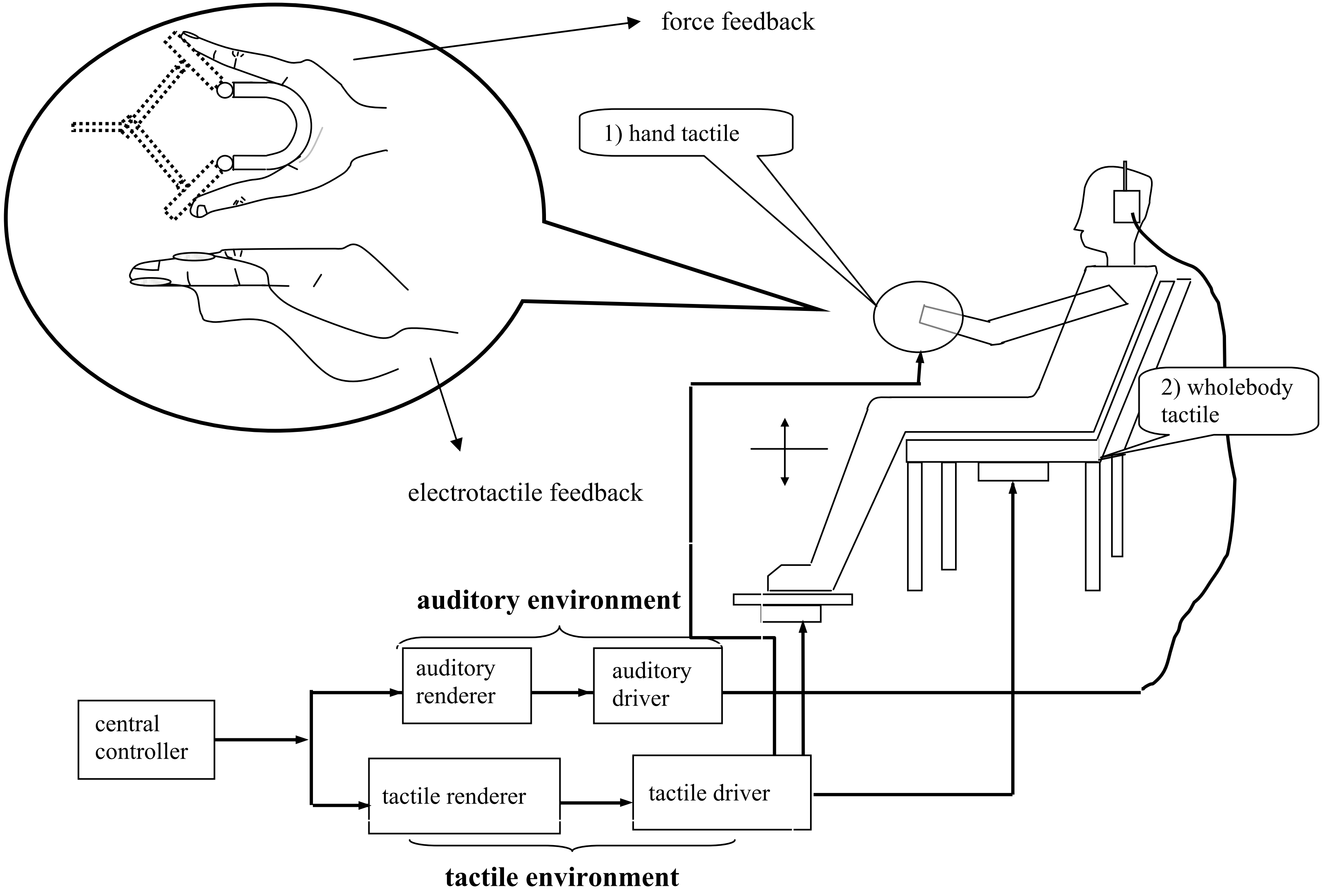


Figure 3.13: An experimental system for auditory-tactile interaction research.

Chapter 4

Segregation of the Auditory-Tactile Event

4.1 Introduction

In our daily life, we perceive most of the time an event through more than one sensory modality (e.g., auditory, tactile, visual). The signals which arise from a common source are transmitted by different sensory channels and are combined in the central nervous system to produce a unified percept. Our brain can easily separate the single-origin co-occurring inputs (multimodal event) from the inputs of independent origins (multiple unimodal events).

Behavioural and neurophysiological studies indicate that human infants can perceive some equivalences between cues in different modalities from a very early age. Experience plays an extremely important role in linking intersensory inputs to form a coherent multisensory representation of the environment (King, 2004; Lewkowicz, 2002). Infants can match the intensity of auditory and visual stimuli by 3 weeks of age (Lewkowicz and Turkewitz, 1980) and intersensory temporal relations based on common synchrony, duration, etc. during the first few months (Spelke, 1979; Lewkowicz, 1996). Prelinguistic infants as young as 4 months of age are able to associate their mother's face with her voice (Spelke and Owsley 1979).

The ability to integrate signals from the separate sensory modalities is based on the fact that the properties of these stimuli are coupled by physical laws (Kohlrausch and van de Par, 1999). Multisensory perception is a wide field of research. In order to approach it in a systematic way, this study concentrates on the influence of the physical coupling on the segregation of the auditory-tactile (multimodal) events.

Sound can be defined as „mechanical vibrations and waves of an elastic medium, particularly in the frequency range of human hearing" (DIN 1320, 1959). From this definition, it is evident that sounds are usually produced by vibrations of the objects. Therefore there is a strong physical relationship between sound and vibration.

Physical properties of the sound and vibration, such as temporal attributes (e.g., onset time, duration, etc.), frequency (spectral features), intensity and location (spatial origin) tell the brain whether the auditory and tactile inputs come from the same physical event or not (Figure 4.1). If both stimuli result from the same event, they should be coupled to each other by physical laws.

One of the most important cues available for the multimodal integration is simultaneity (synchronous onsets). If auditory and tactile events occur at the same time, there is a strong evidence that they come from the same physical event. Separation of the onsets of the auditory and tactile inputs is a powerful tool for the perceptual segregation of the concurrent information.

The second physical attribute which plays a role for multimodal integration is the intensity. If the same event generates the auditory and tactile information, their intensities (sound pressure level, force-feedback or vibration intensity) are directly proportional, as a result of the physical laws. If this proportional relationship breaks down, it results in the segregation of the auditory and tactile inputs.

Another physical attribute which may contribute to the multimodal integration is the frequency. From the physical point of view, if a sound is generated by a vibrating body, there is a correlation between the frequency of the sound and the frequency of the vibration. Therefore, the closer the frequencies of auditory and tactile events are, the stronger our tendency to group them. The degree of perceptual segregation depends on the frequency sepa-

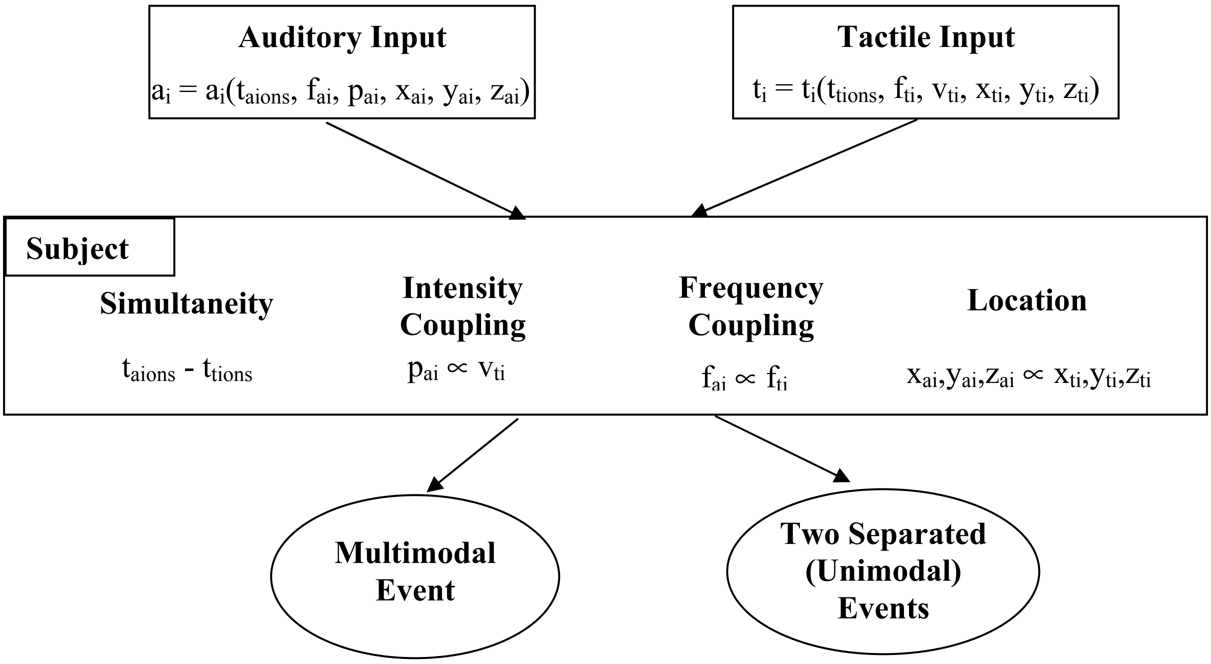


Figure 4.1: A schematic representation of the integration (multimodal event) and segregation (two separated events) processes for auditory and tactile inputs. The physical parameters which contribute to the integration and segregation process are: t_{aions}, onset of the auditory input, t_{tions}, onset of the tactile input, f_{ai}, frequency of the auditory input, f_{ti}, frequency of the tactile input, p_{ai}, sound pressure level, v_{ti}, velocity of the vibration, x_{ai}, y_{ai}, z_{ai}, coordinates of the auditory input, x_{ti}, y_{ti}, z_{ti}, coordinates of the tactile input. i indicates the stimulus number.

ration of auditory and tactile inputs. The larger the frequency difference of auditory and tactile inputs, the more strongly we tend to segregate them perceptually.

If auditory and tactile information result from a single multimodal event, the position of the auditory event and the position of the tactile event (vibrating body that radiates the sound waves) should coincide. Therefore we easily group auditory and tactile inputs which come from the same spatial direction and segregate both input which come from different directions.

It is possible to find some similarities between the physical approach for the crossmodal grouping or segregation as explained above and the Gestalt theory which suggests "grouping" rules for the visual perception related to the geometric statistics of the natural world. According to Gestalt approach, we perceive objects as well-organized patterns rather than separate component parts. The main principles of the Gestalt grouping approach are the rule of proximity, the rule of similarity, the rule of perceived continuity, and the rule of simplicity.

According to the rule of proximity, grouping depends on the closeness of the items. The items which are closer to each other tend to be seen as a group. This principle can be adapted to the multimodal perception by taking into account the comments of Bregman (1990) who has adapted the Gestalt principles to audition. In accordance with this analogy, the spatial dimension of distance in vision has two analogies in audition. One is the separation in time, and the other is the separation in frequency. Following this analogy, auditory and tactile events which are close to another in time and frequency will perceptually be grouped together. As the time separation or frequency separation increases, auditory and tactile events will be segregated and perceived as two separate events. This analogy is applicable also for the spatial origin of auditory and tactile events. The closer the location of the auditory event and tactile event, the stronger we tend to group them. If the separation in location increases, this will result in the segregation of two events.

The rule of similarity indicates that the items which have similar features, such as a similar shape, a similar size, a similar colour, a similar texture, or a similar orientation, tend to be grouped together. According to the similarity rule, it is possible to say that auditory and tactile events which have similar frequencies tend to be grouped together. The rule of similarity confirms the rule of proximity for frequency.

Knowledge of the neural mechanisms underlying the integration of multimodal information is important to confirm the physical approach for multimodal integration. Electrophysiological studies have shown that a presentation of multisensory cues in combination will evoke responses that vary with the relative timing, location, and intensity of the stimuli. In general, response enhancements, that is, higher discharge rates, are observed when stimuli of different modalities are delivered in close temporal and spatial proximity (King 2004; King and Palmer 1985; Meredith and Stein, 1996; Kadunce et al. 1997). These results also show that simultaneity, intensity, and location play an important role for the multimodal integration or segregation.

Separation in time (temporal asynchrony between the auditory and tactile events) and frequency, disappearance of the level coupling, and/or inputs from different locations (differences in spatial origin) can result in the segregation of the auditory and tactile events into two isolated percepts for each modality, instead of a unified multimodal percept. Integration of multi-modal information is an important task for multimedia applications, virtual environments, and for industrial product designers in order to obtain more realistic and compelling products.

In the following sections of this chapter, the physical conditions are investigated which can cause a perceptual segregation of auditory and tactile events. More precisely, the thresholds for the time delay ($t_{aions} - t_{tions}$, $t_{tions} - t_{aions}$), the level difference, the frequency difference ($f_{ai} - f_{ti}$), or the spatial origin difference ($x_{ai} - x_{ti}$, $y_{ai} - y_{ti}$, $z_{ai} - z_{ti}$) between auditory and tactile stimuli are determined which result in the segregation of two inputs. The detection of event singleness (segregation or integration) can be simply measured by asking the subjects whether auditory and tactile stimuli are caused by the “same event” or “not” while the experimenter systematically varies a physical attribute (e.g. time delay etc.). In some experiments, the subjects can be also asked whether the stimuli are synchronous or asynchronous (Van de Par and Kohlrausch, 2000) or whether the stimuli came from the “same location” or from “different locations” (Bertelson and Radeau, 1981). A more detailed discussion on measurement methods can be found in each section.

In **Section 4.2** the investigations on the temporal factors involved in the segregation of auditory and tactile events are presented. **Section 4.3** describes investigations on the relationship between level coupling and segregation. The influence of the separation in frequency on the segregation is presented in **Section 4.4,** and **Section 4.5** addresses the relationship between the spatial origin of the auditory and tactile events and the segregation. In the experiments, the tactile stimulation is applied to the subjects’ hand and also their whole body.

4.2 Audiotactile simultaneity

4.2.1 Introduction

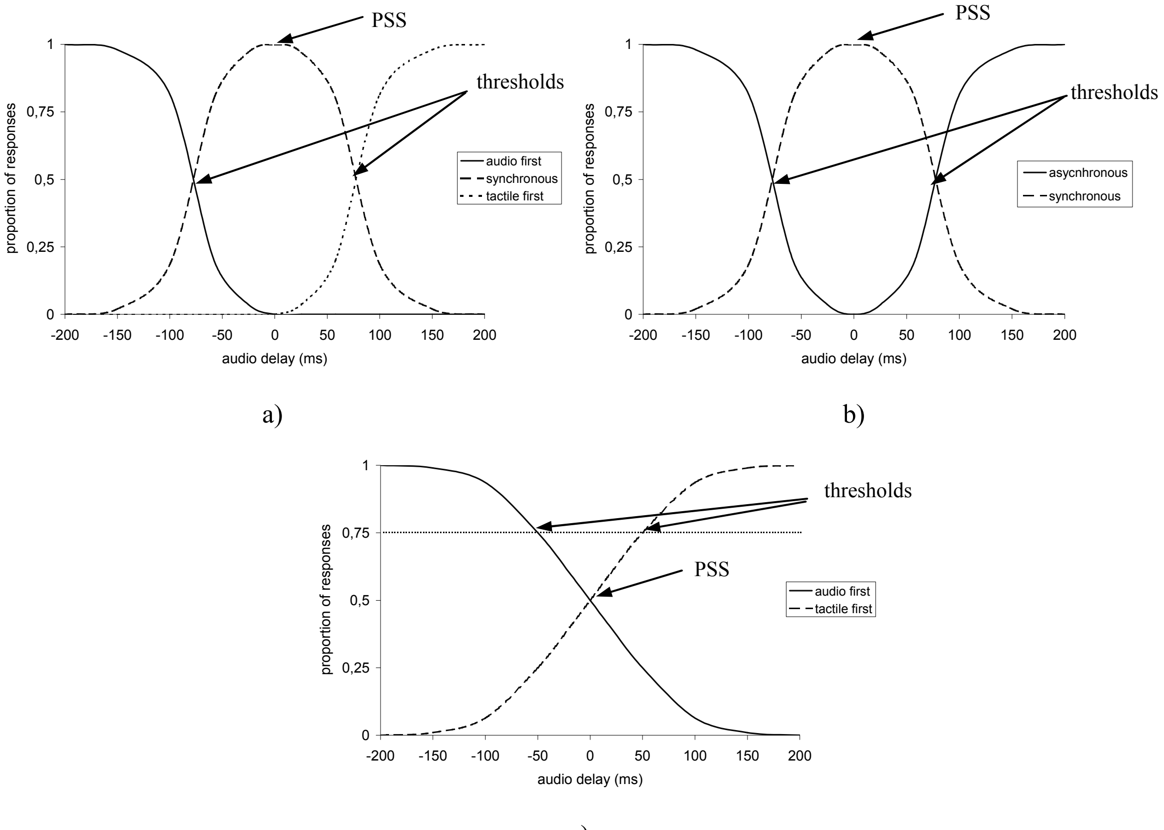
For years or even decades each of us has learned that different physical stimuli which are received simultaneously by various sensory channels (auditory, visual, tactile etc.) is usually caused by one and the same physical event in our environment. Temporal correlation is an important hint for the brain to integrate inputs which are generated by one event and obtained from different sensory channels, and also to differentiate inputs which are related with this event, from other inputs which are not related with this event.

Synchronization of different modalities in multimedia applications is a big problem. Technical constraints such as data transfer time, computer processing time, and delays which occur during feedback generation processes, produce synchronisation problems. As the asynchrony between different modalities increases, the sense of presence and realism of the multi-media applications will decrease. Therefore, an understanding of the human simultaneity detection mechanism and perceptual aspects of multi-modal simultaneity is also a necessary prerequisite for multi-media designers.

Several studies have discussed the perceived simultaneity of multi-modal stimuli. A multi-modal synchronisation threshold has been defined by Altinsoy et. al. (2001) as the maximum tolerable temporal separation of the onset of two stimuli, one of which is presented to one sense and the other to another sense, such that the accompanying sensory objects are perceived as being synchronous. In order to measure this threshold different psychophysical measurement methods have been applied. The schematic response patterns of different methods are shown in Fig. 4.2. One response method asks the subject to make a three-alternative forced-choice judgment as to whether the stimuli are synchronous or which one was presented first (for auditory-visual (AV) asynchrony, Van de Par and Kohlrausch, 2000). The response pattern of the three-alternative forced-choice judgment method is shown in Fig. 4.2a. The intersection between the curves of “audio stimulus preceded the tactile stimulus” and “they were synchronous” judgments defines the threshold for detecting asynchrony in the direction of advanced audio. The intersection between the curves of “tactile stimulus preceded the audio stimulus” and “they were synchronous” judgments defines the threshold for detecting asynchrony in the direction of advanced tactile stimulus. The maximum point on the synchronous curve indicates the point of subjective simultaneity (PSS).

Another measurement method is to ask the subject to judge whether the audio and the tactile stimuli are synchronous or asynchronous. Fig. 4.2b shows the response pattern of this measurement method. In this method, two intersections between the curves of the "synchronous" and "asynchronous" indicate the thresholds for detecting asynchrony in the direction of delayed audio and delayed tactile stimuli. Again, the maximum point on the synchronous curve indicates the point of subjective simultaneity (AV, Dixon and Spitz, 1980, Miner and Caudel, 1998, Van de Par, Kohlrausch and Juola, 2002).

The temporal order judgments (TOJ) is one whereby the subject has to judge the temporal order of an auditory and a tactile stimulus, which results in the minimal multi-modal delay, for which subjects are able to indicate in which temporal order the two different sense are being stimulated (for auditory-tactile (AT) asynchrony; Hirsh and Sherrick, 1961, for AV asynchrony; Jaskowski, Jaroszyk, and Hojan-Jezierska, 1990). The response pattern of the TOJ method is shown in Fig. 4.2c. The intersection between "tactile first" and "audio first" curves gives us the point of subjective simultaneity. The proportion of responses being 25 % and 75 % indicate the thresholds for detecting asynchrony in the direction of delayed audio and delayed tactile stimulus.



a)

b)

c)

Figure 4.2: The response patterns a) 3 categories: audio first, tactile first, synchronous b) 2 categories: asynchronous, synchronous c) 2 categories: tactile first, audio first

Methodological aspects for measuring asynchrony detection in audio-visual stimuli have been reported by Van de Par, Kohlrausch and Juola (2002). They found that the point of subjective equivalence in audiovisual synchrony is shifted towards audio delays by about 35 ms compared to the point of objective equivalence. The TOJ method allows for different decision strategies (for determining whether audio or video was leading even if the stimulus perceived as synchronous) and therefore results of the TOJ method depend on which strategy the subjects chooses. The other methods are rather robust and in agreement with each other.

The aim of this chapter is to measure perceptual threshold values for auditory-tactile asynchrony and to investigate the temporal factors involved in the integration of auditory and tactile information. This investigation will be carried out for two different tactile stimulations: one presented at the tip of the index finger (vibratory finger stimulation), and another presented as whole-body vibration. The results of the different measurement methods will be compared and discussed to establish the point of subjective simultaneity for auditory-tactile asynchrony.

4.2.2 Auditory – tactile asynchrony: Vibratory finger stimulation

4.2.2.1 Subjects

The same six subjects participated in the experiments. They were four right-handed men and two right-handed women with self reported normal hearing ability. Their ages ranged between 22 and 32 years.

4.2.2.2 Experimental set-up and stimuli

The tactile stimulus was a sine wave and presented at the tip of the index finger of the participant via a B&K Type 4810 mini-shaker of the electrodynamic type with a permanent field magnet, with a maximum stroke of 6 mm and a force rating of 10 Newton sine peak in the vertical direction. The shaker delivered the stimuli to the skin via a vibrating probe. The probe was 4 mm in diameter. The shaker was located inside a wooden box, which contained a circular hole on which the participants placed their index finger. A further necessity of the box was to mask the visual information from the shaker. To minimize the structural vibrations generated by the shaker, the floor was isolated from the shaker by using some vibration damping materials.

The auditory stimulus was a burst of white noise presented from a PC. The noise was amplified and delivered diotically through Sennheiser HDA 200 closed-face dynamic headphones which has a very high sound isolation level and therefore masked the background noise of shaker when it generated the signal. The experiments were conducted in a sound-attenuated room.

The durations of the stimuli were 25 ms. It is possible that the intensities of auditory & tactile stimuli has a important influence on perceptual asynchrony. Therefore a cross-modal intensity matching experiment (Stevens 1975) was conducted to determine a suitable sound pressure level and vibration intensity level. In this level-matching experiment the participant's task was to match the apparent loudness of the burst of white noise to the apparent strength of the vibration on their finger. The tactile stimulus were presented randomly at six different levels, 35 dB – 65 dB re 1 µm (0.07, 0.18, 0.6, 0.75, 1.3, 1.6mm) (each stimulus was presented 20 times) and subjects adjusted the level of the sound by using an amplifier until its apparent loudness seemed as great as the strength of the vibration on their finger. In Fig. 4.3, the medians of the sound pressure level are plotted against the vibration amplitude.

The power equations according to Stevens (1975) for the two modalities can be described as follows;

$$\psi_s = \phi_s{}^m \tag{4.1}$$

$$\psi_v = \phi_v{}^n \tag{4.2}$$

where ψ is the subjective magnitude, Φ is the stimulus magnitude, m is the characteristic exponent for noise, n is the characteristic exponent for the vibration, s indicates the auditory modality and v indicates the vibration modality. If the participant equates subjective magnitudes by the cross-modal matching experiment

$$\psi_s = \psi_v \tag{4.3}$$

$$\phi_s = \phi_v{}^{n/m} \tag{4.4}$$

The obtained exponent from the equal sensation function which was determined by the results of the cross-modal matching experiment is n/m = 0.86.

The sound pressure level was set to 56 dB which was shown to match a vibration amplitude of 58 dB, which is a displacement of 0.6 mm.

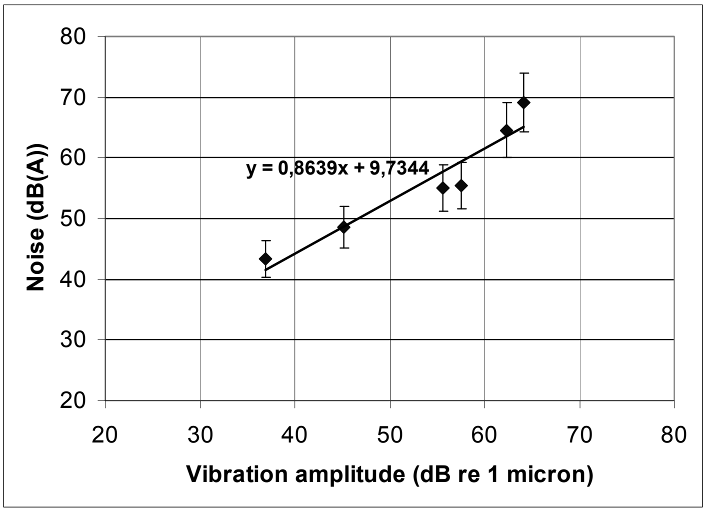


Figure 4.3: Equal sensation functions relating sine vibration on the finger tip to the intensity of a burst of white noise.

4.2.2.3 Methodology and procedure

Altogether four different experiments were conducted. In the first three experiments synchronisation thresholds and the point of subjective simultaneity of auditory-tactile presentations were measured. Stimuli were presented in a random order with an audio delay ranging from -200 to 200ms with varying step sizes (-200 to 150 ms, 50ms steps; -120 to –60 ms, 20ms steps; -60 to 60 ms, 10ms steps; 60 to 120 ms, 20ms steps;150 to 200 ms, 50ms steps). Each condition was presented twelve times. Negative delay values indicate that the auditory stimulus was presented first, and positive delay values indicate that the tactile stimulus was presented first.

In the first experiment response categories were "tactile first", "synchronous" or "audio first". In the second experiment response categories were "synchronous" or "asynchronous. In the third experiment response categories were "tactile first" or "audio first". Condition-order of the three experiments was counter balanced across the subjects according to a Latin square.

The fourth experiment was carried out to measure auditory and tactile reaction times. Subjects were asked to respond as quickly as possible to the stimulus by pressing a button. A warning signal was presented to the subject before each trial. As in the study of Jaskowski, Jaroszyk and Hojan-Jezierska (1990), the stimulus followed the warning signal after a random fore-period. The fore-period was a sum of a fixed interval of 1s and an interval sampled from an exponential distribution with mean equal to 1s. Each modality was stimulated alone. One session consisted of 100 trials, and each subject joined four sessions.

4.2.2.4 Results

The point of subjective simultaneity and synchronisation threshold values are shown according to the measurement method in Table 4.1. The proportions of the responses of all three experiments for each response alternatives are shown for the six subjects that participated in the experiments, in Figure 4.4 - 4.9. The mean reaction times to the auditory stimulus and the tactile stimulus are presented in Figure 4.12.

The results of the first experiment are depicted by thick lines. The black triangles indicate "tactile first" responses, the white diamonds indicate "synchronous" and the black squares indicate "audio first". The synchronous curves seem to peak for slightly positive audio delays. This shift can be seen especially clearly in the results of the subjects S1, S2 and S6. The synchronous judgments of the six subjects participating in Experiment 1 were averaged and shown in Figure 4.10. A psychophysical model was obtained by fitting ogive[5] results using a gaussian fit with exponential background. The goodness of the fitted curves was evaluated for each graph. The R-square value is 0.92 and the sum of squares due to error (SSE) is 0.056. In the first experiment stimuli with audio delays in the range of –23 to 46 ms were judged synchronous.

Table 4.1: PSS's and synchronisation thresholds in milliseconds for six subjects.

	PSS's			Synchronisation Thresholds (JNDs)					
	Exp.1	Exp.2	Exp.3	Exp.1		Exp.2		Exp.3	
S1	10	10	2	-10	36	-10	30	-15	10
S2	8	8	-5	-15	40	-15	35	-14	12
S3	10	10	-2	-60	60	-45	75	-8	10
S4	0	10	12	-35	70	-20	50	0	18
S5	5	-5	-2	-20	53	-28	54	-10	30
S6	2	12	0	-20	52	-23	57	-28	90

---

[5] A distribution curve in which the frequencies are cumulative.

The results of the second experiment are depicted by the thin lines and the grey symbols. Grey diamonds indicate the "synchronous" responses, and the grey circles indicate "asynchronous" responses. Similar to the first experiment, the PSS is shifted toward positive audio delays. The synchronisation thresholds which are found in the second experiment are also very similar to the synchronisation thresholds which are found in experiment 1. Only subject S4 has lower threshold values in the second experiment than in the first experiment.

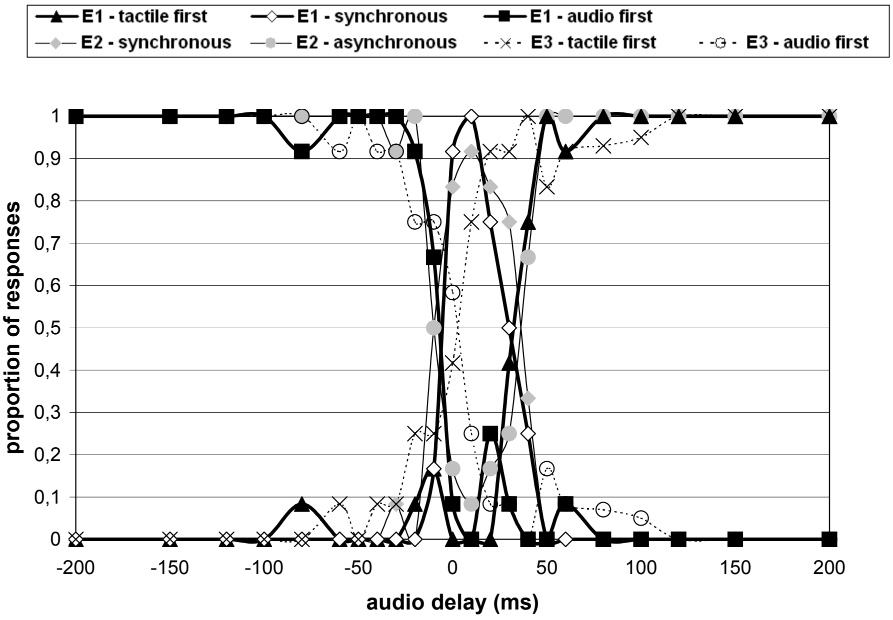


Figure 4.4: Proportion of the subject 1's responses as a function of audio delay in milliseconds. Negative values indicate audio lead and positive values audio lag.

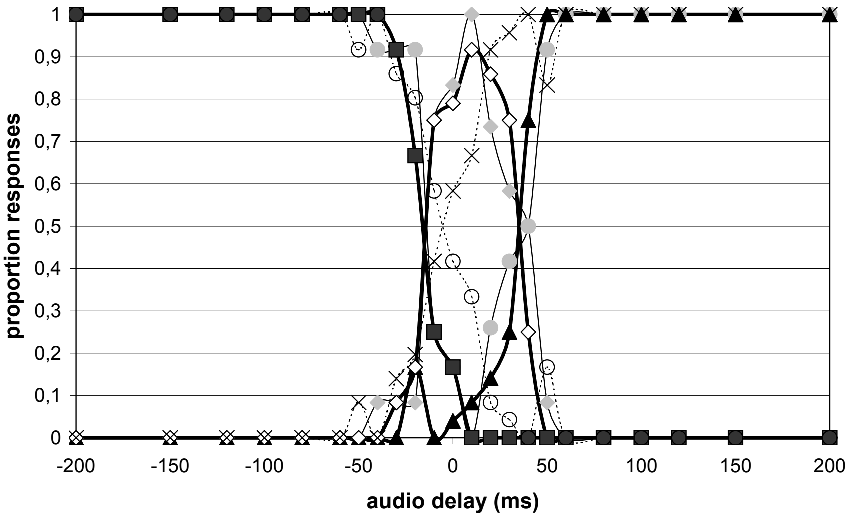


Figure 4.5: Proportion of the subject 2's responses as a function of audio delay in milliseconds.

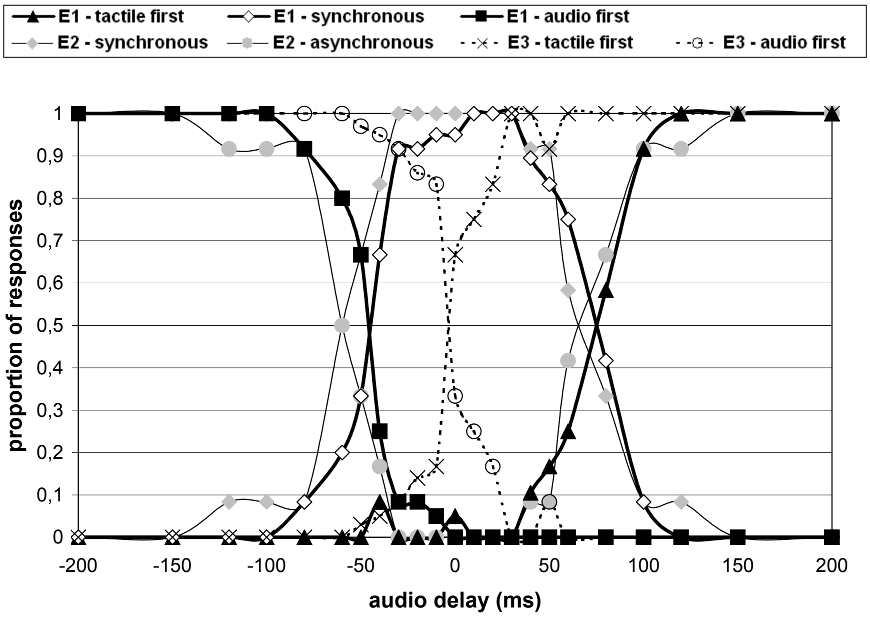


Figure 4.6: Proportion of the subject 3's responses as a function of audio delay in milliseconds.

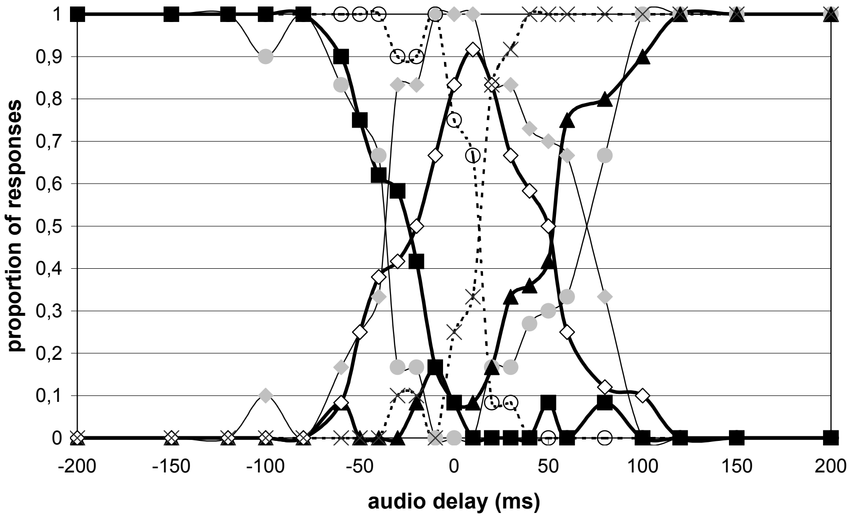


Figure 4.7: Proportion of the subject 4's responses as a function of audio delay in milliseconds.

In both experiments (first and second) the transition between "audio first" and "synchronous" responses is sharper than between "tactile first" and "synchronous". The synchronous judgments of the six subjects participating in Experiment 2 were averaged and are shown in Figure 4.10. A psychophysical model was obtained by fitting the ogive results by using gaussian fit with exponential background. R-square value is 0.91 and the sum of squares due to error (SSE) is 0.067. The results of the second experiment showed that stimuli with audio delays in the range of –25 to 50 ms were judged synchronous.

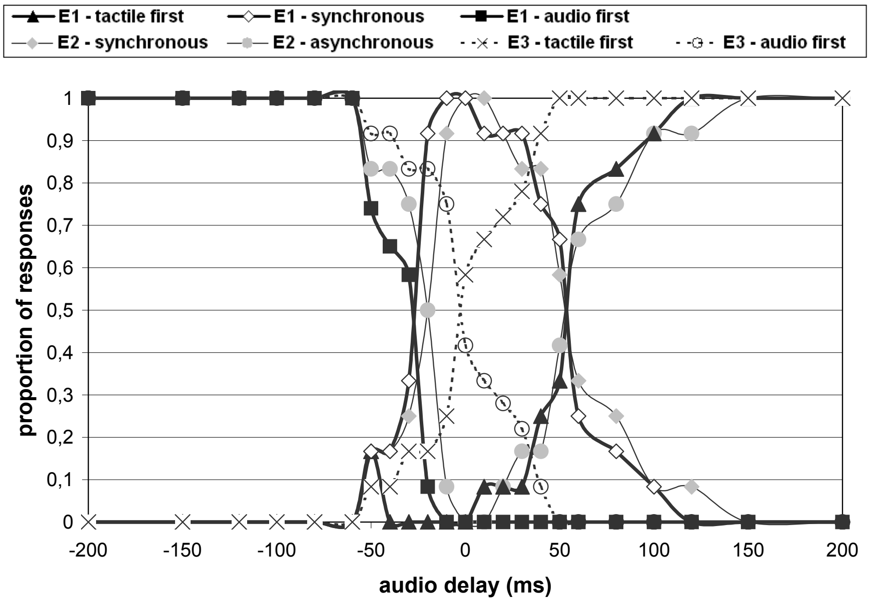


Figure 4.8: Proportion of the subject 5's responses as a function of audio delay in millisecond.

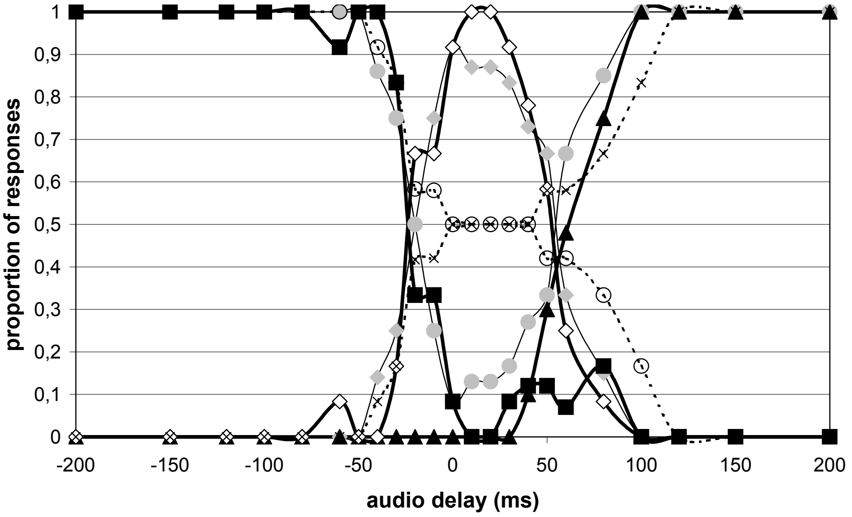


Figure 4.9: Proportion of the subject 6's responses as a function of audio delay in millisecond.

The results of the third experiment are depicted by the dotted lines. The multiplication signs indicate "tactile first" responses and white circles indicate "audio first" responses. Three subjects S2, S3 and S5 show negative PSS, two subjects S1 and S4 show positive PSS, and one subject S6 has a PSS that is zero. For the subjects S3, S4, S5 and S6, intersection of the curves for "tactile first" and "audio first" is in the area where subjects responded with "synchronous" in the first two experiments. For subjects S1 and S2, the transition coincides with the intersection of "audio first" curve with the "synchronous" curve of the first experiment. Psychometric function (S curves) for tactile first responses were obtained from the ogive results by calculating the z scores and applying the least square approximation (Figure 4.11). In the third experiment stimuli with audio delays in the range of –10 to 20 ms were judged synchronous.

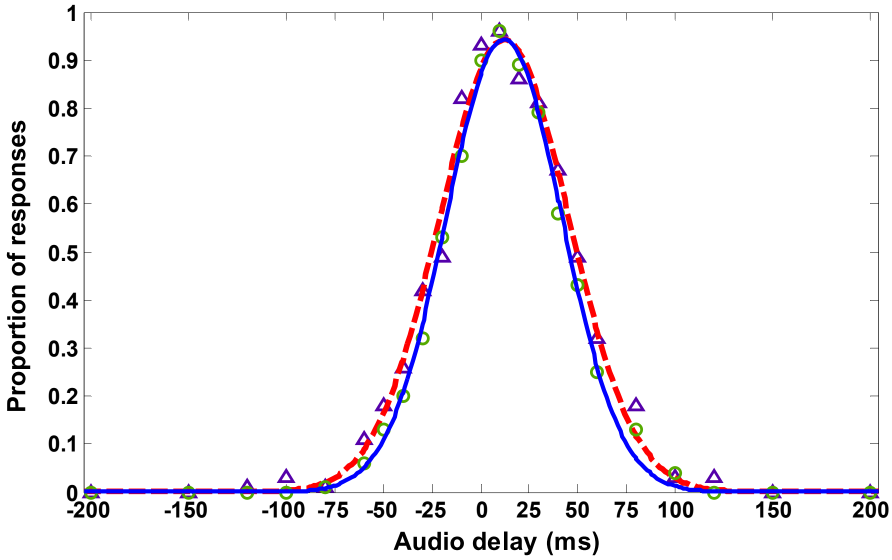


Figure 4.10: The results of the experiment 1 and 2 as a function of audio delay in terms of the proportion of responses. Dashed line indicates the synchronous results of the Experiment 2 and solid line indicates the synchronous results of the Experiment 1.

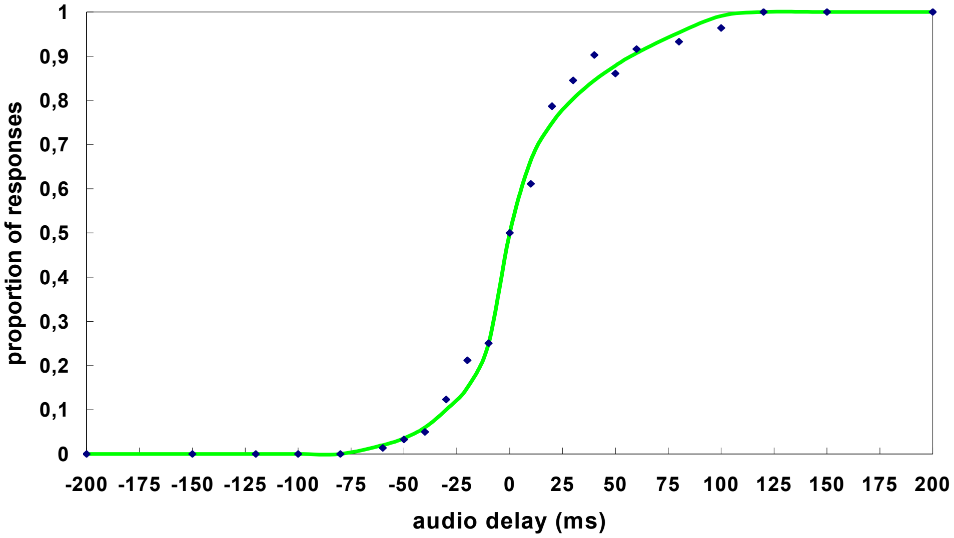


Figure 4.11: The results of experiment 3 as a function of audio delay in terms of the proportion of responses.

The results of the reaction time (RT) experiment show that the participants react 13 ms (SEM 2.17ms) quicker with an auditory stimulus than with a tactile stimulus (Fig. 4.12). A paired sample t-test shows that RT's are significantly shorter with a noise burst compared to tactile stimulation, t (5) = 5.878, p<0.01. The neurophysiological observations also support this result. The transduction time from receptors to the cells are in the range 6-25 ms for auditory stimuli and 18-34 ms for tactile stimuli (Stein and Meredith, 1993; Meredith, Nemitz and Stein, 1987) .

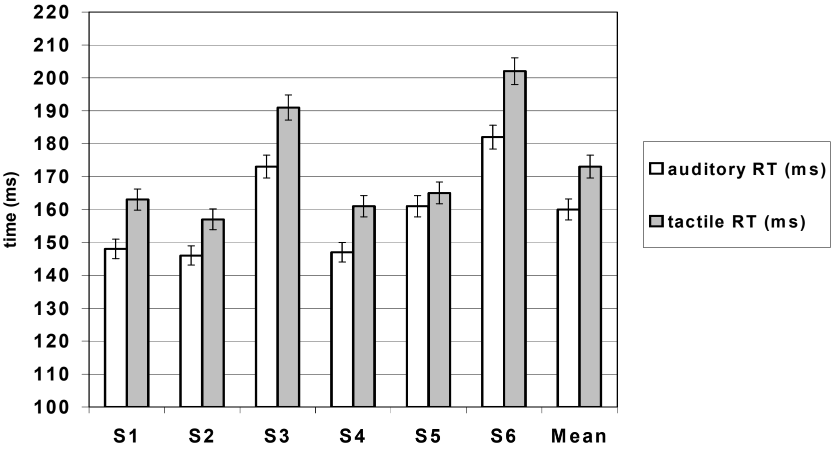


Figure 4.12: Auditory and tactile reaction times (ms)

4.2.2.5 Discussion

The results of the all three experiments show that the point of subjective simultaneity does not coincide with the point of objective simultaneity (0 ms). The PSS is found at an audio delay of about 7 ms. The most interesting finding is that audio advances are detected better than audio delays. These facts may be linked to the physical rules, e.g. speed of sound. The distance between our hands and ears is about 1 m, therefore sound would take about 3 ms longer to reach us than tactile stimulus. Also physiologically, the transduction time along the auditory neural pathway and somatosensory neural pathway is different. The reaction time experiments show this difference. The reaction times are 13 ms shorter for auditory stimuli than for tactile stimuli. Therefore it is possible that the human perceptual system is adapted to tolerate larger audio delays than tactile delays, as suggested for audio-visual asynchrony by van de Par and Kohlrausch (2000), Dixon and Spitz (1980).

The results of the first and second experiments are in agreement with each other. However in the third experiment (TOJ) there is an inconsistency between subjects. The possible reason can be that subjects adopted different decision criteria for determining whether audio or tactile was leading, as suggested for audio-visual asynchrony by van de Par, Kohlrausch and Juola, 2002.

4.2.3 Auditory – tactile asynchrony: Whole-body vibration stimulation

Today, whole-body vibrations play an important role in our life. While driving a vehicle or travelling, we are simultaneously exposed to different forms of whole-body vibrations and noise. Many virtual auditory-tactile environments, e.g. car simulators, require a whole-body tactile stimulation. For this case, literature data are not available. In this section, the investigations related to synchronicity of the auditory and whole-body vibration stimuli shall be given. In order to measure the sensitivity to audio/whole-body vibration asynchrony two types of stimuli, (1) artificial impact-type vibration and noise, (2) the vibration and noise recordings while a car passes a bump will be used.

4.2.3.1 Experiment - Artificial impact-type vibration and noise - Vibration and noise recordings while a car passes a bump

4.2.3.1.1 Subjects

Six men and two women with normal-hearing ability participated in the experiment as subjects. Their age and weight varied between 23 and 50 years (mean 28), 53 and 80 kg (mean 68 kg), respectively. None of them suffered from stomach or back trouble.

4.2.3.1.2 Experimental set-up and stimuli

The whole-body vibrations were produced by an electrohydraulic simulator (Schenk) with a maximum stroke of 250 mm and a dynamic force capability of 25 kN in the vertical direction (z axis). A rigid wooden seat without a backrest was mounted on the simulator platform. In order to obtain the timing relationship (simultaneous and non-simultaneous presentation) between the whole-body vibrations and audio stimuli, two different signals were generated by a sound card of a PC (left channel: audio stimuli, right channel: trigger signal for the vibration generator). The acoustic stimuli were amplified and delivered diotically through headphones. The time delay of the electrohydraulic simulator system which occurs between the arrival time of the trigger signal and the generation of the vibration was measured and compensated with an additional delay in the audio stimulus, when the stimuli were generated.

In order to measure the sensitivity to audio/whole-body vibration asynchrony both realistic and artificial stimuli were used. In the first case a broad band noise for the auditory stimulus and a sine wave for the vibration stimulus were employed. The durations of the artificial stimuli were 25 ms. The sound pressure level was adjusted to 62 dB and the magnitude of the vibration at the platform of the simulator was set to 0.45 m/s^2.

For the realistic stimuli an auditory-vibration recording was made of a car passing a bump. The airborne sound (indoor) was recorded by a dummy head, and the structure-borne sound by a whole-body seat transducer, lying on a car seat. The durations of the realistic stimuli were 300 ms. The sound pressure level was adjusted to 64 dB and the magnitude of the vibration at the platform of the simulator was set to 0.7 m/s^2.

4.2.3.1.3 Methodology and procedure

The subjects were instructed to sit on the rigid wooden seat in a comfortable posture and to wear headphones. Stimuli were presented in random order with an audio delay ranging from –350 to –100 ms in steps of 50 ms, from -100 to 100 ms in steps of 25 ms and from 100 to 350 ms in steps of 50 ms. Each condition was presented four times. The subjects were asked to report on whether the audio signal and the vibration signal were synchronous or asynchronous. Before the start of the experiment, five anchor stimuli were presented to the subjects so that they could become familiar with the system and the stimuli. Each experimental session lasted approximately half an hour including the training session.

4.2.3.1.4 Results

The proportions of the synchronous responses are shown in Figure 4.13 (artificial stimuli) and in Figure 4.14 (realistic stimuli). The psychophysical model was obtained fitting ogive results by using a Gaussian fit with exponential background. The goodness of the fitted curves was evaluated for each graph. For the artificial stimuli, R-square value is 0.98 and the sum of squares due to error (SSE) is 0.023 and for the realistic stimuli, R-square value is 0.947 and the sum of squares due to error (SSE) is 0.098.

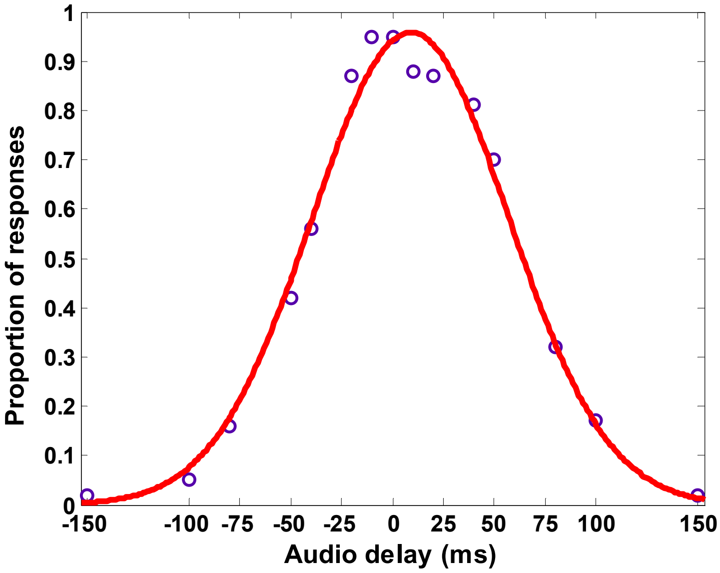


Figure 4.13: The proportions of the responses for synchronous responses (artificial stimuli)

The 50 % synchronisation threshold values of the psychometric functions are -47 and 63 ms for artificial stimuli (PSS = 8 ms) and –58 and 79 ms for the realistic stimuli (PSS = 10ms).

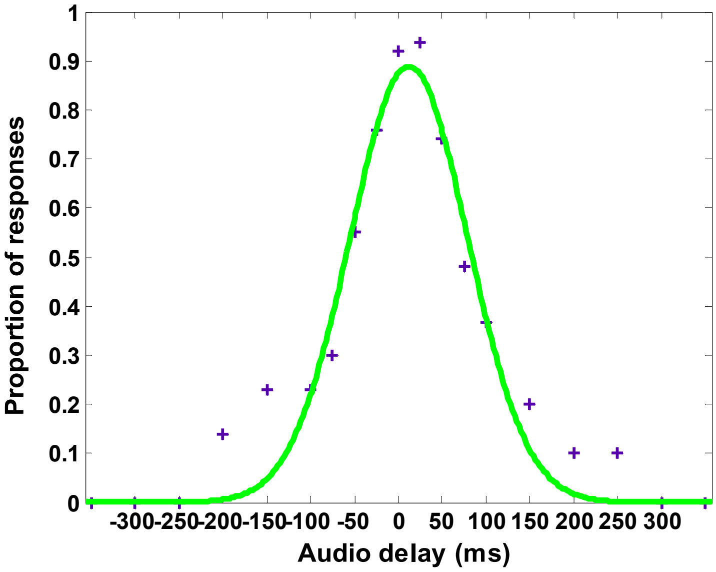


Figure 4.14: The proportions of the responses for synchronous responses (realistic stimuli).

4.2.3.1.5 Discussion

The results show that humans are very sensitive to the synchrony of audio and whole-body-vibration stimuli. The results of the experiments indicate that the synchronisation-threshold values are too critical to be produced by a rather simplistic virtual auditory-tactile environment and that the synchronization has to be at least within an accuracy of 50 ms.

The results of the two experiments show that the PSS does not coincide with the point of objective simultaneity (0 ms). In the condition with the artificial stimuli, PSS was found at a delay of 8 ms. PSS of the realistic stimuli was found at an audio delay of 10ms. Possibly the differences in the conditions for the artificial and realistic stimuli (PSSs and 50 % thresholds) are a result of the physical parameters of the stimuli. The artificial stimuli were very short (25 ms) while the realistic stimuli were considerably longer (300 ms).

4.2.4 General discussion

Simultaneity of multimodal information is the most important parameter of multimodal integration. This section of the current chapter investigated sensitivities to auditory-tactile asynchrony and the relevance of the audiotactile synchrony for the perceptual integration of auditory and tactile stimuli. The investigations were performed with two different tactile stimulations: vibratory finger stimulation, and whole-body vibration.

First of all, methodological aspects for measuring asynchrony detection in auditory-tactile stimuli were discussed. The results of the measurement methods that ask the subjects to make a three-alternative forced-choice judgment as to whether the stimuli were synchronous or to order according to precedence and those that ask the subject to judge whether the audio and the tactile are synchronous or asynchronous are robust and in agreement with each other. However, in the third experiment (TOJ), which asks the subjects to judge the temporal order of an auditory and a tactile stimulus, there is an inconsistency between subjects. Different subjects show different tendencies which results in negative or positive (also zero) PSS values. The possible reason can be that subjects adopted different decision criteria for determining whether audio or tactile stimulus was leading. Because of the robust and consistent results, in the further experiments the measurement method which asks the subjects to judge whether the audio and the tactile are synchronous or asynchronous was applied.

The results of the conducted experiments can be summarized as follows:

- Vibratory finger stimulation: The results of the experiments show that the point of subjective simultaneity does not coincide with the point of objective simultaneity (0 ms). The PSS is found at an audio delay of about 7 ms. Stimuli with audio delays in the range of –25 to 50 ms were judged to be synchronous. The most interesting finding is that audio advances are detected better than audio delays. It is possible that the human perceptual system is adapted to tolerate larger audio delays than tactile delays in accordance with physical and physiological realities.
- Whole-body vibration stimulation: The results of the experiments show that the PSS is found at an audio delay of about 10 ms. Stimuli with audio delays in the range of – 47 to 63 ms were judged to be synchronous. These results show a similar tendency as vibratory finger stimulation, namely, that audio advances can be detected better than audio delays. The synchronization thresholds vary depending on the kind of stimuli. As expected, the impact type stimulus has lower thresholds than the other stimulus. This result is in line with the results of audio-visual asynchrony experiments. For example, the audio-visual synchronisation threshold of speech stimulus is higher than that for an impact stimulus (Miner and Caudell, 1996).

4.3 Physical level coupling

4.3.1 Introduction

Sound generation requires acoustical energy, which is in the most part supplied by the movement of structures, and this movement is a result of tactile interaction with the structures. Therefore, the sound pressure level and the level of force-feedback (by hitting or by scraping) are coupled by physical laws. Sound radiation from plates due to point forces as described by Fahy (1985). The velocity response at point r of a plate structure to single-frequency force excitation at a point (r_0) is given by:

$$v(\mathrm{r},t) = \exp(j\omega t)\sum_n \frac{\tilde{F}\psi_n(\mathrm{r}_0)\psi_n(\mathrm{r})}{\tilde{Z}_n} \tag{4.5}$$

where ψ_n is the non-dimensional modal velocity distribution (mode shape) of mode n, and $\tilde{F}$ is the force amplitude, and $\tilde{Z}_n$ is the modal impedance. The surface pressure transform can be written as:

$$\left[\tilde{P}(k_x, k_z)\right]_{y=0} = \frac{\pm j\omega^2\rho_0}{(k^2 - k_x^2 - k_z^2)^{1/2}} \frac{\tilde{F}(k_x, k_z)}{D[(k_x^2 + k_z^2)^2 - k_b^4]} \tag{4.6}$$

where k is the acoustic wave number ($k=\omega/c$), k_x and k_z are wave number components, k_b is the free structural wave number at the angular frequency ω ($k_b = (k^2c^2m/D)^{1/4}$), c is the speed of sound, ρ_0 is the mean density, h is the thickness of the plate, m is the mass per unit area, E is Young's modulus, υ is the Poisson's ratio and D is the bending stiffness of the plate ($D = Eh^3/12(1-\upsilon^2)$). The power radiated by the total plate is:

$$\overline{P} = \frac{\rho_0\left|\tilde{F}\right|^2}{4\pi c m^2} \tag{4.7}$$

Equation 4.6 indicates the strong relationship between sound pressure and the applied force amplitude, and Equation 4.7 further illustrates the strong relationship between sound power and the applied force amplitude.

Physical realities play an important role in our multi-modal integration mechanism. Therefore level is an important cue for our brain to integrate information from the various sensory modalities, like simultaneity. An example from our daily experience of multi-modal integration, where the level coupling plays an important role, is hitting an object. By hitting an object, reflected force-feedback information by the object (and of course applied force) and loudness of the hitting sound are coupled to each other by physical laws. During perceptual development each of us has learned that if we strike any object stronger (and get stronger force-feedback), the sound becomes louder (reverse is also valid)[6]. If we strike an object and get very strong force-feedback, we wait to hear a very loud sound. In that situation, if we hear a very quiet sound, the situation is not perceptually plausible for us and we will have difficulty integrating a strong force-feedback information with a quiet sound.

The purpose of this chapter is to investigate the influence of stimulus level on the separation of auditory and tactile events. To measure the perceptual threshold values for the level differences, which lead to the separation of auditory and tactile events, the just-noticeable pressure level difference thresholds need to be known. For this reason first just-noticeable pressure level difference thresholds for the sounds, which are generated by hitting event, were measured (on the condition that the hitting force magnitude remains constant and subjects did not lose the integration of auditory and tactile event). Then in the second experiment perceptual threshold values for the level differences between auditory and tactile information were measured, in other words how much level difference can lead to a difficulty in integrating auditory and tactile information and cause separation of these two events.

4.3.2 Just-noticeable differences

4.3.2.1 Subjects

The same eight subjects participated in the experiments. They were five right-handed men and three right-handed women with self reported normal-hearing ability. Their ages ranged from 21 and 30 years.

4.3.2.2 Experimental set-up and stimuli

The Yamaha DD 55 electronic drum was used to measure the applied force level by the subject[7]. The drum transmits applied force information through the midi port of the sound card to the PC in midi format. A computer program was written and used to calculate the force-feedback level and to play the requested drum sample. The sound was amplified and

---

[6] In addition to the applied force magnitude, the elasticity of the object has an influence on the sound pressure level. For example, by hitting a very soft object (huge size), very strong applied forces can cause only very quiet sound. But it should be taken into account that the level of force feedback is also very low, despite the very strong applied force level. To eliminate the effect of object properties, force-feedback level (not the applied force level) will be used in psychophysical experiments as tactile information.

[7] Another reason to use an electronic drum is that the player's beat does not generate as loud a sound as a typical drum.

delivered diotically through Sennheiser HAD 200 closed-face dynamic headphones. The experiments were conducted in a sound-attenuated room. To minimize the noise which was generated by the electro-drum while the subject beats the drum, the electro-drum was placed outside of the sound-attenuated room. Subjects sat in sound-attenuated room and they extended their hand through the cavity of the sound-attenuated room wall to play drum (Figure 4.15). Subjects wore dark eyeglass to avoid any visual information. The auditory stimulus was a virtual "djembe"[8] sound, which was designed to be physically accurate (the drum sound was presented with a level proportional to the calculated force-feedback magnitude).

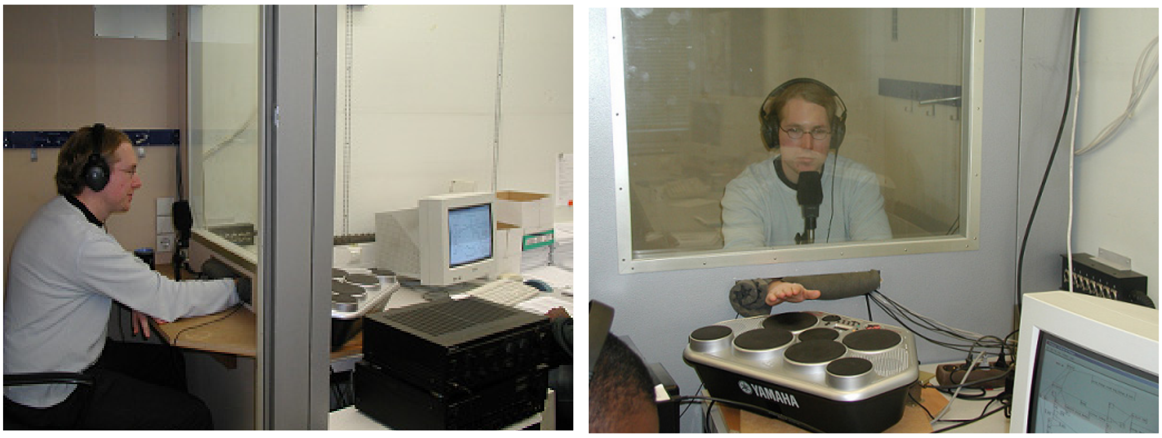


Figure 4.15: Photographs of the experimental set-up.

4.3.2.3 Methodology and procedure

In this experiment, a 2-Interval 2-Alternative Forced Choice procedure was used to determine the minimum level difference, which can be tolerated for a set force level. Subjects were exposed to a number of trials, each trial consisted of a "reference" stimulus (the drum sound was presented with level proportional to the hit-force magnitude) and "test" stimulus (the drum sound was presented with level greater than would be expected by the hit force). Test and reference stimuli were presented randomly. The subjects were instructed to try and hit the drum with a constant force throughout the experiment. The applied force was controlled throughout the experiment, if the variation in the force applied by the subject is more than 10%, the subject received an auditory warning signal and the experiment was repeated again. After each trial, the subject was asked whether the first or the second sound was the loudest.

---

[8] Djembe is a type of African drum.

The amount of decrement was adjusted following a two-down one-up rule. Thus, after two correct answers the sound pressure level was reduced by 0.5 dB, and after one incorrect answer the sound pressure level increased by 0.5 dB. A measurement was terminated after ten reversals (a point where the stimulus level reversed direction, either a peak or a trough). Four threshold values were obtained for all subjects and for both level increment and decrement, from these the mean was derived.

Thresholds were calculated from the mean of the peaks and troughs,

$$\text{Difference threshold} = \frac{\left\lfloor \sum_{i=2}^{i=5} p_i + \sum_{j=2}^{j=5} t_j \right\rfloor}{N} - R \qquad (4.8)$$

where p_i is the sound pressure level of peak I, and t_j is the sound pressure level of trough j; N is the number of reversals; R is the reference magnitude (Morioka and Griffin, 2000). The first two reversals were omitted from the calculation of estimate in order to reduce starting error (as suggested by Levitt and Rabiner 1967).

4.3.2.4 Results

The JND values for individual participants and mean of the JNDs are shown in Figure 4.16.

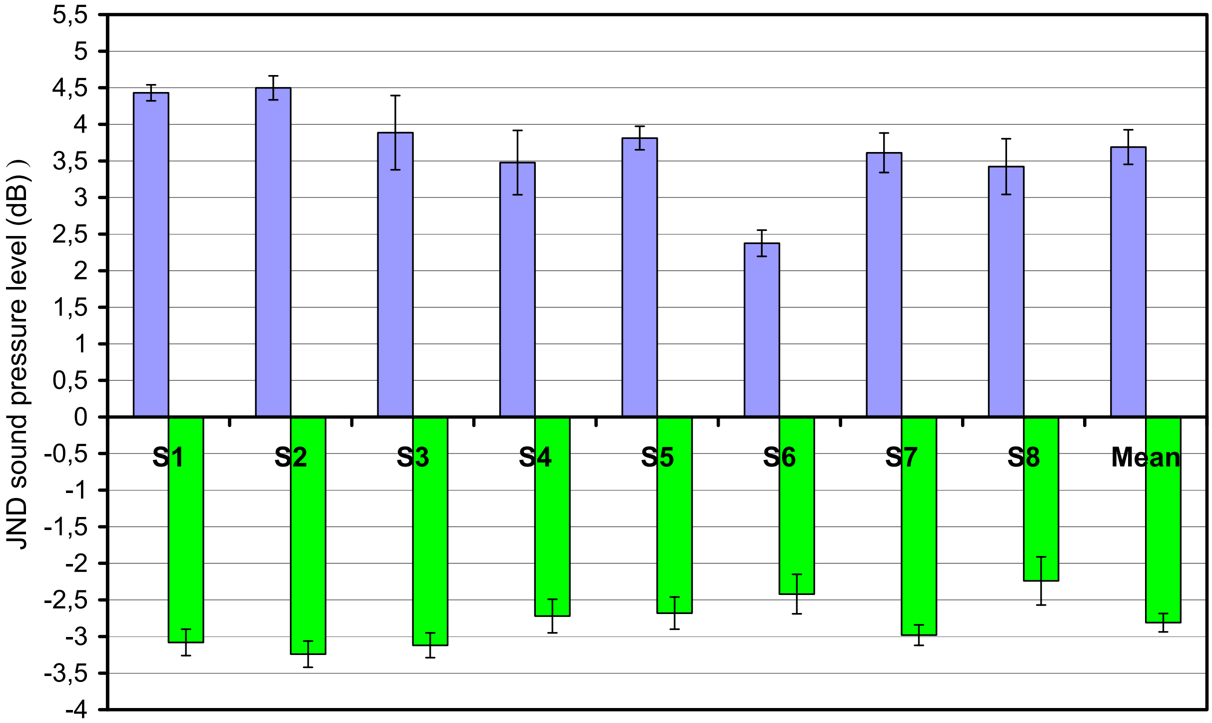


Figure 4.16: Just-noticeable sound pressure level difference thresholds of 8 subjects.

The difference thresholds ranged from 2.37 dB to 4.5 dB for the level increase, and from 2.24 dB to 3.24 dB for the level decrease and the mean values are 3.68 dB for the level increase, and –2.81 dB for the level decrease.

4.3.3 The psychophysical thresholds of level coupling

The purpose of this experiment was to determine the thresholds for level coupling which cause the segregation of auditory and tactile events. Participating subjects, stimuli and experimental set-up of this experiment were the same as in the JND experiment (Section 4.3.1).

4.3.3.1 Methodology and procedure

In this experiment, the method of limits was used. The subjects were instructed to try and hit the drum with constant force throughout the experiment and the applied force by the subject was controlled throughout the experiment. If the force-level deviation was more than 10%, the subject received an auditory warning signal and the experiment was repeated again. In the first part of the experiment, with the first hit, the drum sound was presented with the level proportional to the hit-force magnitude, on the second hit, the level of the drum sound was increased by 3.6 dB, and in each hit subsequently, the level was increased stepwise (3.6 dB, 7.2 dB, 10.8 dB, 14.4 dB etc.). After each stimulus presentation, the subject was asked “Is it acceptable that your beat can result in this sound level? (YES/NO)”. The act of increasing of level continued until the subject says “No”: The subject is no longer able to integrate the two events. In the second part of the experiment the level was decreased stepwise (2.8 dB, 5.6 dB, 8.4 dB etc.). Eight threshold values were obtained for all subjects and for both level increment and decrement. The procedure was the same as in the first part of the experiment.

4.3.3.2 Results

The threshold values for individual participants and the mean of the individual thresholds are shown in Figure 4.17. The thresholds ranged from 14.75 dB to 22.9 dB for the level increase, and from –8.2 dB to –14.8 dB for the level decrease and mean values are 17.6 dB for the level increase, and 11.2 dB for the level decrease.

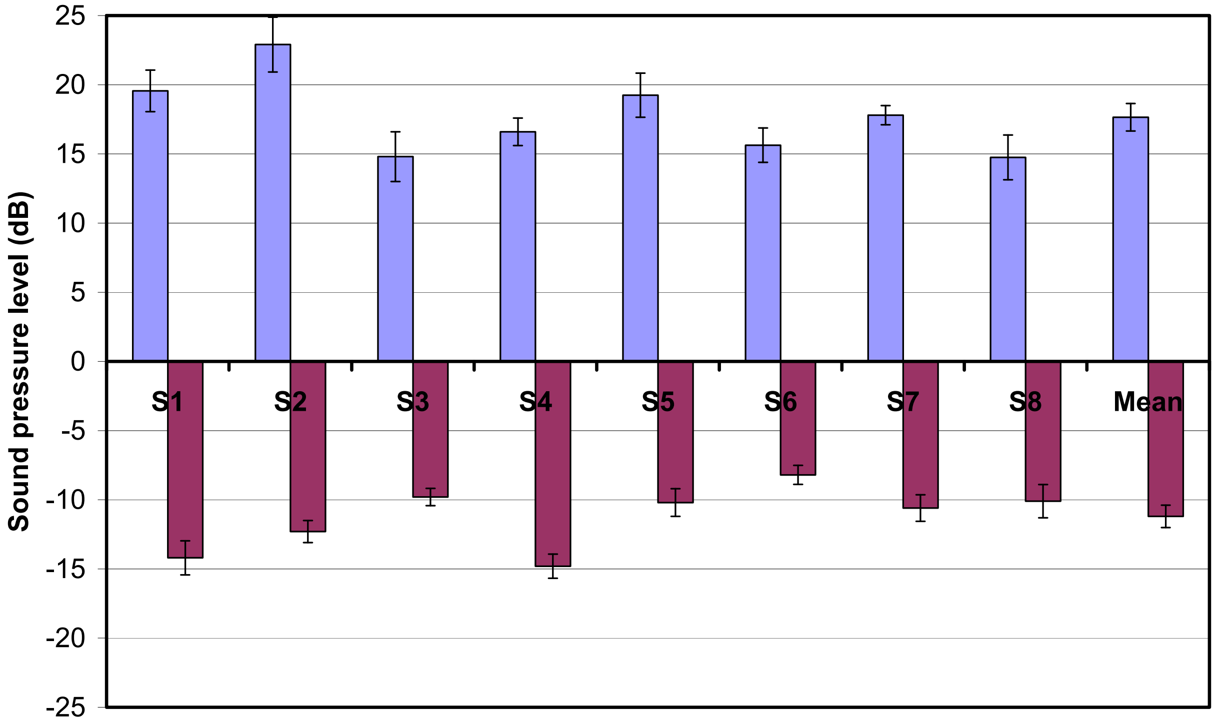


Figure 4.17: The threshold values for the level differences between auditory and tactile stimuli.

4.3.3.3 Discussion

The results of the psychophysical experiments show that humans have a tolerance level for the level difference between auditory and tactile modalities which leads them to perceive one event or two events. The tolerance levels are found as 17.6 dB (S.S.E. 2.03) for the level increase, and 11.2 dB (S.S.E. 1.73) for the level decrease. One of the reasons for these large tolerance levels can be in our daily life, we meet different physical conditions and interact with different physical objects (material, size, and modal properties etc.) and these differences lead us to adapt to the integration of different intensities of the two sensory modalities.

4.4 Frequency

4.4.1 Introduction

Vibrating objects cause the disturbance which moves through the medium and this disturbance generates the sound waves. Therefore, the frequency of the sound and the frequency of the vibration are coupled to each other by physical laws.

Human response to vibration (or to tactile feedback) and sound is strongly dependent on the frequency of the stimulus. Therefore the frequency coupling between auditory and tactile stimuli plays an important role in our integration mechanism of auditory and tactile information (Sugita and Suzuki, 2003; Kohlrausch and van de Par, 1999). The goal of this section is to investigate the influence of frequency on the integration of auditory and tactile information. For this purpose, first of all, the fundamentals of the unimodal (auditory/tactile) frequency perception will be given and then the segregation of the integrated multi-modal (auditory-tactile) event related to the frequency differences will be discussed on the basis of the results of the psychophysical experiments.

Frequency Sensitivity of the Auditory System

Sounds that are audible to the human ear fall in the frequency range of about 20-20,000 Hz, with the highest sensitivity being between 500 and 4,000 Hz. In order to determine the sensory capacity of the auditory system, measurements of the human ability to discriminate the changes in frequency of a pure tone were conducted. Just-noticeable frequency differences for the auditory system were reported by Zwicker and Fastl (1999). They investigated that, at frequencies below 500 Hz, we are able to differentiate between two tone bursts with a frequency difference of only about 1 Hz, and this value increases in proportion to frequency and is approximately 0.002*f (Figure 4.17) above 500 Hz.

Just-noticeable frequency differences are dependent to the tone burst duration. If the burst duration is shorter than 200 ms, JNDs increase. For durations beyond 200ms JNDs remain constant.

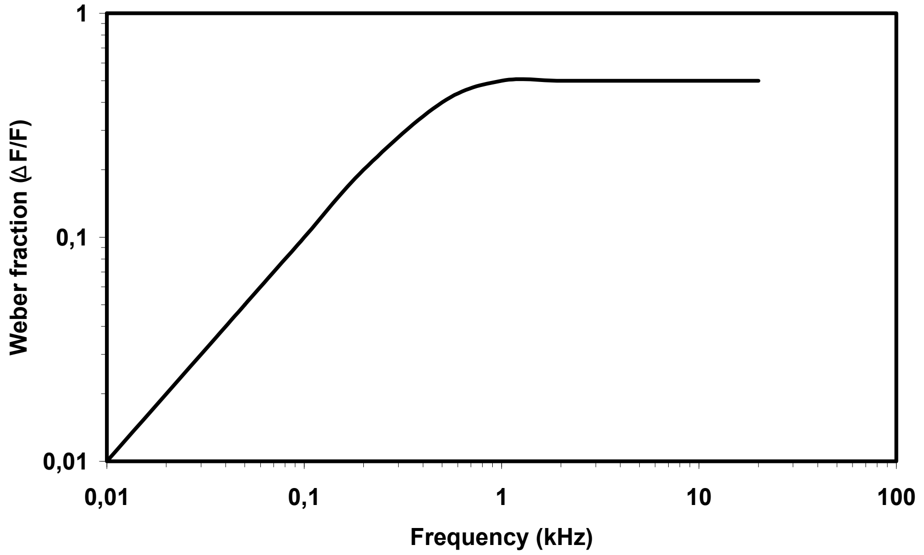


Figure 4.18: Just noticeable frequency differences for pure tones.

Tactile frequency sensitivity

Our skin is sensitive to the frequency range from 8 Hz to 1000 Hz and the highest sensitivity is reached in the range of 200-300 Hz. Just-noticeable frequency differences for sinusoidal vibrations and tactile pulses on the finger and volar forearm were measured by different researchers (Goff,1967; Mowbray and Gebhard, 1957; Rothenberg et al., 1977) (Figure 4.18). Frequency discrimination of the tactile channel is fairly good at low frequencies but deteriorated rapidly as frequency increased (Mowbray and Gebhard, 1957). The results show that the difference limen for sinusoidal vibrations on the finger is about 30 %[9] and pulses can be discriminated better than the sinusoidal vibrations. These results indicate that the skin is rather poor at discriminating frequency in comparison to the ear (auditory system).

Besides of perceptual aspects, it is also interesting to know the vibration frequency range of hand-power tools, because hand-arm vibrations are mostly produced by them. The frequency range of hand power tools is very large extending from a few Hertz to tens of kiloHertz. Griffin (1990) reported that the significant frequency range for some percussive tools is the region of 32 Hz, for chain saws the region of 125 Hz, and for commonly hand-power tools from 8 Hz to 4 kHz.

The range of frequencies most often associated with effects of whole-body vibration in the context of health, activities and comfort is approximately 0.5 to 100 Hz (Griffin, 1990). Just-noticeable frequency differences for whole-body vibrations were measured by Bellmann (2003). Humans are able to differentiate between two vibrations of 5 and 5.4 Hz ($\Delta f = 0.4$).

---

[9] The diffrence limen of visually impaired people is about 10% (Laming, 1986).

Above 5 Hz Δf increases in proportion to frequency and is about 0.34 * f – 1.25 Hz. This equation is applicable for reference frequencies between 5 and 40 Hz.

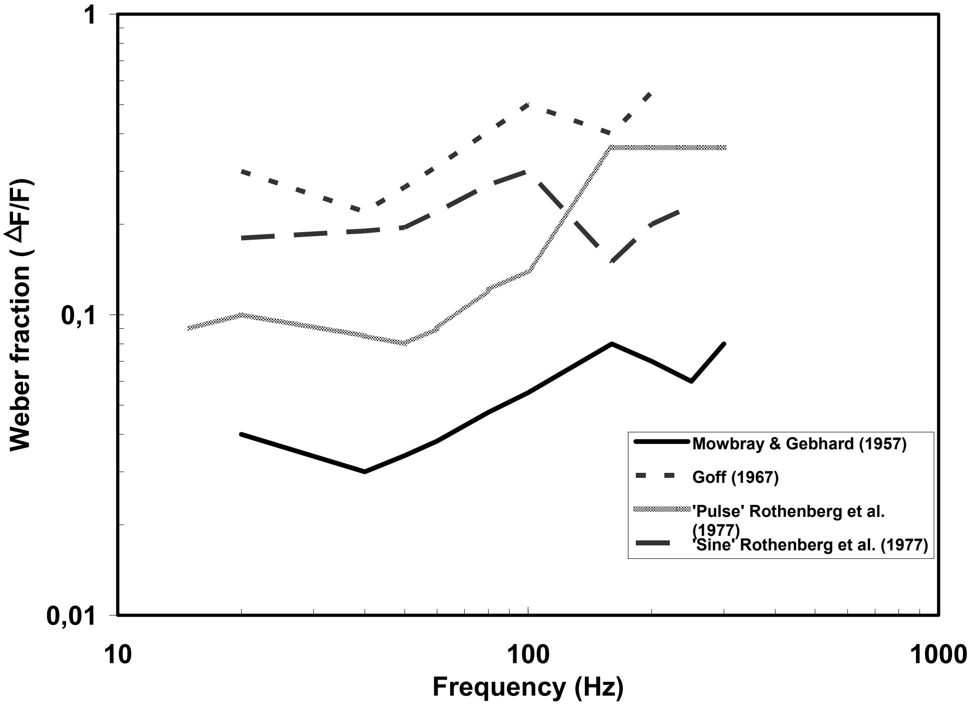


Figure 4.19: Difference limen for frequency discrimination of sinusoids measured on the finger and volar forearm (adapted from Verrillo and Gescheider, 1996).

The role of frequency on multimodal integration

Auditory-visual integration: A number of studies attempted to investigate the role of frequency on the auditory and visual integration. Helmoltz (1954) has suggested an analogy between color and pitch, related to their generation mechanism (they are both caused by waves). He gave the following analogy list between the notes of the piano[10] and the colors of the spectrum:

G	Red
G#	Red
A	Red
A#	Orange-Red
B	Orange
c	Yellow
c#	Green
d	Greenish Blue

---

[10] $G_4 = 392$ Hz, $A_4 = 440$ Hz, $B_4 = 493.88$ Hz, $C_4 = 261.63$ Hz, $D_4 = 293.66$ Hz, $E_4 = 329.63$ Hz, $F_4 = 349.23$ Hz.

d#	Cyanogen-blue
e	Indigo-blue
f	Violet
g	Ultra-violet
g#	Ultra
a	Ultra
a#	Ultra
b	End of the solar spectrum

Auditory-tactile integration: Two recent studies discussed the integration of information in the specific context of haptic-audio texture perception. McGee, Gray and Brewster (2002) suggested that the frequency of haptic and audio stimuli may have influence on multi-modal roughness perception, but they did not provide any experimental results related to the sensitivity to frequency differences between auditory and tactile information.

If we take into account the principles of the Gestalt psychologists in vision, stimuli (auditory/tactile) with the similar frequency will tend to group together related to the proximity analogy.

Measurement method

In order to measure the thresholds of multi-modal integration related to frequency, different experimental methods can be applied. In a response method, a subject is asked whether two stimuli (e.g. auditory and haptic information) are caused by the same product or not. For this type of experiment, it may be useful to define a context or product (e.g. imagine a razor etc.) for the test subjects. It must be explained to the subjects very clearly that they should imagine an event or a product and judge whether the multi-modal information is caused by the same event (the same product) or not.

Another measurement method consists in evaluating a multi-modal attribute which may be influenced by the physical properties of the multimodal stimuli which contribute to the multi-modal event. For example, roughness is such a multi-modal attribute, related to surface texture evaluation. The frequencies of the haptic and the visual stimulus as well as the frequency of the auditory stimulus have an influence on roughness perception. For electrical products (e.g. drill, electric razor, hair-dryer), the performance of the product can be taken as the multi-modal attribute. The disadvantage of this method is the difficulty in analyzing the measurement results to determine the separation thresholds of an integrated multi-modal event.

4.4.2 Experiment

An experiment was conducted to investigate the influence of frequency on the integration of auditory and tactile information.

4.4.2.1 Set-up

The Saitek tactile feedback mouse was used to present the tactile information (vibrations) to the subjects. This mouse contains a motor that relates the vibration or the force-feedback sense to the hand guiding it. The participants were instructed to hold the mouse in their hand and lift it from the table to avoid unwanted structural vibrations which can be generated from the contact between the mouse and the table and also to minimize the noise generated by the mouse.

The auditory stimulus was presented from a PC. It was amplified and delivered diotically through Sennheiser HDA 200 closed-face dynamic headphones which have a very high sound isolation level and therefore mask the background noise of the mouse when it generates the signal. The experiments were conducted in a sound-attenuated room.

4.4.2.2 Subjects

The same nine subjects, four men and five women, aged between 20 and 29 years, participated in the experiment. The subjects were undergraduate students and paid on a hourly basis. All subjects had normal hearing and were right handed, with no known hand disorders. They used their right hand for the experiment.

4.4.2.3 Stimuli and procedure

The tactile stimuli were sinusoidal vibrations varying in frequency (4, 10, 50, 63, 80 and 100 Hz). Auditory stimuli were pure tones at fifteen different frequencies (31.5, 40, 50, 63, 80, 100, 125, 160, 200, 250, 315, 500, 630, 1000 and 2000 Hz).

In this experiment, subjects should imagine that the vibration and auditory information were produced by any device (or product) which they want to imagine.

Tactile and auditory stimulus pairs were presented simultaneously in a random order. Each condition was presented four times. The subjects were asked to report whether the auditory and tactile information caused by same product (same event) or not (yes/no answer option).

Human sensitivity to vibration is highly frequency and also magnitude dependent. To eliminate the effects of magnitude on the experiment, all tactile stimuli were filtered according to ISO 5349 (frequency weighting for hand-transmitted vibration). All auditory stimuli were also filtered using the A filter. The peak-to-peak level of vibration displacement was 0.05 mm (at 80 Hz) and the sound pressure level was 56 dB(A).

Level differences between vibration intensity and sound pressure level may cause dominance of one modality on the other modality, or masking of some perceptive aspects (Forthergill, 1972). By proper controlling the amplitude of the vibrations and loudness of the sounds, it was hoped to avoid masking effects between the modalities.

4.4.3 Results and discussion

The percentages of positive responses are shown in Fig. 4.19 to 4.22 as a function of the acoustic frequency (Fig. 4.19: Vibration frequency 10 Hz, Fig. 4.20: 50 Hz, Fig. 4.21: 63 Hz, Fig. 4.22: 80 Hz).

The maximum of the responses curve (Point of Subjective Equality, PSE) for a 10 Hz vibration is found at 40 Hz pure tone, and the 75 % thresholds (Just Noticeable Differences, JNDs) are 30 and 55 Hz (Fig. 4.19). Fundamental frequency and second harmonics of 10 Hz are assumed to be non-audible for normal hearing subjects, therefore, it is possible that participants try to match the 10 Hz vibration with a 40 Hz pure tone (fourth harmonic).

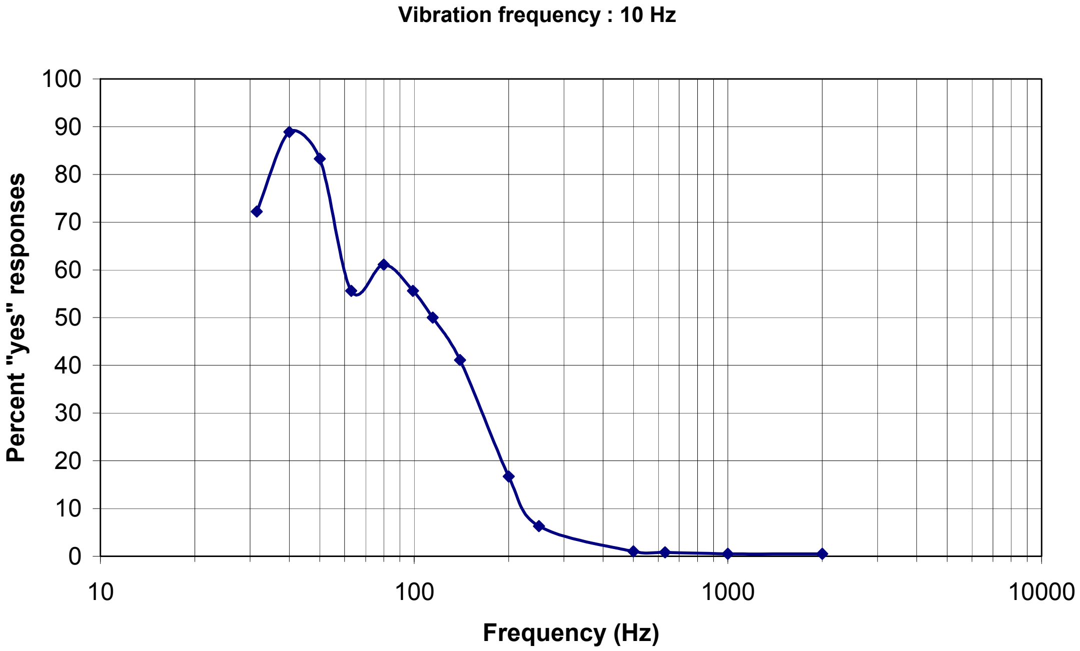


Figure 4.20: The percentage of positive responses for 10 Hz vibration as a function of the acoustic frequency

The PSE value for a 50 Hz vibration is a 63 Hz pure tone, and also the second harmonic of 50 Hz (100 Hz) shows an increase of the percentages of positive responses at neighboring 1/3 octave band frequencies. The 75 % thresholds are 30 and 75 Hz for the first harmonic and 90 and 130 Hz for the second harmonic (Fig. 4.20).

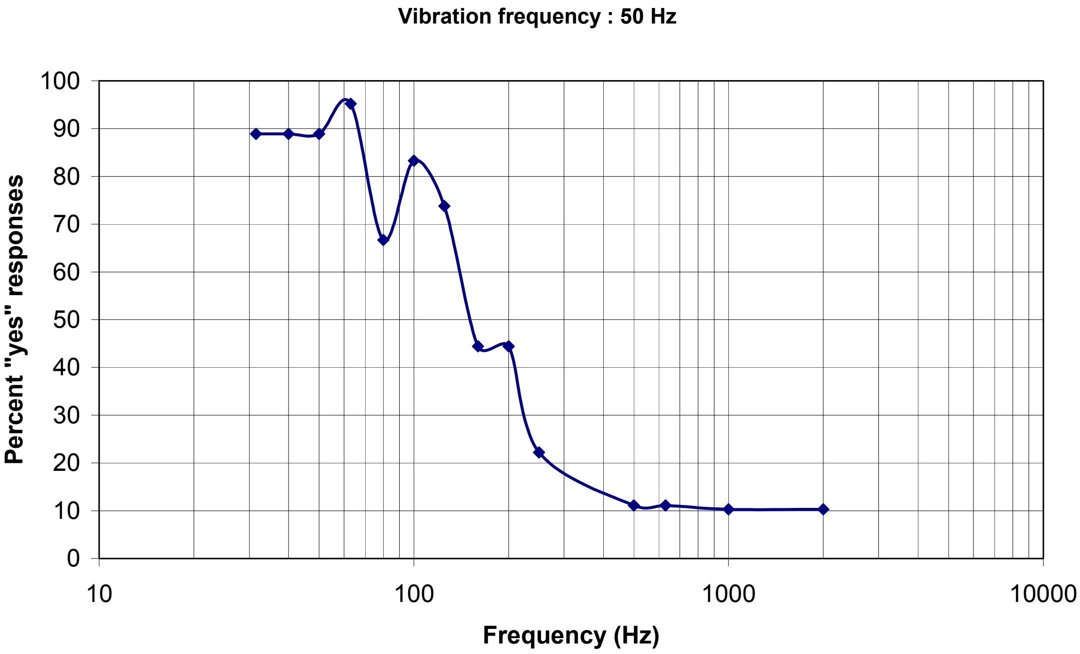


Figure 4.21: The percentage of positive responses for 50 Hz vibration as a function of the acoustic frequency

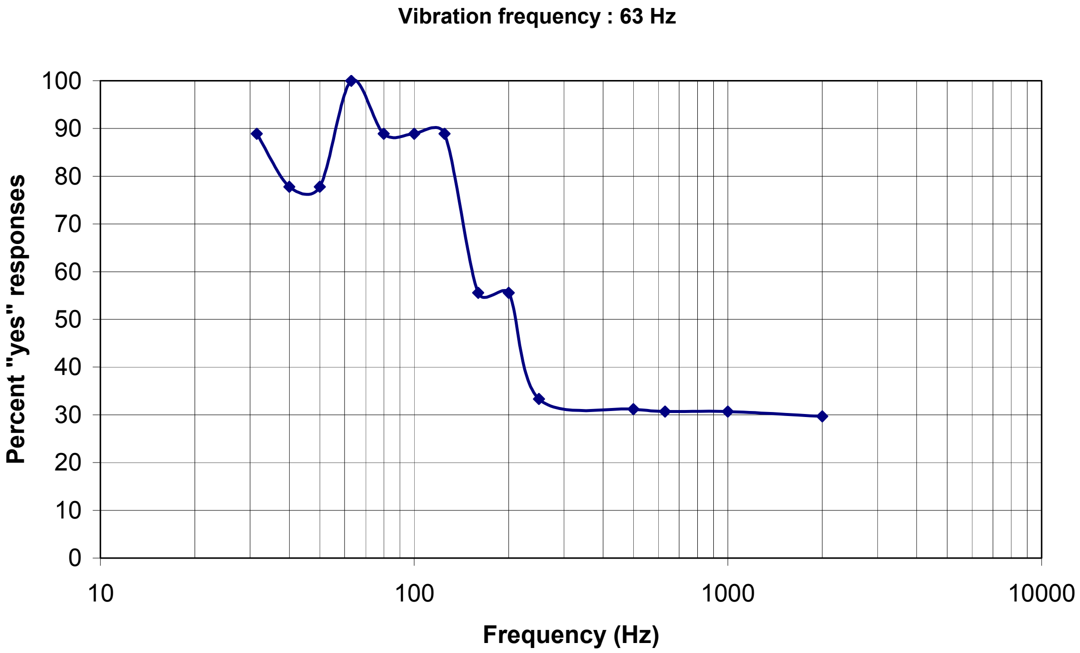


Figure 4.22: The percentage of the positive responses for 63 Hz vibration as a function of the acoustic frequency

The PSE value for a 63 Hz vibration is found at a 63 Hz pure tone, and also the second harmonic of 63 Hz (125 Hz) has a local maximum of positive responses. The 75 % thresholds are 30 and 150 Hz (Fig. 4.21).

The PSE value for an 80 Hz vibration is found at a 100 Hz pure tone, and also at 40 Hz an increase of the percentages of suit responses is observed. The 75 % thresholds are 30 and 115 Hz (Fig. 4.22).

The subjects could not match any suitable pure tone for a 4 Hz vibration (Fig. 4.23). An explanation may be that 4 Hz is a pulsation-type tactile stimulation, and a sinusoidal tone may not integrate with a pulsation.

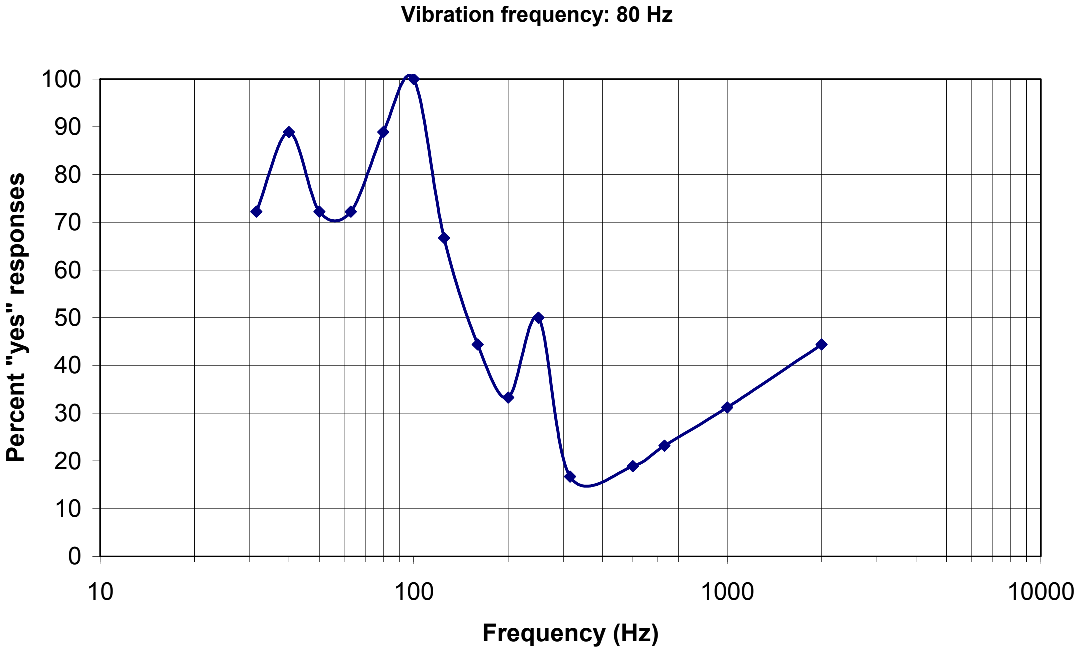


Figure 4.23: The percentage of the positive responses for 80 Hz vibration as a function of the auditory frequency

4.4.4 Conclusions

The results of this experiment suggest that frequency is an important cue for the segregation of auditory and tactile events. To find the most suitable multi-modal stimulus combination for the multi-modal integration, the subjects tend to prefer pairs having the same frequency for the auditory and tactile stimuli. These results show agreement with the proximity rule of Gestalt theory. In most cases, subjects judge also the second or other harmonics of the vibration frequency as being suitable for the auditory frequency, in order to integrate the two perceptual components.

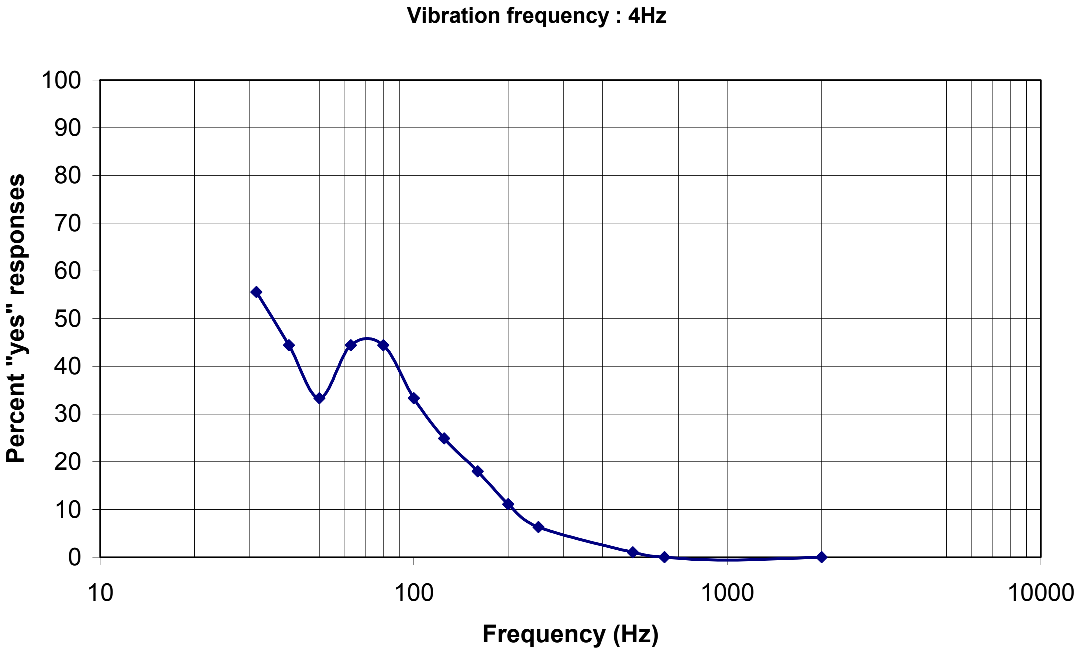


Figure 4.24: The percentage of the positive responses for 4 Hz vibration as a function of the acoustic frequency

A combination of a pulsation type tactile stimulus with a high frequency auditory stimulus (4 Hz vibration situation) results with the segregation of auditory and tactile events. Subjects could not find any suitable auditory stimulus for 4 Hz vibration stimulus. 75 % threshold values show that the subjects meet difficulty approximately above 160 Hz (auditory frequency), to integrate auditory and tactile information for the vibration frequency of 50 Hz, 63 Hz, 80 Hz, and approximately above 50 Hz for the vibration frequency of 10 Hz.

Similar psychophysical experiments were conducted with whole-body vibration as tactile stimulus. The results showed that despite the conflicting information (the frequencies of the auditory stimulus and whole-body vibration stimulus were different), the percept does not segregate into isolated percepts in each modality. Therefore only one example result is shown in Figure 4.24. The subjects have great tolerance for the auditory and whole-body vibration frequency conflict, yes responses were always above the 75 % with the exception that they could not match a pulsation type stimulus with a high frequency stimulus.

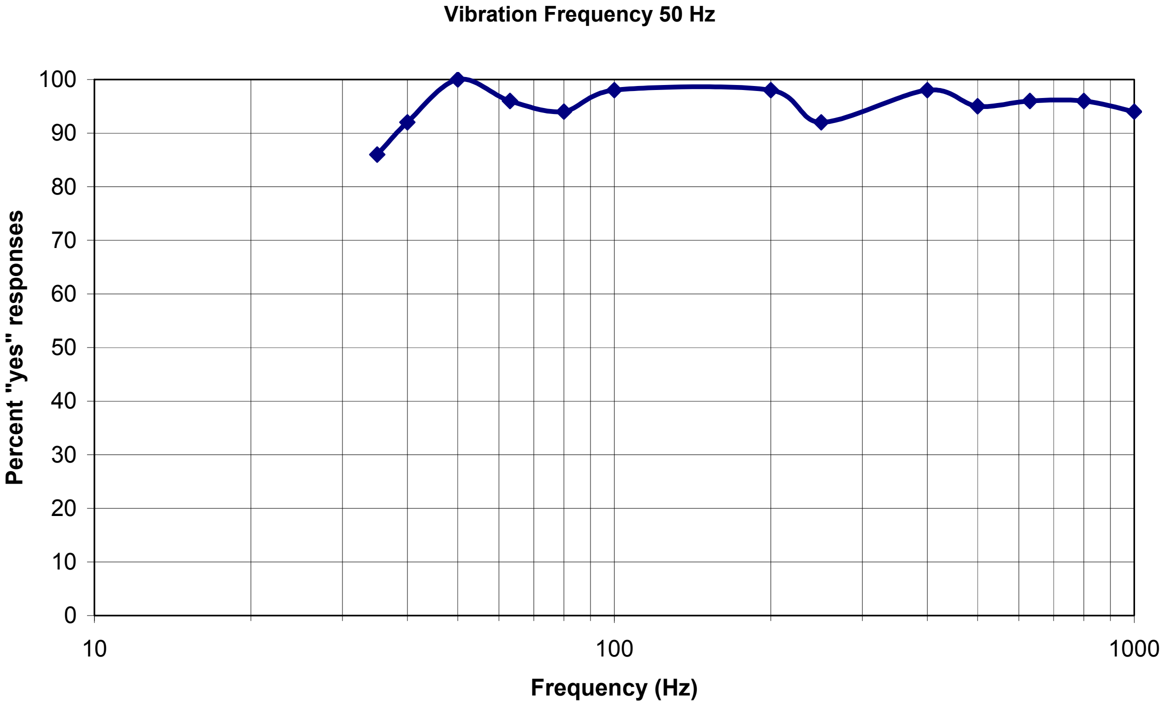


Figure 4.25: The percentage of the positive responses for 50 Hz whole-body vibration as a function of the acoustic frequency.

4.5 Locations of the auditory and tactile events

4.5.1 Introduction

An important physical property which plays a role for the multi-modal integration is the location of the event. Naturally, if auditory, tactile and visual information were generated by (one) same multi-modal event, the locations of the auditory, tactile and visual events should coincide. For example, when manipulating an object, the tactile, auditory, and visual information related to that object will all typically emanate from approximately the same spatial location (Driver and Spence, 2004). Neurophysiological studies showed that the various receptive fields of SC (Superior Colliculus) neurons are organized into overlapping visual, auditory, and somatosensory maps, in effect creating a multisensory map of space (Stein and Meredith, 1993), and the multisensory stimuli which are delivered in close space proximity will evoke higher responses in multisensory neurons in superior colliculus (Stein et al. 2004; King 2004). The purpose of this section is to determine the minimum audible angle between the auditory and tactile events that leads the listener to perceive that the locations of the auditory and tactile events do not coincide and the percept segregates into isolated percepts in each modality (two separate events) instead of forming an integrated multi-modal event.

Before answering this question, a short overview on the localization blur could be useful for the further discussions. Localization blur ($\Delta(\varphi = 0)_{\min}$) is defined by Blauert (1997) as "the smallest change in a specific attribute or in specific attributes of a sound event or of another event correlated to an auditory event that is sufficient to produce a change in the location of the auditory event". A number of measurements were conducted to measure the localization blur for horizontal displacement of the sound source away from the forward direction (see Blauert 1997). The absolute lower limit for the localization blur is about 1° and the localization blur for broadband noise is 3.2°.

In multi-modal interaction research there are several studies regarding different visual-auditory interaction effects on sound-source localization such as the ventriloquist effect (for more details, please see Kohlrausch and van de Par, 1999). Blauert (1970) reported an experiment in directional hearing with simultaneous visual stimulation. He measured the localization blur of speech from the front with and without a simultaneous television image of the person speaking. The loudspeakers were 7m in front of the subject and they were switched on in random order. The subjects were required to say whether their auditory event was above or below, to the left or right of the forward axis. Localization blur of the direction of the auditory event in both the horizontal and median planes proved not to depend on whether the visual image of the person speaking was shown.

However, there are only very few investigations about the influence of tactile perception on the localization of sound sources, and they are not readily applicable for virtual environments (Pick et al.,1969; Fisher, 1968).

Since the location of the whole-body vibration stimuli (seat) is always the same, conflicting spatial information does not cause segregation. Therefore this investigation will be carried out only for tactile stimulation on the fingers.

4.5.2 Experiment

An experiment was conducted to investigate the minimum audible angle between the auditory and tactile events that leads to the segregation of the multi-modal event.

4.5.2.1 Set-up

In this experiment, a loudspeaker array which consists of nine loudspeakers was used to present the auditory stimulus (Fig. 4.25) The loudspeakers were placed 75 cm in front of the subject. To present the tactile stimulus, two electrodes were attached to the index finger of the subject's right hand.

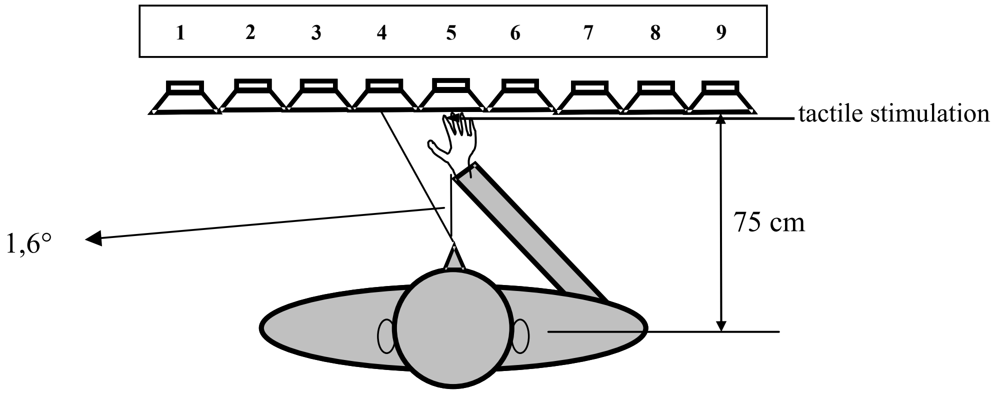


Figure 4.26: Experimental setup.

4.5.2.2 Subjects

Eight subjects, four men and four women, aged between 22 and 29 years, participated in the experiment. The subjects were undergraduate students and paid on an hourly basis. All subjects had normal hearing and were right handed, with no known hand disorders. They used their right hand for the experiment.

4.5.2.3 Stimuli and procedure

Touch-induced sound and tactile feedback such as caused by the scraping on an adhesive paper (sandpaper, grid number 60) were generated in the computer environment. The durations of the stimuli were 0.7 s.

The loudspeakers were switched on in random order. Each condition was presented ten times. The subject's vision was blocked in each condition by an acoustically transparent curtain placed between subject and loudspeaker array. In the first part of the experiment, the localization blur of touch-induced sound was measured without a simultaneous tactile feedback (Exp.1: sound only condition). In the second part of the experiment, the localization blur was measured with a simultaneous tactile feedback (Exp.2: sound & tactile condition).

The subjects were asked whether the sound was perceived from their index finger or not (i.e. whether the position of the auditory event and the position of the tactile information coincide or not-two interval forced choice).

4.5.2.4 Results

The percentages of positive responses for the "sound only" condition are shown in Fig. 4.26. The localization blur ($\Delta(\varphi = 0)_{min}$) of a scraping sound from the front without tactile stimulation is 4° ± 0.1°.

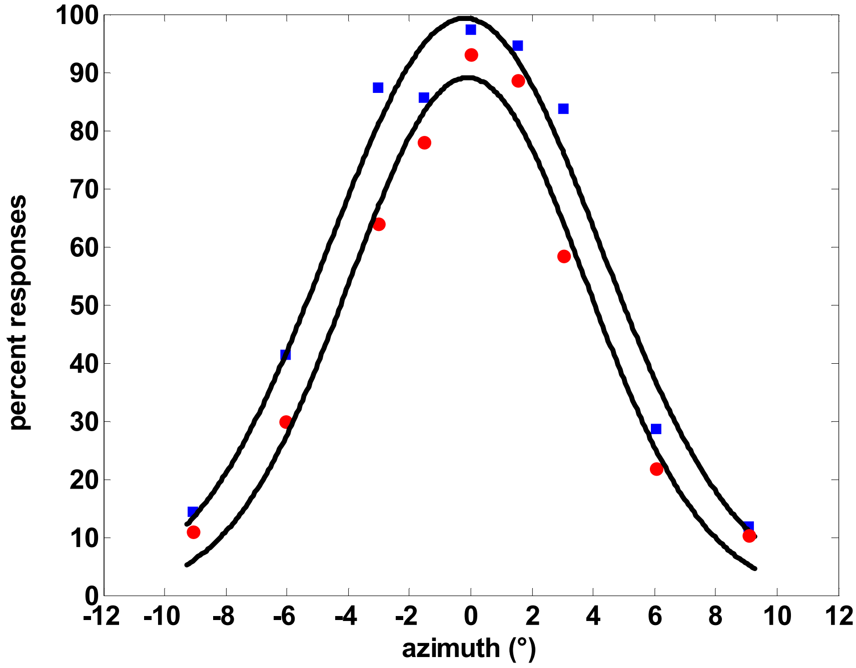


Figure 4.27: The percentages of positive responses for "sound only (◆)" and "sound and tactile stimuli (■)" condition.

The percentages of positive responses for “sound and tactile” condition are shown in Fig. 4.26. The localization blur ($\Delta(\varphi = 0)_{min}$) of scraping sound from the front with simultaneous tactile stimulation is 5.3° ± 0.3°.

4.5.2.5 Discussion of results

The localization blur of scraping sound from the front was found as 3.9° which is in agreement with the literature data. For example the localization blur of broadband noise is 3.2° (Haustein and Schirmer, 1970). The minimum audible angle that allows the subjects to notice the locations of the auditory and tactile events do not coincide is 5.3°. Simultaneously presented electrotactile stimulation enlarges the localization blur in the horizontal plane from 3.9° to 5.3°. Dependent t-test of the means show that both conditions differed significantly form each other ($\underline{t}(9)=-3.25$, $\underline{p} < 0.05$; 2-tailed). This result shows that tactile stimulation has an influence on the localization of the sound sources and it is likely that the tactile stimulation pulls the auditory source to the direction of its location.

Chapter 5

Auditory-Tactile Interaction

5.1 Introduction

Multimodal information which are obtained from different sensory channels are integrated in the central nervous system to produce a single unified percept. The integration is a complex process and depends on the interaction between sensory channels. Two or three modalities can be combined and the resulting multimodal percept may be a weaker, stronger, or altogether different percept (McGee, Gray and Brewster, 2002). Knowledge on the rules of the multimodal interaction process is a prerequisite for designing virtual reality or multimedia displays.

Different hypothesis were suggested to explain the multimodal interaction process. According to the *modality appropriateness* hypothesis, the sensory modalities are differently suited to process incoming information. The modality which is most appropriate for the specific task demands will be favoured (Lederman, Thorne and Jones, 1986; Welch and Warren, 1980). Vision is the superior modality for spatial tasks, therefore it is expected that vision should dominate the tactile sense and audition when discrepancies in object size, shape, and spatial location are involved. Audition is the superior modality for temporal tasks, therefore it is expected that audition should dominate the tactile sense and vision when temporal discrepancy such as rate and duration arises.

The modality-appropriateness hypothesis does not allow quantitative predictions about the relative weightings of the different modalities. In multidimensional tasks, it is possible that information from different modalities will contribute to the resulting percept with relative weightings (Lederman and Klatzky, 2004). For example, in a psychophysical experiment, subjects weighted haptic and visual inputs equally when they were asked to judge the perceived "texture" of abrasive surfaces (Lederman, Thorne and Jones, 1986). In another experiment with bimodal judgments, both haptic and visual information contributes to the perceived size of the objects (McDonnelln and Duffett, 1972).

The degree to which multimodal integration takes place will depend, in part, on the level of congruency between the multimodal stimuli. These were categorized into three groups as conflicting, redundant, and complimentary by McGee (2002). In the redundant condition, each sense obtains the same (congruent) information, and observers might process only one modality of information from the available ones in a multimodal percept. However, observers may report an increase or a reduction in the mental representation of the information. In the complementary situation, the resulting multimodal percept is more than the sum of the individual parts (superadditivity[11]). If each sense obtains contradictory information, the resulting multimodal percept may become distorted or completely lost in the process. This condition is called as conflicting. However, if the multimodal percept is completely lost in the process, this means that information from different sensory modalities are not integrated, but on the contrary segregated.

In summary, if information is obtained from multiple modalities (Figure 5.1):

- One modality can be dominant on the multimodal percept and the observer ignores information from other modalities. If it is assumed that the unified multimodal percept is a weighted combination of multiple sensory inputs, then the weight of the dominant modality is 100 % (or simply 1) and the weights of other modalities 0 % (or 0).
- Information provided from different modalities can be the same (redundant), and the observer processes one modality of information related to the his/her attention, personal preference, or physical/perceptual capabilities
- Different modalities contribute to the multimodal percept with different relative weightings
- The multimodal percept can be more than the sum of the individual parts

The research questions related to multimodal interaction can be grouped in two categories: (1) Investigations on the alteration of a certain percept in one sensory modality due to the presence of a stimulus in another sensory modality, and (2) determination of the relative weights of the different sensory modalities on the multimodal percept. Multimodal interaction is studied in the laboratory mostly by creating a stimulus which provides conflicting information in two sensory modalities. Despite the conflicting information, the percept should not segregate into isolated percepts in each modality (Kohlrausch and van de Par, 1999).

This section addresses the auditory-tactile interaction by means of two different examples of auditory-tactile interaction. In the **Section 5.2,** investigations on the effect of loudness on the haptic force-feedback perception are reported. In **Section 5.3,** auditory-tactile texture perception is investigated. In each section some examples from auditory-visual interaction studies are given, because a large part of the literature is devoted to auditory-visual interaction while only very few investigations address auditory-tactile interaction.

---

[11] A function f(x) is superadditive, if f(x+y) > f(x) + f(y) (Polya and Szegö, 1976).

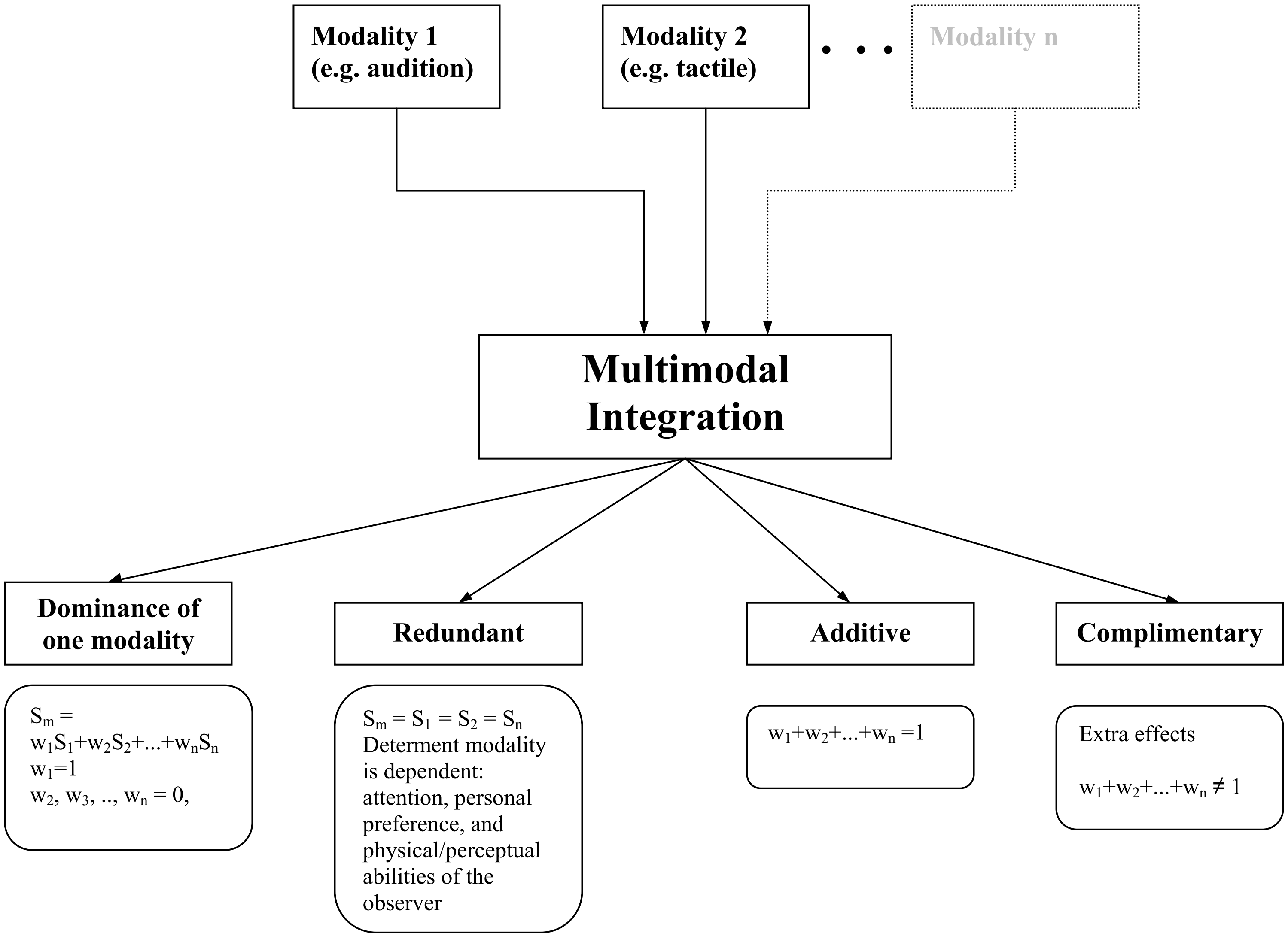


Figure 5.1: Multimodal integration and resulting multimodal percept. S_i is the unimodal response, S_m is the multisensory response, w is the weighting of the unimodal responses

5.2 Effect of loudness on haptic force-feedback perception

5.2.1 Introduction

Characterization and understanding of how humans recognize and manipulate objects are fundamental issues for the design of multimodal user interfaces. Force feedback, which is the sensation of weight or resistance, is an important source of information that enables us to interact with objects. This section describes experiments related to the effect of loudness on haptic force-feedback perception. The physical coupling between sound pressure level and the level of force-feedback, which were generated by beating, was introduced in Section 4.3. In a normal environment, force-feedback would be perceived in our hands as a consequence of the beating event combined with the loudness of the beating sound giving us the required information about how much force we have applied.

A number of studies attempted to investigate the cross-modal interactions related to the level coupling. An example of such auditory-visual interaction is the effect of size (as a visual information) upon perceived loudness. In a study reported by Höger and Greifenstein (1997), the effects of the size of heavy-goods vehicles (HGV) on the perceived loudness were investigated. In their experiment, size and emitted sound level of trucks were systematically varied and subjects rated the relative loudness of HGVs with different size and sound pressure level. The results showed that though keeping the sound pressure level constant small HGVs were rated quieter than large HGVs.

The classical example of a visual-haptic interaction is the size-weight illusion. The perceived size of an object appears to affect weight perception. Koseleff (1957) reported that the perceived weight of an object changed when subjects were required to view the object through reducing or enlarging lenses. Subjects perceived the larger volume to be lighter than the smaller volume, although the two objects had equal weight.

The effects of visual cues on haptic stiffness perception were investigated by Srinivasan, Beauregard and Brock (1996). They measured human performance in discriminating the stiffness of two virtual springs. The results showed that graphically manipulated visual information could give rise to compelling haptic illusions about stiffness of the object.

The effect of auditory cues on the haptic perception of stiffness was investigated by DiFranco, Beauregard and Srinivasan (1997). Their investigation consisted of a series of psychophysical experiments designed to examine the effect that various impact sounds have on the perceived stiffness of virtual objects felt by tapping with a force reflecting device. Auditory cues affect the ability of human to discriminate stiffness. It was found that auditory cues are used for ranking surfaces when there is no difference in haptic stiffness between the surfaces. When the haptic stiffness paired with the sound of a pen striking a cloth, it received only an average of 2.3 points (out of 9 possible points, 9 point scale, 9 = stiffest, 1 = least stiff). However, when the same surface was paired with the sound of a metal screwdriver striking a metal plate, it received an average of 8.6 points.

Simulation of hard contact examples in VR environments is technologically very difficult. There is a general consensus that virtual walls are never as rigid as real walls due to hardware limitations (Colgate, Grafing and Stanley, 1993). Realistic hard contact requires a stiff interface with very small compliance and stiff mechanical design means a heavy interface which can tire the user. Another limitation is that large feedback forces which are required for simulating hard contact can cause user fatigue, therefore most of the haptic interfaces provide only small feedback forces (Burdeau, 1996). To overcome the limitations of the haptic interfaces and provide sufficient realism, corresponding auditory or visual feedback could be useful.

The aims of the experiments, described in this section, are to find out if there is an influence of loudness on the haptic force-feedback perception and to determine the potential benefits of auditory-tactile interaction to overcome the limitations of the haptic interfaces (haptic-alone presentation). Therefore, in a series of experiments, uni-modal and multimodal (congruent and incongruent) presentation is assessed.

5.2.2 Experiments 1, 2, 3

5.2.2.1 Set-up

The Cyber-Grasp force-feedback system was used to present force-feedback information generated from a virtual drum to the subjects (see Fig. 5.2). The force-feedback system consists of a force-reflecting exoskeleton and a hand-tracker glove, and provides force-feedback to each finger of the user relative to the palm of their hand.

The auditory stimulus was presented from a PC. The sound was amplified and delivered to both ears simultaneously through closed-face dynamic headphones, which have a very high sound-isolation level and therefore masked the background noise generated by the force-feedback system. A Silicon Graphics Onyx 2 workstation was used to simulate the virtual environment.

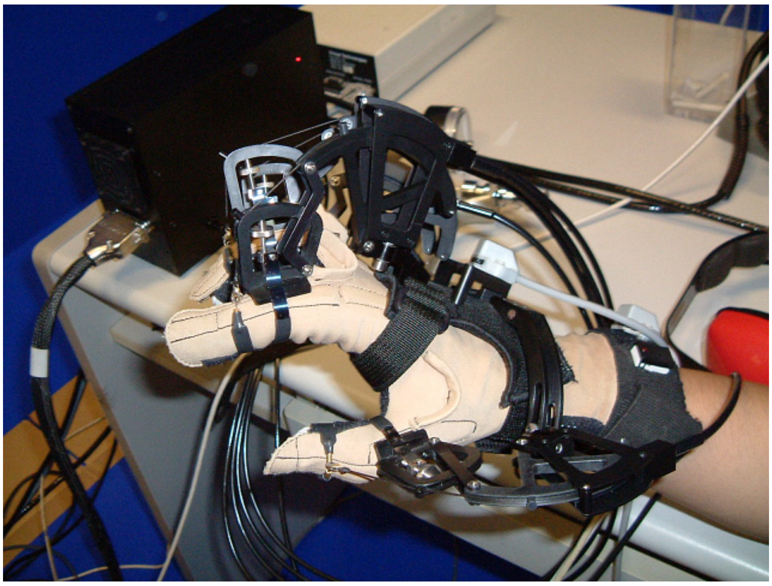


Figure 5.2: Cyber-Grasp force-feedback exoskeleton

5.2.1.2 Subjects

Six subjects, four right-handed men and two right-handed women with self-reported normal-hearing ability and normal tactual/motoric capabilities, participated in the experiments. Their ages ranged from 22 to 34 years.

5.2.1.3 Stimuli and procedure

Example sounds and feedback forces were recorded by hitting a drum (Djembe). The auditory stimulus was a drum sound with level proportional to the hitting-force magnitude. The tactile stimulus was a force feedback, which was applied to the fingers of the subject. Subjects were presented with: Exp. 1) only haptic force-feedback information, Exp. 2) only auditory information, Exp. 3) both auditory and haptic information. The stimulus-pairs were designed to be physically accurate (the drum sound was presented with a level proportional to the hitting-force magnitude). In each experiment, subjects were asked how much force they had applied when playing the virtual drum by assigning numbers to the test stimuli. Strength magnitude was estimated using a magnitude estimation with a standard stimuli. In each trial a standard stimulus was presented and the participant was told that the strongness sensation it produced has a certain numerical value (i.e. 10). After the standard stimuli, a test stimulus was presented and the participant's task was to assign numbers proportional to his/her subjective impression of the strongness related to the standard stimuli. Each trial was presented twelve times in a random order. Before the start of the experiment, thirty anchor stimuli were presented to the subjects so that they could become familiar with the system and the stimuli.

5.2.1.4 Results

Experiment 1,2,3

The responses for all subjects are shown in Figure 5.3 in a log-log scale. Geometric mean and standard error values were computed for the one hundred twenty magnitude estimates obtained from all subjects for each stimulus combination in each condition. The method of least squares technique was used to determine the psychometric functions. The r^2 values for haptic only, auditory only, and auditory plus haptic conditions are 0.89, 0.97, 0.98, respectively. These values show that the fitted psychometric functions are in good agreement with the measurement data.

The auditory only condition produced higher estimates than did either the haptic only and auditory plus with haptic conditions. The mean values of the auditory & haptic condition are between the means of the audition-only and haptic-only conditions.

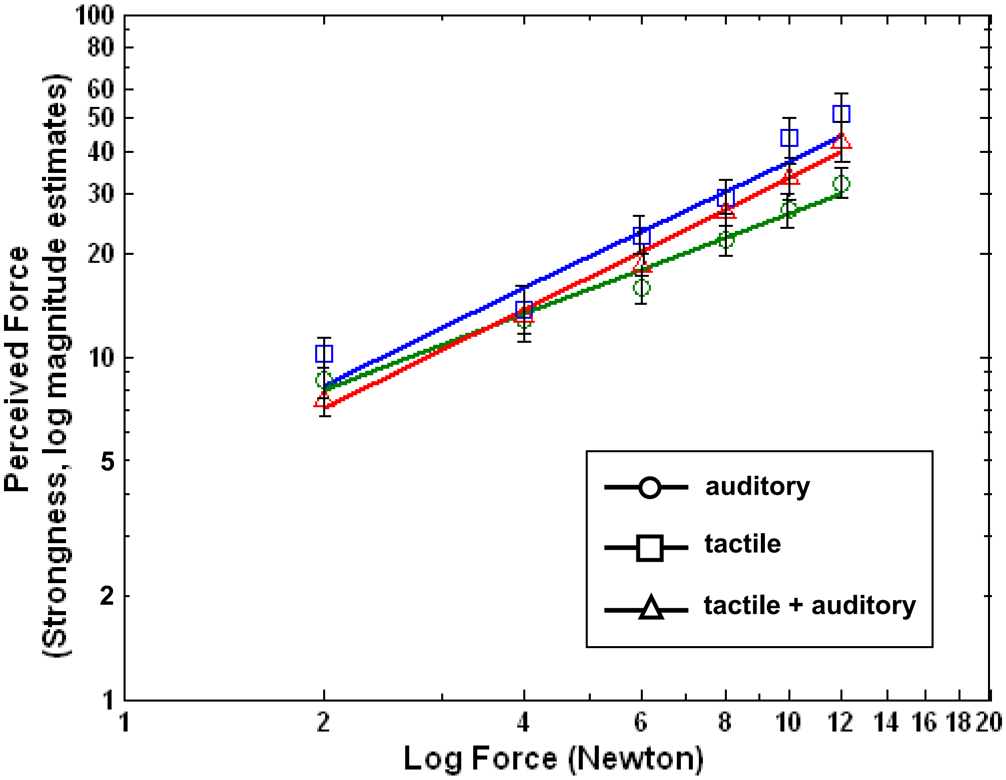


Figure 5.3: Perceived force as a function of log force and the sensory mode of the judgment (auditory, tactile, and tactile & auditory together). The data are averaged across the subjects.

Dependent t-tests of the means showed that all three conditions differed significantly (auditory only vs. tactile only, t(6) = 2.93, $p<0.05$, 2 tailed; auditory only vs. multimodal, t(6) =-2.79, $p<0.05$, 2 tailed; auditory only vs. multimodal, t(6) = 2.88, $p<0.05$, 2 tailed). To calculate a measure of the relative contributions of haptic and audition conditions to the bimodal (auditory & haptic) estimates, a technique, which is proposed by Lederman, Klatzky, Morgan and Hamilton (2002) for the multi-sensory perception of the surface roughness, was used. This statistic, which indicates the percent weighting of the haptic information in the bimodal judgments, was calculated related to the distances between three different conditions:

$$\%\mathrm{T}_{\mathrm{dominance}} = \left[\frac{\left(\mathrm{Mean}_{\mathrm{Haptic+Audition}} - \mathrm{Mean}_{\mathrm{Audition\ only}}\right)}{\left(\mathrm{Mean}_{\mathrm{Haptic\ only}} - \mathrm{Mean}_{\mathrm{Audition\ only}}\right)} \right] \tag{5.1}$$

In this equation, 100% indicates that haptic information is dominant on the bimodal judgments and 0% indicates that auditory information is dominant on the bimodal judgments. The relative weighting of the haptic information is 45%, and accordingly the relative weighting of the auditory information is 55%. These results indicate that the auditory and haptic information were approximately equally weighted and both information contributed to the bimodal judgments of “strongness”.

5.2.3 Experiment 4

Setup and subjects were the same as in the previous experiments.

5.2.3.1 Stimuli and procedure

Auditory and haptic information were presented together. Some stimulus-pairs were designed to be physically accurate, and in some stimulus-pairs the drum sounds were presented with sound pressure levels greater than would be expected from the hitting force. This explanation allows us to investigate the role of loudness increment on the haptic force-feedback perception.

In each experiment, subjects were asked to judge on how much force they had applied when playing the virtual drum by assigning a number to the test stimuli. Strength magnitude was estimated using magnitude estimation with a standard stimulus as reference. The subjects were specifically instructed to ignore the beat sounds they heard, and to base their judgments on only tactile information.

5.2.3.2 Results

The mean values of the logarithmically normalized magnitude estimates are plotted as a function of the sound pressure level (dB(A)) for three different constant force- feedback conditions (4N, 6N, and 12N) (Figure 5.4).

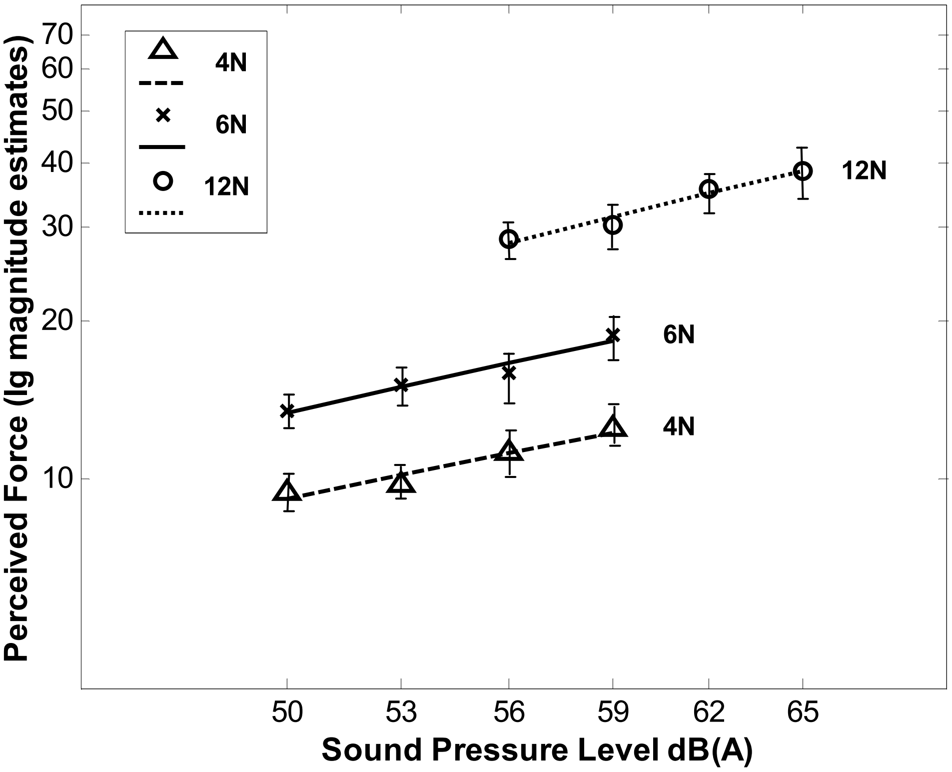


Figure 5.4: Perceived force as a function of sound pressure level and the constant force feedback conditions (4N, 6N, and 12N). The data are averaged across the subjects.

The method of least squares technique was used to determine the psychometric functions similarly to the former experiments. The r^2 values for constant 4N, 6N and 12N conditions are 0.95, 0.95 and 0.96, respectively.

For the constant 4 N force-feedback condition, if sound pressure level increases, it results in an increase in the perceived force-feedback magnitude. For example, for the 50 dB(A) sound pressure level, perceived force magnitude is 9.4 (SE 1.0), and it increases up to the value of 12.4 (SE 1.4) for the 59 dB(A) sound pressure level. The perceived increase in force magnitude as a function of sound pressure level also holds for the 6 N and 12 N applied force condition. The exponents of the 4N, 6N, and 12N conditions are a_{4N} =1.78, a_{6N} = 1.90, a_{12N} =2.1. For the constant 12 N force-feedback condition, the slope of the increase in the perceived force-feedback magnitude as a consequence of increasing sound pressure level is slightly higher than in the 4 N and 6 N conditions.

5.2.4 Discussion

The first observation related to the experimental results is that subjects were able to evaluate the magnitude of force strength on the basis of auditory-alone, force-feedback alone or sound-and-force-feedback together conditions. Auditory judgments were as discriminating as haptic alone judgments.

The results of the Experiment 1, 2 and 3 clearly show that in the bimodal judgments, both haptic and auditory information contribute to the perceived strength of the applied force by hitting a virtual drum with the auditory and haptic information being approximately equally weighted. Observers do not completely ignore one of the available information senses (auditory/haptic) in their judgments, which is in contrast to a number of intersensory bias studies which report that the dominance of the one sensory modality over another modality, such as vision, strongly dominates touch, proprioception, and audition on the size, shape, and spatial-location judgments (e.g.,Walker, 1972; Teghtsoonian and Teghtsoonian, 1970; Pick et al., 1969) or audition dominates vision in temporal tasks involving judgment of rate or duration (e.g.,Welch et al.; 1986).

Related to the physical coupling between sound and force-feedback in this event, both sensory cues are equally informative and information may be obtained easily and quickly from both modalities. This similarity is probably one of the reasons for the equal weighting.

The results of the fourth experiment indicate that the magnitude of strength increases with increasing level in spite of no change in force-feedback as generated by the virtual drum and applied to the subjects hand. Therefore, it appears that participants weight loudness to a greater degree than haptic information if there is no change in force-feedback when trying to discern information from two modalities. But there also appears to be an interaction between applied force and sound-pressure level, with the effect of increasing sound-pressure level having a slightly greater effect upon perceived force if the standard applied force is greater. One interpretation of these findings can be found in the “integration-of-information” hypothesis previously suggested by McGee, Gray and Brewster (2002). They suggested that, if the audio stimulus and haptic stimulus are incongruent but complementary, then judgments will move along in the direction predicted by the direction of the incongruency. The results of the fourth experiment show that when an audio and a haptic stimulus are combined such that the magnitude of the auditory strength information is higher than the haptic stimulus then the judgment of the magnitude of strength is moved along in the direction of increasing strength.

The psychophysical data obtained in the current experiment provide information relevant to the design of multimodal interfaces and contribute to our knowledge on how to effectively combine haptic and auditory information in virtual environments. The results of the current study indicate that auditory information can be useful in overcoming the limitations of haptic interfaces and provide further realism. Multimodal interaction can be a helpful possibility for the multimodal interface designers to achieve required results.

So in conclusion, in multimodal conditions, humans do not exactly feel what the haptic sense tells them, but, rather they integrate the two modalities of hearing and touch, and what they feel will be dependent upon the level of the stimuli and the force-feedback which it creates. Our findings show that auditory information can change the percept of a tactile stimulus. In fact, a tactile illusion which is induced by sound, has been discovered, namely, when a constant haptic force-feedback stimulus is accompanied by an auditory stimulus of different sound-pressure level, the auditory stimulus modulates the tactile perception and the magnitude of strength increases with increasing loudness in spite of no change in the force-feedback.

5.3 Virtual texture perception: Roughness

5.3.1 Introduction

Texture perception is an important exploration mechanism of humans to identify objects and their properties. For example, in our daily life texture information is useful to evaluate the quality of clothes, in the field of medicine, doctors use it to investigate the abnormalities of tissue in a patient, and in the field of geology, scientists use texture to help to identify rocks and determine their history etc.

Roughness of the surfaces is the most important physical and perceptual determinant of texture perception. Therefore most studies related to human response to textures were concentrated upon the investigation of roughness perception. The physical roughness of any surface can be defined as the height of the surface along a line across the surface (Figure 5.5).

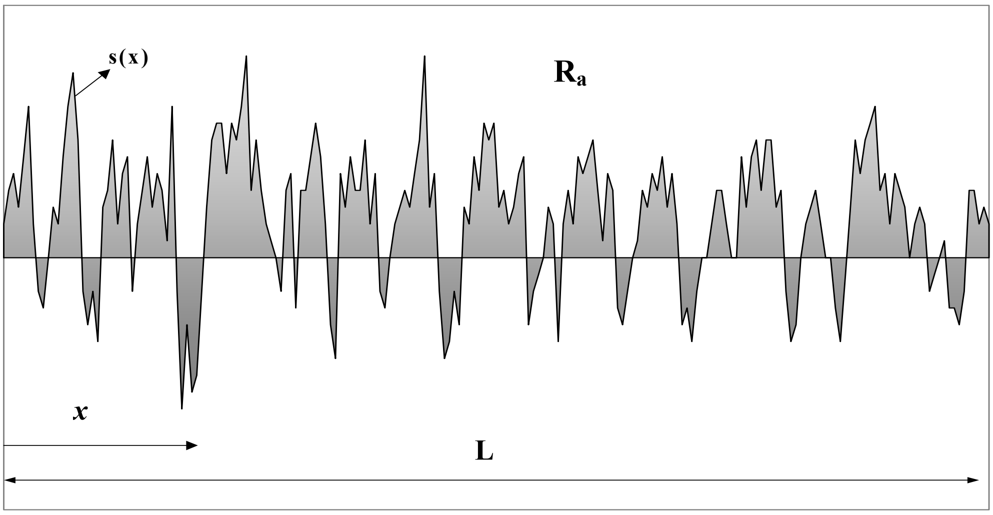


Figure 5.5: The profile of a sample surface.

The most common and general measures of roughness are the average roughness (R_a), which is the area between the texture profile and its mean line, or the integral of the absolute value of the roughness profile height over the evaluation length, and root mean square roughness (R_q), which is

$$R_a = \frac{1}{L}\int_0^L |s(x)|\,dx\,; \tag{5.2}$$

$$R_q = \sqrt{\frac{1}{L}\int_0^L s^2(x)\,dx}\,. \tag{5.3}$$

Aside from the average roughness or the root-mean-square roughness, some other measures are used to define roughness of the particular surfaces. For example the grit numbers of the sandpaper is a measure of their physical roughness. It is a reference to the number of abrasive particles per inch of sandpaper. The lower the grit-number, the rougher the sandpaper and visa versa. In the following table, the common names of the different grit-number intervals are shown (Table 5.1). This scale has only physical meaning related to the proper uses of the sandpapers.

Table 5.1: Classification of sandpaper by its grit-number.

Grit-number	Common name
40-60	Coarse
80-120	Medium
150-180	Fine
220-240	Very fine
280-400	Extra Fine
400-1000	Super Fine

People are capable of evaluating the roughness of surfaces moved across their fingertips. The investigation of the relationship between physical-roughness-descriptors and roughness perception is an interesting research topic for virtual-environment designers who want to mimic different textures in their environment. Perceiving the texture of a surface by touching it (scraping with the fingertips) is a multimodal task in which information from auditory, tactile and visual sensory channels are available. What are the relative contributions of the various systems (tactile, auditory, visual) on the multimodal percept, how does incongruent sensory information interact and how can the combination of multimodal output of information be designed better?

The aim of this section is to investigate design guidelines for multimodal (auditory-tactile) textures for virtual environments. In order to achieve this aim, experiments with unimodal and multimodal stimulus presentations were conducted and, especially, the effects of the perceptual discrepancy between the auditory and the tactile sensory modalities on the multi-sensory roughness judgment were investigated.

This section is divided into three parts. In the first part, the simulation of the realistic tactile textures in a virtual environment is discussed regarding tactile roughness perception. In the second part, the relationship between physical-roughness descriptors and the auditory attributes is established and, then, the influence of auditory attributes on the auditory-texture (roughness) perception was investigated by using synthesized touch-induced scraping sounds. In the final section, relative contributions of the virtual auditory and tactile texture information on the multimodal roughness percept are evaluated and the influence of the incongruent auditory-tactile roughness information on the multimodal percept is investigated.

5.3.2 Tactile texture perception: Roughness

Realistic texture profiles are mostly non-linear and randomly characterized, therefore, to eliminate the difficulties (analysis) and to control the conditions, in most psychophysical studies regular (i.e. linear) surfaces were used. Regarding the psychophysical studies on the roughness perception, textures can be categorized and simplified into two different stimulus-categories: raised dots, e.g. abrasive surfaces such as sandpaper etc. and grooved surfaces, e.g. grammophone plaque (Figure 5.6).

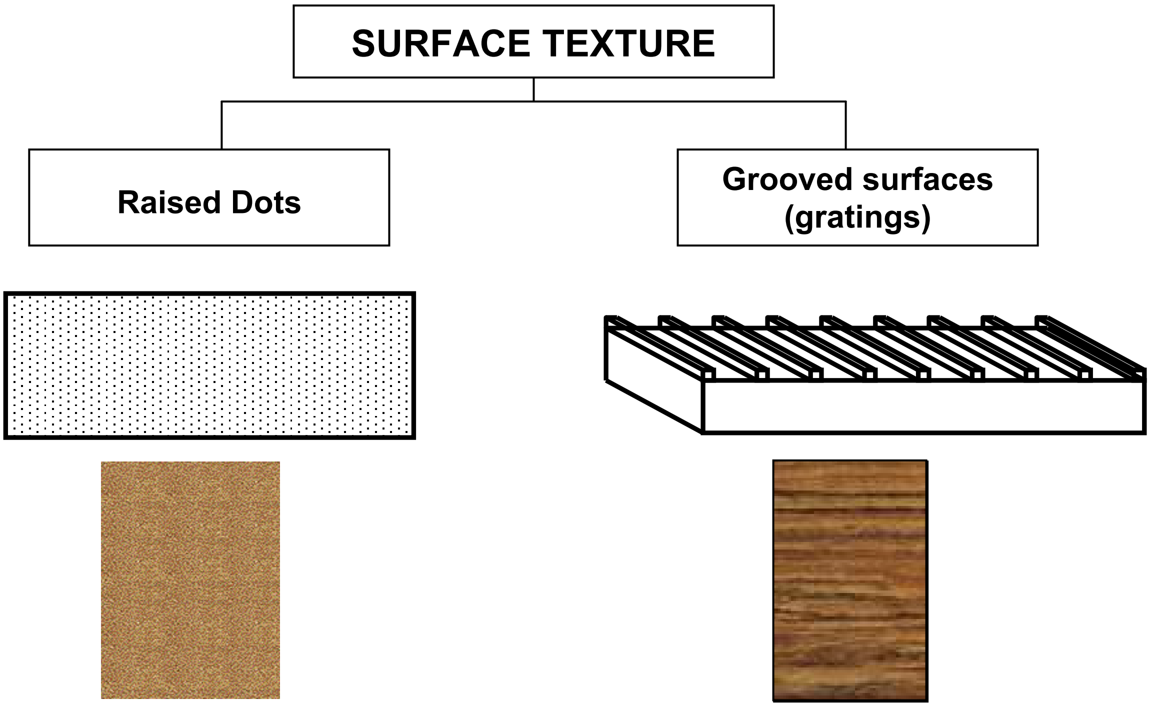


Figure 5.6: Categorization of the textures for the psychophysical studies.

In the first psychophysical study on tactile roughness perception, sandpapers in various grades were used as stimuli (Stevens and Harris, 1962). Stevens and Harris found that the perceived roughness of sandpapers increases with decreasing grit number. In their results, the magnitude estimates of the roughness of sandpaper on log-log coordinates yielded power functions with slopes of about 1.5 (Figure 5.7).

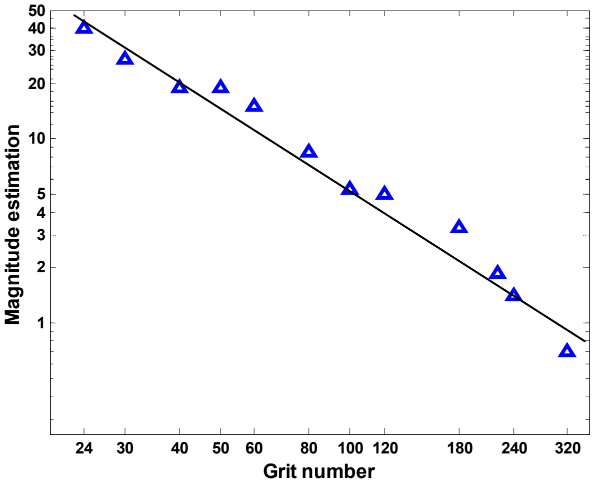


Figure 5.7: The geometric means of the estimations of roughness against grit number in log-log coordinates (adapted from Stevens and Harris, 1962).

As previously mentioned, the other type of stimulus which was used in psychophysical studies is the grooved surfaces. In Lederman and Taylor's (1972) experiment, subjects made magnitude estimates of the perceived roughness of grooved aluminum plates by actively moving three fingers across the surfaces under conditions with controlled-finger-force. Their results indicated that apparent roughness tends to increase as the grooves widen, as the finger force increases, and as the spacing between the grooves (“ridge[12]”) narrows. However, when ridge had any influence at all, it produced a considerably more modest effect than groove width. In addition, neither the groove-to-ridge ratio nor the spatial period (inverse of spatial frequency) of the gratings affected perceived roughness.

Perceived roughness of irregular virtual surfaces was measured and the physical surface parameters which influence the roughness perception were investigated by Costa and Cutkosky (2000). They selected a fractal technique to simulate the surface profiles, with

$$R_q^2(x-x_0)=C\tau \tag{5.4}$$

where C is the amplitude coefficient, and τ is the sampling resolution. They used the Fourier filtering method and $1/f^{\beta}$ power spectral-density function ($\beta = -2D+5$, D is the fractal dimension). Their results indicate that for all numbers of fractal dimension, the RMS amplitude is the overriding factor in determining surface roughness perception.

After these fundamental studies, a number of further studies were conducted to investigate the effects of applied fingertip force and scraping velocity, etc., on the roughness perception using similar stimuli. Based on their earlier results, Taylor and Lederman (1975) proposed a

---

[12] An elevated body part.

quasi-static model of perceived roughness based on a mechanical analysis of the skin deformation resulting from changes in groove width, fingertip force and ridge width. The effects of speed were not modeled, as they were negligible relative to those just listed. Their model for perceived roughness of gratings suggested that perception mapped best to the mean deviation of the skin from its initial resting position, summed over the total area of skin contact. Such variation in the magnitude of skin deformation proved to be the best candidate parameter for predicting the empirical estimates of perceived roughness. Taylor and Lederman described the representation of roughness as "intensive", since the most viable skin deformation parameter varied in magnitude along a single continuum. At the time, technological difficulties prevented them from further assessing the contribution of any spatial attributes pertaining to the skin deformation pattern.

Due to the development of haptic devices and the fact that tactile sense is included in multi-media applications, the research on roughness perception via a haptic interface was becoming important.

Caused by the difficulties of the fingertip stimulation and necessities of virtual-CAD and medical applications (telemedicine, plastic surgery), in most haptic interfaces it was tried to present tactile information to the user through a probe, e.g. phantom pen (McGee, 2002), or a joystick handle (Minsky and Lederman, 1996). Lederman, Klatzky, Hamilton, and Ramsay (1999) examined psychophysically how people perceive surface texture via a rigid probe. The results of this study indicate that, although perceiving surface roughness via a rigid probe is not quite as precise as when the bare finger is used, it can still be reasonably effective in discriminating amongst textures. Their study showed that when a rigid probe was used, a quadratic equation generally describes the psychophysical functions (perceived roughness by inter element spacing on log scales) better than a linear equation did. Previous studies had indicated that a linear equation fits the psychophysical roughness functions best when the bare finger was used (Stevens and Harris, 1962; Lederman and Taylor, 1972).

Although the main stimulus for textures is force normal to the fingertip's skin surface, Minsky and Lederman (1996) have tried to simulate haptic surface textures using only lateral forces which were applied through a force-feedback joystick (two-degree-of-freedom). Their results showed that there is a relationship between lateral force and perceived roughness, and that surface textures that can be effectively simulated using lateral forces only.

In summary, in psychophysical studies mostly three different stimulus types, namely, raised dots, grooved surfaces, and fractal irregular surfaces, were used to investigate the perception of surface roughness. Perceived roughness of sandpapers increases with decreasing grit number, and roughness of grooved surfaces increases with increasing groove width and decreasing ridge width. To present the haptic surfaces, real textures and also artificial textures (by using a rigid probe or joystick handle) were mostly used[13]. A linear equation fits the psychophysical roughness functions best when the bare finger was used, and a quadratic equation fits the psychophysical roughness functions best when a rigid probe was used.

The aims of the experiment, described in this section, are to find out whether it is possible to simulate haptic surface textures by using an electrotactile display, and using this display technique, to investigate the relationship between stimulation current, pulse frequency and roughness perception. These results are needed among others to investigate the auditory-tactile interaction issues on roughness perception by using electrotactile display technique.

5.3.2.1 Experiment 1

5.3.2.1.1 Set-up

5.3.2.1.2 Subjects

Nine subjects, five men and four women, aged between 22 and 29 years, participated in this experiment. The subjects were undergraduate students and paid on an hourly basis. All subjects were right handed, with no known heart and hand disorders and they used their right hand for the experiment.

5.3.2.1.3 Stimuli and procedure

Current magnitude (mA) and pulse frequency (Hz) of the electrotactile stimulus are the parameters which allow to represent the texture profiles for different roughnesses. To investigate the effects of the current magnitude and pulse frequency on roughness perception, the stimuli were presented with various current magnitudes and pulse frequencies. The textures had current magnitudes of 2, 4, 6, 8, and 10 mA and one out of four frequencies; 10, 30, 75, 100 Hz.

By representing the texture information, the electro-tactile feedback unit (see section 3.5.2) was used. The virtual textures were presented and roughness was estimated using an absolute estimation method (Zwislocki & Goodman, 1980). The subjects task was to report the degree of perceived roughness using numbers. For the first stimulus, they were asked to assign any positive, non-zero number (decimal, fraction or whole-number) that they think to be

---

[13] There are also some different artificial texture-presentation techniques such as the voice-coil motor etc., but there are no psychophysical roughness data available for these display techniques.

appropriate. For the next stimulus, they were required try to give an appropriate number in relation to the previous stimulus (rational). In other words if the texture feels 2 times as rough as the previous stimulus, they should assign a number which is two times the number which they had assigned to the previous stimulus, e.g. if they assigned the number 5 for the previous stimulus, they should now assign 10. The subjects were asked not to worry about being consistent.

In the training phase, which took about 15 minutes, firstly all participants were presented with different stimulus combinations from across the full stimulus range, and then they were familiarized with the magnitude-estimation procedure using six different stimulus combinations. To prevent participants devising a fixed response range, they were informed that they might experience rougher or smoother stimuli in the actual experiment than in the training (as in Lederman, Klatzky, Hamilton, and Ramsay, 1999). In the actual experiment, each stimulus was presented in random order and four times.

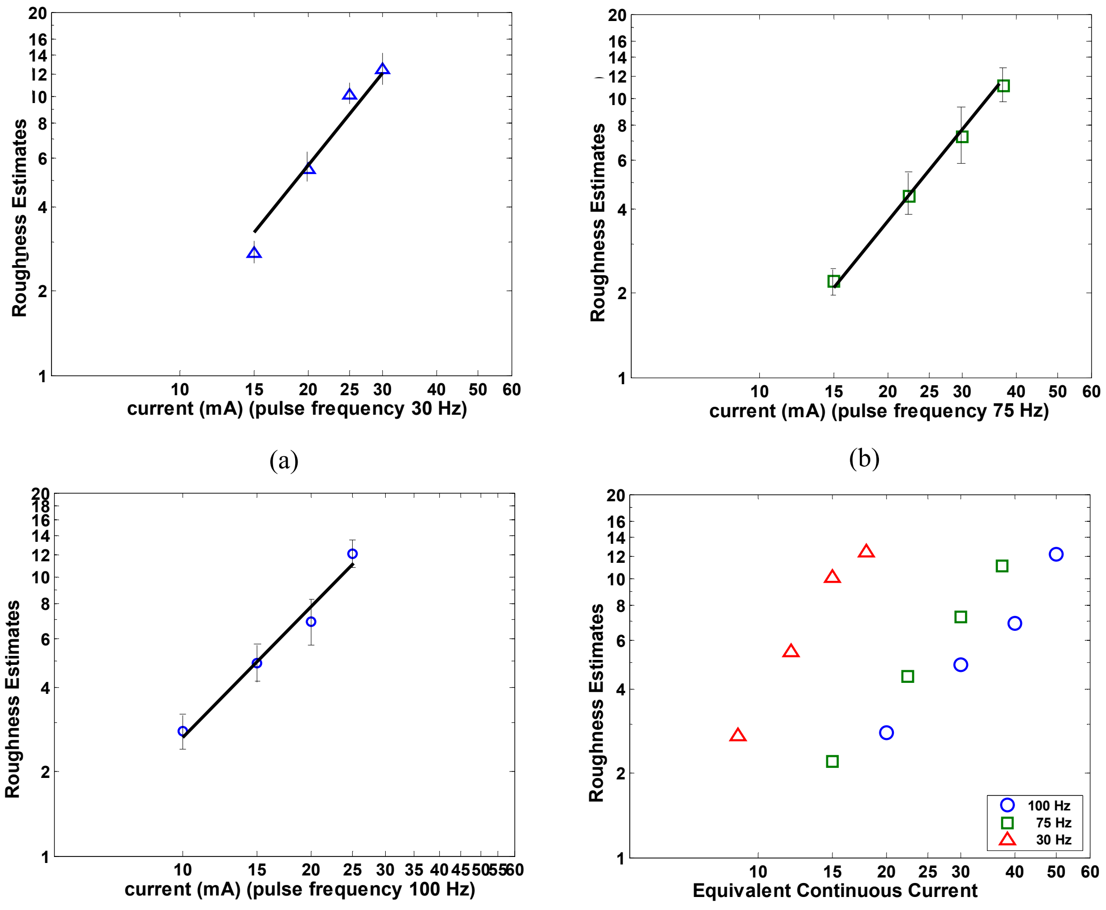
5.3.2.1.4 Results

The psychophysical roughness functions for the pulse frequencies 30 Hz, 75 Hz, and 100 Hz as a function of current magnitude are shown in Figure 5.8 a)b)c respectively. In all figures, the x-axis indicates the log current magnitude (mA) and the y-axis indicates the log roughness estimates. The data points represent log means (geometrical) and are based on 36 responses.

The method of least squares technique was used to determine the psychometric functions. The r^2 values for 30 Hz, 75 Hz, and 100 Hz conditions are 0.94, 0.97 and 0.93, respectively.

The results show that perceived roughness increases with increasing current magnitude at all frequencies. For example, for the 100 Hz pulse frequency the value of roughness estimation is 2.7 (SE 0.65) for 10 mA current, and it increases up to the value of 12.13 (SE 2.47). This increase is also valid for the 75-Hz and 30-Hz conditions.

If we compare the roughness estimates for the 25-mA-current magnitude condition in all three frequencies, it is observable that the roughness estimate of the 100-Hz condition is higher than the roughness estimate of the 30-Hz condition (it is also valid for the comparison of 75 Hz and 30 Hz conditions and 15 mA and 20 mA conditions). This observation indicates that the physical intensity plays a role on the roughness estimates. Therefore an equivalent continuous current, which is the area under the current curve and the function of the current and time (consequently frequency), was calculated (as shown in Figure 5.9) and roughness estimates are given in Figure 5.8 d as a function of the equivalent continuous current for each frequency condition.



(a) (b)

(c) (d)

Figure 5.8: Perceived roughness as a function of log current: (a) for 30-Hz pulse frequencies, (b) for 75-Hz pulse frequency, (c) for 100-Hz pulse frequency and as a function of log equivalent continuous current (d).

5.3.2.1.5 Discussion

The results show that subjects can judge the roughness using an electrotactile stimulus. An increase in the current magnitude results in an increase in perceived roughness. Same tendency is observed for an increase in the pulse frequency. Here, perceived roughness increased with increasing frequency, but the amount of the ratio or the slope of the line was not high as much as amount of the current magnitude – perceived roughness slope. One of the explanations of this finding can be found in the human response to the electrotactile stimulus, the sensitivity to the same amount of current magnitude increases with frequency similarly to the vibrotactile perception. For example, an electrotactile stimulus with the current magnitude 10 mA and 50 Hz frequency feels more intensive (stronger) than an electrotactile stimulus with the current magnitude of 10 mA and 30 Hz frequency. It can be seen from Figure 5.8d, although an increase in current or pulse frequency results in an increase of perceived roughness, the weight of the pulse frequency is not as high as that of the current and the equivalent continuous current is not explanatory to describe the relationship between perceived roughness and physical attributes of the electrotactile stimulus (current and pulse frequency).

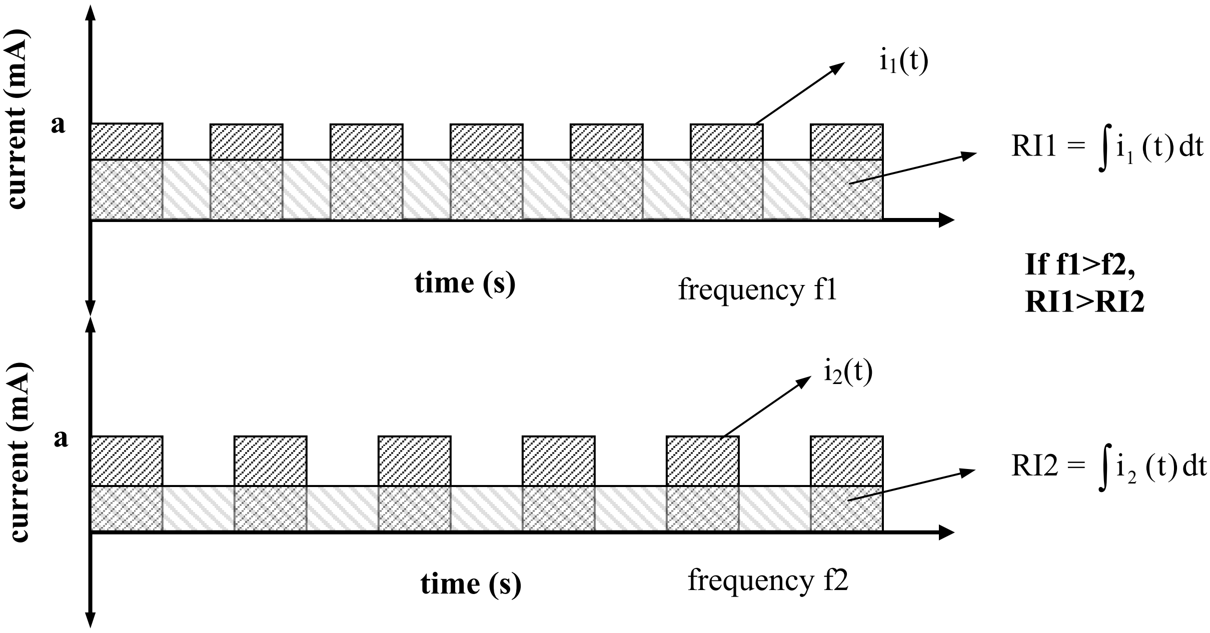


Figure 5.9: The influence of frequency on the equivalent continuous current.

A few subjects complained that they have confused their judgments because of intensity & pulse-frequency variation. They reported that if they imagined realistic surfaces and their roughness, for the roughest surfaces the stimulus was very intensive but at the same time it had low frequency. These complaints lead the author to conduct further experiments to investigate the perceived roughness of the electrotactile stimulus as compared to realistic surfaces. This investigation will help to interpret the results of the first experiment and also supply useful data for the multimodal roughness-perception experiments.

5.3.2.2 Experiment 2

Taking into account the results of the first experiment, which aimed at investigating the relationship between stimulation current, pulse frequency and roughness perception, these further experiments were conducted. As previously explained, the aim of these experiments was to investigate the perceived roughness of the electrotactile stimulus compared to realistic textures. Therefore texture profiles commonly used in psychophysical studies, i.e. sandpaper (raised dots) and grooved woods, (gratings), were selected as stimuli.

5.3.2.2.1 Set-up

The setup of this experiment was same as in the previous experiment.

5.3.2.2.2 Subjects

Ten subjects, four men and six women, aged between 22 and 27 years, participated in the experiment. The subjects were paid on an hourly basis. All subjects were right handed, and had no known heart and hand disorders. They used their right hand in the experiment.

5.3.2.2.3 Stimuli and procedure

The stimuli used in the first part of the experiment (Exp. 2a) were 8 different sandpapers with varying grit numbers: 60, 120, 150, 220, 320, 500, 800 and 1000. In the second part of the experiment (Exp. 2b), the stimuli were rectangular wood pieces, 14 x 4 x 1.5 cm, each with a set of linear grooves (0.25, 0.5, 0.75, 1.00, 1.50 mm) with constant 1.00 mm ridge width.

The method of adjustment was applied as the measurement method. The electrotactile stimulus was presented to the subject's right hand. One of the self-adhesive electrodes was attached to the subject's right index finger and the other one was attached to the upper side of the hand, between index and thumb fingers (Figure 5.10). Realistic textures could be explored by the subject moving the tip of his/her left index finger across the surfaces. Both stimuli were presented/explored simultaneously and the subjects were instructed to adjust the intensity and frequency of the electrotactile stimulus until it was perceived equally to the reference realistic texture (sandpaper or grooved wood). The subjects should first adjust the frequency and then the intensity of the stimulus.

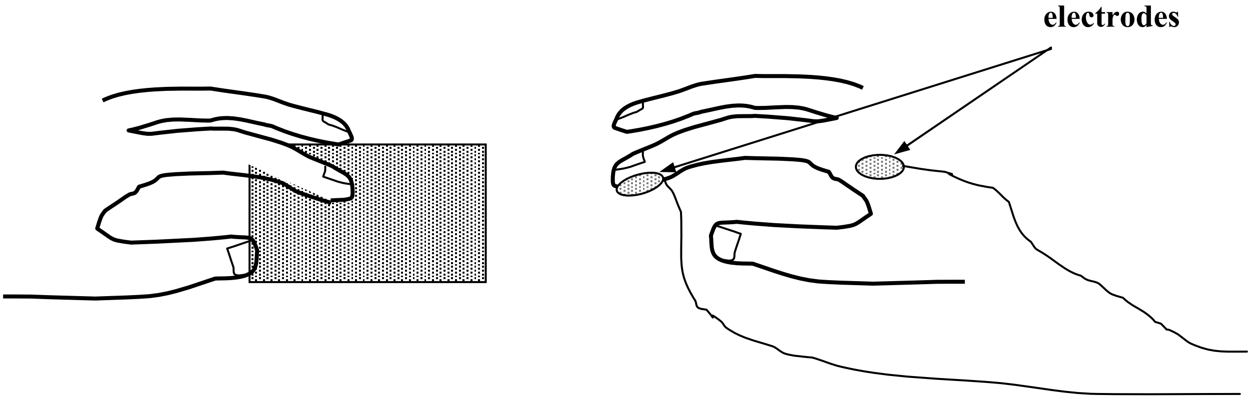


Figure 5.10: Virtual and realistic texture adjustment experimental setup.

Subjects were blindfolded (the experimenter helped the subjects to reach realistic surfaces) and wore closed damped headphones to eliminate the touch produced sounds.

In the training phase which took about 15 minutes, firstly all participants were presented with some electrotactile stimuli with different frequencies and intensities to make the subjects familiar with the electrotactile stimulus and to introduce the range of frequency and intensity variations. After the training, each realistic stimulus was presented four times and in a random order.

5.3.2.2.4 Results

The PSE values of the sandpaper stimulus with their standard errors are shown in Figure 5.11 (Exp. 2a). In this figure, the x-axis indicates the pulse frequency and the y-axis indicates the current magnitude. Each grit number was represented with a different symbol. Current magnitudes and pulse frequencies as adjusted by the subjects during the test were averaged across all subjects and trials for each grit number. The ANOVA test shows that the grit number has a significant effect on the current and pulse frequency ($p<0.0005$).

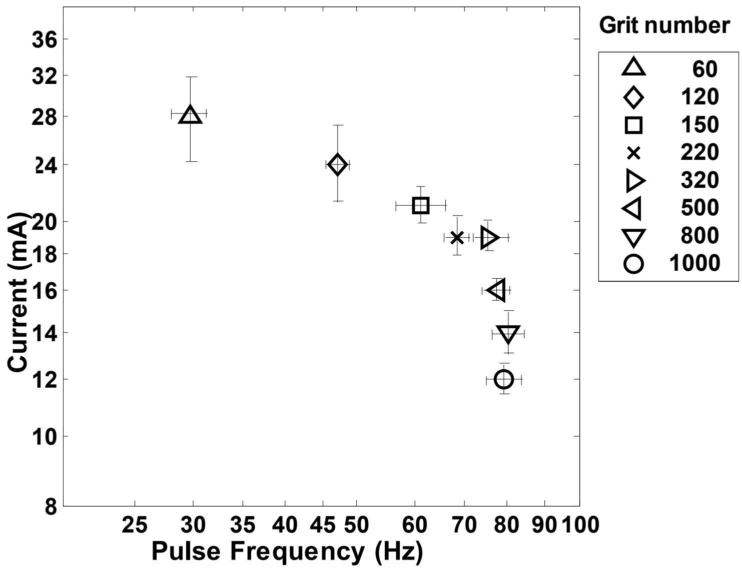


Figure 5.11: Adjusted current (mA) and pulse frequency (Hz) values for the sandpapers with varying grit numbers.

The PSE values of the grooved wood stimulus with their standard errors are shown in Figure 5.12 (Exp. 2b). Also in this figure, each groove width is represented with different symbol. Adjusted current magnitudes and pulse frequencies are averaged across all subjects and trials for each groove width. The ANOVA test indicates that the groove width has a significant effect on the current and pulse frequency ($p<0.0005$).

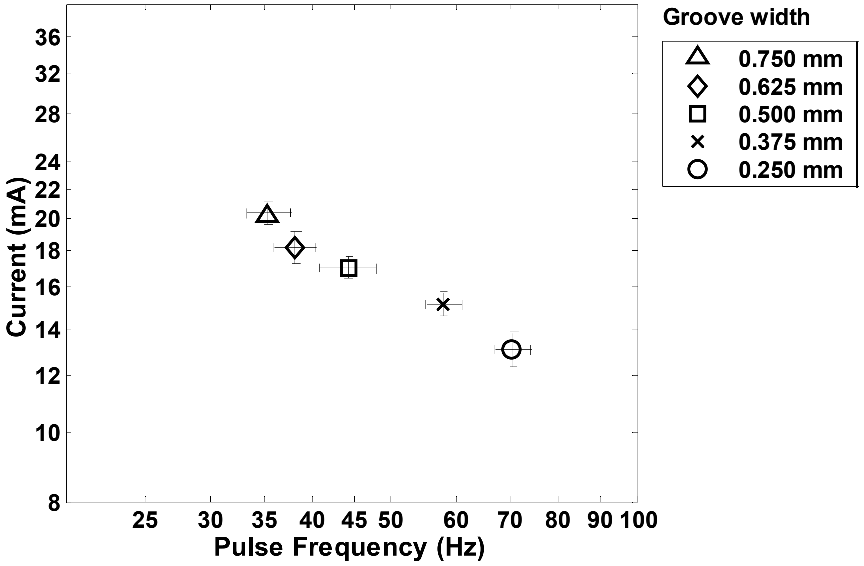


Figure 5.12: Adjusted current (mA) and pulse frequency (Hz) values for the grooved woods with varying groove widths.

5.3.2.2.5 Discussions

The results of both experiments indicate that the electrotactile stimuli can be used to simulate realistic surfaces. Physical and perceived intensity of the electrotactile stimulus are a function of the current level and the pulse frequency. The results of Experiment 2a show that an increase in grit number results in a decrease in current magnitude and in an increase in pulse frequency. For example, while for grit number 60 the current is 28 mA and the pulse frequency is 29 Hz, for grit number 800 the current is 14 mA and the pulse frequency is 81 Hz.

The results of Experiment 2b show that an increase in groove width causes an increase in current magnitude and a decrease in pulse frequency. For example, while for 0.750 mm groove width, the current is 20 mA and the pulse frequency is 35 Hz, for 0.375 mm groove width, the current is 15 mA and pulse frequency is 57 Hz.

Let us compare the results of the first and second experiments and try to provide information relevant to the design of haptic interfaces. In case subjects do not have a realistic criterion, they tend to feel rougher if current or pulse frequency of the electrotactile stimulus or both of them increase. But if they have any realistic criterion, they tend to find an electrotactile stimulus which has a high current magnitude and a low pulse frequency more suitable for rough realistic surfaces (such as grit number 60 sandpaper or grooved wood with the groove width 0.75 mm).

Following Experiment 2b, the subjects were presented a wood plate without any groove (very smooth) and asked if they could find a suitable electrotactile stimulus for that. The results of this small survey showed that subjects tended to find just perceptible current magnitudes suitable for very smooth surfaces and they did not show a preference for a certain frequency.

5.3.3 Auditory texture perception: Roughness

In our daily life, sound is often the result of human-object interaction (touch, striking, scraping, etc.) and mostly informative. It conveys to the listener required information about physical attributes of the interaction, spatial properties of the sound event, e.g. location, geometry and sound generation event.

Katz has shown that people are highly skilled in using touch-produced sounds to identify material of the various objects (1925). By touching an object, the tactile sense informs us about the roughness of the object which raises the question whether auditory information is used as well. The first and only study on auditory texture perception was conducted by Lederman (1979). She gives a good example depicting the importance of the sound to convey information about the roughness if tactile information is not available:

"A recent television advertisement begins with a close-up of a young man's face. Half is shaved with one brand of shaving cream, the other half with a competing brand. To show how much closer the shave is with the cream being advertised, a credit card is drawn along the skin of first one and then the other side of the face. The sounds produced make it quite clear which side is smoother."

Lederman's results somehow confirm this example, the subjects can judge the roughness of plates of varying groove and ridge width by using only the sounds produced by touching the surfaces. The auditory judgments were similar to, but not as discriminating, as those made by touch alone or by touch plus audition. Although subjects are able to evaluate roughness by sounds alone, when these are presented in the company of the tactile cues which produce them, the auditory cues seem to be ignored.

Lederman suggested that the loudness and modulation frequency of the touch-induced scraping sounds can be used by subjects in their judgment and that they can be important psychoacoustical determinants of the texture perception. If we look into the physical generation mechanism of the touch-produced sounds, groove width and, consequently, groove frequency or grit number have an influence on the modulation frequency of the sound, and the friction force between finger and texture has an influence on the acoustic level. If we remember that the groove width and the friction force are two determinants of the tactile roughness perception, Lederman's approach is in harmony with the physical background of the event. Taking Lederman's approach into account, psychoacoustical experiments were conducted to investigate the relationship between acoustic properties (level and modulation frequency) and texture perception in this section. In order to control the psychoacoustical parameters, synthesized scraping sounds were used in the experiments.

Before explaining the conducted experiments, it can be useful to clarify that the sound which is generated by touching any grooved surfaces or raised dot like surfaces, is amplitude-modulated broad band noise (modulated by grooves or dots) and modulation frequency (f_{mod}) and depth may give some information related to the surface (Figure 5.13). But auditory roughness sensation and roughness of the surfaces are two different attributes. Detailed information about the auditory roughness sensation can be found in Section 2.2.2.

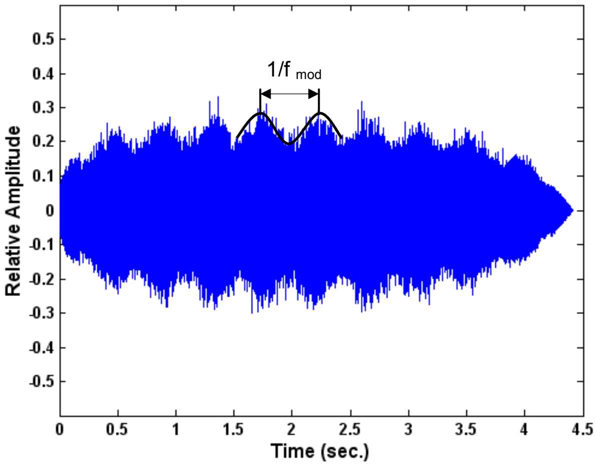


Figure 5.13: A time series of a scraping sound example.

5.3.3.1 Experiment 1

5.3.3.1.1 Set-up

The auditory stimulus was presented from a PC. It was amplified and delivered diotically through Sennheiser HDA 200 closed-face dynamic headphones. The experiments were conducted in a sound-attenuated room.

5.3.3.1.2 Subjects

Ten subjects, four men and six women, aged between 22 and 27 years, who had already participated in Experiment in 2a and 2b took part in the experiment. The subjects were undergraduate students and were paid on a hourly basis. All subjects had self-reported normal hearing.

5.3.3.1.3 Stimuli and procedure

The stimuli were sounds such as those generated by touching rectangular wood pieces, 14 x 4 x 1.5 cm, each with a set of linear grooves (0.25, 0.5, 0.75, 1.00, 1.50 mm) and constant 1.00 mm ridge width. In order to control modulation frequency and loudness of the scraping sounds, they were synthesized in a computer environment (see 3.5.3). In the first part of the experiment to investigate the influence of the modulation frequency (which is related to groove numbers) on the roughness perception, the loudness of the stimulus was equalized to the level of the 1.50-mm groove width. In the second part of the experiment, the same stimuli without any loudness equalization were used.

The sounds were presented and the roughness of the surfaces which result these sounds was estimated using an absolute estimation method. The subjects task was to tell how rough they feel by assigning numbers to them. For the first stimulus, they were asked to assign any positive, non-zero number (decimal, fraction or whole-number) they thought was appropriate. For the next stimulus, they were instructed to give an appropriate number regarding the previous stimulus (rational) and not to worry about being consistent.

In the training phase which took about 5 minutes, subjects could touch real grooved woods with varying groove widths, they were able to hear touching sounds and also feel with their fingertip the tactile information (roughness). Then different sounds from across the full stimulus range were presented through headphones to all participants. To prevent participants devising a fixed response range, they were informed that they might experience rougher or smoother stimuli in the actual experiment than in the training. In the experiment, each stimulus was presented in random order and ten times.

5.3.3.1.4 Results and discussion

The perceived roughness by listening touch-induced sounds for two conditions is shown in Figure 5.14 as a function of groove width. The data points represent magnitude estimates and are based on 100 responses. As usual, the method of least squares technique was used to determine the theoretical psychometric function. The r^2 value for the loudness equalized condition was 0.73 and the r^2 value for the condition without loudness equalization was 0.86.

In both conditions, perceived roughness increases with an increase in groove width. The results of the first condition show that the subjects can judge the roughness of grooved surfaces using only modulation-frequency information. For example, while the roughness estimate was 15.7 for 0.5 mm groove width, for 1.5 mm groove width it increased to values up to 26.6. Only for the very narrow groove width conditions like 0.25 mm and 0.50 mm, the difference between roughness estimates is not as high as in the larger groove width conditions. This observation is in line with the results of Lederman (1979). A reason for the low increase may be that it was not very easy for subjects to feel frequency differences for such high modulation frequencies (94 Hz and 112 Hz).

The results of the second condition show that besides the modulation-frequency information, loudness supports the subjects' roughness estimation and makes it easier to feel differences in the groove width. Results confirm the hypothesis that subjects can judge the roughness of surfaces using only the sounds produced by touch action.

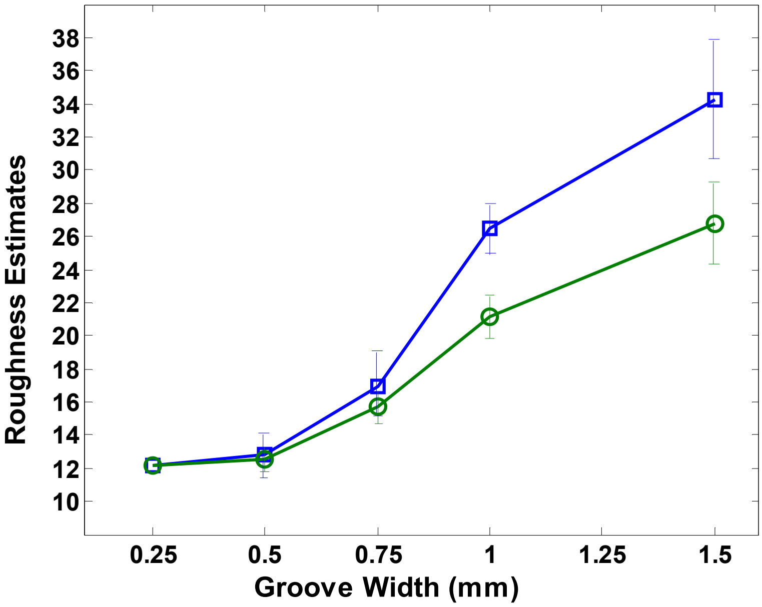


Figure 5.14: Perceived roughness as a function of groove width for two different conditions (o: with loudness equalization, □: without loudness equalization). The data are averaged across subjects.

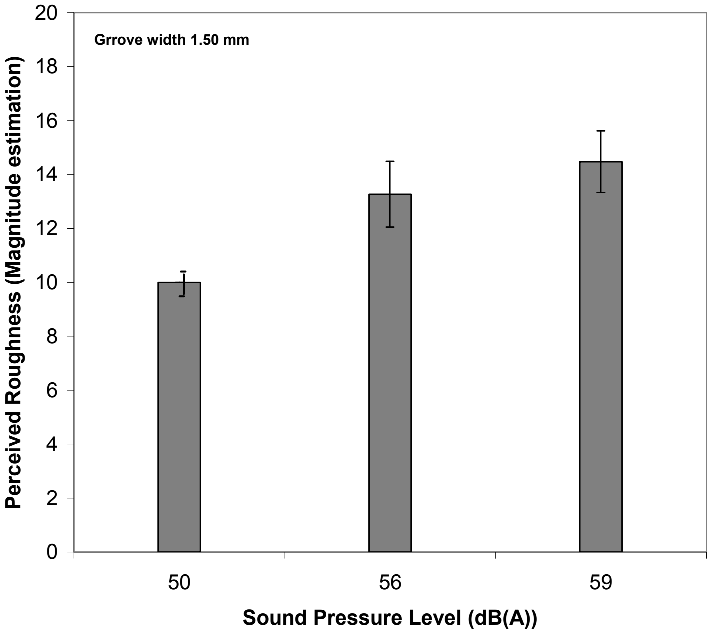
5.3.3.2 Experiment 2

The results of the first experiment indicate that besides the modulation-frequency information, loudness is an important cue for the subjects to judge the roughness of the grooved surfaces. The aim of Experiment 2 is to investigate how different loudness conditions affect the subjects' roughness perception.

Subjects, set-up and procedure were the same as in the first experiment. Stimuli were also identical but in two stimulus conditions (0.5 and 1.75 mm groove width), the sound pressure level was increased by 6 and 9 dB.

5.3.3.2.1 Results

The roughness estimates (and SE's) for the 1.50 mm groove width as a function of sound pressure level is shown in Figure 5.15. The data consist of geometric means of 100 magnitude estimates. To eliminate the influence of the choosen numerical scale (which could be freely selected by the subjects), the resulting mean magnitude estimates (each participant's ten magnitude estimates) were subsequently normalized by dividing each score by the individual participant's mean, then multiplying it by ten.



.

Figure 5.15: Perceived roughness as a function of sound pressure level for the scraping sound of the wood piece with 1.50 mm groove width.

The roughness estimates (and SE's) for the 0.50 mm groove width as a function of sound pressure level is shown in Figure 5.16. Similar statistical processing was done for the magnitude estimates as for the 1.50 mm groove width.

5.3.3.2.2 Discussion

Increasing loudness results in an increase on the perceived roughness. This observation is true for two different groove widths (0.50 mm and 1.50 mm). Especially the increase in the roughness estimates for the 0.50 mm groove width was higher than the increase for the 1.50 mm groove width condition. The results show that subjects use loudness increase as a cue for the roughness increase of surfaces like they do with modulation frequency decrease.

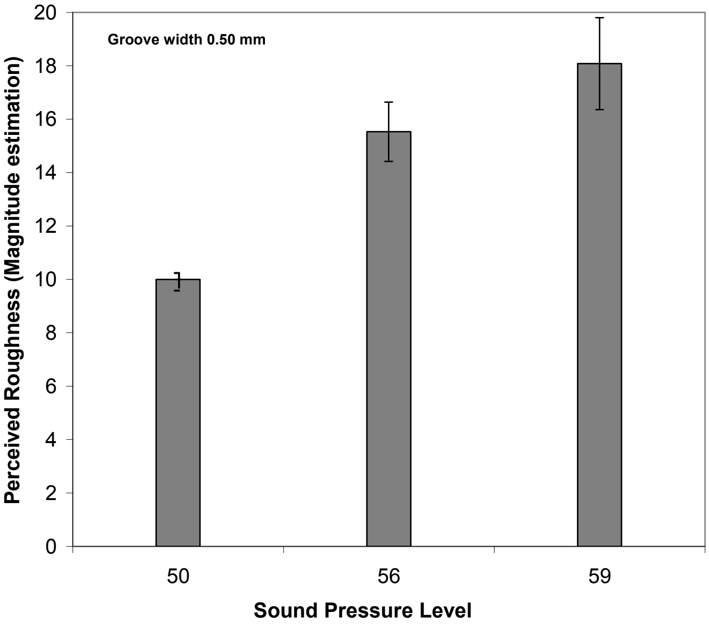


Figure 5.16: Perceived roughness as a function of sound pressure level for the scraping sound of the wood piece with 0.50 mm groove width.

5.3.4 Multimodal texture perception: Roughness

Scraping a surface with the finger tip is a multimodal event. We obtain information about the texture, i.e. roughness of the surface, at least through three different sensory channels, i.e. tactile, visual and auditory. The main questions for the multimodal research are: How do these modalities interact and what are the effects of the perceptual discrepancy between the modalities on the multi-sensory roughness judgment.

In a study by Lederman (1979), congruent auditory and tactile texture information was presented to the subjects and they were asked for the roughness of the surfaces. She found that, if tactile and auditory sources of information are available, subjects tend to use tactile cues to judge surface roughness. This result indicates that tactile texture cues completely dominate the auditory cues in determining texture perception. Her explanation for this result is that in daily life sound cues, which are generated by touching the texture of a surface, are masked by background noises due to their low level. Therefore, our attention is directed to the tactile modality. This argument was somehow confirmed in another study by Lederman et al. (2002). They experimentally assessed the relative contributions of tactile and auditory information to bimodal judgments of surface roughness using a rigid probe. The sounds generated due to contact between a rigid probe and a rigid surface are louder than those generated by bare finger. Their results indicated that when a subject explores the surfaces by a rigid probe, she/he uses both tactile and auditory information to make their estimates. (the tactile weight is 62 % and the auditory weight 38 %).

Under certain conditions, auditory cues which are generated by bare finger can also influence tactile roughness judgments (Jousmäki and Hari, 1998). Jousmäki and Hari have named this effect as "parchment-skin illusion". In their experiment, subjects have rubbed their hands together and listened simultaneously to modified sounds as generated by rubbing. After the stimuli presentation, they were asked to rate roughness and moistness of the palmar skin of their hands. The results showed that when overall sound pressure level increased (20 dB or 40 dB), or when the frequency components within the frequency range of 2-20 kHz were amplified, subjects have felt smoother and dryer. If the sound pressure level decreased, they felt rougher and moister. Later on Guest et al. (2002) has demonstrated that the same effect is valid for the sandpaper stimulus also.

The studies of Jousmäki and Haki (1998) and Guest et al. (2002) indicate that under certain conditions, auditory and tactile information can interact by determining the roughness of the textures. Increasing loudness can result in a decrease in perceived roughness. From the view of a virtual environment designer, the following question arises: If loudness can play such a role on the multimodal texture perception, what can be the influence of auditory modulation frequency (fundamental frequency related groove or grit number) on the multimodal texture perception. To investigate the interaction of incongruent auditory and tactile stimulus presentation, multimodal roughness experiments were conducted.

Experiments

The aim of the first experiment was to investigate the relative contributions of the auditory and tactile information on the bimodal judgments if sound pressure levels are 10 dB higher than physically accurate. The second experiment was conducted to investigate the influence of the modulation frequency information on the tactile roughness judgments.

5.3.4.1 Experiment 1

5.3.4.1.1 Set-up

The tactile stimuli were presented through electrotactile electrodes. The auditory stimuli were presented from a PC. They were amplified and delivered diotically through Sennheiser HDA 200 closed-face dynamic headphones. The experiments were conducted in a sound-attenuated room.

5.3.4.1.2 Subjects

Ten subjects, four men and six women, aged between 22 and 27 years, who had already participated in auditory- and tactile-roughness experiments took part in the experiment. The subjects were undergraduate students and paid on an hourly basis. All subjects had self-reported normal hearing and were right handed, with no known hand disorders. They used their right hand for the experiment.

5.3.4.1.3 Stimuli and procedure

The stimuli were rectangular wood pieces, 14 x 4 x 1.5 cm, each with a set of linear grooves (0.25, 0.5, 0.75, 1.00, 1.50 mm) and constant 1.00 mm ridge width. Taking into account the statement of Lederman (1979) which is "*In daily life sound cues, which are generated by touching the texture of a surface, are masked by background noises due to their low level. Therefore tactile texture cues completely dominate the auditory cues in determining texture perception*", the sound pressure levels of the auditory stimuli were amplified 10 dB as compared to the physically accurate value. The procedure was same as in the auditory-roughness-perception experiment.

5.3.4.1.4 Result

Roughness judgments for the conditions: auditory only, tactile only and auditory and tactile together are shown in Figure 5.17 as a function of the log groove width.

The data points represent log magnitude estimates and are based on 100 responses. To eliminate the influence of the choosen numerical scale (which could be freely selected by the subjects), the resulting mean magnitude estimates (each computed from participant's ten magnitude estimates) were subsequently normalized by dividing each score by the individual participant mean, then multiplying it by ten. Responses are normalized to the value 10 for 0.25 mm groove width.

Dependent t-tests of the means show that all three conditions differed significantly (auditory only – tactile only: $\underline{t}(9) = -5.64$, $p<0.05$; tactile only – audiotactile: $\underline{t}(9) = 6.74$, $p<0.05$; auditory only – audiotactile: $\underline{t}(9) = -5.85$, $p<0.05$). The percent weightings were calculated by using overall means of three conditions and equation 4.2.1 (as suggested by Lederman et al. 2002),. The relative weighting of the tactile only condition is approximately 60% and the relative weighting of auditory only condition is approximately 40%.

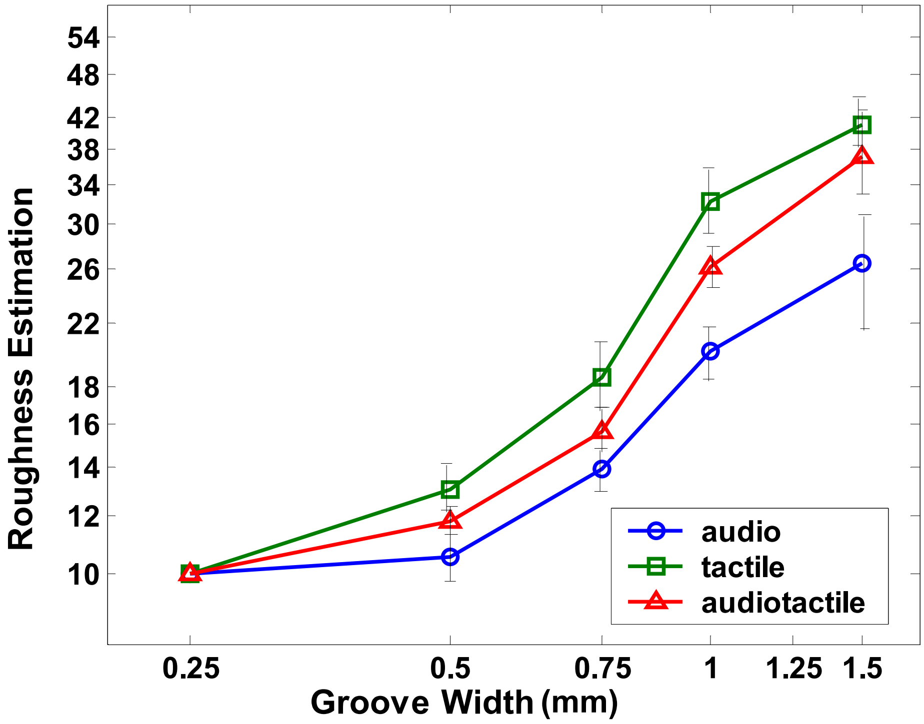


Figure 5.17: Perceived roughness as a function of groove width and sensory mode of the judgment (audio, tactile, audiotactile). The data are averaged across subjects.

5.3.4.1.5 Discussion

In all three conditions, subjects could judge the roughness of wooden plates for varying groove width. Perceived roughness increases with increasing groove width. In all three conditions the roughness estimates differed from each other.

The slope of the auditory only condition shows slower acceleration as seen in the other conditions. This result is in line with the results of Lederman (1979). In the tactile only condition, the roughness estimates are higher than in the auditory-only and auditory-and-tactile-together conditions.

The curve of the auditory-tactile roughness judgments and the results of the relative weightings show that the subjects take into account both tactile and auditory information. These result do not agree with the results of the Lederman (1979), who found that touch based auditory cues do not play any role on the bimodal judgments. Recall that she has argued that low-level sound cues are frequently masked by the general background noise in many everyday situations. One of the reasons for the difference between the results of the present study and the results of Lederman (1979) could be that in the present study all sound-pressure-levels are 10 dB above the physically accurate value and this amplification results an increase of the contribution of the auditory information on the bimodal judgments. The results of Lederman et al. (2002) on the assessment of bimodal roughness judgments using a rigid probe confirm this arguent. With rigid contact between surface and end effector, the amplitude of the accompanying sounds is usually considerably greater and their results show that the subjects used not only tactile information, but also auditory information on the bimodal judgments. These results also confirm the results of the Jousmäki and Hari (1998), and Guest et al. (2002) that under certain conditions auditory cues which are generated by the bare finger can also influence tactile roughness judgments.

5.3.4.2 Experiment 2

The aim of this experiment was to investigate the role of the modulation-frequency information on the tactile-roughness judgments. Subjects, set-up and procedure were the same as in the first experiment. In this experiment, some congruent (modulation frequency and tactile frequency) and incongruent stimuli (Table 5.2) pairs were presented.

Table 5.2: Stimuli list of the experiment 2

Incongruent Stimuli Number	Auditory Stimulus	Tactile Stimulus
1	0.25 mm groove width	0.5 mm groove width
2	0.5 mm groove width	0.5 mm groove width
3	1 mm groove width	0.5 mm groove width
4	0.25 mm groove width	0.75 mm groove width
5	0.75 mm groove width	0.75 mm groove width
6	1.5 mm groove width	0.75 mm groove width

5.3.4.2.1 Procedure

Similarly to other experiments, the absolute-magnitude-estimation method was used in this experiment. The subject's task was to report how rough they felt by assigning numbers regarding the roughness of the tactile stimulus. They were specifically instructed to ignore the touch sounds they heard, and to base their judgments only on tactile information.

5.3.4.2.2 Results

The roughness estimates (and SE's) for the stimulus numbers 1, 2, and 3, as a function of the auditory modulation frequency (groove width) are shown in Figure 5.18. Figure 5.19 shows the data for the stimulus numbers 4, 5, and 6, as a function of auditory modulation frequency. The data points represent geometrical means of the 100 responses.

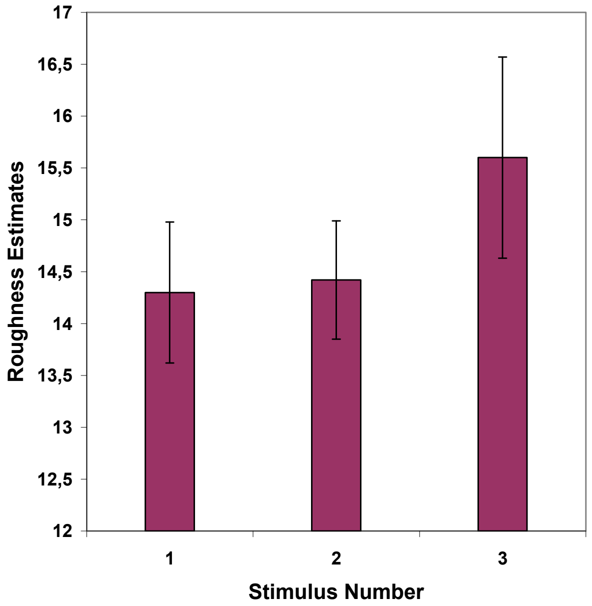


Figure 5.18: Perceived roughness of the stimuli 1, 2, 3 (see Table 5.2). The data are averaged across subjects.

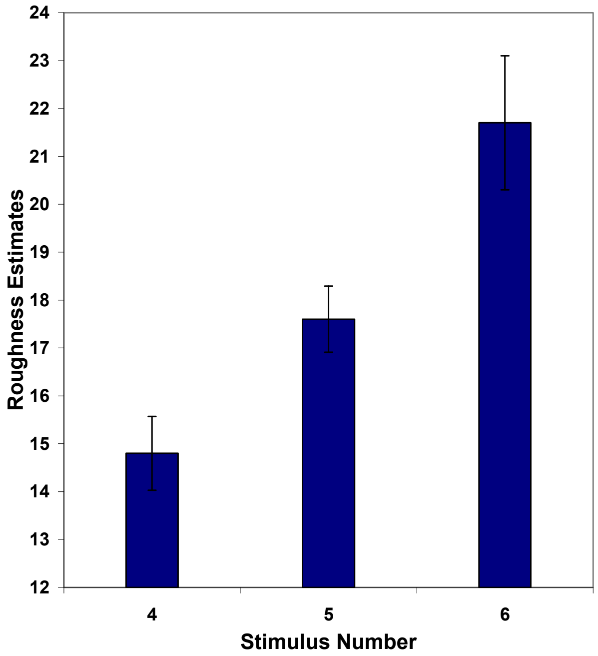


Figure 5.19: Perceived roughness of the stimuli 4, 5, 6 (see Table 5.2). The data are averaged across subjects.

5.3.4.2.3 Discussion

The results show that in incongruent stimuli presentations, the auditory modulation frequency can alter the tactile information. Decreasing modulation frequency results in an increase in perceived tactile roughness, even though tactile information is smoother than the auditory information. This effect is also observable for the increasing-modulation-frequency condition. Here the increasing modulation frequency results in a decrease of the roughness estimate, even though tactile channel indicates that it is rougher. The effect can be seen very clearly for the stimulus numbers 4, 5 and 6, but less clear for the stimulus numbers 1, 2, and 3. One reason may be that the small difference between 0.5 mm and 0.25 mm groove width is not enough to alter the tactile information.

5.3.5 General conclusion

Texture perception of a surface by touch is a complex, multimodal process, and it is difficult to simulate realistic multimodal textures using virtual-reality displays. Generation of multimodal textures for virtual reality applications also requires knowledge about the nature of the realistic and virtual textures. In this chapter, the design of multimodal textures for virtual-reality applications was discussed by investigating unimodal and multimodal (auditory-tactile) roughness-perception issues.

Tactile-roughness-perception experiments show that it is possible to simulate realistic textures using the electrotactile-stimulation technique in virtual environments. Intensity (current) and pulse frequency of the electrotactile stimulus are two parameters which enable people to feel the differences between textures regarding their roughness. Increasing intensity and increasing pulse frequency result in an increase of the perceived roughness. But the simulation of very rough sandpaper or grooved materials requires an electrotactile stimulus which has high intensity and low frequency. Simulation of very smooth surfaces is possible with just-perceptible current magnitudes. Frequency variability is not a criterion.

Auditory roughness perception experiments show that the subjects can judge the roughness of wooden plates with varying groove width by utilizing only the sounds produced by touching the surfaces. The amplitude-modulation frequency which is related to groove number (or interelement spacing) is the most important auditory attribute which enables people to feel the differences between textures regarding their roughness. Increasing groove width causes a decrease in the modulation frequency, therefore the perceived roughness increases with decreasing modulation frequency as expected. Besides modulation frequency, loudness is a useful and also important cue which informs us about the roughness of the textures. Increasing loudness results in an increase in the perceived roughness. If the modulation frequency decreases and simultaneously loudness increases, the increase in the perceived roughness is greater than in the condition where only modulation frequency decreases. This finding shows that the influences of both auditory cues on the roughness perception are superimposed.

The results of the multimodal roughness judgment experiment show that subjects take both tactile information and auditory information into account in their bimodal judgments. The auditory modality (40 % weighting in the bimodal judgments) is nearly as informative regarding the roughness of the textures as the tactile modality is (60 % weighting in the bimodal judgments). Auditory attributes, e.g. modulation frequency and loudness, can alter significantly the tactile-roughness-perception in bimodal stimulus presentation conditions, if they are incongruent with the tactile information. Decreasing modulation frequency results in an increase in the perceived tactile roughness as in some conditions (small increment, such as 6 or 9 dB) increasing loudness results an increase in the perceived tactile roughness, while, on the contrary, in other conditions (great increment, such as 20 or 40 dB) increasing loudness results in a decrease in the perceived tactile roughness. Interaction on the tactile roughness perception related to auditory and tactile attributes is complex, and the effects depend on the conditions. Therefore, designers should be aware of this complexity if they want to use benefits of the bimodal stimuli presentation. The results of the study indicate that the auditory information can be useful in overcoming the limitations of the electrotactile texture presentation and provide further realism, keeping in mind that high current magnitudes are very uncomfortable for the subjects.

When the results of the multimodal roughness judgment experiments are interpreted from the intersensory-organization perspective, the *modality superiority* hypothesis, as is suggested by Welch et al. (1979) and Lederman, and Abbot (1981) is in line with the current findings. The measures of the modality superiority hypothesis are accuracy, sensitivity, discrimination, precision, and other aspects of performance. The results of the current study show that if the sound pressure level of the scraping sounds is 10 dB above the physically accurate value, auditory and touch performances related to roughness of the textures are very similar. Auditory judgments are nearly as precise and discriminative as the touch judgments, and information can be obtained easily and quickly by both modalities. Therefore both information were used by the subjects and somehow they superimpose both information in their bimodal judgments. This argumentation can be confirmed by another intersensory-organization hypothesis, namely, *ecological validity*, which is suggested by Lederman (1979). Lederman argued that one reason that tactile sense may bias auditory sense in a texture-related task, (in situations where audition is less discriminating than touch) is that tactual cues to texture are more ecologically valid than auditory cues. In the current experiment the auditory sense was nearly as influencial as touch when judging the roughness of surfaces, therefore the task (roughness judgment) may be considered ecologically valid for both modalities.

Chapter 6

Audiotactile Interactions in Product Quality Perception and Evaluation

6.1 Introduction

In our daily life we permanently obtain information about products through all our senses during product use. Nonetheless, product design and product-quality evaluation often lack an integrated concept of multimodal perception related to product use. For good reasons, most effort has so far been put to understanding the processes of unimodal perception alone, regarding the questions of when, what, and why we perceive and how we judge unimodal perceptual events. However, both product design and quality evaluation of products would perform even better if more intense efforts were made to approach multi-modal aspects of product perception systematically. This is a difficult task. Anyhow, in a first step to approximate this goal isolated aspects of product quality can be put in relative terms to each other. In this study it is auditory and tactile perception of industrial products in relation to product use.

The aim of the present chapter is to investigate the combined influence of auditory and tactile information (i.e. vibration) on product quality. First of all, the concept of ‘product quality’ as a design issue and a value in use is introduced, then product quality is reflected on, the development of product-sound-quality research and its main issues are briefly summarized, entailing limitations to product-sound design are sketched, an outline of the present state of tactile perception is given, and finally, a new definition of product-vibration quality is introduced. The combined influence of auditory and tactile information on product quality is discussed on the basis of the results of the conducted experiments.

6.2 Product quality

Every now and then product designers face the fact that users behave differently than expected. Although product specification has been built on market research (i.e., on the situation and the extent of the demand) as well as product analysis, they fail to gauge the real demand for a product in the end. User expectations are not meet. One reason obviously is that market success is often built on tenets central to product functionalism and safety of products and processes. An industrial product has to be appropriate for use (function): performance, accessibility, practicability, reliability, creditability, safety are the main quality characteristics. Without doubt, functionalism and safety are central design issues of industrial products. They can analytically be expressed as mathematical formulas, the modes of description heavily relying on standardized formalisations. There are often catalogues of standard requirements specific to different types of products which serve as the expected norm. The task of the designer is to create an object which fulfils all regulations with regard to functional requirements, capabilities and limitations in such a way that the product becomes an object of desire. The product becomes an object of desire when it is sensually appealing, when its form invites the customer to use it. In other words: the design goal is to create an object which – when it becomes an object of perception – stimulates and satisfies customer interests.

Customer interests are satisfied when the designed product leads to a harmonic perceptual entity. This is the case when the form of the product perceived, and the function it conveys are coherent. Apart from the form-function relation, coherence is also related to the different information channels, i.e. sensory inputs.

We perceive our world in a multimodal way, and what seems to be important is that – when using a product – there are no contradicting information given by the different senses, and that conventions are not violated. Our expectations are based on experience, and it is experience which frames conventions. We judge a product as being of high quality when all our expectations are met or even exceeded, when we perceive it as a closed harmonic entity. Quality features of industrial products are, of course, functionalism, safety, usefulness, but it is also aesthetic aspects, emotional reactions. They all and others have to be designed in such a way that they unambiguously direct at the product as such, which means that all perceived quality features have to comply with the general product idea.

This discussion introduces the basic perspective taken in this study: Whilst driving at physical perfectionism, industrial design has for a long time been concentrating on constructing products according to, e.g., functional aspects and safety. This approach viewed design mainly from a technological perspective. In contrast, here the primary weight is put on the perceptions of the product users. If it is the users who decide whether perceived features carry that sort of information they expect, and whether the extracted information and the associated meaning are supportive to the general product idea or not, it is beneficial to analyse users with regard to their perceiving, interpreting, and judging. Users perceive products by all their senses, i.e. in a multimodal way. In that sense, multimodally perceived aspects of products are carriers of information, it is them which hint at functional aspects, at safety, practicability, performance etc. In other words: We do not perceive functionalism as such, but we perceive features of the product which we interpret as pointers to functionalism.

Consequently, in the specification phase of an industrial product the first question to answer is: What are the most important design issues, which kind of information should the product convey? Based on the answers given, the means to achieve these design goals have to be decided on, posing the following questions: How do these individual types of information like safety sound, how do they look like, taste, smell or feel? Are there dominant senses responsive to specific types of product information? Are there generic carriers of these types of product information? What are their physical counterparts? Can they be determined independently from each other or is there a mutual interaction between them? If so, is there a dominance of one over the other or are they equally important?

This is wide field of research. In order to approach it in a systematic way, this article concentrates on audio-tactile interactions in product quality perception. The questions in focus are: How do we perceive product quality globally? How do we judge auditory, how tactile product quality each? How do we perceive interactive audio-tactile product quality? Which consequences can be drawn for product design and quality evaluation?

6.3 Product-sound quality

For a long period, acoustic engineers made an effort to reduce the acoustic energy of the products. In the middle of the 80's, it was stated that the A-weighted sound pressure level was not sufficient to describe the character of the product sound (Blauert, 1986).

The acoustic emission from a product leads to a perception of an auditory event upon which the hearer makes a quality judgement (Blauert and Jekosch, 1996). Sound as a product attribute simultaneously plays a functional and an aesthetic role. The product sound contributes significantly to the character of the product. It can be regarded as an acoustic fingerprint (Bednarzyk, 1997). Since people become increasingly sensitive with regard to the auditory perception of product sounds, product designers have to include the quality element "product sound" as an essential part of their design considerations.

Blauert and Jekosch's (1996) product-sound-quality definition can be generalized for all product features including *vibration*. They defined product sound quality as "*a descriptor of the adequacy of the sound attached to a product. It results from judgements being performed with reference to the set of those desired features of the product which are apparent to the users in their actual cognitive and emotional situation*". In the context of this definition, the designers should know the desired features of the product (user expectations), and try to meet or exceed user expectations with their design.

The development of the *"product vibration quality"* approach follows the sound-quality approach step by step. Until the last decade, in most cases the task of the product vibration designer was to reduce the vibration energy emitted by the product. But vibration can also play a functional and aesthetic role as a product feature. In fact, every vibration can be wanted or unwanted regarding to different expectations. The importance of expectations of the user for the evaluation of vibration was also confirmed by ISO 2631 as a statement. In the following section a review of the product-vibration evaluation from the historical perspective will be given and the main issues of the product-vibration quality will be introduced, based on the knowledge of the product sound quality.

6.4 Product vibration quality

People are exposed to many forms of vibration from different products by using them, e.g. vibrations from vibrating tools (drill, electric-razor, hand mixer, vacuum cleaner, etc.), vibrations from vehicles, or vibrations from musical instruments (guitar, drum, etc.). Vibration of a product is a product parameter such as sound, visual image (aesthetic, colour, form), weight, price, etc. Due to the increasing usage of the vibrating products (household appliances, vehicles, etc.) and increasing demands for the improvement of the life quality, the research related to human response to product vibration is becoming increasingly important. Vibration of the products may be categorized related to which part of the body comes into contact with the vibrating product, mainly in three types: hand-transmitted vibrations, whole-body vibrations and foot vibrations. The force-feedback, which is generated by a product, may also be added in this group, although it is not a type of vibration.

The study of human response to vibration has a long history. As early as 1834, Ernst Heinrich Weber published an extensive investigation, which is related to fundamental aspects of the human tactile sense. From then until now, a number of studies were conducted to obtain thresholds of the vibrotactile perception, tolerance, annoyance and unpleasantness limits, and contours of equivalent sensation. Besides fundamental research, many studies related to human response to hand-transmitted vibrations are concerned with some disorders and injuries associated with hand-transmitted vibrations. A very detailed overview on this type of disorders and injuries is given by Griffin (1990).

One of the first suggestions for the measurement method of human response to whole-body vibrations (building vibration) were published by Reiher and Meister in 1931. They used a semantic scale to measure approximate strengths of perception for various magnitudes of vibration, and this scale consists of the labels “not perceptible”, “weakly perceptible”, “easily perceptible”, “strongly perceptible”, “unpleasant, believed dangerous for long periods”, and “very unpleasant, believed dangerous for short periods”. The boundaries were evaluated using the experimental results. Most of the whole-body vibration research was concentrated upon the investigation of discomfort levels of the whole-body vibrations. They applied different psychophysical measurement methods to determine the discomfort levels and mostly used a unidimensional subjective rating scale that consists of the descriptors “very uncomfortable”, “uncomfortable”, “mildly uncomfortable”, “noticeable, but not uncomfortable.” The International Standard ISO 2631 offered guidance to the evaluation of discomfort produced by whole-body vibration. This standard categorized the discomfort conditions with six different semantic labels, “extremely uncomfortable”, “very uncomfortable”, “uncomfortable” “fairly uncomfortable”, “a little uncomfortable”, and “not uncomfortable” with an important statement that the reactions at various magnitudes depends on passenger expectations with regard to trip duration and the type of activities passengers expect to accomplish (e.g. reading, eating, writing, etc.). A subjective rating scale which contains not only discomfort conditions but also comfort conditions is used by the automotive industry (Table 6.1). Originally this scale which consists of the ten descriptors is suggested by the society of automotive engineers to evaluate the vehicle tires.

Table 6.1: A rating scale for the quality evaluation according to the VDI 2563 (1990)

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
very poor	poor	less than mediocre	mediocre	borderline	acceptable	fair	good	very good	excellent

VDI’s rating scale was also used to evaluate the whole-body vibrations of cars (VDI 2574). Bellmann (2002) has also used this scale and evaluated vibrations of different types of cars (e.g. diesel, gasoline etc.). They have used in their work also a term called “car-vibration quality”, but the term was not discussed apart from the scale labels.

Most studies which deal with the human response to hand-transmitted or whole-body vibration have concentrated upon the unwanted effects of vibration (discomfort etc.), and the main approach of the product designers was the reduction of the vibration energy which was emitted by the product. However in many products, the vibration-intensity levels were already reduced under the hazardous levels and most of the household products are used for short durations which may not cause any disorders (except of some hand-held power tools, e.g. chipping hammer, pressure hammer, chain saw etc. and some vehicles, e.g. earth moving machineries, forklifts etc. have harmful vibration levels and with long duration usage, they can cause some disorders). This achievement gives product designers a chance to take into consideration the multidimensional characteristic of the vibration as a quality parameter of a product, instead of only concentrating upon the vibration-level reduction. This progress shows extensive similarities with the progress of the product-sound design.

In the following we want to discuss the concept of product-vibration design by using the knowledge on the product-sound quality.

Comfort is a dimension of quality

Comfort is one aspect of the overall quality assessment of cars and it is used both for interiour car sound and vibration evaluation. Comfort can be defined as “a conscious well-being, ease and relaxation.” Most researchers have used a unidimensional meaning of comfort in their studies and have asked the subject to judge on a bipolar scale ranging from “extremely discomfortable” to the “extremely comfortable” continuously (as explained in more detail in the introduction).

Can we summarize the quality of product vibration with the comfort or well-being ? Although comfort is one of the important dimensions of quality, the product vibration has a multidimensional characteristic. Dimensions of the product vibration have some similarities with dimensions of the product sound. If we look at the dimensions of the product sound “suitability, or stimulus-response compatibility”, “pleasantness (at least no unpleasantness)”, and “identifiability of sounds or sound sources”. (Blauert and Bodden, 1994, see also Guski, 1997), “typicality” is added to the former dimensions by Blauert and Jekosch (1996),“Strength or magnitude”, “Annoyance value”, “Amenity value”, and “Information content” (Lyon, 2003). Product vibrations should be pleasant, informative, typical, and give positive associations related to the product. All of these dimensions contribute to product-vibration quality evaluations of the user, but they are not all aspects of the product vibration. Of course cognition, action, and emotion play an important role in the portion determination of these dimensions on the quality evaluation.

Vibration as a language between product and user

The user of the product gets information from different sensory channels by using a product. Information (vibration), which comes from the tactile channel, informs the user about functional features of the product and also on what the designer wants to transmit to the user. We can assume that the vibration is one of the languages of the product to communicate with its user. For example, while driving a car we can feel the acceleration of the car clearly from the gas pedal vibrations. In many cases we correlate the increasing speed with the increasing vibration level.

Vibration does not only carry information from the product, it also informs us about the environmental conditions. An example from our daily life is that while driving a car or riding a bicycle, vibrations inform us about the road conditions, and this information helps us to change our driving style to drive safely. Another example is that most of the time we don’t know exactly the stiffness of the wall material of our house and the vibrations which we get from a drilling machine while drilling a hole inform us about the material of the wall.

Another necessary information which is transmitted to the user by the product is that of whether the product operates properly or not. The user can feel not only if there is a problem or not in the machine, but can also identify the problem (location on the product, which part, reason etc.) as well. Of course to identify functional problems, the user should understand the language of the product and this is only possible with experience (identifiablity and familiarity of product vibrations will be discussed in the following paragraph). In the manufacturing industry, trained experts use vibration and noise for the quality control of products.

Each vibration is an information carrier, and even more, it is a communication code between the product and the user.

Identifiability and familiarity of product vibration

The user experiences different information by using a product and tries to understand the information which is provided by the product. Very similar to the human language learning process, the users need a long time of experience (the duration may be different for each user) to learn the language of the product and to create a dictionary for this product in their memory. By designing a vibration, the designers should take into account the existing vibrations and the meaning of the vibration in the users’ dictionary, to avoid misunderstanding by the users.

Typicality of product vibration

Besides the identification or the meaning of vibration information, another important aspect of the product vibration is its typicality. The typicality of the product information which we mean here, is not related to the functional information.

Typicality provides the product to get an artistic identity in the market between the same class of products and also it may form of enthusiasts group for the product. Anyone of the product attributes may be responsible for the typicality, of course the vibration too. An important aspect for the designer is that typicality plays an important and leading role for the fan-group customers and the other aspects of the vibration (e.g. discomfort related to the level or frequency) may not play a role in the overall quality evaluation any more. Designing a typical vibration is an artistic work and therefore it is difficult (or impossible) to find some guidelines for it.

6.5 Combined influence of sound and vibration on the product quality

In user interaction with complex products various information reaches the user from different modalities. Consequently, the cross-modal information has a substantial influence on the product quality evaluation of the user (Altinsoy and Jekosch, 2004). The majority of empirical research on the product quality has focused on isolated unimodal information. However multimodal interactions may have a strong impact on the evaluation of the products. Bednarzyk (1997) has reported an experiment which shows that how important the influence of the multimodal information (non-auditory) on the product-sound quality evaluation can be. Twelve binaurally recorded car sounds (from different brands) were presented to the subjects. Each owner of the cars were asked to identify her/his car sound from the 12 different sounds. However none of them were able to identify it. Only auditory information was not enough for the subjects to recognize their car.

If two modalities are combined, the resulting multimodal percept may be a weaker, stronger, ("additive or subtractive" interaction, Västfjäll, 2003) or an altogether different percept, and of course it is also possible that one modality can be dominant over the overall assessment related to the physical/perceptual ability, the nature of the task, or personal preference (McGee, 2002). Only few investigations address the combined influence of sound and vibration on the overall-quality judgment, and mostly they have concentrated upon annoyance, similar as in product-vibration-quality research. The total annoyance of various combinations of noise and vibration was investigated by Howarth and Griffin (1990). Their results show that when railway noise and railway-induced building vibration occur together, the overall annoyance depends on the magnitudes of both stimuli. A reasonable approximation of the annoyance caused by combinations of noise and vibration may be determined from a summation of the individual effects ("additive interaction"). They have given also a formula for the overall annoyance. Another study on the same subject was conducted by Paulsen and Kastka (1995). On the basis of their results, noise is a dominant information on the overall annoyance judgment, and vibration has only a smaller influence ("additive interaction"). The influence of combined noise and vibration (as stressor) on the subject performance were measured in different studies. While Sandover and Champion (1984) have found subtractive interaction, Innocent and Sandhover (1972) have found no effect. Västfjäll (2003) has found that both sound and vibration stimuli decreased annoyance and increased pleasantness as compared to the sound only condition.

Quehl et al. (1999) developed a semantic differential for aircraft interior sound and vibration. They presented combined noise and vibration pairs in their experiments: However in this study they did not evaluate the unimodal contributions and interaction issues of the noise and vibration on the product (aircraft and helicopter) quality. In another study, Quehl (2001) has investigated the interaction between sound and vibration level and their individual contributions to the vibroacoustic comfort. An increased sound-pressure level combined with an increased vibration magnitude received the least favourable assessment. A decreased sound pressure level combined with a decreased vibration magnitude caused relatively the most frequent comfortable ratings. The results of Quehl (2001) validated the empirical findings of Howarth and Griffin (1991).

Dempsey et al. (1978) have developed noise-and-vibration ride discomfort criteria and they suggested that a simple summation of noise and vibration effects may not accurately predict overall discomfort.

In view of the growing use of multimedia applications (e.g. video telephony, video-conferencing, satellite TV), the perceived quality of auditory-visual stimuli is an exciting research topic. Two recent studies investigated the influence of the presentation of visual stimuli on the loudness evaluation of sounds from traffic noise (Patsouras, 2003; Fastl, 2004). It was shown that the colour of the visual stimulus can influence loudness evaluation in a way that – for the same acoustics stimuli – the loudness of a red train can be rated 15 % higher than the loudness of a green train. It was also reported that the additional visual input reduces the perceived loudness. While still pictures can reduce the perceived loudness on the average by about 2,5 %, moving pictures can induce reductions in perceived loudness around 5 %. In a study reported by Abe et al (1999)., picture of a waterfall improved the ratings of a white noise towards more positive adjectives as compared to the white noise only condition.

An important aspect for the product vibration designer is that sound is usually produced by the vibration. Therefore there is a strong relationship between the physical attributes of the sound and the physical attributes of the vibration. The result of the designed vibration will/can influence the product sound. Some control techniques, e.g. active noise control, give the designer some freedom to overcome the limitations which are coming from the physical coupling between sound and vibration. But in each situation the designer should take into account the interactions between auditory and tactile stimuli on the overall product-quality assessment. In the following section, the auditory-tactile interaction issues in the product design are discussed.

In this section, psychophysical experiments were conducted to investigate combined the influence and interaction issues of auditory and tactile (hand-transmitted vibrations) stimuli on overall product-quality assessment. In these experiments, drill machine noise and vibrations were presented to the subjects.

6.6 Experiments

6.6.1 Set-up

The Saitek tactile-feedback mouse was used to present the tactile information (vibrations) to the subjects. The participants were instructed to hold the mouse in their hand and lift it from the table to avoid unwanted structural vibrations which can be generated from the contact between the mouse and the table and, also, to minimize the noise generated by the mouse. The auditory stimulus was presented from a PC. It was amplified and delivered diotically through Sennheiser HDA 200 closed-face dynamic headphones which have a very high sound-isolation level and therefore mask the background noise of the mouse when it generates the signal. The experiments were conducted in a sound-attenuated room.

6.6.2 Subjects

Twenty subjects, eleven men and nine women, aged between 20 and 62 years, participated in the experiments. All subjects had normal hearing and were right handed, with no known hand disorders. They used their right hand for the experiment.

6.6.3 Stimuli and procedure

To evaluate the quality of electrical drilling machines, five different drilling machines of different price and power range have been selected. Their noise and vibrations were recorded. The recorded drilling machine noises were presented together with the recorded and designed (artificial) vibrations to the subjects via headphones and through the force-feedback mouse simultaneously.

A semantic differential (SD) list was developed for the quality evaluation of the drilling machine. This SD list consists of 15 adjective-pairs: powerless – powerful, irregular – regular, loud – quiet[14], big – small, slow – fast, acute – grave, unpleasant – pleasant, boring – exciting, untypical – typical, intolerable – tolerable (for user), troublesome – untroublesome, rattling – non-rattling, heavy – light, aggressive – calm, not recommendable – recommendable.

In the experiments, subjects had to imagine that they use an electrical drill machine and they should judge on each feature on the SD list. The subjects indicated the intensity of their association on a seven-point scale. The sound and the vibration of five different drilling machines were evaluated using this semantic differential. The experiment consisted of three sessions: In the first session, only sounds of drilling machines were presented. In the second session only vibrations were presented, and in the third session sound and vibration of drilling machines were presented together. Session-order of the experiment was counter balanced across the subjects according to a Latin square.

6.7 Results

Session 1: Drilling machine sounds

The judgements on five drilling machine sounds by 20 subjects were averaged and mean scores are shown in Figure 6.3. The provided attributes cover a wide range of verbal distinctions, clear differences can be observed in polarity profiles of the different drilling machine sounds. Dependent t-tests of the means showed that all five sounds differed significantly from one another ($\underline{p} < 0.05$).

The results of the semantic-differential test show that there is a strong relationship between the attributes loudness, powerfulness and pleasantness. Quiet drilling machine sounds can get higher pleasantness ratings but, at the same time, they are found powerless (see line ◇). Loud drilling machine sounds can get higher power ratings; at the same time, they are found unpleasant (see line (×), and line (▷) in Figure 6.2). Loudness is an information carrier for the user who makes connection between the power of the drilling machine and the loudness. In most of the cases increasing loudness results in an increase in powerfulness judgments. However the drilling machines are noisy machines and high sound pressure levels cause annoyance. It disturbs the users, their family and neighbours, and their activities. The task of the designers is to design a pleasant sound which also gives the impression of ample power. However this task is not so easy. Bisping (1995) has reported a similar problem for car interior sounds. He has studied the dimensions of the car interior sound quality applying the semantic differential technique for evaluating the sounds of luxury, sporty, middle and small

---

[14] The adjective pair „strong-weak“ has replaced the adjective pair „loud–quiet“ for only tactile stimulation condition, according to the absence of the auditory stimulus.

sized cars. He stated that pleasantness and powerfulness form a two-dimensional space for evaluation of car interior sound quality. The relationship (in other words, dependency) between the attributes powerfulness and pleasantness was studied by the manipulation of the sound pressure level (SPL) of car interior sounds. In a magnitude estimation experiment the powerfulness and pleasantness of the car interior sounds were evaluated by the subjects. Figure 6.3 shows the change of the powerfulness and pleasantness ratings according to increase of the SPL.

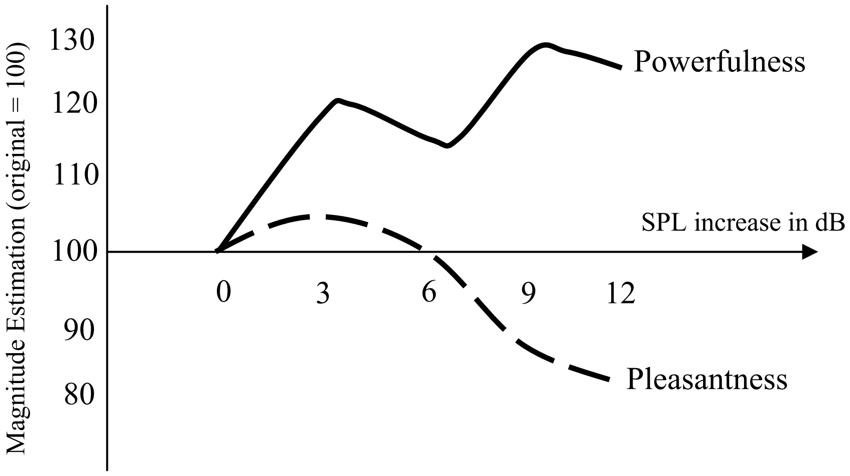


Figure 6.1: Influence of the SPL increase on the powerfulness and pleasantness ratings.

The results of the manipulation confirm the dependency of the attributes powerfulness and pleasantness. Only 3 dB SPL increase results in a simultaneous increase of the powerfulness and pleasantness ratings. However further increments result in a decrease of the pleasantness judgments while the powerfulness rating increases further.

Irregularity and rattling are other factors that make the drilling machine sounds more intolerable (for example Sound B (☐)). They tell the user that there is a problem in the mechanical components of a drilling machine (see the relationship between troublesome and rattling judgments). The attention of the listeners is focused directly on the irregularities of the sound and this causes a reduction of the pleasantness ratings of the sounds.

The results of the semantic differential test shows that listeners associate the size (small-big) and weight (light-heavy) of the drilling machines with the loudness. While the drilling machines which have quiet sounds were perceived as small and light, the loud sounds were associated with big and heavy. Similar type association was reported by Höger and Greifenstein (1997) for the the size of heavy goods vehicles (HGV) and the perceived loudness.

There is a correlation between the judgments of grave-acute and the judgments of pleasantness. Grave sounds were perceived as pleasant.

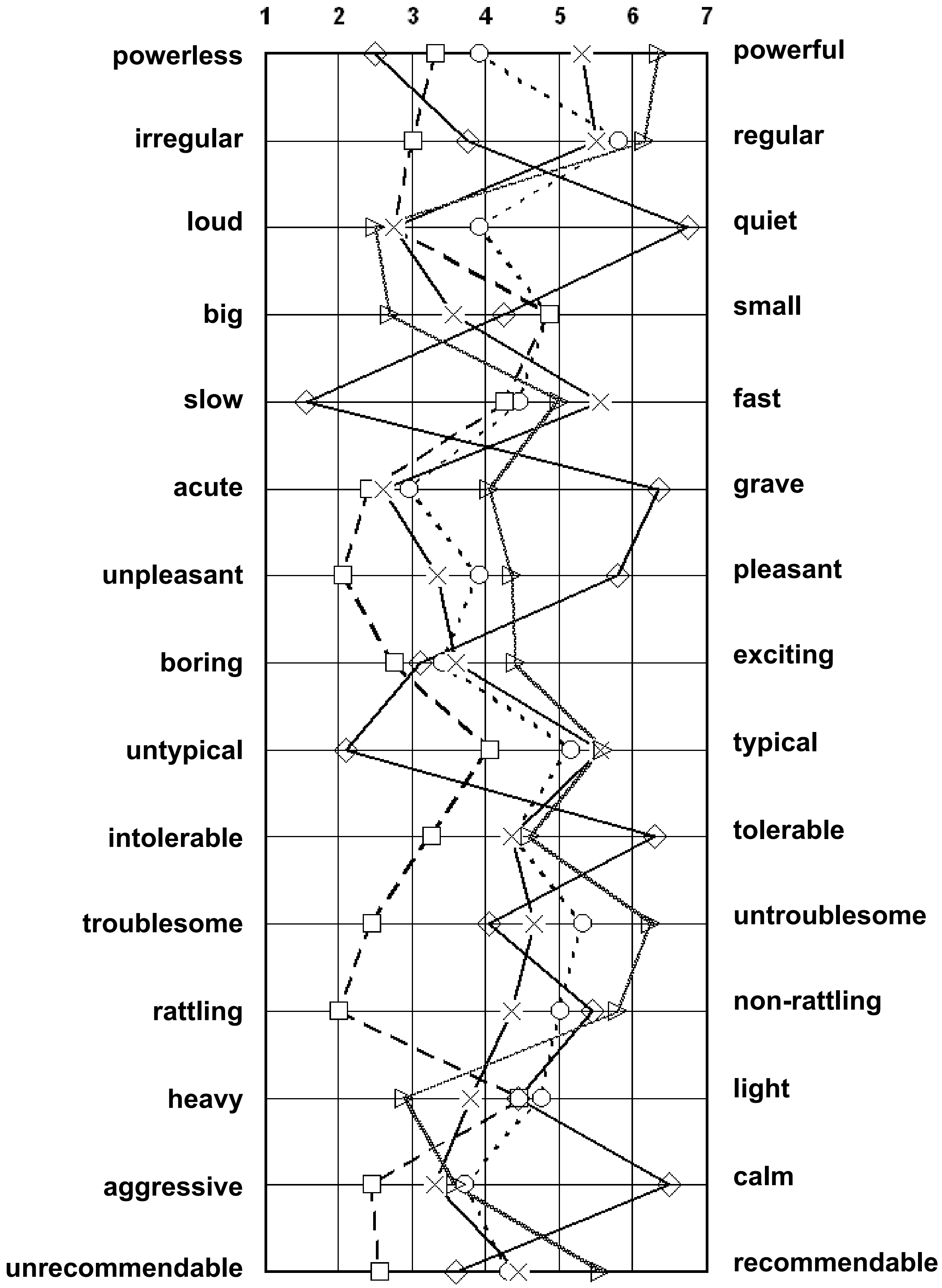


Figure 6.2: Semantic profiles of the drilling machine sounds. ◇ Sound A □ Sound B ▷ Sound C ✕ Sound D ○ Sound E

Sessions 2 and 3: Drilling Machine Sounds and Vibrations

In the second session of the experiment the vibrations of the five drilling machines were evaluated by the subjects (only tactile stimulus). Although the clear differences can be observed in polarity profiles of the different drilling machine vibrations (e.g. powerful vs. powerless, fast vs. slow, rattling vs. non-rattling), the pleasantness ratings ranged only from 3.7 to 5. The subjects made connection between the power of the drilling machine and the intensity of the vibrations, strong drilling machine vibrations can get higher powerfulness ratings. The size (small-big) and weight (light-heavy) of the drilling machines were associated with the magnitudes of vibration.

In the third session sound and vibration of drilling machines were presented together simultaneously and evaluated by the subjects using the same SD list. To investigate the combined influence of auditory and tactile information on the product quality, the same auditory stimuli were presented with different tactile stimulus combinations. The results of the factor analysis of the semantic differential leads to three independent factors explaining 72.83 % of the variance (Table 6.2). Pleasantness and powerfulness are two important dimensions of the drilling-machine sound and vibration.

The semantic profiles for the drilling machine "A"'s sound with different tactile stimulus combinations are shown in Fig. 6.5. As far as cross-modal interaction is concerned, there is no simple overall relationship (i.e., dominance of one modality independent of the attributes), but one that is dependent on the attribute the judgment is based on. As an example for the attributes "power" and "pleasant" two individual modalities interact and lead to a combined perceptual event. With regard to the attributes "fast" and "regular", audition is the dominant modality for both experimental conditions. However vibration is the dominant modality with regard to the attribute "small-big". There is little influence of the tactile stimulus on the auditory loudness judgments.[15] For the attribute "aggressive", we have a case where the tactile information obviously does not play any role (vib A), and where two individual modalities interact (vib C). Comparable relations can be seen for the attribute 'rattling'.

---

[15] An additional pilot experiment was performed on five individuals to measure the influence of tactile information on the loudness of the drilling machine sounds. The results of the pilot experiment showed that an increasing intensity of the tactile stimulus cause an increase (5-8 % of the loudness judgments, related to the stimuli) in the loudness of the drilling-machine sound.

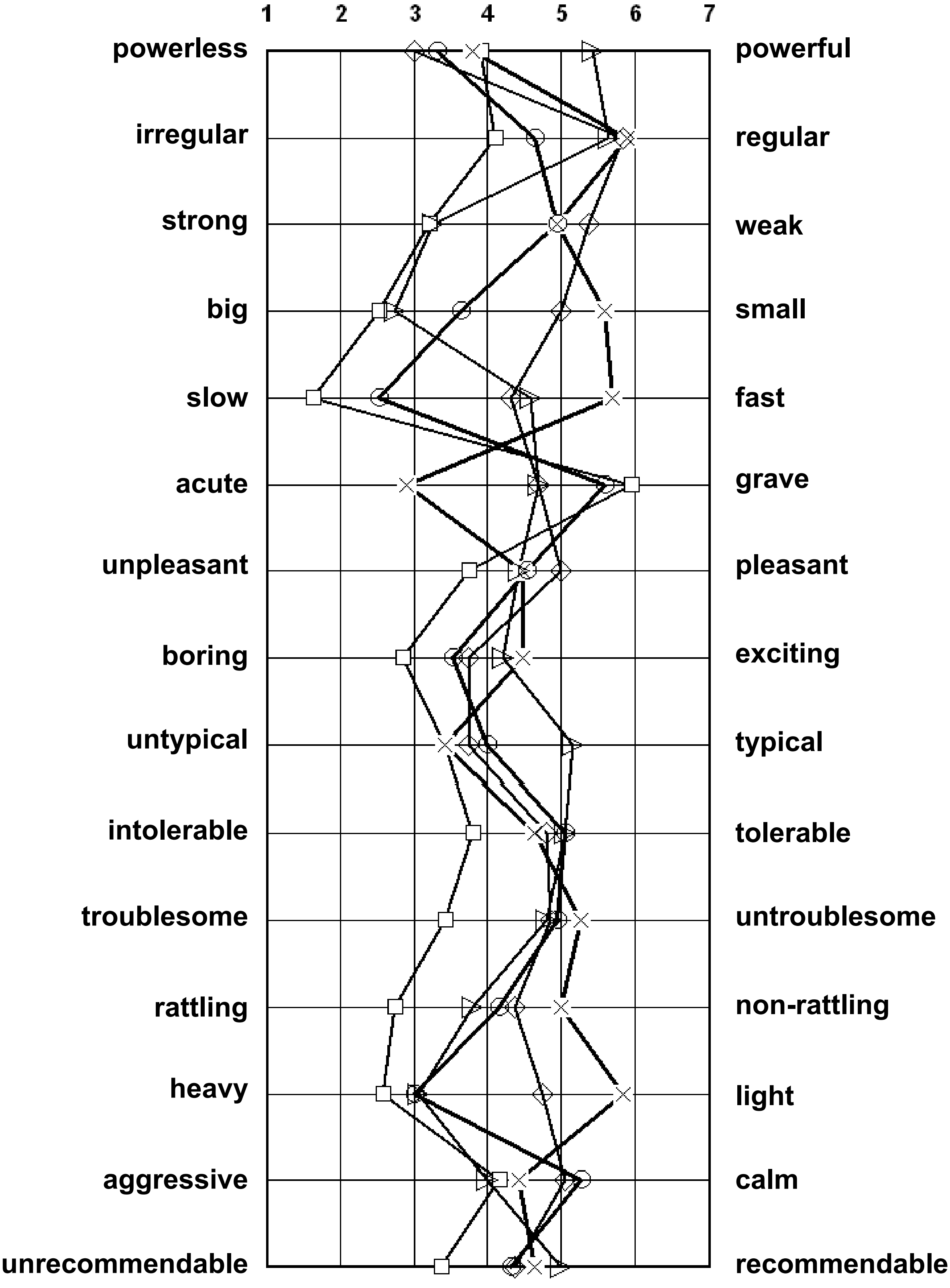


Figure 6.3: Semantic profiles of the drilling machine vibrations. ◇ Vibration A □ Vibration B ▷ Vibration C ✕ Vibration D ○ Vibration E

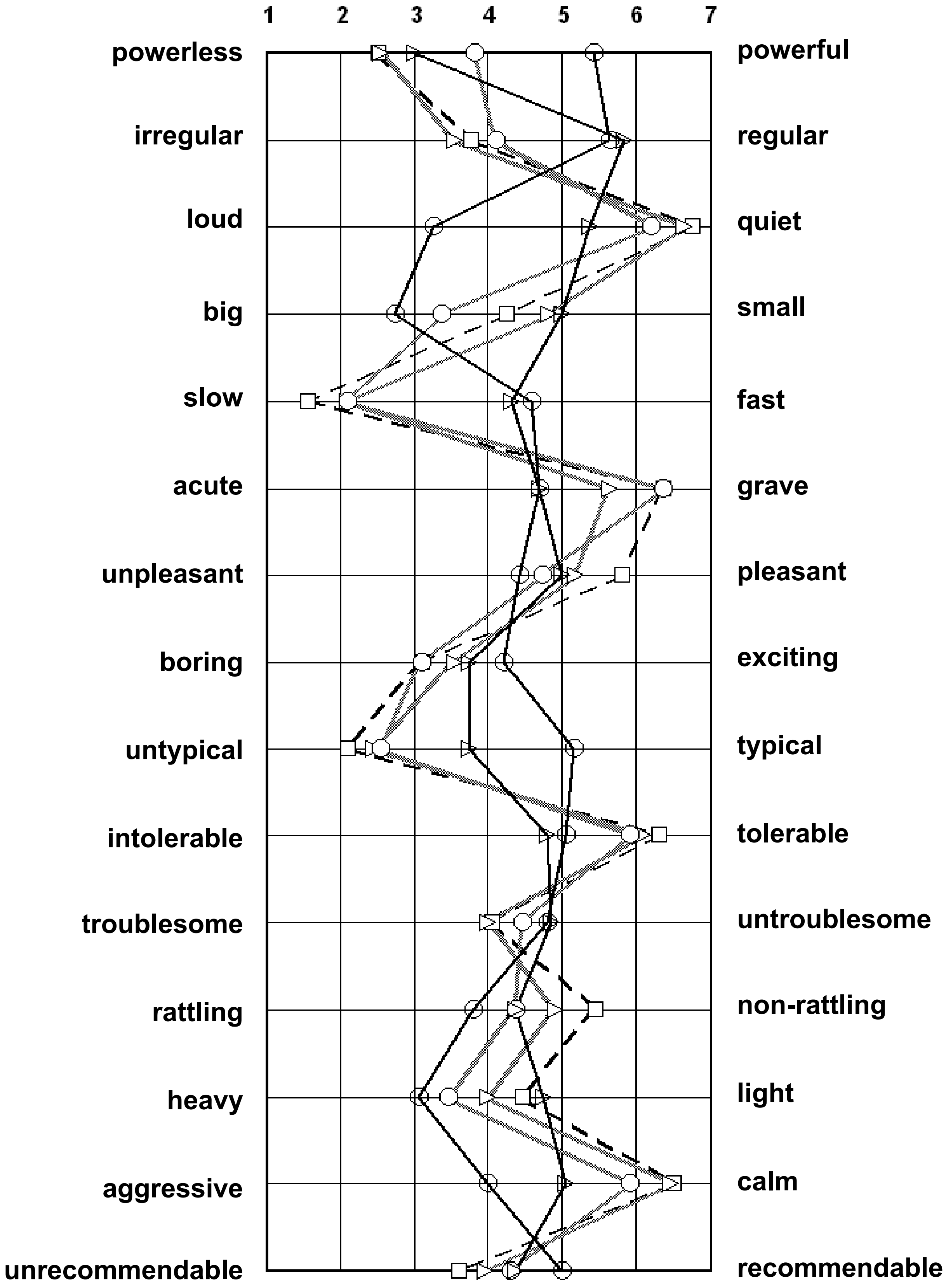


Figure 6.4: Semantic profiles of the drilling machine vibration and sounds. ☐ Sound A ⊖ Vibration C ▷ Vibration A ○ Sound A & Vibration C ▷ Sound A & Vibration A

The results show that tactile information can be a good solution for the optimisation of powerfulness and pleasantness judgments. The drilling machine “A” has a quiet sound which gets high pleasant rating and low power rating. But if it is combined with the “vib C”, the powerfulness judgments increases from the value of 2.5 to the value of 3.8. This additional tactile input causes only little decrease of the pleasantness judgments. In this case, tactile input assumes the task which is informing the user about the functional stage (power of the drilling machine), while the drilling machine sounds pleasant. However, it should not be forgotten that if the sound-pressure level is decreased, it may cause the masking of useful auditory information which leads to avoidance of it and an increase of the disturbance regarding the sound, e.g. if people feel that a sound could be avoided, it is judged more annoying (Guski, 1997). Therefore the level may be adjusted such that user also can sense changes in the speed of rotation from the auditory information.

Table 6.2: Factor analysis results for the drilling machine sound and vibrations.

Factor number	Adjectives	% Variance
Factor 1	Pleasant, loud, tolerable, rattling, grave, calm, untroublesome	42.429
Factor 2	Powerful, fast, regular, recommendable	20.626
Factor 3	Small, light	12.782

6.8 Discussion

Interest in multimodal aspects of product-quality perception has grown rapidly in the past few years. One of the reasons for that is the recent development in virtual-reality technology which makes the user feel immersed in the simulation or application he is running by providing a multimodal, rich and real-time sensorial interaction. Based on the conducted experiments in this study, it is possible to state that virtual environments which enable designers to generate complex, interactive multimodal scenarios are potent tools for multimodal product-quality research.

The results of the present study show that auditory-tactile perception of the products differs from unimodal (auditory or tactile modality separately) perception. Assessment of product quality is a multidimensional task. The availability of different sensory information about the multidimensional aspects of the product may result in complex judgment processes. Therefore it was not possible to observe the dominance of one modality over the overall-quality judgments. Users choose to weight the various modality inputs according to attributes of the product. Different interaction levels were observed for different attributes. Auditory dominance over tactile information can be observed as well as tactile dominance over auditory information depending on the particular attributes with interaction levels which were similar to levels which were presented in Figure 6.4.

From the point of view of product designers, multimodal interaction can be useful in overcoming the limitations of the unimodal product design. The task of the designers is to share the design issues between different sensory information adequately and in harmony to achieve a target product.

The results of the experiments indicate that multimodal design may be a good solution for the problem of the optimisation of the powerfulness and pleasantness for the unimodal stimulus. The correct combination of the auditory and tactile stimuli can result in a powerful and pleasant product image.

Irregularity and rattling in both the auditory and tactile stimulus indicate to the user that there is a mechanical problem and make the product intolerable and unpleasant. The results of the experiments show that when an auditory and a tactile stimulus are combined such that the rating of the tactile rattling is higher than the rating of the auditory rattling then the judgment of magnitude of rattling is moved along in the direction of increasing rattling rating (see Sound A & Vibration C, and Sound A & Vibration A). To lower the high rattling rating of a unimodal stimulus, an idea can be to combine this unimodal stimulus with another unimodal stimulus which has lower rattling rating. Of course the best solution is the elimination of the rattling in both stimuli.

According to the modality-appropriateness hypothesis (Lederman and Klatzky, 2004), audition is especially well suited to temporal tasks (Myers, Cotton, & Hilp, 1981). This hypothesis can be an explanation for the dominance of the auditory information with regard to the attributes like “fast” and “regular”.

Chapter 7

Conclusions

The aim of this thesis was to gain a better understanding of the interaction of auditory and tactile information that is presented in virtual environments by considering the implications of this knowledge for the design of multisensory displays and interfaces. To achieve this aim, it was necessary to be able to specify which criteria have to be met with respect to auditory-tactile integration. Therefore, it was first investigated which physical factors of the auditory and tactile stimuli and which physical conditions can cause a perceptual segregation of auditory and tactile events (Chapter 4). In the second part of this work, two fundamental auditory-tactile interaction examples were chosen and the mechanisms underlying multisensory perception regarding auditory-tactile stimuli was investigated (Chapter 5). The last part focuses on the perceived quality of auditory and tactile stimuli (Chapter 6).

An experimental system was developed and used while performing the auditory-tactile-interaction investigations in this study. Both finger and whole-body stimulations were subject of interest. Also, touch-induced scraping sounds were physically modelled and synthesized to conduct interactive auditory-tactile interaction experiments.

Simultaneity, spatial origin, frequency, and intensity are the key factors which contribute to the perceptual binding of the auditory and tactile stimulus. In Chapter 4.2, the sensitivity of the subjects to asynchrony between auditory and tactile stimuli was investigated and the thresholds for detecting asynchrony were measured. Single impact events were selected as auditory-tactile stimuli. For the vibratory finger stimulation, an auditory delay of 50 ms and a tactile delay of 25 ms were just detectable. Tactile delays are detected better than audio delays. The thresholds were 47 ms (sound earlier) and 63 ms (sound delayed) for auditory and whole-body vibration stimuli. The results indicate that the margin between the thresholds and the required time to generate real-time interactive applications is small. Of course, in complex virtual environments users may not directly focus on the individual impact events, and the tolerated levels may slightly increase. In any case, designers should optimise the system delays. Subjects are more sensitive to tactile delays than to auditory delays.

The levels of the auditory and tactile stimuli are coupled to each other by physical laws when they are generated by one and the same event. In Chapter 4.3, it was investigated how much level difference between auditory and tactile information can lead to a difficulty in integrating auditory and tactile information. In these experiments the level of tactile stimuli were held constant, while the sound pressure level was in- or decreased. The tolerance levels were found to be 17.6 dB for the level increase and 11.2 dB for the level decrease. The results show that subjects have large tolerances for level differences.

There is a strong relationship between the frequency of the auditory stimulus and the frequency of the tactile stimulus, which simply results from the physical processes that generate the stimuli. Interestingly, there are also some similarities between the information processing in the auditory system and the tactile system related to the transformation of frequency. For example, the Pacinian Corpuscle channel functions as a linear integrator of stimulus energy, similarly to a critical band filter in audition (Verrillo and Gescheider, 1975; Marks, 1979; Gescheider et al., 1994; Makous et al., 1995). Chapter 3.4 reports an investigation related to the influence of frequency differences between auditory and tactile stimuli on the segregation of the auditory-tactile events. In this investigation, sinusoidal sounds and vibrations were used. As expected, the subjects tend to prefer pairs having the same frequency for the auditory and tactile stimuli, when looking for the most suitable multi-modal stimulus combination for the multi-modal integration. In most cases, subjects also judge the second harmonic of the tactile frequency to be suitable for the auditory frequency. The tolerated frequency shift ranges of the auditory stimulus were from 30 Hz to 75 Hz for the first harmonic and from 90 Hz to 130 Hz for the second harmonic of tactile frequency of 50 Hz. The tolerated frequency range was from 30 Hz to 150 Hz (auditory frequency) for 63 Hz (tactile frequency). The tolerated frequency ranges were from 30 Hz to 115 Hz (auditory frequency) for 80 Hz (tactile frequency). Subjects had difficulties to integrate a pulsation-type tactile stimulus (4 Hz vibration) with a high-frequency auditory stimulus.

Spatial origin is an important cue for humans to determine whether auditory and tactile signals originate from the same event/object or not. In Chapter 4.5, it was investigated at which levels spatial origin differences can be noticed by subjects. For this purpose a virtual sound source was presented at different locations in the virtual environment via loudspeakers and a tactile stimulus was presented via a tactile device which was held by the subjects. The minimum perceptible angle that allows the subjects to notice differences in the locations of the auditory and tactile events was 5.3°.

Despite the availability of technologies which allow multimodal interfaces to be implemented (at a realistic cost), there is a lack of applied knowledge on how our senses interact when using multimodal interfaces (McGee, 2002). When multiple sensory channels provide information about the same physical event and multisensory integration does appear, what are the relative weights of the different sensory modalities on the final unified multimodal percept? Are there any potential benefits of auditory-tactile interaction to overcome the limitations of the haptic or auditory interfaces? How can auditory and haptic information can be combined effectively in virtual environments? These questions and the underlying principles of the auditory-tactile interactions were addressed in the second part of this thesis (Chapter 5). The first multimodal event experiment was based on hitting an object. If we hit an object, this object will be excited and this excitation will mostly result in an auditory and a tactile feedback. Related to our evaluation and experience in the world, we know that there is a physical relationship between sound pressure level and the level of force-feedback, which were both generated by hitting the object. In this section, psychophysical experiments were conducted to investigate the relative contributions of tactile and auditory information to the perceived strength of the applied force (bimodal; auditory + tactile strongness estimates) and the effect of loudness on tactile force-feedback perception ("strongness") by playing a virtual drum. In the first experiment, subjects participated in three modality conditions: auditory only, tactile only, auditory + tactile. In the bimodal condition, auditory and tactile stimuli were presented physically accurate (the drum sound was presented with loudness proportional to the beat-force magnitude). Results of the experiment show that the auditory and haptic information were approximately equally weighted by subjects and both information contributed to the bimodal judgments of "strongness". In the second experiment, subjects again participated in three modality conditions, but in the bimodal condition, some stimulus-pairs were not physically identical, the drum sounds were presented with loudness greater than would be expected by the beat force. The results indicate that the magnitude of strength increases with increasing loudness in spite of no change in force-feedback as generated by the virtual drum and applied to the subject's hand. This result is promising for virtual environment designers. When considering haptic interfaces, one of the problems is to generate virtual walls as rigid as real walls. Related to the technical limitations, it is not possible to simulate very rigid contact surfaces. Appropriate usage of the auditory information can be useful in overcoming this type of haptic interface limitations.

The second auditory-tactile interaction example was the multisensory roughness perception. Roughness is one of the important physical and perceptual dimensions of the texture. Perceiving the texture of a surface by touching it (scraping with the fingertips) is a multimodal task in which information from auditory, tactile and visual sensory channels are available.

Reproduction of the daily textures in the virtual environments is a difficult task. In Chapter 5.3, firstly, the simulation of realistic tactile textures in a virtual environment was discussed, and then the relationship between physical roughness descriptors and the auditory attributes of the scraping sound was established by using synthesized touch-induced scraping sounds. In the next stage, subjects examined the roughness of different surfaces using tactile and auditory information simultaneously, and the relative contributions of the auditory and tactile information to bimodal roughness estimates were assessed by using unimodal and bimodal roughness estimates. Finally, the possible auditory-tactile interaction issues were discussed related to the roughness perception.

By representing the tactile textures, the electro-tactile stimulation technique was used (more detail can be found in Chapter 2). Current magnitude and pulse frequency of the electro-tactile stimulus were the parameters which allow representing the texture profiles for different roughnesses. An increase in the current magnitude results in an increase in perceived roughness. Subjects tend to find an electrotactile stimulus which has a high current magnitude and a low pulse frequency more suitable for realistic rough surfaces. They tend to find just perceptible current magnitudes suitable for very smooth surfaces and did not show a preference for a certain frequency.

When the relationship between touch-induced scraping sounds and the roughness estimates were studied according to the physical generation mechanism of the sound, the modulation frequency and the modulation depth are information carriers for the texture roughness. The results of the experiments show that perceived roughness increases with a decrease in the modulation frequency. It has also revealed that aside from the modulation frequency, loudness influences the subjects’ roughness estimation and the variation of this parameter in virtual environments makes it easier to feel differences on the roughness of the surfaces. Increasing loudness results in an increase of the perceived roughness.

Subjects bimodally (auditory and tactile simultaneously) explored the same surfaces that were presented using tactile information only and auditory information only. For the bimodal roughness estimations, subjects chose to use both information. The relative contributions of the auditory and tactile information to bimodal roughness estimates were approximately 40% (auditory) and 60% (tactile).

In the multimodal roughness experiments, there was an additional configuration using a stimulus which provided conflicting information in two sensory modalities (the conflicting information did not cause segregation of the percept). In this experiment the roughness of the tactile stimulus and the roughness of the auditory stimulus were slightly different. The perceptual consequences were studied by varying modulation frequency and loudness of the auditory stimulus. The perceived tactile roughness was substantially altered towards the roughness which the auditory stimulus alone perceived. Decreasing modulation frequency results in an increase in perceived tactile roughness, even though the tactile information is smoother than the auditory information. Increasing sound pressure level (approximately 4 or 6 dB) results also an increase in the perceived tactile roughness.

People are exposed to auditory and tactile information simultaneously, when operating a machine (household appliances, hand-power tools, etc.), travelling in a car, or in an airplane. Consequently, the cross-modal information has a substantial influence when evaluating the product quality. However, product design and product quality evaluation often lack an integrated concept of multimodal perception related to product use. The last chapter of this thesis focuses on the combined influence and relative contributions of auditory and tactile information on product quality. In the experiments, drill machine noise and vibrations were presented to the subjects. A semantic differential list which consists of 15 adjective-pairs was developed for the quality evaluation of the drilling machine. The sound and the vibration of five different drilling machines were evaluated using this semantic differential. The subjects performed three sessions, one in which they judged sound quality alone, one in which they judged vibration quality alone, and one in which overall quality was judged. The contribution of changes in presented audio and tactile quality on the perceived overall quality was discussed on the basis of the results of the conducted experiments.

The results of the experiments show that there is no dominance of one sensory modality (auditory or tactile) on the overall product quality judgments. The availability of the different sensory information about the multidimensional aspects of the product may result with the complex judgment processes which are described in Section 5.1. Auditory and tactile information interacts. In some cases the subjects integrate the different sensory sources in some form of compromise that is not identical to any one of them. In some cases subjects uses only one modality of information to make their judgments from multimodal cues. Considering the product design and quality evaluation of products, multimodal interactions must be taken into account, in addition to the separate measurement of product sound and vibration quality.

The results of the experiments have highlighted and encouraged us that the well adjusted combination of auditory and tactile stimuli can be useful to overcome a number of common product design problems. In many cases, for example, users make an association between loudness of the product sound and powerfulness of the product. In the drilling machine sound quality example, this phenomenon was observed. The results suggest that there is a relationship between the attributes: loudness, powerfulness and pleasantness. Increasing loudness results in increasing powerfulness ratings. However, high sound pressure level can cause annoyance and a decrease on the acceptability. The task of the designer is to design a pleasant sound which gives also the impression of the ample power. This is not an easy task, but the results of the experiments show that accompanying tactile information can be a solution for this problem and lead to an improvement of the overall product quality. A quiet drilling machine sound which gets high pleasantness ratings and low powerfulness ratings, received higher powerfulness ratings when it was accompanied by a vibration. In our case, the tactile information gave the user confidence that the drilling machine is more powerful.

Those parameters which indicate to the user that there is a problem with the machine, e.g. rattling, or squeak, should be eliminated in both stimuli (auditory and tactile). Otherwise the attention of the subject focuses on this component and it degrades the overall product quality.

The conducted experiments and their results proved that multimodal virtual environments are potent tools for the multimodal product quality research.

Considering the results of the numerous investigations in this study, it is clear that benefits of the auditory-tactile interaction are very promising for engineers who design multi-modal user interfaces or industrial products. Especially the investigations which were introduced in Chapter 4 give some guidelines to the designers to create more plausible virtual auditory-tactile environments. The results of the Chapter 5 give some hints for optimization of the auditory and tactile attributes in multimodal displays and also indicate the value of cross-modal displays to overcome the limitations of the unimodal interfaces. Chapter 6 shows the clear influence of the auditory and tactile information on the perceived product quality and potential benefits in presenting information through both auditory and tactile modalities instead of through only one of them.

Future Research

The influence of the differences in spatial origin of auditory stimulus and tactile stimulus on the segregation of the auditory and tactile events was investigated by presenting participants synchronous but spatially discordant auditory and tactile stimulus-pairs. However this investigation can be extended by presenting participants both asynchronous and spatially discordant auditory and tactile stimulus-pairs. A pilot experiment, which was not included in this thesis, gave some hints that there may be an audiotactile precedence effect. Whether this is indeed the case, should be investigated by further perception experiments.

The influence of frequency differences between auditory and tactile stimuli on the segregation of the auditory-tactile event was investigated by using sinusoidal sound and vibration signals. However, in our daily life we are exposed broadband and/or time-varying vibration and sounds. Future investigations will need to replicate and extend this work to include other conditions.

Concerning the auditory-tactile interaction, the first multimodal event experiment dealt with hitting an object and the influence of loudness on tactile force-feedback perception was investigated. However hitting causes not only change in sound pressure level, but also changes in other auditory attributes. In further investigations, it may be examined if these attributes have an influence on auditory-tactile interaction. If so, the contributions of these attributes to the auditory-tactile interaction need to be investigated.

The use of haptic devices is just emerging. Some of the potential promising applications are virtual environments for remote medical surgery and e-commerce. These applications require knowledge about different object properties in addition to the texture (roughness). A potential subject for further research may be identification and perception of materials from contact sounds and tactile attributes, and intersensory material perception.

In this study, the vibration axis of interest was vertical axis. Complete freedom of motion in space admits of six Degrees of Freedom (6DOF), which include the possibility of three directions of displacement and three angular gyrations (Martens and Woszczyk, 2004). To increase the knowledge which was gained in this thesis, future investigations will have to consider other available types of vibrations and motions.

For people who are visually impaired, the tactile and auditory senses are very important in everyday life. Auditory-tactile interaction can be a promising tool for blind and visually impaired multimedia users (e.g. computer, virtual environment etc.). The results of the earlier psychophysical experiments show that blind people are better in discriminating the roughness of textures than sighted people. In future the potential benefits of auditory-tactile interaction for visually impaired persons has to be investigated by conducting more psychophysical experiments.

Everybody has sometimes experienced the vibrations generated by the performance of music. The floor or the chair can vibrate because of the resonance or the structure-borne sound stimulated by instruments (Daub and Altinsoy, 2004). Daub (2004) has investigated the cross-modal relationship between auditory and tactile (whole-body vibration) perception of musical events. Knowledge on the interaction between vibration and airborne sound regarding the music may be important to develop new multimedia display systems.

The results of this thesis indicate the importance of the multimodal product quality perception. The modalities used in this work were the tactile and the auditory modality. It would be beneficial to investigate the effects of third or fourth modalities (such as vision, olfactory, etc.) repeating this work.

Literature

Abe, K., K., Ozawa, Y., Suzuki and T. Sone (1999): The effects of visual information on the impression of environmental sounds. In: *Proc. Internoise'99*, 1177-1182.

Altınsoy, M.E., J. Blauert, and C. Treier (2001): Inter- modal effects of non-simultaneous stimulus presentation. In: A. Alippi, eds., *Proc. 17th I.C.A.*, **4**, 8-10.

Altınsoy, M. E. and U. Jekosch (2004):. The influence of non-stationary vibrations on the quality of non-stationary product sounds," In: *Proc.. 7. Congrès Français d'Acoustique / 30. Dtsch. Jahrestg. F. Akust. (CFA/DAGA '04),* **164**, 263-264.

Altınsoy, M. E. (2003):. Perceptual aspects of auditory-tactile asynchrony. *Proc.Tenth Int. Cong. Sound and Vib.*, 3831-3838.

ANSI S1.1 (1994): *American national standard acoustical terminology*. Acoustical Society of America, New York.

Barlow, H. B. and J. D. Mollon (1982): *The senses*. Cambridge University Press, Cambridge.

Bednarzyk, M. (1997): *Qualitätsbeurteilung der Geräusche industrieller Produkte. Der Stand der Forschung, abgehandelt am Beispiel der Kfz-Innenraumgeräusche*. Ph.D. Thesis. VDI Reihe 12 Nr. 396, VDI-Verlag, Düsseldorf

Bekesy, G. von (1960): *Experiments in hearing*, McGraw Hill, New York.

Bellmann, M. (2002): Perception of whole-body vibrations: From basic experiments to effects of seat and steering-wheel vibrations on the passenger's comfort inside vehicles. Ph.D. Thesis. Carl von Ossietzky Universität Oldenburg.

Bensmäıa, S. J. and M. Hollins (2000): Complex tactile waveform discrimination. *J. Acoust. Soc. Am.*, **108**, 1236–1245.

Benson, A.J. (1990): Sensory functions and limitations of the vestibular systems. In: R. Warren and A.H. Wertheim, eds., *Perception and control of self motion*. Lawrence Erlbaum Assoc.: Hillsdale, NJ.

Bertelson, P. and M. Radeau (1981): Cross-modal bias and perceptual fusion with auditory-visual spatial discordance. *Percept. Psychophys.*, **29**, 578-584.

Blake, D. T., S. S. Hsiao and K. O. Johnson (1997): Neural coding mechanism in tactile pattern recognition: The relative contributions of slowly and rapidly adapting mechanoreceptors to perceived roughness. *J. Neurosci.*, **17**, 7480–7489.

Bisping, R. (1995): Emotional effect of car interior sounds: pleasantness and power and their relation to acoustic key features. *Proc. Noise and Vibration Conf.* Soc. Automotive Eng., Traverse City (MI). SAE paper **951284**, 1203-1209.

Blauert, J. (1970): Ein Versuch zum Richtungshören bei gleichzeitiger optischer Stimulation. *Acustica*, **23**, 118-119

Blauert, J. (1986): Cognitive and aesthetic aspects of noise engineering. *Proc. Inter-noise*, Cambridge, 5-13.

Blauert, J. and M. Bodden (1994): Gütebeurteilung von Geräuschen– Warum ein Problem? In: Q.-H. Vo, eds., *Sound engineering*. Expert, Renningen.

Blauert, J. and U. Jekosch (1997): Sound-quality evaluation—A multilayered problem. *Acustica United with Acta Acustica*, **83**, 747–753.

Blauert, J. (1997): *Spatial hearing*. MIT Press, Cambridge.

Blauert, J., H. Lehnert, J. Sahrhage and H. Strauss (2000): An interactive virtual-environment generator for psychoacoustic research. *Acustica united with Acta Acustica*, **86**, 94-102.

Blauert, J. (2003): Akustik II. Lecture Notes. Ruhr-University Bochum.

Boerger, G., P. Laws and J. Blauert (1977): Stereophone Kopfhörerwiedergabe mit Steuerung bestimmter Übertragungfaktoren durch Kopfderhbewegungen. *Acustica*, **39**, 22-26.

Bolanowski, S. J. Jr., G. A. Gescheider, R. T. Verillo, and C. M. Checkosky (1988): Four channels mediate the mechanical aspects of touch. *J. Acoust. Soc. Am.*, **84**, 1680–1694.

Booth, G.B. (1958): Random motion test techniques. *Proc Inst Env Engrs.*

Bregman, A.S. (1990): *Auditory scene analysis*. MIT Press, Cambridge.

Burdea, G. C., and P. Coiffet (1994): *Virtual reality technology*. Wiley, New York.

Burdea, G. C. (1996): *Force and touch feedback for virtual reality*. John Wiley and Sons, Inc, New York.

Cohen, J., M. Lin, D. Manocha and M. Ponamgi, (1995): I-COLLIDE: An interactive and exact collision detection system for large-scale environments. *Proc. ACM Interactive 3D Graphics Conference,* ACM, NY, 189-196.

Cook, P. (2002): *Real sound syhnthesis for interactive applications*. A K Peters Ltd. Natick, MA.

Colgate, J. E., P. E. Grafing, M. C. Stanley and G. Schenkel (1993): Implementation of stiff virtual walls in force-reflecting interfaces. *IEEE Virtual Reality Annual International Symposium*, Seattle, WA, 202-207.

Connor, C. E., S. S. Hsiao, J. R. Philips and K. O. Johnson (1990): Tactile roughness: Neural codes that account for psychophysical magnitude estimates. *J. Neurosci.*, **10**, 3823–3836.

Connor, C. E. and K. O. Johnson (1992): Neural coding of tactile texture: Comparisons of spatial and temporal mechanisms for roughness perception. *J. Neurosci.*, **12**, 3414–3426.

Corkin, S., B. Milner and T. Rasmussen (1970): Somatosensory thresholds. contrasting effects of postcentral gyrus and posterior parietal-lobe excisions, *Arch. Neurol.*, **23**, 41-58.

Costa, M.A. and M.R.Cutkosky (2000): Roughness perception of haptically displayed fractal surfaces, *Proc.ASME IMECE DSC Haptics Symposium*, Orlando, Florida.

Daub, M. (2003): *Kreuzmodale Korrelation – Zusammenhang zwischen musikalisch produzierten Ganzkörpervibrationen und der auditiven Wahrnehmung*. Diplomarbeit. Institut für Kommunikationsakustik. Ruhr-University, Bochum.

Daub, M. and M. E. Altınsoy (2004): Audiotactile simultaneity perception of whole-body vibrations produced by musical presentations. In: *Proc. 7. Congrès Français d'Acoustique / 30. Dtsch. Jahrestg. F. Akust. (CFA/DAGA '04)*, **175**, 111-112.

Dempsey, T.K., Leatherwood, J.D. & Sherman, A.C. (1978): Development of noise and vibration ride comfort criteria. *J. Acoust. Soc. Am.*, **65**, 124-132.

DiFranco, D.E., G.L. Beauregard and M. Srinivasan (1997): The effect of auditory cues on the haptic perception of stiffness in virtual environments. *Proc. ASME Dyn. Sys. and Cont. Div.*, DSC **61**, 17-22.

DIN 1320 (1959): *Allgemeine Benennungen in der Akustik*. Beuth-Vertrieb, Berlin.

Driver, J. and C. Spence (2004): Crossmodal spatial attention: evidence from human performance. In: C. Spence and J. Driver, eds., *Crosmodal space and crossmodal attention*. Oxford University Press, New York.

Dix, A., J. Finlay, G. Abowd and R. Beale (1993): *Human-computer interaction.* Prentice Hall, New York.

Dixon, N.F. and L. Spitz (1980): The detection of auditory visual desynchrony. *Perception*, **9**, 719-721.

Fastl, H. (2004): Audio-visual interactions in loudness evaluation. *Proc. 18th ICA*, Kyoto. Tu3X1.

Fisher G.H. (1968): Agreement between the spatial senses. *Perceptual and Motor Skills*, **126**, 849-850.

Fletcher, N. and T. Rossing (1991): *The physics of musical instruments*. Springer-Verlag, New York.

Foley, J., A. van Dam, S. Feiner, J. Hughes and R. Philips (1994): *Introduction to computer graphics*. Addison-Wesley Publishing Co., Menlo Park, CA.

Forthergill, L.C. (1972): *A study of the subjective response to whole-body vibration*. M.Sc. Thesis, University of Southampton.

Gardner, E.P. and E.R. Kandel (1999): Touch. In: E. Kandel, J.H. Scwartz and T.M. Jessell, eds., *Principles of neural science*. McGraw-Hill.

Gerhard, D. (2003): Pitch extraction and fundamental frequency: History and current techniques. Technical Report TR-CS. Department of Computer Science University of Regina, Regina, Saskatchewan.

Gescheider, G. A., K. E. Hoffman, M. A. Harrison, M. L. Travis and S. J. Bolanowski (1994): The effects of masking on vibrotactile temporal summation in the detection of sinusoidal and noise signals. *J. Acoust. Soc. Am.*, **95**, 1006–1016.

Goff, G.D. (1967): Differential discrimination of frequency of cutaneous mechanical vibration. *J. Exp. Psychol*, **74**, 294-299.

Graziano, M.S.A., C.G. Gross, C.S.R. Taylor and T. Moore (2004): A system of multimodal areas in the primate brain. In: C. Spence and J. Driver, eds., C*rosmodal space and crossmodal attention*. Oxford University Press, New York.

Griffin, M.J. (1990): *Handbook of Human Vibration*, Academic Press, London.

Guest, S., C. Catmur, D. Llyod and C. Spence (2002): Audiotactile interactions in roughness perception. *Exp. Brain Res.*, **146**, 161-171.

Guski, R. (1997): Psychological methods for evaluating sound quality and assessing acoustic information. *Acustica united with Acta Acustica*, **83**, 765–774.

Haustein, B. G. and W. Shirmer (1970): Messeinrichtung zur Untersuchung des menschlichen Gehörs. *Hochfrequenztech. u. Elektroakustik*, **79**, 96-101.

Helmholtz, H. L. F. (1954): *On the sensations of tone as a physiological basis for the theory of music* (2nd ed., A. J. Ellis, Trans.). New York: Dover. (Original work published 1885).

Hirsh I.J. and C.E. Sherrrick (1961): Perceived order in different sense modalities. *J. Exp. Psychol.*, **62**, 423-432.

Howarth, H.V.C. and M.J. Griffin (1988): The frequency dependence of subjective reaction to vertical and horizontal whole-body vibration at low magnitudes. *J. Acoust. Soc. Am.*, **83**(4): 1406-1413.

Howarth, H.V.C. and M.J. Griffin (1990): Subjective response to combined noise and vibration: Summation and interaction effects. *J. Sound and Vibration*, **143**, 443-454.

Höger, R., and P. Greifenstein (1988): Zum Einfluß der Größe von Lastkraftwagen auf deren wahrgenommene Lautheit. *Zeitschrift für Lärmbekämpfung*, **35**, 128-131.

Hudspeth, A.J. (1999): Sensory transduction in the ear. In: E. Kandel, J.H. Scwartz and T.M. Jessell, eds., *Principles of neural science*. McGraw-Hill.

Innocent, P.R. and J. Sandhover, (1972): A pilot study of the effects of noise and vibration acting together; subjective assessment and task performance. *Proc. United Kingdom Informal Group Meeting on Human Resp. to Vib.*, University of Sheffield.

ISO 2631 (1997): *Evaluation of human exposure to whole-body vibration, International Organisation for Standardisation*. Geneva.

Jaskowski, P., F. Jaroszyk and D. Hojan-Jezierska (1990): Temporal-order judgment and reaction time for stimuli of different modalities. *Psychol. Res.*, **52**, 35-38.

Jex, H. (1988): Four critical tests for control-feel simulators. *23rd annual Conference on Manual Control, Cambridge*, MA.

Johansson, R.S. and A.B. Valbo, (1979): Tactile sensibility in the human hand: Relative and absolute densities of four types of mechanoreceptive units in glabrous skin. *J. Exp. Psychol.*, **286**, 283-300.

Jones, L. (1989): The assessment of hand function: A critical review of techniques. *J. Hand Surgery*, **14**, 221–228.

Jones, L. (1997): Dextrous hands: Human prosthetic, and robotic. *Presence*, **6**, 29-56

Jones, L. and M. Berris (2002): The psychophysics of temperature perception and thermal-interface design. *10th Symp. Haptic Inter. Virt. Env. and Teleope. Sys*. Orlando, Florida.

Jousmäki, V., R. Hari (1998): Parchment-skin illusion: Sound-biased touch. *Current Biology*, **8**, 190.

Kaczmarek, K. and P. Bach-y-Rita (1995): Tactile displays. In W. Barfield and T. Furness III eds., *Virtual Environments and Advanced Interface Design*. pp.349-414. Oxford University Press. New York.

Kadunce, D.C., J.W. Vaughan, M.T. Wallace, G. Benedek and B.E. Stein (1997): Mechanisms of within- and cross-modality suppression in the superior colliculus. *J. Neurophysiol.*, **78**, 2834-2847.

Kandel, E. R., J. H. Schwartz and T. M. Jessell (1991): *Principles of neuroscience*, 3rd ed. Elsevier, New York.

Katz, D. (1925): *Der Aufbau der Tastwelt.* Barth Verlag, Leipzig.

King, A.J. (2004): Development of multisensory spatial integration. In: C. Spence and J. Driver, eds., C*rosmodal space and crossmodal attention*. Oxford University Press, New York.

King, A.J. and A.R. Palmer (1985): Integration of visual and auditory information in bimodal neurons in the guinea-pig superior colliculus. *Exp. Brain Res.*, **60**, 492-500.

Klatzky, R.L. and S.J. Lederman (2000): Modality specificity in cognition: The case of touch. In: H.L. Roediger, J.S. Nairne, I. Neath and A.M. Suprenant, eds. *The Nature of Remembering: Essays in Honor of Robert G. Crowder*. American Psychological Association Press, Washington, D.C.

Klatzky, R.L., D.K. Pai and E.P. Krotkov (2000): Perception of material from contact sounds. *Presence: Teleoperators and Virtual Environments*, **9**, 399-410.

Klatzky, R.L. and S.J. Lederman (2002): Touch. In: A.F. Healy and R. W. Proctor, eds., Experimental Psychology. In: I.B. Weiner (Editor-in-Chief): *Handbook of Psychology*. Wiley, New York.

Kohlrausch, A. and S. van de Par (1999): Auditory-visual interaction: From fundamental research in cognitive psychology to (possible): applications. In: *Human Vision and Electronic Imaging IV*, *Proc. Soc. Photo-Optical Instrumentation Engrs*., 3644, 34-44.

Kohlrausch A. (2000): Perceptual consequences of asnchrony in audio-visual stimuli. *IPO Ann. Progr. Rep.*, **35**, 140-149.

Koseleff, P. (1957): Studies in the perception of heaviness. I. *Acta Psychologica*, **13**, 242-252.

Kuhl, P.K. and A.N. Meltzoff (1982): The bimodal perception of speech in infancy. *Science*, **218**, 1138-1141.

Ladavas, E. and A. Farne (2004): Neurophysiological evidence for multimodal representations of space near specific body parts. In: C. Spence and J. Driver, eds., C*rosmodal space and crossmodal attention*. Oxford University Press, New York.

Lakatos, S., S. McAdams and R. Caussé (1997): The representation of auditory source characteristics: simple geometric form. *Percept. Psychophys.*, **59**, 1180–1190.

Laming, D. (1986): *Sensory analysis*. Academic Press, London.

Lederman, S.J. and M.M. Taylor (1972): Fingertip force, surface geometry and the perception of roughness by active touch. *Percept. Psychophys.*, **12**, 401-408.

Lederman, S.J. (1979): Auditory texture perception. *Perception*, **8**, 93-103.

Lederman, S.J. and S.G. Abbott (1981): Texture perception: Studies of intersensory organization using a discrepancy paradigm and visual versus tactual psychophysics. *J. Exp. Psychol.: Human Perception and Performance*, **7**, 902-915.

Lederman, S.J., G. Thorne and B. Jones (1986): Perception of texture by vision and touch: Multidimensionality and intersensory integration. *J. Exp. Psychol.: Human Perception and Performance*, **12**, 169-180.

Lederman, S.J., R.L. Klatzky, C.L. Hamilton and G.I. Ramsay (1999): Perceiving roughness via a rigid probe: Psychophysical effects of exploration speed and mode of touch. *Haptics-e (Electronic Journal of Haptics Research)*, **1**, 1-20.

Lederman, S.J., R.L. Klatzky, C. Hamilton and T. Morgan (2002): Integrating multimodal information about surface texture via a probe: Relative contributions of haptic and touch produced sound sources. *10th Annual meeting of Haptic Interfaces for Teleoperator and Virtual Environment Systems. Satellite meeting of the Annual IEEE VR '02 meeting*. pp 97-104.

Lederman, S.J. and R.L. Klatzky (2004): Multisensory texture perception. In G. Calvert, C. Spence and B. Stein, eds., *Handbook of multisensory processes*. MIT Press, Cambridge.

Levitt, H. (1970): Transformed up-down procedures in psychoacoustics. *J. Acoust. Soc. Am.* **49**, 467–477.

Lewkowicz, D.J. and G. Turkewitz (1980): Cross-modal equivalence in early infancy: auditory-visual intensity matching. *Developmental Psychology*, **16**, 597-607.

Lewkowicz, D.J. (1986): Developmental changes in infants' bisensory response to synchronous durations. *Infant Behaviour and Development*, **9**, 335-353.

Lewkowicz, D.J. (1996): Perception of auditory-visual temporal synchrony in human infants. *J. Exp. Psychol.*, **22**, 1094-1106

Lewkowicz, D.J. (2002): Heterogeneity and heterochrony in the development of intersensory perception. *Cogn. Brain Res.*, **14**, 41-63.

Lyon, R. (2003): Product sound quality – from perception to design. *Sound and Vibration*, **3**, 18-22.

Makous, J. C., R. M. Friedman and Jr. C. J. Vierck (1995): A critical band filter in touch. *J. Neurosci.*, **15**, 2808–2818.

Marks, L. E. (1979): Summation of vibrotactile intensity: An analog to auditory critical bands? *Sensory Process.*, **3**, 188–203.

Martens, W. L. and W. Woszczyk (2004): Subspace projection of multichannel audio data for automatic control of motion-platform-based multimedia display systems. *Proc. IEEE Intl. Conf. Acoustics, Speech, Signal Processing (ICASSP).*

McDonnell, P.M. and J. Duffett, (1972): Vision and touch: A reconsideration of conflict between the two senses. *Canadian J. Psychology*, **26**, 171-180.

McGee, M.R., P.D. Gray and S.A. Brewster (2002): Mixed feelings: Multimodal perception of virtual roughness. In: *Proc. Eurohaptics 2002* (Edinburgh, UK): 47-52 Edinburgh University.

McGee, M.R. (2002): *Investigating a multimodal solution for improving force feedback generated textures*. Ph.D. Thesis. University of Glasgow.

Meredith, M.A. and B.E. Stein (1996): Spatial determinants of multisensory integration in cat superior colliculus neurons. *J. Neurophysiol.*, **75**, 1843-1857.

Minsky, M. and S.J. Lederman, (1996): Simulated haptic textures: Roughness. In *Proc. ASME Int. Mech. Eng. Cong.: Dyn. Sys. and Cont. Div.*, **58**, 421-426.

Miner N. and T. Caudell (1998): Computational requirements and synchronization issues for virtual acoustic displays, *Presence*, **7**, 396-409.

Moore, M. and J. Wilhelms (1988): Collision detection and response for computer animation. *Computer Graphics*, **22**, 289-298.

Morioka, M. and J.M. Griffin (2000): Difference thresholds for intensity perception of whole-body vertical vibration: Effect of frquency and magnitude, *J. Acoust. Soc. Am*. **107** 620 – 624.

Morley, J.W., A.W. Goodwin and I. Darian-Smith (1983): Tactile discrimination of gratings. *Exp. Brain Res.*, **49**, 291-299.

Mowbray, G.H. and J.W. Gebhard (1957): Sensitivity of the skin to changes in the rate of intermittent mechanical stimuli. *Science*, **125**, 1297-1298.

MUVII (Multi-User Virtual Interactive Interface) (2004): Final Report. The European Commission Information Society Technologies Project. IST-2000-28463

Myers, A.K., B. Cotton and H.A., Hilp (1981): Matching of the rate of concurrent tone bursts and light flashes as a function of flash surround luminance. *Percept. Psychophys.,* **30**, 33-38.

Patsouras, C. (2003): *Geräuschqualität von Fahrzeugen – Beurteilung, Gestaltung und multimodale Einflüsse*. Ph.D. Thesis. Shaker Verlag, Aachen.

Paulsen, R. and J. Kastka (1995): Effects of combined noise and annoyance. *J. Sound and Vibration*, **181**, 295-31.

Pearson, K. and J. Gordon (2000): Spinal Reflexes. In: E. Kandel, J.H. Scwartz and T.M. Jessell, eds., *Principles of neural science*. McGraw-Hill.

Pellegrini, R.S. (2002): *A virtual reference listening room as an application of auditory virtual environments*. Ph.D. Thesis. dissertation.de.

Philips, J. R. and K. O. Johnson (1981): Tactile spatial resolution: II. Neural representation of bars, edges and gratings in monkey primary afferents. *J. Neurophysiol.*, **46**, 1204–1225.

Pick, H.L., D.H., Warren and J.C. Hay (1969): Sensory conflict in judgments of spatial directions. *Percept. Psychophys.*, **6**, 203-205.

Polya, G. and G. Szegö. (1976):. *Problems and theorems in analysis.* Springer-Verlag, New York

Quehl, J.J., A. Schick, V. Mellert, B. Schulte-Fortkamp and H. Remmers (1999): Combined effects of helicopter and aircraft interior sound and vibration on passengers' well-being and comfort sensation. In: *Forum Acusticum Berlin, CD-ROM, Collected Papers from the Joint Meeting Berlin 99*. DEGA, Oldenburg.

Quehl, J.J. (2001): *Comfort studies on aircraft interior sound and vibration*. Ph.D. Thesis. Shaker Verlag, Aachen.

Quilliam, T.A. (1978): The structure of finger print skin. In: G. Gordon, eds., *Active touch: The mechanism of recognition of objects by manipulation*. Pergamon, Oxford.

Reiher, H. and F.J. Meister (1931): The sensitivities of the human body to vibrations [Orig. Empfindlichkeit des Menschen gegen Erschütterung]. *Forschung VDI*, **2**, 381-386.

Rothenberg, M., R.T. Verrillo, S.A. Zahorian, M.L. Brachman and S.J.Jr. Bolanowski (1977): Vibrotactile frequency for encoding a speech parameter. *J. Acoust. Soc. Am.*, **62**, 1003-1012

Salisbury, J. K. and M. A. Srinivasan, (1997): Phantom-based haptic interaction with virtual objects, *IEEE Computer Graphics and Applications*, **17**.

Sandover, J. and D.F. Champion, (1984): Some effects of a combined noise and vibration environment on a mental arithmetic task. *J. Sound and Vibration*, **95**, 203-212.

SCATIS (Spatially-Coordinated Audio/Tactile Interactive Scenario) (1996): European Union Project, ESPRIT #6358

Shams, L., Y. Kamitani, S. Shimojo (2000): What you see is what you hear. *Nature*, **408**, 788.

Shams, L., Y. Kamitani, S. Shimojo (2002): Visual illusion induced by sound. *Cognitive Brain Res.*, **14**, 147-152.

Silzle A, Novo P, Strauss H (2004): IKA-SIM: A system to generate auditory virtual environments, *116th AES Convention, Berlin, Germany*, preprint number #6016

Singer, W. & Gray, C.M. (1995): Visual feature integration and the temporal correlation hypothesis. *Annu. Rev. Neurosci.*, **18**, 555-586.

Spelke, E.S. (1979): Perceiving bimodally specified events in infancy. *Developmental Psychology*, **15**, 626-636.

Spelke, E.S. and C.J. Owsley (1979): Intermodal exploration and knowledge in infancy. *Infant Behaviour and Development*, **2**, 13-27.

Srinivasan, M.A., G.L. Beauregard and D.L. Brock (1996): The impact of visual information on the haptic perception of stiffness in virtual environments. *Proc. of the ASME Dynamic Systems and Control Division, DSC*, **58**, 555-559.

Stein, B.E., E. Labos and L. Kruger (1973): Sequence of changes in properties of neurons of superior colliculus of the kitten during maturation. *J. Neurophysiol.*, **36**, 667-679.

Stein, B.E., and M.A. Meredith, (1993): *The merging of senses*. MIT Press, Cambridge, Massachusetts.

Stein, B.E., T.R. Stanford, M.T. Wallace, J.W. Vaughan and W. Jiang, (2004): Crossmodal spatial interactions in subcortical and cortical circuits. In: C. Spence and J. Driver, eds., C*rosmodal space and crossmodal attention*. Oxford University Press, New York.

Stevens, J. C. and J. D. Mack (1959): Scales of apparent force. *J. Exp. Psychol.*, **58**, 405–413.

Stevens, S.S. and J. Harris (1962): The scaling of subjective roughness and smoothness. *J. Exp. Psychol.*, **64**, 494-498.

Stevens, S.S. (1975): *Psychophysics: Introduction to its perceptual, neural and social prospects*. Wiley, New York.

Sugita, Y. and Y. Suzuki (2003): Audiovisual perception: Implicit estimation of sound-arrival time, *Nature*, **421**, 875-973.

Sutherland, I. (1965): The ultimate display. *Proc. Int. Fed. Inf. Process.'65*, 506-508.

Tan, H. Z., M. A. Srinavasan, B. Eberman and B. Cheng, (1994): Human factors for the design of force-reflecting haptic interfaces. In: C. J. Radcliffe, eds., *Dynamic systems and control* (DSC 55–1). The American Society of Mechanical Engineers. New York.

Taylor, M.M. and S.J. Lederman, (1975): Tactile roughness of grooved surfaces: A model and the effect of friction. *Percept. Psychophys.*, **17**, 23-36.

Teghtsoonian, R. and M. Teghtsoonian, (1970): Two varieties of perceived length. *Percept. Psychophys.*, **8**, 389-392.

Terhardt, E. (1968): Über akustische Rauhigkeit und Schwankungsstärke (On the acoustic roughness and fluctuation strength). *Acustica*, **20**, 210-214.

Van de Par, S. and A. Kohlrausch (2000): Sensitivity to auditory-visual asynchrony and to jitter in auditory-visual timing. In: Bernice E. Rogowitz, Thrasyvoulos N.Pappas, eds., *Human Vision and Electronic Imaging V, Proc. SPIE*, **3959**, 234-242.

Van de Par, S., A. Kohlrausch and J.F. Juola (2002): Some methodological aspects for measuring asynchrony detection in audio-visual stimuli. In: *Proc. Forum Acusticum Sevilla 2002*. Sociedad Espanola de Acustica, Hirzel.

Västfjäll, D. (2003): *Affect as a component of perceived sound and vibration quality in aircraft*. Ph.D. Thesis. Chalmers University, Göteborg.

VDI 2574 (1981): *Hinweise für die Bewertung der Innengeräusche von Kraftfahrzeugen*. Beuth Verlag, Berlin.

Verrillo, R.T. (1963): The effect of contactor area on the vibrotactile threshold. *J. Acoust. Soc. Amer.*, **35**, 1962-1966.

Verrillo, R. T. and G. A. Gescheider (1975): Enhancement and summation in the perception of two successive vibrotactile stimuli. *Percept. Psychophys*. **18**, 128–136.

Vogels, I.M.L.C. (2001): Selective attention and the perception of visual-haptic asynchrony. In: Baber, C., Faint, M., Wall, S. and Wing A.M (eds.). *Proc. EuroHaptics 2001*.

Walker, J.T. (1972): Natural visual capture in bilateral length comparisons. *Percept. Psychophys.*, **11**, 247-251.

Wallace, M.T. and B.E. Stein (1997): Development of multisensory neurons and multisensory integration in cat superior colliculus. *J. Neurosci.*, **17**, 2429-2444.

Weber, H. (1996): On the tactile senses. By Helen E. Ross and David J. Murray. Psychology Press (UK):; 2nd Ed edition.

Weber, E. H. (1851): *Die Lehre vom Tastsinn und Gemeingefühl, auf Versuche gegründet*. Braunschweig: Vieweg.

Welch, R., M. Widawski, J. Harrington and D. Warren (1979): An examination of the relationship between visual capture and prism adaptation. *Percept. Psychophys.*, **25**, 126-132.

Welch, R.B. and D.H. Warren (1980): Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, **88**, 638-667.

Wenzel, E. (1997): Analysis of the role of update rate and system latency in interactive virtual acoustic environments. 103rd AES Convention. NewYork.

Wheat, H. E. and A. W. Goodwin, (2001): Tactile discrimination of edge shape: Limits on spatial resolution imposed by parameters of the peripheral neural population. *J. Neurosci.*, **21**, 7751– 7763.

Youngblut, C., R. Johnson, S. Nash, R. Wienclaw and C. Will (1996): Review of virtual environment interface technology. Institute for Defense Analysis IDA Paper P-3186. Available: http://www.hitl.washington.edu/scivw/IDA.

Zwicker, E. and H. Fastl (1999): *Psychoacoustics: Facts and models*. Springer, Berlin, 2nd edition.

Zwislocki, J. and D. Goodman (1980): Absolute scaling of sensory magnitudes: A validation. *Percept. Psychophys.*, **28**, 28-38.

Appendix A:

Einleitung

Schall wird im Alltag gewöhnlich von Körperschwingungen erzeugt. Viele dieser Schwingungen werden von uns sowohl auditiv als auch taktil wahrgenommen. Typische Beispiele hierfür sind Lebenssituationen wie Autofahren, Bohren oder Spielen eines Instruments.

Wahrnehmung ist ein multisensorischer Vorgang, und die Integration des auditiv und taktil Wahrgenommenen ist eine fundamentale Funktion unseres Gehirns. Das Gehirn ist nahezu ständig damit beschäftigt, unterschiedliche multimodale Signalen zu verarbeiten und auszuwerten. Die Signale, die von einem bestimmten multimodalen Umweltereignisstammen werden von solchen getrennt, die von einer anderen Ereignisquelle stammen. Diese Trennung basiert auf Erfahrungen von physikalischen Ereignissen und den damit verbundenen Wahrnehmungsvorgängen. Das Verstehen eben dieser Erfahrungen und den daraus resultierenden Verhaltensreaktionen macht sich die Systementwicklung und Modellbildung zunutze. So spielt zum Beispiel das sog. haptische Feedback bei der Entwicklung von neuen multimodalen Applikationen eine zunehmende Rolle. In vielen Applikationen, z.B. bei der Konstruktion einer virtuellen Realität, eines Flugsimulators, eines Web-basierten Simulationssystems oder auch bei Anwendungen in der medizinischen Chirurgie, erfahren Benutzer auditive und taktile Information gleichzeitig. Die Entwicklung von solchen Systemen, die erwartete Wahrnehmungen hervorrufen, ist heute immer noch durch vielfältige technische Gegebenheiten eingeschränkt. So kann etwa die Prozesszeit von Rechnern eine Latenz bei der Feedbackreproduktion verursachen, es gibt Schwierigkeiten bei der Erzeugung von starkem Kraftfeedback, welches für die Simulation von starrem Kontakt benötigt wird, und es gibt nicht unerhebliche Beschränkungen, die auf Grund der mechanischen Kraftfeedback-Bandbreite auftreten. Um als multimodaler Applikations-Designer mit solchen Tatsachen und Gegebenheiten besser umgehen zu können, braucht man u.a. ein besseres Verständnis über die auditive und taktile Integration und Interaktion.

Entsprechend ist es das Ziel der vorliegenden Arbeit, die auditive und taktile Integration und Interaktion grundlegend zu untersuchen. Zuerst werden die physikalischen Faktoren und Bedingungen ermittelt, auf denen die Trennung der auditiven und taktilen Ereignisse beruht. Es zeigt sich, dass Simultanität eine der wichtigsten Faktoren ist, anhand derer der wahrnehmende Mensch entscheidet, ob zwei Signale unterschiedlicher Modalitäten von einem einzigen Ereignis oder von zwei unterschiedlichen Umweltereignissen stammen.

Neben der Simultanität spielt der sog. „spatial origin“ eine entscheidende Rolle bei der multimodalen Integration. Wenn multisensorische Signale von einem einzelnen Ereignis stammen, dann müssen ihre wahrgenommenen Ereignissorte identisch sein.

Durch Messungen von menschlichen Toleranzschwellen haben umfangreiche vorherige Studien gezeigt, wie wichtig der „Spatial-Origin“ für die audiovisuelle Integration ist. Da anzunehmen ist, dass dies auch für audiotaktile Ereignisse gilt, wird im Rahmen dieser Arbeit die Möglichkeit der Trennung der audiotaktilen Ereignisse durch „Spatial Origin“ ermittelt.

Außerdem werden zwei weitere Faktoren, nämlich Frequenz und Intensität vorgestellt, welche die Trennung der audiotaktilen Ereignisse beeinflussen können. Des Weiteren werden jeweils die physikalischen Bedingungen erfasst, auf denen die Trennung offensichtlich beruht.

Obwohl sich einige Studien mit dem Einfluss der Frequenz auf die audiotaktile Interaktion beschäftigt haben, bleibt die Frage offen, ob die Frequenz tatsächlich einen Einfluss auf die Trennbarkeit von auditiven und taktilen Ereignissen ausübt.

Werden unserem Gehirn multisensorische Ereignistypen angeboten, die von einer Ereignisquelle stammen, dann findet ein vereinigter Wahrnehmungsvorgang statt, in dessen Verlauf Verarbeitungsvorgänge stattfinden. Dabei kann man den multisensorischen Wahrnehmungsgegenstand immer auch als eine „Konstruktion“ des Gehirns verstehen, die keineswegs willkürlich ist, sondern auf einer gewichtigen Kombination von multisensorischen Signalen beruht. Während des Verlaufs der multimodalen Bildung von Wahrgenommenem findet stets eine gewisse Interaktion zwischen auditiv und taktil motivierten Ereignissen statt. So können z.B. durchaus zwei Modalitäten kombiniert werden, und der dann resultierte multimodale Wahrnehmungsgegenstand kann entweder ein schwaches, ein starkes oder ein gänzlich qualitativ neues Perzept werden. Vor diesem Hintergrund stehen folgende Fragen im Raum: Wie gewichtet unser Gehirn die Angebote, die von unterschiedlichen Sinnen stammen, um einen endgültigen Wahrnehmungsgegenstand zu bilden? Oder, mit anderen Worten: Was sind die relativen Beiträge von unterschiedlichen Sinnesmodalitäten auf ein multimodales Perzept? Kann ein Perzept, das allein auf einer Modalität beruht, durch ein gleichzeitiges Angebot einer weiteren Sinnesmodalität beeinflusst werden?

Um dies analytisch zu untersuchen, werden Tests mit Versuchspersonen durchgeführt: So werden ihnen bspw. Stimuli präsentiert, die widersprüchliche Wahrnehmungsgegenstände in zwei oder mehreren Modalitäten hervorrufen. Die Experimente, die Ergebnisauswertung und die Schlussfolgerungen daraus werden im ersten Teil dieser Arbeit detailliert geschildert.

Im zweiten Teil werden zwei grundlegende audio-taktile Interaktions-Beispiele gewählt. Dabei werden Regelhaftigkeiten bei deren Wahrnehmunguntersucht. Das Ziel eines ersten Experiments bezieht sich auf die Untersuchung des Einflusses von Lautheitswahrnehmung auf die haptische „force-feedback“ Wahrnehmung. So ist ja z.B. das Schlagen auf ein Objekt (z.B. mit der Handfläche auf eine Tischoberfläche) ein alltägliches multisensorisches Ereignis. Da dabei der physikalische Zusammenhang zwischen dem Schallereignis und dem Taktilereignis wohl bekannt ist, wird der Vorgang des Schlagens als erstes Untersuchungsbeispiel gewählt. Ein weiteres gewöhnliches multimodales Ereignis ist die Berührung einer Oberfläche mit den Fingern, die ich visuell beobachte und/oder kontrolliere. Während der Berührung der Oberfläche bekommen wir gleichzeitig taktile, auditive und visuelle Informationen über eben jenes Objekt. Dabei ist Rauigkeit eine der wichtigsten Eigenschaften der Oberfläche. Das Ziel eines zweiten Untersuchungsbeispiels ist die Ermittlung der relativen Beiträge von auditiven und taktilen Wahrnehmungsereignissen auf das Perzept der multimodalen Rauigkeit, sowie die Untersuchung des Einflusses von auditiven Wahrnehmungsereignissen auf die taktile Rauigkeitswahrnehmung.

Im Alltag erhalten wir gewöhnlich Informationen von unterschiedlichen Sinnesorganen während der Produktbenutzung, z.B. beim Fahren in einem Auto, beim Saugen eines Teppichs mit einem Staubsauger, oder wenn wir mit dem Mixer Eiweiß zu Eischnee schlagen. Daher haben kreuzmodale Informationen einen beträchtlichen Einfluss auf die Frage, wie wir bei oder nach Produktbenutzung die Qualität z.B. eines Autos, eines Staubsaugers oder eines Mixers einschätzen. Leider basieren Qualitätsuntersuchungen von Produkten heute immer noch auf unimodalen Experimenten oder Befragungen dazu.

Die Bedeutung der multimodalen Aspekte im Zuge der Produktqualitätsbeurteilung wurden bereits von Bednarzyk (1996), Blauert und Jekosch (1997), Kohlrausch und van de Par (1999), und Quehl (2001) hervorgehoben. Trotzdem fehlt eine systematische Untersuchung, die sich mit den multimodalen Aspekten der Produktqualitätsbeurteilung beschäftigt. In dieser Arbeit wird der kreuzmodale Einfluss von auditiven und taktilen Stimuli auf die gesamte Produktqualitätsbeurteilung ermittelt.

Heutzutage bekommen virtuelle Realitätsumgebungen zunehmende Bedeutung. Deshalb werden die hier beschriebenen Untersuchungen in einer virtuellen audiotaktilen Umgebung durchgeführt. Die taktilen Komponenten dieser Umgebung wurden im Verlaufe dieser Arbeit entwickelt. Die Umgebung besteht aus einem haptischen Handschuh, der an die Hand des Benutzers vibrotaktile und Kraft feedback darbietet, und einem Ganzkörperschwingungssystem. Die interaktiven Geräusche, die vom haptischen Kontakt mit virtuellen Objekten (z.B. durch Schlagen, Reiben, Streichen) hervorgerufen werden, bekommen u.a. zunehmende Bedeutung in Echtzeit multimedia Applikationen. Für die in dieser Arbeit beschriebenen Untersuchungen werden unterschiedliche Berührungsgeräusche physikalisch modelliert und synthetisiert.

Im Folgenden werden die Themen der einzelnen Kapitel schwerpunktmäßig kurz erläutert:

Das **Kapitel 2** beschreibt die physiologischen, neurophysiologischen und psychophysikalischen Grundlagen der auditiven und taktilen Modalitäten.

In **Kapitel 3** werden bestehende virtuelle auditive und taktile Displays besprochen. Anschließend wird ein in dieser Arbeit entwickeltes System zur Durchführung von Experimenten vorgestellt.

In **Kapitel 4** sind die physikalischen Faktoren zusammengetragen, die eine Rolle bei der Trennung von auditiven und taktilen Ereignissen spielen. In diesem Zusammenhang werden psychophysikalische Experimente und ihnen zugrunde liegende physikalische Bedingungen erläutert.

Kapitel 5 hat audiotaktile Interaktionen zum Hauptschwerpunkt. Zwei grundlegende audiotaktile Ereignisbeispiele und die folgenden Untersuchungen werden vorgestellt:

- Den Einfluss von Lautheit auf haptische „force feedback“ Wahrnehmung
- Der Einfluss von auditiven Ereignissen auf die taktile Rauigkeitswahrnehmung

Kapitel 6 beschäftigt sich mit den audiotaktilen Interaktionen in der Produktqualitätswahrnehmung.

Eine Zusammenfassung der Erkenntnisse findet sich in **Kapitel 7**. Anschließend werden einige Fragen für zukünftige Arbeiten vorgestellt.

LEBENSLAUF

Persönliche Daten

Geburtsdatum : 25.05.1974
Geburtsort : Eskisehir / Türkei
Staatsangehörigkeit : türkisch
Familienstand : verheiratet

Ausbildung und Arbeitsverhältnisse

1980 - 1985	Besuch der Grundschule, Kütahya, Türkei.
1985 - 1991	Besuch des Gymnasiums, Kütahya Atatürk Lisesi, Kütahya, Türkei.
1991 - 1996	Bachelor-Studium an der TU Istanbul, Fachrichtung: Maschinenbau Abschluss des Studiums und Erwerb des Grades 'B.Sc.', Istanbul, Türkei.
1996 - 1999	Master-Studium an der TU Istanbul, Erwerb des Grades 'M.Sc.' am Lehrstuhl für Maschinentheorie und Steuerung, Istanbul, Türkei.
1996 – 1999	Wissenschaftlicher Mitarbeiter am Lehrstuhl für Maschinentheorie der Fakultät für Maschinenbau der TU Istanbul, Türkei.
2000 - 2004	Wissenschaftlicher Mitarbeiter am Institut für Kommunikationsakustik, Fakultät für Elektrotechnik und Informationstechnik, Ruhr-Universität Bochum, Germany.
Seit 2004	Consulting-Ingenieur im Bereich Consulting NVH bei HEAD acoustics GmbH, Herzogenrath, Germany.

Mitgliedschaften

DEGA (Deutsche Gesellschaft für Akustik) e.V.
TAS (Türkische Gesellschaft für Akustik)
ASA (Amerikanische Gesellschaft für Akustik)