Quality of Auditory-Tactile Virtual Environments

M. Ercan Altinsoy

Chair of Communication Acoustics, Dresden University of Technology, Germany

ercan.altinsoy@tu-dresden.de

Abstract

In our daily life, we mostly perceive an event by more than one sensory modality (e.g., vision, audition, touch). Therefore multimodal integration and interaction play an important role when we use objects in our environment, e.g. for object or event recognition. Burdea and Coiffet define the virtual environment as a simulation in which computer graphics is used to create a realistic looking interactive world [1]. Virtual environments support communication with the user not only through the visual channel but also through other modalities, taking into account the multisensory nature of humans. In addition to the visual communication channel, the auditory and tactile senses are the most used communication channels. In recent years a variety of customer products which have additional tactile input and output capabilities have been developed (for example: Apple iPhone, different touch-screen applications, Wiimote, etc.). Some of these devices bring new possibilities to interact with a computer. These trends are also valid for the auditory modality. Binaural synthesis and rendering are becoming a key technology for multimedia products. “Virtual Environments” are not any more only the subject of academic research, there are also commercial applications, particularly in medicine, game and entertainment industry. This trend results that the quality of virtual environments becomes more and more important. The user interactivity of virtual environments is a key issue for the quality perception. Several studies have discussed the quality of displays, input and output devices (for different modalities) as well as software and hardware issues. However the multimodal user interaction should be taken into account in addition to the mentioned parameters to assess the quality of virtual environments. The focus of this paper is on the parameters influencing the perceptual quality of audio-tactile virtual environments.

Index Terms: Virtual environments, audio-tactile interaction, human factors.

1. Introduction

Functionalism and safety are main quality issues of virtual environments. Without doubt, they should be guaranteed for the satisfaction of the user. According to the German standard DIN 55350, the quality is defined as “physical nature of an entity with regards to its ability to fulfill predetermined and fixed requirements” [2]. Jekosch has extended this definition as *“Quality is a descriptor of the adequacy of the perceived characteristics of an entity with regard to required features…The required features are formed by the totality of the features of individual expectancies and/or social requirements and/or proper demands*” [3]. Taking into account these definitions, it is possible to say that the quality assessment of virtual environments is based on the expectations of the user from the virtual environment and the functional requirements for the specific application or task. The layers of the quality, which were described by [4,5], are also valid for the quality assessment of virtual environments. The process of quality judgment contains physical (quality elements) and perceptual (quality features) layers [6].

2. Quality Aspects

The principle architecture of an auditory-tactile virtual environment is shown in Figure 1 [7]. The classical components are the input devices, virtual reality engine, auditory and tactile drivers and output devices. The physical properties of the tracking systems are the lag (overall latency), the update rate, the interference (sensitivity to environmental factors, e.g. lighting conditions, magnetic noises, etc., the accuracy, the resolution and the working space. The virtual reality engine computes the virtual world and contains auditory and tactile renderers. Auditory and tactile renderers mostly use the same world representation.

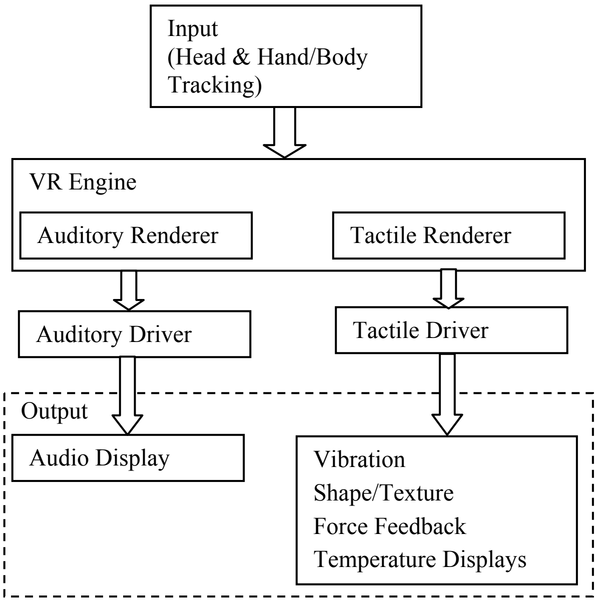

Figure 1: *The principle architecture of an auditory-tactile virtual environment (is adapted by [7]).*

From the quality point of view, an abstraction of this architecture into the components hardware (input/output devices), software (VR engine), and databases is possible [1]. The detailed architecture of the auditory part is shown in Figure 2 [8]. As input information, position of the user’s head is required for Auditory Virtual Environments (AVE). Sound source and sound field are two main modules of the architecture. The sound source signals can be obtained through recording or synthesis. A detailed overview on the sound source synthesis can be found in [9,10]. The behavior of the sound field, in which user and sound sources stay, should be physics-based or perception-based modeled. A complex model might result in more authentic reproduction, but also high computer processing time (reproduction delay). The perception-based models are based on psychoacoustical investigations (for an overview see [11]). Another important part of the auditory rendering is the reproduction-based renderer, e.g. Wave-field synthesis, Ambisonics, etc. Depending on the reproduction technique, head-related transfer functions of the user can be necessary.

The sound source signals, the reflection and directivity filters, the filter parameters of the HRTFs and the algorithms of the reproduction renderer are the input parameters of the signal processing module.

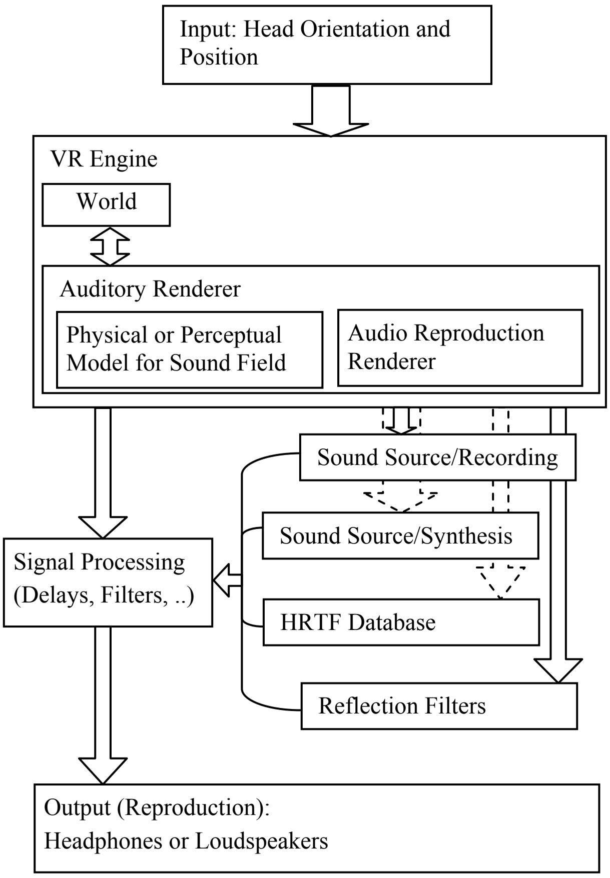

Figure 2: *The architecture of the auditory part of the virtual environment.*

The quality elements of the auditory virtual environments were defined by [11] as the frequency resolution and the bandwidth, the spatial resolution, the temporal resolution and the dynamic behavior. An alternative list was proposed by [12] as HRTFs, the number of mirror sources, the position of the mirror sources, the reflection filter and the late reverb tail generation. The quality features of the auditory virtual environments are the loudness, the auditory spaciousness, the timbre, the localization accuracy, the reverberance, the dynamic accuracy and the artefacts [12].

A number of elicitation experiments were conducted to define the quality attributes for multichannel audio, which is a part (reproduction and rendering) of the auditory virtual environment [13, 14, 15, 16]. These attributes are summarized by [17] as the localization, the source width, the envelopment, the distance to events and the depth, the space perception and the naturalness. A model, which contains the above mentioned attributes, was developed to predict multichannel audio quality [18].

Figure 3 shows the modules of the tactile virtual environment as well as the data-flow paths. The required input variables are the position and the orientation of the hand and fingers, the applied finger/hand force, temperature and in some applications, the position and orientation of the body. Similar to auditory virtual environments, the tactile interaction can be modeled physically or perceptually. Physical models are based primarily on the Newtonian physical law [18]. Depending on the interaction, various components of physical modeling, e.g. collision detection, surface deformation, grasping, texture, gravity and friction, determine the dynamic behavior of virtual objects and tactile feedback features [18]. The calculated forces, accelerations, positions and temperature are sent to the interface controller which drives the tactile interfaces. Tactile interfaces can be categorized based on the feedback type as force feedback, vibratory feedback (finger or whole-body), texture feedback and temperature feedback.

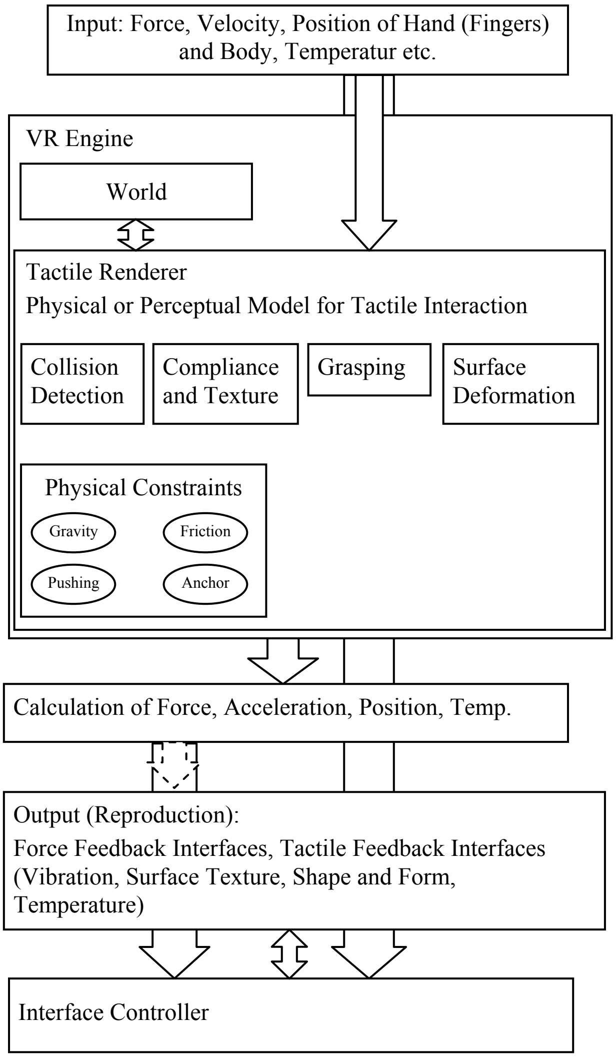

Figure 3: *The architecture of the tactile part of the virtual environment.*

The criteria for a good force-feedback interface were defined by [19] as:

- The interface should be able to simulate a piece of light salsa wood, with negligible inertia, friction or vibrations
- It should be able to simulate a crisp hard stop

- It should simulate Coulomb friction without sponginess or jitter
- It should simulate a mechanical centering detent with crisp transitions and no lag.

Similar criteria can be defined for the tactile interfaces which simulate surface textures as:

- The interface should be able to simulate a sandpaper in all possible grit numbers from 40 (coarse) to 1000 (super fine)
- It should be able to simulate very smooth surfaces, e.g. the surface of a glass or laminated wood
- It should simulate grooved surfaces in all possible groove width dimensions from daily life.

The physical parameters (quality elements), which play a role on the quality of hand vibration, whole-body vibration and texture reproduction interfaces, are the system linearity, the flat frequency response, the cross-talk between axes and the attenuation of harmonic vibration components. Although a strong cross-talk between axes causes degradation of the quality, in some cases low cross-talk can increase the feedback plausibility [20].

If we summarize the previous thoughts, important physical properties of almost all tactile interfaces, which have strong influence on the feedback quality, are the bandwith of the device, the linear frequency response, latency, resolution, maximum feedback amplitude, power-to-weight ratio and power-to-volume ratio. An optimum interface should:

- be able to simulate the human sensing and control bandwidth
- be linear
- have a flat frequency response
- have very low latency
- have Just Noticeable Difference adequate resolution
- be negligible light
- be able to simulate sufficient output
- be sufficient stiff.

Most of the studies measure the task performance to evaluate a tactile interface or system [some examples 21, 22, 23]. Task completion time and task error rate are two criteria for the quality evaluation. Recognition of an object or a texture or direct comparison of the simulated tactile feedback with tactile feedback, which is the result of the interaction with daily objects, are also used research methods [some examples 24, 25].

3. Quality Layers of Auditory-Tactile Virtual Environments

In previous section, the architecture of the auditory-tactile virtual environments and the quality elements and features of the auditory and tactile parts were presented. In this section, we discuss the quality layers of the whole auditory-tactile virtual environment.

The user's activities cause real time changes in the virtual environment. In other words, the user communicates with the virtual environment during her/his active exploration. The expectations, the experience, the motivation, the memory, the emotions, the familiarity and the attitude of the user are user dependent factors in respect to quality evaluation of auditory-tactile virtual environments (Figure 5). This is similar for quality evaluation of telecommunication services [26]. The information exchange between the user and the virtual environment should be optimal designed. This point is valid for an authentic but also for a plausible reproduction.

Depending on the virtual environment application, the references of the user vary for quality assessment. The aim of the most auditory-tactile virtual environments is to reproduce the physical behavior of a desired real environment. Educational training, e.g. machine manufacturing, etc., medical training, driving simulators, entertainment, e.g. a tennis game, virtual music instruments or virtual museums, applications use VR technology to simulate as well as possible the real world human interactions. In such kind of applications, the reference of the user is her/his daily experiences. In our daily life, sound is usually produced by the vibrations of a body. Therefore sound and vibration are coupled to each other physically. The experiences, which we did in our childhood and play role in our rest of life, are based on this physical relationship. If we interact with an object in a virtual environment, we expect a feedback to our action. Feedback belongs to the communication with the virtual environment and refers to the process of sending back information to the user about what has been done. A virtual tennis game can be a good example for auditory-tactile virtual environment applications. If the user hits a tennis ball with his tennis racket, she/he wants to hear and feel a feedback based on the ball-racket contact. The features of this feedback play an important role on the quality perception of the application. Some exemplary features are the delay between the hitting event and the sound, the changes of the sound parameters like loudness, timbre, etc. based on hitting parameters like velocity, force, etc., the locations of the hitting event and of the auditory event.

Synchronization - Perceived Synchrony

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 synchronization 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.

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 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 , L1, L2, and L3 are the latencies of the each unimodal subsystems; visual, auditory, tactile, respectively. L4, L5, and L6 are the latencies between modalities; visual-auditory, auditory-tactile, visual-tactile, respectively. The time t_{act} is the moment which the action occurs, t_{vis} is the arriving time of the visual information, t_{aud} is the arriving time of the auditory information, t_{tac} is the arriving time of the tactile information. An approximate latency for an auditory-tactile virtual environment can be estimated to be in the range of 20 ms to 40 ms.

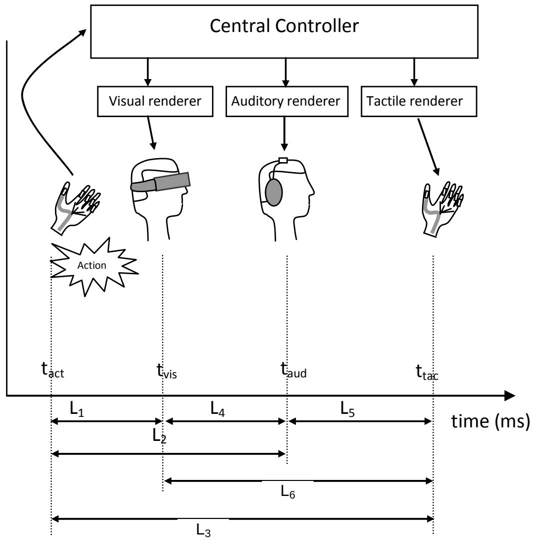

Figure 4. System latencies in virtual environments.

Several studies have discussed the perceived simultaneity of multi-modal stimuli. A multi-modal synchronisation threshold has been defined by [27] 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 (For detailed review see [9]). The obtained results vary, depending on the kind of stimuli and the psychometric methods employed.

Perceptual threshold values for auditory-tactile asynchrony were reported by Altinsoy [27]. The tactile stimulus was a sine wave and presented at the tip of the index finger of the participant via a mini electrodynamic shaker (passive tactile stimulation). 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. The auditory stimulus was a burst of white noise presented from a PC. The perceptual threshold values are 50 ms for audio lag and 25 ms for audio lead. The results of the psychophysical experiments indicate that the synchronization between auditory and tactile modalities has to be at least within an accuracy of 25 ms. Thus, the auditory-tactile delay is even more critical than the auditory-visual delay.

In most of the daily life situations, such as playing a piano or writing a text with a keyboard, we touch objects actively. This active action causes audio and tactile feedback. Levitin et al. [28] and Adelstein et al. [29] investigated the perceptual asynchrony threshold values for an active tactile interaction situation (playing a drum). The threshold value of 42 ms was reported by [28]. The thresholds vary between 18 ms to 31 ms depending on the stimulus duration [29]. The different values between both studies are possibly caused by the psychophysical measurement method.

All these studies report that some of the participants had very low threshold values (app. 10 ms) [27, 28, 29]. Particularly musicians have smaller thresholds than the mean population, possibly because of the training [9]. Therefore we suggest the synchronization requirement of 10 ms for auditory-tactile virtual environments.

The results of the investigations show that the point of subjective simultaneity (PSS) does not coincide with the point of objective simultaneity (0 ms) [27]. 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 [27]. Therefore it is possible that the human perceptual system is adapted to tolerate larger audio delays than tactile delays. The pilot experiments show that a short audio delay (between 1-7 ms) can lead to a higher perceived quality than synchronous reproduction. Similar tendencies were also found for auditory-visual perception [28].

The perceptual threshold values for auditory-whole body vibration asynchrony are 39 ms for audio lag and 35 ms for audio lead [31].

Location - Localization

Spatial origin is an important cue for humans to determine whether auditory and tactile signals originate from the same event/object or not. Naturally, if auditory, tactile and visual information have been generated by (one) same multi-modal event, the locations of the auditory and tactile events should coincide.

An investigation was conducted to investigate the minimum angle between auditory and tactile events that leads the listener to perceive that the locations of the auditory and tactile events do not coincide [32]. Scraping a surface with the finger tip, which is a very common multisensory event in our daily life, was the stimulus condition. The localization blur of scraping sound from the front was measured to be 3.9°. The minimum angle that allows the subjects to notice the locations of the auditory and tactile events do not coincide is 5.3°. The results show that humans are very sensitive to the spatial source differences.

Frequency – Pitch

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 we expect conformity between the frequency of auditory and tactile stimuli in auditory-tactile virtual environment applications.

The sensitivity of human to the frequency discrepancy between the auditory and the tactile stimuli (hand) was investigated by [33]. 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). A tactile feedback mouse was used to present the tactile information (vibrations) to the subjects. This mouse contains a motor that relates the vibration sense to the hand guiding it. 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). 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 as expected. The threshold for the frequency deviation of the tactile stimuli was about 60 % of the auditory stimuli. In most cases, subjects judge also the second or other harmonics of the vibration frequency as being suitable for the auditory frequency. The results for the sensitivity of human to the frequency discrepancy between the auditory and the whole-body vibration stimuli were very similar (app. 60 %) [34].

Intensity – Loudness & Strength

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. 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) . 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.

An investigation was conducted to measure the perceptual threshold values for the level differences, which lead to the separation of auditory and tactile events [9]. The stimuli condition was the playing of a virtual drum. Sound pressure level of the auditory stimuli and the force-feedback level of the tactile stimuli varied in the experiment. The tolerance levels are found as 17.6 dB for the level increase, and 11.2 dB 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. Improvement of the Quality of the Auditory-Tactile Virtual Environments through Audiotactile Interaction

Besides of the feedback, object or event recognition are essential for the user's interaction with the virtual environment. While in real environments the cross-modal cues are related to each other in a meaningful way, virtual environments have to assure that all modalities are fed consistently [12]. The big advantage of virtual environments is that the information in different modalities can be generated independently. Information in one modality can be used to replace or alter information that is perceived using an other modality. Audio-tactile illusions can be used in the conception of virtual environments. They can even improve the quality of audio-tactile virtual environments. The currently available interfacing technologies to human senses are insufficient and of low quality for most senses, compared to the capabilities of humans' perception. 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 [9]. Psychophysical experiments were conducted to investigate the effect of loudness on tactile force-feedback perception ("strongness") by playing a virtual drum [35]. The investigations 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 haptic perception and the magnitude of strength increases with increasing loudness in spite of no change in the force-feedback.

Similar effects were also observed for the multisensory texture exploration. 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. The perceptual consequences were studied by varying modulation frequency and loudness of the auditory stimulus [36]. 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.

5. Summary & Conclusions

The quality model for auditory-tactile virtual environments, which contains above-mentioned criteria, was presented in Figure 5. The quality judgment is strongly based on the user properties. These properties are grouped into the module user factors. The quality elements and features of the auditory and tactile parts of the virtual environment are presented in two separate modules. These elements and features are the preliminary stage of the module auditory-tactile virtual environment.

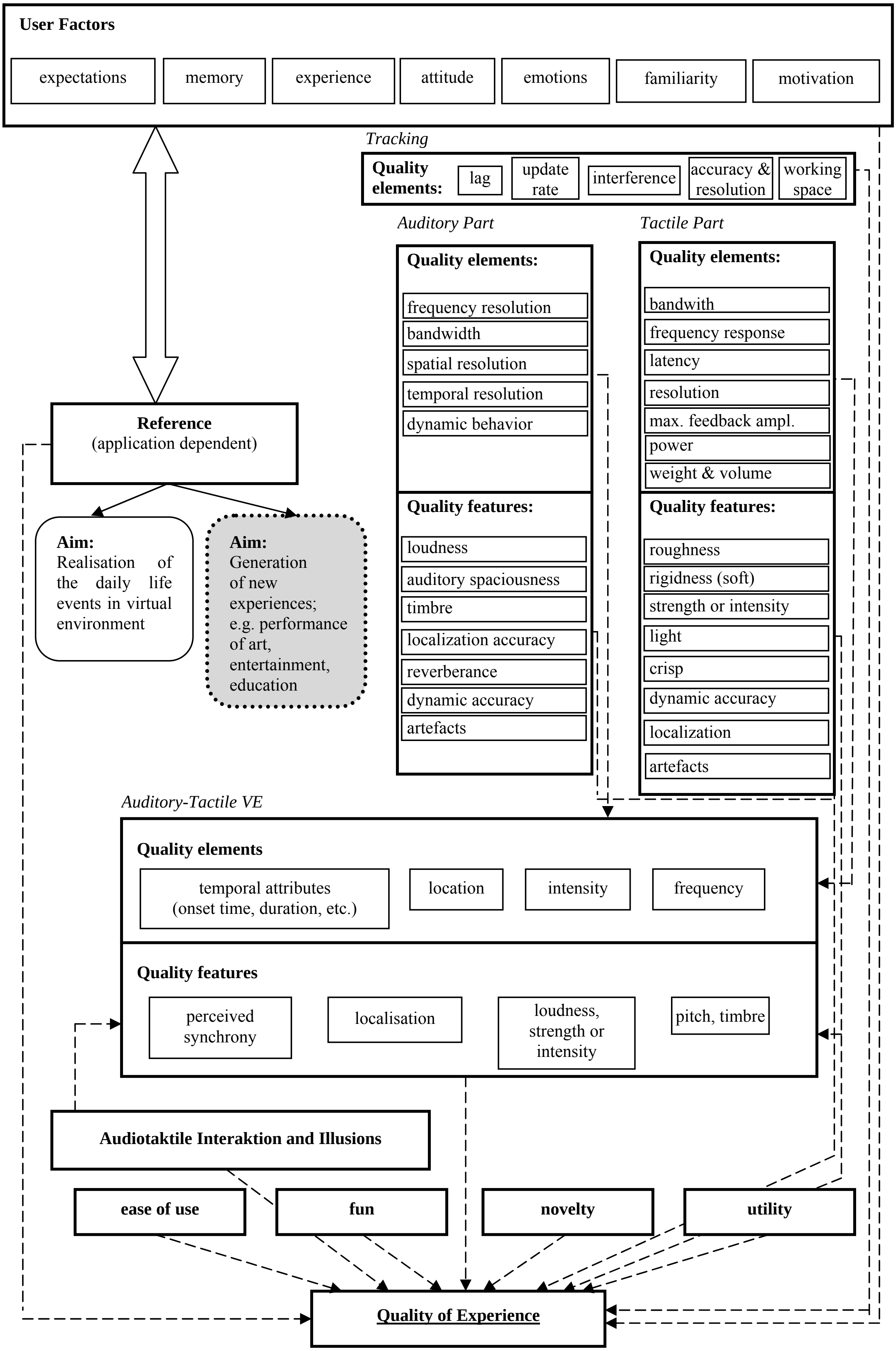

Figure 5. *The model for the quality judgment of auditory-tactile virtual environments.*

If and in which degree the factors of these modules play a role on the quality judgments is dependent on the reference of the user, in other words the aim of the auditory-tactile virtual environment. If the aim is to reproduce the physical behavior of a desired real environment, most of the elements play a pronounced role on the quality judgment. Particularly the synchronicity and the spatial origin of auditory and tactile stimuli are very important quality features. Some guidelines were given based on experimental data in Sections 4 and 5.

Auditory-tactile interaction and illusions are very promising to improve the quality of the environment. Therefore they are presented as a module.

If the aim of the virtual environment is to generate new experiences; e.g. performance of art, entertainment, education. Most of the factors in these modules do not play any pronounced role. Because the users are open for new experiences which are not conform with physical laws. Therefore the factors, such as the fun, the novelty and the ease of use, obtain an importance.

6. References

[1] Burdea, G., and Coiffet, P. (1994) "Virtual Reality Technology", John Wiley & Sons, New York NY.

[2] DIN 55350: Begriffe der Qualitätssicherung und Statistik. Part 12, Merkmalsbezogene Begriffe. Beuth-Vertrieb, Berlin, 1989.

[3] Jekosch, U. (1999) "Meaning in the context of sound-quality assessment" ACUSTICA united with ACTA ACUSTICA 85, 681-684.

[4] Blauert, J. & Jekosch, U. (1996) "Sound-quality evaluation − a multi-layered problem" ACUSTICA united with ACTA ACUSTICA 83, 747−753

[5] Blauert, J. & Jekosch, U. (2010) "A Layer Model of Sound Quality" in Proc. of Third International Perceptual Quality of Systems Workshop, Bautzen, Germany

[6] Jekosch, U. (2004) "Basic Concepts and Terms of "Quality", Reconsidered in the Context of Product-Sound Quality", ACUSTICA united with ACTA ACUSTICA 90, 999-1006

[7] 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

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

[9] Altinsoy, E. (2006) "Auditory-Tactile Interaction in Virtual Environments" Shaker Verlag, Germany

[10] Cook, P. (2002) "Real sound syhnthesis for interactive applications" A K Peters Ltd. Natick, MA.

[11] Pellegrini, R. S. (2002) "A Virtual Reference Listening Room as an Application of Auditory Virtual Environments" dissertation.de, ISBN 3-89825-403-8.

[12] Silzle, A. (2008) "Generation of Quality Taxonomies for Auditory Virtual Environments by Means of Systematic Expert Survey" Shaker Verlag, Germany

[13] Berg, J., Rumsey, F. (2006) "Identification of Quality Attributes of Spatial Audio by Repertory Grid Technique" Journal of Audio Eng. Soc. Volume 54 Issue 5 pp. 365-379

[14] Zacharov, N., Koivuniemi, K. (2001) "Unravelling the Perception of Spatial Sound Reproduction: Language Development, Verbal Protocol Analysis and Listener Training" Presented at 111th AES Convention at New York, September 2001 Preprint number 5424

[15] Choisel, S., Wickelmaier, F. (2005) "Extraction of Auditory Features and Elicitation of Attributes for the Assessment of Multichannel Reproduced Sound" 118th AES international convention, Barcelona, Spain, Preprint Number: 6369

[16] Bech, S. and Zacharov, N. (2006) "Perceptual Audio Evaluation - Theory, Method and Application" GB−Chichester, Wiley

[17] George, S. (2009) "Objective models for predicting selected multichannel audio quality attributes" Ph.D. Thesis, University of Surrey.

[18] Burdea, G. C. (1996) "Force and touch feedback for virtual reality". John Wiley and Sons, Inc, New York

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

[20] Rosenthal, M. (2009) "Reflex™ Technology: Actuators Capable of Direct Control of Haptics From Audio Signals" in Proceedings of Fourth International Workshop on Haptic and Audio Interaction Design, HAID 2009, Dresden, Germany

[21] Kontarinis, D., and Howe, R. (1994) "Tactile Display of Vibratory Information in Teleoperation and Virtual Environments,"Presence, Vol. 4., No. 4, pp. 387-402.

[22] Power, C. (2006) "On the Accuracy of Tactile Displays" in Miesenberger, K. et al. (Eds.): ICCHP 2006, LNCS 4061, pp. 1155–1162. Berlin, Germany. Springer-Verlag

[23] Altinsoy, M.E. and Merchel, S. (2009) "Audiotactile Feedback Design for Touch Screens", in Altinsoy, M. E., Jekosch, U. & Brewster, S. (Eds.), Haptic and Audio Interaction Design 2009, LNCS 5763, pp. 136-144. Berlin, Germany. Springer

[24] Kirkpatrick, A.E. and Douglas S.A. (2001) "A Shape Recognition Benchmark for Evaluating Usability of a Haptic Environment" Haptic Human-Computer Interaction, LNCS 2058, pp. 151-156. Berlin, Germany. Springer

[25] Stamm, M., Altinsoy, M.E. and Merchel, S. (2010) "Idenification Accuracy and Efficiency of Haptic Virtual Objects Using Force-Feedback" in Proc. of Third International Perceptual Quality of Systems Workshop, Bautzen, Germany

[26] Möller, S. (2005) "Quality of Transmitted Speech for Humans and Machines", Communication Acoustics. J. Blauert (ed). D-Heidelberg-New York NY, Springer

[27] Altinsoy, M.E. (2003) "Perceptual aspects of auditory-tactile asynchrony," in Proceedings of the Tenth International Congress on Sound and Vibration, International Institute of Sound and Vibration, Stockholm, Schweden, pp. 3831-3838

[28] Levitin, D.J., MacLean, K., Mathews, M., Chu L, and Jensen, E. (2000) "The perception of cross-modal simultaneity" in Proc. of Computing Anticipatory Systems. AIP Conf. Proc. 517, pp. 323-329

[29] B. D. Adelstein, D. Begault, M. Anderson, and Wenzel. E. (2003) "Sensitivity to haptic-audio asynchrony" in Proc. Int. Conf. on Multimodal Interfaces, New York, NY, USA, pp. 73-76

[30] Kohlrausch, A. & van der Paar, St (2005) "Audio-visual interaction in the context of multi-media applications". In: Blauert, J. (ed.) Communication Acoustics, Springer, Berlin

[31] Altinsoy, M.E., Blauert, J., and Treier, C. (2001) "Inter- Modal Effects of Non-Simultaneous Stimulus Presentation", in A. Alippi (Ed.), Proceedings of the 17 th International Congress on Acoustics. Rome, Italy

[32] Altinsoy, M.E. (2010) "The Effect of Spatial Disparity on the Integration of Auditory and Tactile Information" in Nordahl, R., Fontana, F., Serafin, S. & Brewster, S. (Eds.), Haptic and Audio Interaction Design 2010, LNCS. Berlin, Germany. Springer

[33] Altinsoy, M.E. (2004) "The influence of frequency on the integration of auditory and tactile information," in Proc. of the 18th International Congress on Acoustics (ICA), Kyoto, Japan

[34] Altinsoy, M.E. and Merchel, S. (2010) "Cross-Modal Frequency Matching: Sound and Wholebody vibration" in Nordahl, R., Fontana, F., Serafin, S. & Brewster, S. (Eds.), Haptic and Audio Interaction Design 2010, LNCS. Berlin, Germany. Springer

[35] Altinsoy, M. E., Blauert, J., and So, R.H.Y. (2003) "Effect of loudness on the haptic force-feedback perception in virtual environments," J. Acoust. Soc. Am. 114, 2330-2331

[36] Altinsoy, E. (2008) "The Effect of Auditory Cues on the Audiotactile Roughness Perception: Modulation Frequency and Sound Pressure Level" in Haptic and Audio Interaction Design, edited by Antti Pirhonen, Stephen A. Brewster. Lecture Notes in Computer Science 5270, Berlin, Germany: Springer