EuroHaptics 2010

PROCEEDINGS OF THE EUROHAPTICS 2010
SPECIAL SYMPOSIUM:

HAPTIC AND AUDIO-VISUAL STIMULI:
ENHANCING EXPERIENCES AND INTERACTION

Amsterdam, July 7, 2010

Anton Nijholt, Esko O. Dijk, Paul M.C. Lemmens, Steven Luitjens (eds.)

CIP GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Nijholt, A., Dijk, E.O., Lemmens, P.M.C., Luitjens, S.

Special Symposium at Conference EuroHaptics 2010
Proceedings of Special Symposium at EuroHaptics 2010
Haptic and Audio-Visual Stimuli: Enhancing Experiences and Interaction
A. Nijholt, E.O. Dijk, P.M.C. Lemmens, S. Luitjens (eds.)
Amsterdam, Universiteit Twente, Faculteit Elektrotechniek, Wiskunde en Informatica

ISSN 0929–0672
CTIT Workshop Proceedings Series WP10-01

trefwoorden: Audio, video and haptic stimulation; Influence of temporal and spatial patterns of (haptic) stimuli, Relaxation, easing the mind, comforting; Applications of biofeedback; Intelligence and algorithms to optimize the user experience.

Book orders:
Ms. C. Bijron
University of Twente
Faculty of Electrical Engineering, Mathematics and Computer Science
P.O. Box 217
NL 7500 AE Enschede
tel: +31 53 4893740
fax: +31 53 4893503
Email: bijron@cs.utwente.nl

Druk- en bindwerk: Ipskamp Drukkers, Enschede.

Preface

Abstract

The intention of the symposium on Haptic and Audio-visual stimuli at the EuroHaptics 2010 conference is to deepen the understanding of the effect of combined Haptic and Audio-visual stimuli. The knowledge gained will be used to enhance experiences and interactions in daily life. To this end, a number of key lectures have been organized and accompanying papers can be found in this proceedings. With the lectures and the interaction between the researchers at the symposium, we aim to initiate a community interested in multimodal stimulation involving haptic elements with an emphasis on experiences for entertainment, well-being and relaxation.

Background

Multimodal stimulation is capable of creating strong effects on users, because the effects of the various stimuli can enforce each other. It can be used to enhance entertainment experiences, as well as well-being and relaxation experiences. By default, multimodal stimulation often only considers visual with auditory stimulation, because these senses are most prominent in our environment. However, humans have at least three more senses that can be used to create multimodal sensations: touch, taste, and smell. The latter two are technologically difficult to implement but stimulation and feedback using the tactile sense is rapidly becoming more prevalent. An example application is to use haptic and tactile actuator elements to provide the player of a game with a more thrilling experience. In this case, tactile stimulation is provided that is linked to the visual and auditory information in the game and together, these stimuli create very strong experiences.

A deep understanding of the requirements to create a convincing multimodal experience is needed to create an experience that is more than just the sum of the elements. This understanding is needed on the level of individual sensory modalities but also on the interactions of information processing in each modality and covers aspects of the relative contribution of individual modalities, aspects of timing and synchronization, et cetera. This especially applies to the tactile modality that is relatively unexplored in the context of multimodal stimulation. Work on these topics is being carried out for separate modalities [1-3] and also for multimodal stimulation [4] and covers topics as diverse as intensity, spatial distribution, timing, tactile perception [1,5], tactile displays [6], et cetera. However, the effects multimodal stimulation including haptic elements on the user experiences and interactions has not yet been thoroughly studied in the context of entertainment, well-being, and relaxation applications.

About This Symposium

In this special symposium we address the specific effects of combined (multi-sensory) stimuli that aim to achieve total effects that are more than just the sum of their elements. Topics range from basic elements such as mutual timing in audio, video, and haptic stimuli, through actuator technologies, to how such "more than the sum of the elements" effects of multimodal stimuli are created in a user’s perception and how to evaluate these experiences and perceptions.

Our guiding hypothesis is that an optimal user experience will be obtained by taking into account human perception, careful personalization, and intelligent optimization. The latter should be based on both general knowledge of human perception, and on (measured or inferred) knowledge of the individual user. Research on human perception will provide information on the basic capabilities and limitations of individual modalities but also on how combined information processing in multiple modalities operates. To this end we have planned a number of key lectures on the technologies employed, the psychological and physiological sensitivities of people and the algorithms used to optimize the effect of multimodal stimuli. We have been able to invite researchers working on the following topics:

- Haptic illusions
- Design
- Relaxation using haptic stimulation
- Mediated social touch
- Audiotactile interaction
- Personalized tactile feedback
- Tactile stimulation for entertainment

These presentations and the interaction between researchers could initiate a community of researchers who are interested in multimodal stimulation involving haptic elements with particular emphasis on experiences in entertainment, well-being, and relaxation. In these proceedings of the symposium you can find the contributions of most of the key speakers. Short summaries of these contributions follow below.

These proceedings start with a (preliminary) position paper *("Audio-tactile Stimuli to Improve Health and Well-being")* by Esko Dijk and his co-authors. The paper aims at defining a research area where auditory and tactile stimulation, possibly enhanced with visual information and stimuli, is combined and applied to improve people's health and well-being. It is argued that these combined stimuli can have effects on the human body and mind by, for example, reducing stress, improving alertness or promoting sleep. Presently there is a variety of low-cost and miniature tactile actuators on the market. They find application in mobile phones, but also in jackets that provide dynamic and spatial tactile patterns on the human body. Audio-tactile patterns can be designed for many applications, for example, for navigation, for entertainment or for health and well-being purposes. The paper briefly surveys research results on audio-tactile stimuli, available technology, and audio-tactile composition. Scientific challenges are identified that need to be explored in order to design personalized audio-tactile systems that adapt to their users either off-line or online. The aim of this position paper is to create a research community for answering these challenges.

Clearly, before being able to interpret the effect of audio-tactile stimuli it is necessary to know the effect of uni-modal stimuli. For example, how can tactile stimuli induce emotions? In *"Tactile Experiences"* Paul Lemmens and his co-authors take William James' viewpoint that every emotion has a distinct bodily reaction. They reversed this observation and studied whether providing bodily stimuli while watching appropriate video clips could induce or enhance an emotion. The design of an emotion jacket is described. The jacket provides tactile sensations on the torso with the help of sixty-four actuators (eccentric rotating-mass motors) embedded in stretchable fabric. Various tactile emotion patterns were designed for video clips that were chosen to elicit certain emotional responses. In a user study participants viewed the clips with and without the emotion patterns projected onto their bodies. Questionnaires were used and psychophysiological responses were recorded in order to obtain information about the emotional experience and immersion. The results convinced the authors that adding the tactile emotion patterns enhanced the emotional experience of the viewers.

Hendrik Richter's contribution (*"Multi-Haptics and Personalized Tactile Feedback on Interactive Surfaces"*) builds further on the recent trends of using haptic feedback for touch screen interaction. In this application area, the touch and visual senses come together. While current systems can mostly provide haptic feedback for only a single point of interaction (i.e. finger), he proposes a first extension to multi-touch surfaces. A second extension is also proposed to take away one important restriction of current solutions, namely that the haptic feedback is always given at the location of interaction on the screen. It is proposed to spatially disunite the body-part of interaction (finger, hand) and the resulting tactile feedback, potentially leading to completely new touch screen interaction paradigms using haptics. Firstly, feedback can be given at multiple body locations and/or using multiple actuation means (an approach called multi-haptics) and secondly the haptic feedback can be personalized to each user in collaborative scenarios where multiple users are interacting on the same touch surface. Three prototypes that have been used for initial explorations in this domain are described.

Valeria Occelli (*"Assessing Audiotactile Interactions: Spatiotemporal Factors and Role of Visual Experience"*) provides a well founded overview of her work on the interaction of hearing and touch; an interaction that happens often in daily life but that has received relatively little attention in scientific literature. She has studied monkeys, patients with brain damage, blind people, and a non-patient population and shows that the location at which crossmodal audio-tactile stimulations are presented strongly influences how much attention is given to the stimulus. Locations directly behind the head attract most attention. These stimuli in peri-personal space attract less attention and, with increasing distance from the body, the number of resources allocated to the stimuli also decreases. Moreover, Occelli shows that certain types of sound interact with the effects of spatial location: pure tones have different effects than white noise has.

Antal Haans and Wijnand IJsselsteijn (*"Combining mediated social touch with vision: From self-attribution to telepresence?"*) investigate the topic of mediated social touch, id est interpersonal touch over a distance by means of tactile display technology. They investigate combining mediated touch with vision, allowing people simultaneously to both feel and see how a remote partner is touching them. Adding another sensory modality (in this case vision) for the person receiving the touches, can potentially increase a user's sense of "being in the same environment" with the remote partner. The paper confirms this effect and also shows that adding vision can increase the perceived naturalness of the mediated touches. This serves as a good example of how perceived quality in one modality can be increased by adding congruent stimuli in another modality. The experimental findings illustrate that visual feedback, especially when the visual shows a resemblance to a human body being touched, can improve mediated social touch. As such, the results of this work could be seen as ingredients for future systems that improve people's well-being, by facilitating closer contact with loved ones even though they may be far away.

In the paper *"'Breathe with the Ocean': A System for Relaxation using Combined Audio and Haptic Stimuli"* by Esko Dijk and Alina Weffers, a breathing guidance system is introduced that uses audio, haptic, and visual stimuli and that was created for the purpose of relaxing a user. The authors provide evidence from the literature that audio stimuli (in particular music), haptic stimuli (in the forms of vibrations), and visual stimuli can induce relaxation. The breathing guidance system makes use of a Touch Blanket, an actuation device developed by Philips that can provide haptic patterns on body parts. The blanket contains 176 small vibration motors arranged in a 2D matrix. 'Haptic waves' synchronized with audio can move up and down the body in various cycles. These cycles can be fixed (e.g., taking a rate similar to a breathing pace that is associated with relaxation), they can follow the breathing behavior of the user or they can guide the user to an optimal breathing behavior taking into account respiration and heart rate. Results of a first evaluation of these approaches are shown.

Stefania Serafin et al., in *"Identification of virtual grounds using virtual reality haptic shoes and sound synthesis"*, report on an experiment using a combination of haptic and auditory stimuli to simulate the sensation of walking on different kinds of surfaces, for example beach sand, gravel, metal etc.. Haptic stimulation was provided by actuators mounted in the soles of shoes. Audio stimulation was generated by using physical models of walking combined with sounds recorded during walking on various kinds of surfaces. Both stimuli were coupled to the physical action of walking by sensors in the soles of the shoes. The aim of this interesting experiment was to find out about the enhancement of the sensation of walking by adding the haptic feedback to the auditory one. From the user tests it appeared that the main role in creating the sensation and recognizing the kind of surface was the auditory stimulation. Although in some cases haptic stimulation added significantly.

Finally, Saskia Bakker et al. (*"Design for the Periphery"*) work on the topic of designing for the periphery which revolves around design technology interactions in such a way that only peripheral attention is needed to process and carry out these interactions. As a foundation for their work, they discuss the notion of calm technology, and attention theory as developed in psychological literature. Calm technology is technology that works in the background, not demanding our attention, and that can be attended by peripheral attention. Bakker et al. propose that interaction design for calm technology should be guided by principles from psychological theories of attention such that the interaction with the technology can be done without requiring major attentional effort. Because humans effortlessly interact with tangible objects, the haptic modality seems a candidate with a lot of potential for interaction design in the periphery.

During the symposium some presentations were given that could not be included in these proceedings. George VanDoorn gave a talk entitled "Haptics Can Lend a Hand to a Bionic Eye." and Maud Marchal gave a talk on "Pseudo-Haptics". In addition to the oral presentations there were demonstrations of tactile vests by Sense-Company, Tilburg, in the Netherlands, and by the Hogeschool voor de Kunsten, Utrecht, in the Netherlands.

Acknowledgements

We are grateful to the EuroHaptics 2010 organizers for allowing us to organize this special symposium under the flag of the EuroHaptics conference. A special word of thanks goes to Hendri Hondorp who took on the role of technical editor of these proceedings.

References

[1] G.A. Gescheider, J.H. Wright, and R.T. Verrillo, Information-Processing Channels in the Tactile Sensory System, New York: Psychology Press, 2009.

[2] D. Marr, Vision: a computational investigation into the human representation and processing of visual information, Cambridge MA: MIT Press, 2010.

[3] E. Zwicker and H. Fastl, Psychoacoustics: Facts and Models, Berlin: Springer-Verlag,1990.

[4] C. Spence, J. Ranson, and J. Driver, "Cross-modal Selective Attention: On the Difficulty of Ignoring Sounds at the Locus of Visual Attention," Perception & Psychophysics, vol. 62, pp. 410-424, 2000.

[5] R.W. Cholewiak, "The Perception of Tactile Distance: Influences of Body Site, Space, and Time," Perception, vol. 28, pp. 851-875.

[6] J.B.F. Van Erp, "Tactile displays for navigation and orientation: perception and behavior", Ph.D. thesis, Utrecht University, 2007.

Anton Nijholt, Esko O. Dijk, Paul M.C. Lemmens & Steven Luitjens   Enschede/Eindhoven, July 2010

Committees

Program and Organizing Committee

Anton Nijholt	Human Media Interaction, University of Twente, The Netherlands
Esko O. Dijk	Philips Research Eindhoven, The Netherlands
Paul M.C. Lemmens	Philips Research Eindhoven, The Netherlands
Steven Luitjens	Philips Research Eindhoven, The Netherlands

Scientific Advisers

Dirk Brokken	Philips Research Eindhoven, The Netherlands
Jan van Erp	TNO Human Factors, Soesterberg, The Netherlands

Administrative/Technical/Financial Support

Hendri Hondorp	Human Media Interaction, University of Twente, The Netherlands
Lynn Packwood	Human Media Interaction, University of Twente, The Netherlands
Charlotte Bijron	Human Media Interaction, University of Twente, The Netherlands

Programme

09.30	Opening
09.45	Design for the Periphery
	Saskia Bakker, Elise van den Hoven, Berry Eggen
	Eindhoven University of Technology, The Netherlands
10.15	Pseudo-Haptics (preliminary title)
	Maud Marchal, Anatole Lécuyer
	INRIA, Rennes Cedex, France
10.45	Break
11.00	Combining Mediated Social Touch with Vision: From Self-attribution to Telepresence?
	Antal Haans, Wijnand A. IJsselsteijn
	Eindhoven University of Technology, The Netherlands
11.30	Haptics Can Lend a Hand to a Bionic Eye
	George VanDoorn, Barry Richardson
	Monash University, Churchill, Australia
12.00	Multi-Haptics and Personalized Tactile Feedback on Interactive Surfaces
	Hendrik Richter
	University of Munich, Germany
12.30	Lunch
13.45	Assessing Audiotactile Interactions: Spatiotemporal Factors and Role of Visual Experience
	Valeria Occelli
	University of Trento, Italy
14.15	Demonstrations and Talks by Sense-Company and HKU
	Ewoud Kuyper, Sense-Company, Tilburg, The Netherlands
	Gerard van Wolferen, Hogeschool voor de Kunsten, Utrecht
15.00	Break
15.30	Tactile Experiences
	Paul M.C. Lemmens, Dirk Brokken, Floris M.H. Crompvoets, Jack van den Eerenbeemd, Gert-Jan de Vries
	Philips Research, Eindhoven, The Netherlands
16.00	’Breathe with the Ocean’: A System for Relaxation using Combined Audio and Haptic Stimuli
	Esko Dijk, Alina Weffers-Albu
	Philips Research, Eindhoven, The Netherlands
16.30	Identification of virtual grounds using virtual reality haptic shoes and sound synthesis
	Stefania Serafin, Luca Turchet, Rolf Nordahl, Smilen Dimitrov
	Medialogy, Aalborg University, Copenhagen, Denmark
17.15	Discussion, Conclusions, Future
17.30	Closing

Contents

Audio-tactile stimuli to improve health and well-being

A preliminary position paper

Esko O. Dijk[a] Anton Nijholt[b] Jan B.F. van Erp[c]
Ewoud Kuyper[d] Gerard van Wolferen[e]

[a] Philips Research, 34, Eindhoven, The Netherlands
[b] University of Twente, Enschede, The Netherlands
[c] TNO Human Factors, Soesterberg, The Netherlands
[d] Sense Company BV, Tilburg, The Netherlands
[e] Utrecht School of the Arts (HKU), Hilversum, The Netherlands

Abstract

From literature and through common experience it is known that stimulation of the tactile (touch) sense or auditory (hearing) sense can be used to improve people's health and well-being. For example, to make people relax, feel better, sleep better or feel comforted. In this position paper we propose the concept of combined auditory-tactile stimulation and argue that it potentially has positive effects on human health and well-being through influencing a user's body and mental state. Such effects have, to date, not yet been fully explored in scientific research. The current relevant state of the art is briefly addressed and its limitations are indicated. Based on this, a vision is presented of how auditory-tactile stimulation could be used in healthcare and various other application domains. Three interesting research challenges in this field are identified: 1) identifying relevant mechanisms of human perception of combined auditory-tactile stimuli; 2) finding methods for automatic conversions between audio and tactile content; 3) using measurement and analysis of human bio-signals and behavior to adapt the stimulation in an optimal way to the user. Ideas and possible routes to address these challenges are presented.

1 Introduction

1.1 Improving health and well-being through touch and hearing

People perceive the world through their senses: sight, smell, taste, hearing and touch. The tactile (touch) sense is an important one: it is in fact the first sense to develop in the womb. Touch can give people strong emotional experiences [1] and is vital for health and well-being [1]. Tactile stimulation or *somatosensory* stimulation, applying touch to the human body, is often used as a way to make people feel better or to reduce stress. Examples range from basic touch to comfort someone, massaging techniques, whole-body vibration training and physiotherapy to alternative treatments such as acupressure, Reiki and vibro-acoustic therapy.

Besides touch, the sense of hearing is also used to make people feel better: one can listen to spoken encouragements, relaxing music, or nature sounds to sleep better. Music therapy [2,3] is an established practice and has been extensively investigated in the scientific community.

The scientific literature shows evidence that specific methods of stimulation of the auditory (hearing) or tactile senses can indeed effectively reduce stress and muscle tension, increase well-being, or promote sleep. Furthermore, there are indications [4-7] that stimulating the two senses of hearing and touch *at the same time* can have stronger effects on the human body and mind than stimulating only one of these senses at a time. Hence a promising area of scientific research is the use of a combination of sound heard and touch felt by a user at the same time to influence the user's body and mind in a positive way.

1.2 Scope

The research area discussed in this paper we refer to as *combined auditory-tactile stimulation and its effects on human health, well-being, body state and mental state*. However, little scientific work has been done so far in this field. The aim of this paper is to present our vision on this research field of auditory-tactile stimulation, and to present research challenges and opportunities that we have identified.

Although we often refer to the term *health* as a goal of the systems we investigate, we do not mean to replace established treatment methods with new ones. Instead, in healthcare contexts the goal of our approach is to augment the existing care and treatment methods where possible by stimulating well-being, relaxation or sleep.

1.3 Example applications

1.3.1 Healthcare

One particular use case is a small relaxation room in a care institute where a user can sit in a comfortable chair with their eyes closed. Light, music and sounds are played in the room, and the user feels gentle taps on the body and calming oscillations. The chair senses how the individual user reacts to these stimuli in real-time. During a session the stimuli are composed by an intelligent system in such a way that the combined effect is optimally relaxing for this user. Maybe one user prefers taps, while another prefers gentle vibrations. And each user may have a personal level of intensity and patterns that he/she likes best.

A similar use case could be envisioned for people with autism, analogous to a multisensory environment investigated earlier in the MEDIATE project [8].

1.3.2 Home

Another use case example is in the home (consumer) environment. Imagine a user at home, who wants to relax after a busy day at work. He owns a multisensory relaxation/entertainment system that consists of a blanket with integrated tactile actuators (e.g. [5,6]) and headphones. The system provides a combination of sounds, music and tactile stimulation that is designed to relax. After a session of 20 minutes, the user feels much more relaxed than before.

1.3.3 Public transport

In public transport, it is vital that train drivers are alert during their work shift. However, the working hours in this profession are often irregular, inducing the risk of decreased alertness at times when it is most needed. The largest Dutch railway company NS has already experimented [9] with special power-nap relaxation rooms, which have the multi-sensory stimulation product AlphaSphere [7] installed. The goal is to enable personnel such as train drivers to have a quick, 25-minute rest e.g. during their break, in order to increase alertness during their work.

1.4 Structure of the paper

To be able to clearly outline our vision and the research challenges ahead in Section 3, we first provide an overview in Section 2 of the current relevant state of the art and its limitations. Section 4 ends with discussion and conclusions.

2 Current state of the art

The present section does not aim to be a complete overview or review of the state of the art. Rather, we briefly sketch the research and application fields that are considered relevant, with the help of a few key references. We expect that this Special Symposium at EuroHaptics 2010, *Haptic and Audio-Visual*

Stimuli: Enhancing Experiences and Interaction, or possible follow-up events will contribute to a more complete overview and hence an improved vision for the future of auditory-tactile stimulation.

2.1 Stimulating the sense of touch

Stimulating the tactile sense can give people strong emotional experiences and is vital for health and well-being [1]. Interpersonal touch is known to be an important element of human love and social bonding. Tactile stimulation is used today in methods to reduce stress or muscle tension, train the body, or to make people feel good, feel cared for, happy, energized, sleep better or simply more relaxed [10-14]. There are studies on the subjective pleasantness of touch [14] and studies on the mental, health-related and bodily effects of low-frequency vibration [10-13,15].

These methods can involve a human performing the stimulation, a machine, or a human helped by a machine. Of course it cannot be expected that touch by a machine, in general, will have as similar effect to touch applied by a human. But on the other hand, the properties of machine-generated or machine-mediated touch are still being actively researched. A recent result [16] suggests that the effect size of machine-produced touch in a specific experimental situation could be similar to that of touch performed by a human, although more research would be needed to substantiate such hypotheses.

2.2 Touch actuators

To fully understand the opportunities in the field of auditory-tactile stimulation, it is helpful to look at tactile actuation (i.e. touch stimulation) technology. In recent years, advances in actuators and embedded computing have enabled a wide range of machine-driven methods for tactile stimulation. The strong growth of haptic (tactile feedback) technology in the mobile phone market has brought a variety of small mechanical actuators onto the market. Such actuators are used in for example jackets ([4] or Figure 1) that can stimulate different points on the upper body or a blanket [5] that can provide tactile stimuli to the whole body. Miniature actuators can be combined [17] with larger actuators, which enables interesting compositions of effects. Today, a large variety of tactile effects can be achieved relatively easily, at low cost and be suitable for daily use situations.

The types of tactile actuators that we currently consider for our purposes are:

1. Miniature ERM (Eccentric Rotating Mass) vibration motors, used in many cell phones. These do not offer precise independent control of frequency or amplitude of tactile effects. Used in [4,5].
2. Small tactile transducers, capable of playing effects with precise frequency/amplitude control. Used in some cell phones and in [17].
3. Larger tactile transducers, used in certain home cinema products and theme parks for powerful bass effects (called "rumblers" or "shakers"). Also used in [17].
4. Common bass loudspeakers, sometimes used as an alternative to option 3 above. Used in [7].
5. Actuator systems for providing mechanical displacement or pressure on the body. For example solenoids, rotary driven pistons or pneumatic/hydraulic actuators. Motion is used by [18].

See Figure 1 (left side) for an example product: the *Feel The Music Suit* created by Sense Company [sense-company.nl] with the Utrecht School of Arts [hku.nl] and TNO [tno.nl].

2.3 Stimulating the sense of hearing

Like touch, the human sense of hearing is often used in methods for health and well-being. Examples are relaxation music, nature sounds, or self-help audio guides. In the literature, the effects of music and therapeutic use of music have been well investigated (e.g. [2,3]). Various audio products exist for well-being and mental state influencing, including so-called brainwave entrainment methods such as *binaural beats* or *isochronic tones*.

Figure 1: Dutch minister of Economic Affairs wearing a Feel The Music Suit at a public event (top-left). On the bottom left, design sketches of the suit. On the right, a demonstration of the TNO tactile dance suit showing coordinated dance movements by three users.

A good example of innovative audio content with well-being application is *Meditainment* [19] (= Meditation + Entertainment), which combines relaxing music, ambient soundscapes, nature sounds, voice coaching and guided meditation and visualisation techniques. One particular use of this content which is reportedly being investigated is pain management for hospital patients. Typically, the audio content in existing products such as Meditainment is static, id est not interacting with the user nor automatically adapting to the user. One of our hypotheses is that making this content more adaptive to the user could make audio stimulation much more effective and attractive.

2.4 Combined stimulation of touch and hearing: *auditory-tactile stimulation*

An interesting concept is to combine stimulation of the sense of hearing with the sense of touch. If each one alone can have positive effects, can the combination be even more effective or more enjoyable? See Figure 2 for an impression of this stimulation approach. As an example, the Meditainment audio content presented in the previous section does not include tactile stimuli – could tactile stimuli significantly increase the effectiveness of such content? Next, we look at what the scientific literature tells us about the health and well-being effects of combined stimulation.

Some work has been done on a specific method of combined auditory-tactile stimulation called vibro-acoustics [2,4,5]. Here, a tactile (vibration) effect is directly derived from the lower frequencies of music and played by one or two tactile actuators. Experimental results suggest that this combination may work well for relaxation or sleep. An experiment [15] with playing the didgeridoo, which evokes vibrations in the upper body along with sounds, shows promise as a specific medical treatment.

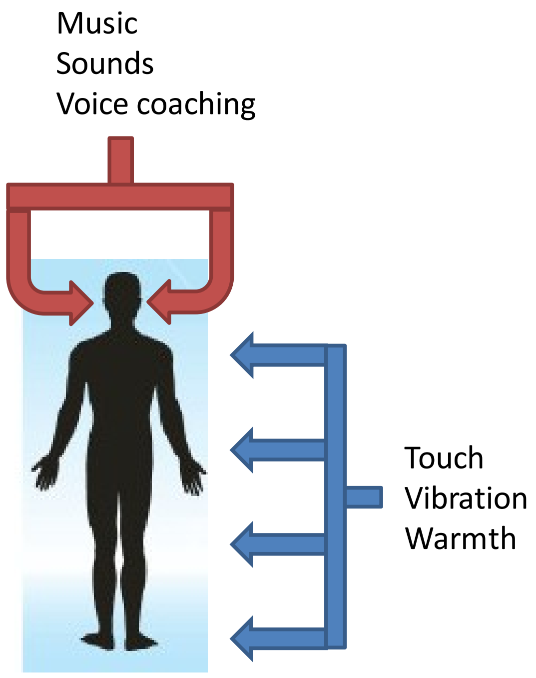

Figure 2: Impression of auditory-tactile stimulation of a user

Initial experiments at Philips Research with a combination of ambient sounds, relaxing music and patterns on a tactile blanket [6] also appear to be promising. The results suggest that the different modalities can mutually strengthen each other to provide a total experience that people really like and find relaxing.

On the other hand, a great deal of knowledge does exist in literature about the mutual interactions of the auditory, tactile and visual modalities, so-called *cross-modal effects*. This knowledge can be roughly categorized into in number of sub-areas [20,21]: perception and sensory thresholds, information processing performance, spatial attention, navigation/orienting in spaces, synaesthesia, neural plasticity, aging, perceptual illusions and sensory substitution. But the aspects of health, well-being and pleasantness are only addressed to a limited degree in this existing work.

2.5 Audio-tactile composition

A related area of research and creative work is *audio-tactile composition* [17,22,23], which refers to composing a musical piece or audible (ring) tone and at the same time composing tactile vibration effects that a user can feel. This is starting to become a commercially viable area, due to the rapid growth of haptics features in cell phones. However, in this field a number of topics have not yet been addressed in scientific work:

1. A link to human health and well-being has not been investigated;
2. Automatic audio to tactile conversion methods, suitable for driving multiple tactile actuators at the same time have not yet been investigated in a well-being context;
3. There is lack of well-founded composition tools to compose audio-tactile experiences, especially when considering applications, like ours, that are outside the limited context of mobile phone haptic ringtone composition.

3 Beyond the state of the art

3.1 Introduction

The aim of the scientific research identified in the previous section is to employ a combination of sound heard by a user and touch felt by a user at the same time, to influence the user's state of body and mind in a positive way.

We conclude from the previous section that stimulating the human senses of hearing and touch at the same time has great potential but needs further scientific study. We also found that existing approaches for auditory, tactile and auditory-tactile stimulation for health and well-being use fixed *content* that does not adapt to, nor interact with, the individual user. Content here refers to the combination of sounds and touch effects and how these are arranged in time and how touch stimulation patterns are arranged across the body. The adaptivity that a software-based solution for driving the auditory-tactile stimulation would provide, has not yet been exploited anywhere. To make our vision more concrete, we have provided example use cases in Section 1.3.

In this section, we first present in Section 3.2 three key scientific challenges/questions that have been identified. In the subsequent sections 3.3-3.5 the research topics are presented in somewhat more detail. Section 3.7 concludes by sketching a vision of the type of system that we believe is interesting for the research community to work towards, combining the results that should come out of the three research topics.

3.2 Scientific challenges

Within the wider area of auditory-tactile stimulation, we specifically want to highlight the following three scientific challenges/questions:

1. What are the mechanisms of human perception of combined auditory-tactile stimuli, and how can these mechanisms be modeled and used by a software-based system to influence the state of the human body and mind towards a desired state?
2. What are good methods for *conversion* between the audio domain and the tactile domain, in this context? *Conversion* here refers to converting content, for example music, from one domain to another, but also at a meta level to converting methods or paradigms from one domain to another. For example, how could a paradigm from music composition be converted into a paradigm for tactile composition? Tactile to music conversion is also something we consider.
3. How should measurement and interpretation of a user's biosignals and behavior during a stimulation session be done, to adapt the stimuli in an optimal way? Adaptation should help to better and faster achieve the users' well-being goals.

To start addressing these scientific challenges, we will outline a number of related potential research directions in the remaining text of Section 3.

3.3 Topic 1: Effects of multimodal stimuli on health and well-being

In the first proposed research topic, the goal is to study multi-actuator tactile stimulation of the human body and the effect of this stimulation on human health and well-being, alone and in combination (*cross-modal* effects) with the sense of hearing. Also other senses such as smell and vision may have to be taken into account here.

For a user, desired states can be (depending on the application) relaxed, peaceful, sleepy, engaged, dreamy, satisfied, active, et cetera. The mechanisms of human perception here may include auditory-tactile *sensory illusions*. Sensory illusions - perceiving things that are not really physically there - can be a very powerful way to evoke emotions.

A first step is to investigate existing literature on auditory and tactile perception and the related stimulation methods that use touch and hearing. Based on experimental tests, requirements from the application field, and findings from literature, one could apply an iterative, user-centered design and research process to come up with auditory-tactile stimuli that are likely to have a certain health or well-being related effect. These effects will then have to be investigated in user tests. Artificial Intelligence (AI) techniques such as rule-based systems, machine learning, personalization and (real-time) adaptation need to be investigated and employed to design models and systems that make it possible to link user characteristics and user experience to the properties of temporal and spatial patterns of auditory-tactile stimuli. Results of this type of research can then be applied in the work described under Topic 3 in Section 3.5.

The models just mentioned may use or incorporate existing models of human mental state, known from literature. One example is the well-known valence/arousal model proposed by Russell [24]. This model could be used to represent the known arousal-decreasing effects of certain tactile stimuli as described in [10,13].

3.4 Topic 2: Audio to tactile and tactile to audio conversion methods

The second research topic focuses on conversions between the audio domain and the tactile domain. One purpose is for example automatic conversion of existing music and non-music audio content to corresponding tactile stimuli. The audio and generated tactile stimuli can then be played simultaneously, creating a combined auditory-tactile user experience. By using the music content as the basic ingredient, a potentially large number of auditory-tactile compositions can be created from existing music.

Automatic translation methods of audio to corresponding tactile stimuli will have to take into account the (well-being) effects on the human and the methods will have to be suitable for multi-actuator tactile stimulation systems. Translation should be done in such a way that the user's health goal (e.g. muscle relaxation, sleep, energizing, etc.) and other goals (e.g. pleasantness, compositional coherence) are achieved. Conversion methods may include detecting the structural and symbolic expression of a piece of music, and using this information such that the tactile composition will reflect the same expression.

At a more general level, we also consider the possibility of conversion of methods and paradigms between the audio and tactile domains. For example, the existing knowledge on music composition and musical expression and communication of meaning could possibly be “translated” into approaches for tactile composition and tactile expression and communication of meaning. For music to tactile conversion and vice versa, a musical ontology can be used as a basis. Specifically, the system of Schillinger [25] is a candidate. Schillinger explored the mathematical foundations of music, and was particularly inspired by Fourier analysis and synthesis.

The topic of studying conversion methods that translate tactile stimuli into audio or music seems less obvious at first sight. However, we also envision useful applications here such as translating an existing tactile massage pattern that works well into matching music, in order to strengthen the psychological effect of the stimulation on the user.

3.5 Topic 3: Audio-tactile systems that adapt to the user based on sensor information

The third research topic involves so-called *closed-loop systems* or *biocybernetic loop* [26]: a sensory stimulation system, in which biosignals and behavior from a user are measured and used to adapt the stimuli that this user receives. To do the adaptation properly, relevant user influencing strategies should be used. Based on measurements on the user state, the system can then select the optimal influencing strategy.

User-adaptive methods have an added potential to be more effective, and at the same time more appealing to the user. This potential is still untapped today. Besides having the possibility of explicit multimodal interaction [27,28,29] with a user, interactive content can also be created by *implicit personalization* of auditory-tactile experiences. This is very useful in cases where the user cannot be expected to actively interact a lot with a system, foe example for elderly, people with impairments, or hospitalized people with temporary impairments. Research into multimodal (vision, hearing, touch, speech, gesture) user interfaces that optimally combine explicit and implicit interaction to provide the best level of personalization during a session could be a part of this research theme.

The biosignals that can be sensed and used for an adaptive system may include brain signals (EEG), heart signals (ECG), respiration, or skin conductivity (SCL); but also behavioral signals such as the user's movements, speech utterances or facial expressions during a session.

A particular research challenge for the use of biosignals in an automated system is that there are large inter-person variations. A system using a person's biosignals, would first have to learn about the user, id est calibrate its interpretation of signals towards the current user. This calibration challenge is also part of the research area that we propose.

3.6 Synergies in the three research topics

The above three research activities may require close cooperation mutually. Also they involve a cooperation between the field of artistic content creation on the one hand, and on the other hand scientific areas such as haptics, perception, brain and cognition. The perceptual mechanisms studied in topic #1 are on the one hand linked to lower-level brain mechanisms and to cognition, but on the other hand also to topics such as aesthetic perception. The audio-tactile conversion in topic #2 is primarily linked to composition and to the arts, but can only succeed if the physical and mental health goals are respected – topics related to multi-sensory perception, brain and cognition. Similarly for topic #3: although measurement and interpretation of user state mainly links to psychophysiology, perception, brain and cognition, it is also necessary to have influencing strategies in place that guide a user towards a desired state. These influencing methods will probably have a strong creative/artistic component in them.

3.7 Vision: Sensing + algorithms + content = optimal personalized experience

Combining the work proposed in the above three topics, we can sketch a vision to work towards. With recent advances in ICT such as low-cost embedded data processing, solid-state storage growth, ubiquitous networking, and recent progress in unobtrusive brain and biosignal sensors, a novel type of sensory stimulation system becomes feasible. This type of system will in real-time adapt a stimulation session towards an optimal, personalized experience for the current user. This personalization can be based on a generic model of a user and his/her mental state, which is continuously updated based on sensor interpretation and data mining. Here, the data is sensed (preferably in an unobtrusive way) from the user during a stimulation session.

The measured signals and their interpretation can be used to construct a software model of the current user state, which may describe current estimated levels of relaxation, sleepiness and comfort. Based on the model, influencing strategies can be chosen to help achieve the user's health and well-being goals.

Artificial Intelligence methods could be used effectively in construction of the user state model and in the optimal selection of influencing strategies. This would be a novel approach beyond the current state of art for auditory-tactile or tactile stimulation. In addition, this approach could be extended in the future to include also stimulating the visual sense (with images, video or light) or olfactory sense.

4 Discussion and next steps

4.1 General conclusions

In this position paper we have introduced auditory-tactile stimulation as a possible means to increase health and well-being, applicable in various application areas. The existing state of the art has been briefly addressed and based on our findings so far we conclude that there is clear potential for innovation in auditory-tactile stimulation approaches. Three specific areas for further research have been identified. Finally, a vision is presented of a software-based learning system that can automatically or semi-automatically adapt to the individual user based on general knowledge plus sensor information obtained from the user during a session. The system will then decide which specific auditory or tactile stimuli to render, to optimally achieve the user's goals.

4.2 Relevance of the proposed topics

If we take a broader look at society as a whole, one trend is that due to an aging population, Western economies are increasingly struggling with increasing healthcare costs and shortage of healthcare personnel. Therefore, there is a growing need for preventive healthcare. Preventive care is a useful instrument, not only to improve the quality of people's lifes, but also to partly avoid the cost of expensive regular treatments. The results of the research we propose could be applied for preventive healthcare, and can therefore have a positive impact on society.

Other potential users could be the healthcare workers themselves. Due to the ageing trend and economic constraints, their work will become ever more efficiency-oriented, time-pressured and stressful. Looking outside the domain of healthcare, other potential user groups can be identified who increasingly have to cope with highly stressful events occurring at work or who have to work under pressure. For example public transport personnel, school teachers, fire fighters or police officers. All these user groups could benefit from innovative new ways of coping with stress, or ways of inducing relaxation or sleep, quickly and on-demand. Auditory-tactile stimulation holds this promise.

One concrete example where results can be applied on the shorter term is the small enterprise Sense Company, an active supplier of sensory stimulation solutions to care organizations. Their portfolio includes tactile stimulation products. Other examples would be providing enjoyable relaxation solutions for hospital patients or medical personnel, or products that help to manage pain for chronically ill users at home.

As another example of application of results, the Utrecht School of the Arts (HKU) has already done various projects with partner organizations over the past years, aiming at people with special needs such as people who are deafblind. They investigated how these people can benefit from musical and rhythmic tactile stimulation.

4.3 Creating a research community

By organizing the Special Symposium "Haptic and Audio-Visual Stimuli: Enhancing Experiences and Interaction" as part of the EuroHaptics 2010 conference, the authors aim to start the process of gathering a research community around the topic of tactile (haptic) stimulation combined with other modalities, for applications in healthcare, well-being, entertainment and user interaction. We plan to organize a follow-up event around these topics in the future, possibly as a workshop linked to an existing conference.

Acknowledgements
This work was partially supported by the ITEA2 Metaverse1 (www.metaverse1.org) Project.

References

[1] Gallace, A., Spence, C. The science of interpersonal touch: an overview, *Neurosci Biobehav Rev.* Feb;**34**(2):246-59, 2010.

[2] Wigram, A.L. *The effects of vibroacoustic therapy on clinical and non-clinical populations*, PhD thesis, St. George's Hospital Medical School, London University, 1996.

[3] de Niet G, Tiemens B, Lendemeijer B, Hutschemaekers G. Music-assisted relaxation to improve sleep quality: meta-analysis. *J Adv Nurs.* 2009 Jul;**65**(7):1356-64, 2009.

[4] Lemmens, P., Crompvoets, F., Brokken, D., van den Eerenbeemd, J., and de Vries, G.-J. A body-conforming tactile jacket to enrich movie viewing. In the *Proceedings of the 3rd joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems* (Salt Lake City, UT, Mar. 18-20), or *WorldHaptics '09*, 2009.

[5] Dijk, E.O., Weffers-Albu, A., & De Zeeuw, T., A tactile actuation blanket to intensify movie experiences with personalised tactile effects. *Demonstration papers proc. 3rd International Conference on Intelligent Technologies for Interactive Entertainment* (INTETAIN), Amsterdam, 2009. www.eskodijk.nl/haptics or wwwhome.ewi.utwente.nl/~dennisr/intetain/INTETAINsupplementary.pdf

[6] Dijk, E.O., Weffers, M.A., Breathe with the Ocean: A System for Relaxation using Combined Audio and Haptic Stimuli, *Proc. Special Symposium on Haptic and Audio-Visual Stimuli: Enhancing Experiences and Interaction*, of EuroHaptics 2010, July 7th, Amsterdam, The Netherlands, 2010, (these proceedings).

[7] Sha-art, AlphaSphere relaxation chair or "Spaceship for the inner journey", website http://www.sha-art.com, May 2010.

[8] Parés, N., Masri, P., van Wolferen, G., Creed, C. Achieving Dialogue with Children with Severe Autism in an Adaptive Multisensory Interaction: The "MEDIATE" Project, *IEEE Trans. on Visualisation and Computer Graphics*, **11**(6), November 2005.

[9] van Panhuis, B. "Bijtanken met een dutje", newspaper *Trouw* July 17th, 2008. http://www.trouw.nl/incoming/article1724624.ece

[10] Wigram, A.L. *The effects of vibroacoustic therapy on clinical and non-clinical populations*, PhD thesis, St. George's Hospital Medical School, London University, 1996.

[11] Prisby, R.D., et al. Effects of whole body vibration on the skeleton and other organ systems in man and animal models: What we know and what we need to know, *Ageing Research Reviews* **7** 319-329, 2008.

[12] Kvam, M.H. The Effect of Vibroacoustic Therapy, *Physiotherapy*, June, **83**(6) 290-295, 1997.

[13] Patrick, G. The effects of vibroacoustic music on symptom reduction, *IEEE Engineering in Medicine and Biology Magazine*, **18**(2) 97-100, 1999. DOI 10.1109/51.752987

[14] Essick, G.K., et al. Quantitative assessment of pleasant touch, *Neuroscience and Biobehavioural Reviews* **34** 192-203, 2010.

[15] Puhan, M.A. et al. Didgeridoo playing as alternative treatment for obstructive sleep apnoea syndrome: randomised controlled trial, *BMJ* Feb 4;**332**(7536):266-70, 2006. PMID: 16377643

[16] Haans, A. and IJsselsteijn. W.A. The Virtual Midas Touch: Helping Behavior After a Mediated Social Touch, *IEEE Trans. on Haptics* **2**(3) 136-140 July-Sept, 2009.

[17] Gunter, E. Skinscape: *A Tool for Composition in the Tactile Modality*, M.Sc. thesis, Department of Electrical Engineering and Computer Science, MIT, 2001.

[18] Virtual Relaxation Solutions website, http://www.vrelaxation.com/ , May 2010.

[19] Meditainment ltd., website http://www.meditainment.com, May 2010.

[20] Calvert, G. (ed.), *The handbook of multisensory processes*. MIT Press, MA, USA, 2004.

[21] Grünwald, M. (ed.), *Human Haptic Perception – Basics and Applications*. Birkhäuser Verlag, Switzerland, 2008.

[22] Chang, A. and O'Sullivan, C. An Audio-Haptic Aesthetic Framework Influenced by Visual Theory, *Proc. HAID 2008*, LNCS Vol. 5270/2008, pp. 70-80, 2008.

[23] van Erp, J.B.F. and M.M.A. Spapé. Distilling the underlying dimensions of tactile melodies. Proceedings of Eurohaptics 2003, pp. 111-120.

[24] Russell, J. A. A circumplex model of affect. *Journal of personality and social psychology*, **39**, 1161 – 1178, 1980.

[25] Schillinger, J. The Schillinger System of Musical Composition. Da Capo Press music reprint series, 1977. ISBN: 0306775522. Reprint of 1941 ed. published by C. Fischer, New York, 1941. See also http://en.wikipedia.org/wiki/Schillinger_System

[26] Serbedzija, N.B., S.H. Fairclough. Biocybernetic loop: from awareness to evolution. Proc. *11th Conf. on Evolutionary Computation*, Trondheim, Norway, pp. 2063-2069, 2009.

[27] Cao, Y. and Theune, M. and Nijholt, A. Towards Cognitive-Aware Multimodal Presentation: The Modality Effects in High-Load HCI. In: Engineering Psychology and Cognitive Ergonomics. 8th International Conference, EPCE 2009, Held as part of HCI International 2009, 19-24 July, San Diego, USA. pp. 3-12. Lecture Notes in Computer Science 5639. Springer Verlag, 2009.

[28] Pantic, M. and Nijholt, A. and Pentland, A. and Huang, T.S. Human-Centred Intelligent Human-Computer Interaction (HCI²): how far are we from attaining it? *Int. J. of Autonomous and Adaptive Communications Systems*, 1 (2). pp. 168-187, 2008.

[29] van Gerven, M. and Farquhar, J. and Schaefer, R. and Vlek, R. and Geuze, J. and Nijholt, A. and Ramsay, N. and Haselager, P. and Vuurpijl, L. and Gielen, S. and Desain, P. The Brain-Computer Interface Cycle. *Journal of Neural Engineering*, 6 (4). pp.1-10, 2009.

Tactile Experiences

Paul M.C. Lemmens, Dirk Brokken, Floris M.H. Crompvoets,
Jack van den Eerenbeemd, Gert-Jan de Vries

Philips Research, High Tech Campus 34, NL-5656 AE Eindhoven, The Netherlands

Abstract

A simple touch can result in a profound and deep experience. Tactile communication has been used in information displays or to increase the entertainment value of arcade and pc games. The study of communication of emotion via tactile stimulation started only recently. We have built an emotion jacket as a research prototype to study the communication of emotions with vibrotactile stimulation. We recreated bodily feelings related to emotional experiences (e.g., a shiver down one's spine) as tactile stimulation patterns and showed that these emotion patterns, in a movie-viewing context, can increase emotional immersion with the content being viewed.

1 Introduction

Being touched can be a very powerful experience and its effects can range from a scary and unnerving unexpected hand on a shoulder, to a soothing, relaxing massage, or to a reinvigorating hug, with many more gradients in between [1]. The tactile modality, contained within the human skin, is the largest sensory modality (in terms of surface) that humans have. In the womb it starts to develop the earliest of all other senses and is most developed at birth [2]. This strong and old connection between tactile sensations and the intimacy and safety of the womb could be considered one of the reasons why a touch can evoke such powerful emotions. Touches sooth and arouse infants, and a touch can also regulate an infant's state [3]. Touches may even be a basic need like food, water, or sleep [4]. Following the principle of equipotentiality, it may be that touch is a particularly strong medium for children to communicate their (emotional) state [3].

Compared to modalities like vision and audition, scientific study of the properties of the tactile sensory system, and its communicative abilities in particular, was scarce for a long period of time [3]. Recently, Van Erp and colleagues used the information processing properties of the tactile modality to create informative tactile displays to provide additional navigational cues to airplane and helicopter pilots. They successfully used tactile stimulation to prevent overloading the visual and auditory senses that are already highly taxed in a cockpit context [4,5].

To improve the quality of the experiences that they generate, the entertainment industry has been using the tactile modality for some time. One can think of the arm-wrestling machine or a racing simulator with force feedback in its steering that can both be found in any arcade hall. More recently, the availability of tactile stimulation systems for personal entertainment systems like game consoles and pc's has increased [5,6]. The available technology ranges from simple rumblers and force feedback systems in joysticks to full torso vests containing multiple air bladders that quickly inflate upon impact in, for instance, first-person shooters [7]. Unfortunately, however, both the research into the informational properties of tactile information as well as studies investigating the effects of added tactile stimulation in entertainment context [8] neglect the aforementioned strong and intimate link between tactile stimulation and their effect on emotions.

Some work into this area can be found in interaction and design research on virtual mediated touch (see Haans *et al.* [1] for a review). A lot of that work focused on using the inherent intimacy or closeness of a touch to improve virtual communication between humans on a personal level [9-11]. The Lovebomb [12], on the other hand, provided the ability to anonymously indicate one's emotional state in public spaces and among strangers. Cramer showed that embodied agents and robots were judged less credible when their empathic responses were incongruent with that of their user and that touches from pro-active agents resulted in a less machine-like character [13]. Another work that involves connecting emotion and tactile stimulation is the Emoti-chair that uses a model of the human cochlea to provide tactile actuation based on the processing of, for instance, music [14].

The code of what makes a touch communicate a happy, sad, angry, or other emotional message is far from clear. Hertenstein and his group at the DePauw University have been working on code of tactile stimulation and emotion [2,15]. They showed that strangers could accurately (ranging from 48% to 83%) decode distinct emotions when touched by another person. Moreover, when Hertenstein *et al.* video recorded these touches and showed them to another group of participants, these participants also recognized the intended emotion with high accuracy. For specific emotions, percentages of correctly recognized emotions ranged from 38% to 71% [2].

In our own work we have taken a different approach to the study of the communication of emotion via tactile stimulation. We worked from William James' observation that every emotion has a distinct bodily reaction [16-18] and have listed various bodily reactions that are a result of an emotional experience. For instance, we considered responses like a shiver down one's spine, having butterflies in your stomach, a racing heart beat, etc. We then reversed James' idea and studied whether providing one of these bodily reactions could actually induce an emotion.

Note that it is important to stress the difference between enhancing emotional experiences compared to enhancing movie effects. For instance, in a movie scene in which Bruce Lee is surrounded by evil henchmen, one can try to enhance the experience by converting the visually presented punches and kicks into tactile sensations. This is the movie-effects approach. On the other hand, one could also try and provide a tactile experience of the anxiety that Bruce Lee feels in such a tight situation and the relief once he won the fight and survived. This latter approach of enhancing emotional experiences is the one that we have taken in the current study.

We asked whether tactally recreating these bodily reactions and using them as stimuli could enhance the emotional experience of watching movie content. By measuring psychophysiological signals that change due to changes in emotional state as well as taking questionnaire responses regarding emotional state, we measured the emotional state of viewers before, during, and after movie viewing. We expected that these responses would show responses indicative of deeper immersion when comparing a film clip without and with tactile actuation.

In the next section, we describe the emotion jacket that we developed to be able to project tactile sensations on the torso of our viewers. In section 3, the user test to evaluate our idea is briefly described and we round up with some conclusions based on our findings.

2 Body–Conforming Jacket With Tactile Actuators

The main design criteria for the emotion jacket were: ability to stimulate back and front of the human torso and arms, being battery powered, having smooth integration of electronics with the fabric for good aesthetics, good accessibility of electronics, and being light weight. The design aimed to enable projection of tactile patterns on the entire torso while keeping the electronical design low on complexity. This resulted in a jacket with 64 uniformly distributed actuators in a layout covering the entire torso with roughly 15 cm distance between neighboring actuators (see Figure 1).

A stretchable fabric was chosen to create a tight fit. This ensured that the actuators were close enough to the skin for the best tactile sensation possible. Small, medium, large, and extra large vests were built to accommodate different sizes.

2.1 Electronics Design

For the actuators we opted for pancake–shaped (coin type) generic eccentric rotating–mass (ERM) motors because they were light weight, thin, and inexpensive compared to other offerings. A disadvantage is that we were limited to vibrotactile stimulation only. The ERM motors are glued onto the back of custom-made PCB's that were connected to the segment-driver PCB's using thin flexible wires. Each segment-driver PCB controlled 4 motors.

The driver segments were daisy chained to form a serial bus that starts and terminated at a custom–made interface PCB. This PCB combined the SPI-bus from the external USB-to-SPI interface with the

power supply line from the two AA batteries. The electronics design was a compromise between number

Figure 1. Outer lining (left) and jacket turned inside-out showing the inner lining with electronics and wires (right-hand panel). Red (thicker) wiring is the serial bus that connects all segment drivers and provides communication and power; white (thinner) wiring connects motor PCB's to the respective segment driver.

of cables, number of drivers and the limit of the (flexible) cabling in terms of throughput of current (Watts).

The jacket was operated on 2 AA-sized batteries. With rechargeable batteries that each deliver 2500 mAh, the jacket had an operational lifetime of 1.5 hours when continuously driving 20 motors at the same time.

2.2 Textile Integration

The jacket consisted of two layers: an outer lining and an inner lining. The PCB's were sewn onto the inner lining using holes along their outer perimeter and were then covered by the outer lining for protection and aesthetics. At the bottom side of the torso and at the end of the sleeves the linings were not sewn together for easy access to the electronic components. The photograph on the right of Figure 1 shows the shirt inside-out, exposing the PCBs and wires. The total weight of one vest, including electronics and batteries, was approximately 700 grams.

2.3 Designing Tactile Stimuli

The actuators in the jacket are controlled from a PC using a LabVIEW™ (National Instruments, Austin, TX, USA) software interface that was developed in–house. The whole system has been principally designed to be able to use a 10 ms resolution for specifying changes in the tactile stimuli; in practice, we often reverted to a 20 ms resolution.

The LabVIEW application allowed us to generate tactile stimuli on various levels of granularity. First, we created different types of shapes. Shapes have, in principle, an unlimited duration and the amplitude specified for each 10 ms step reflects the intensity of the vibration of the motors. Example shapes are sine waves, block waves, sawtooth waves, etc. Thus, shapes define the vibration intensity over time, and are the building blocks for patterns.

Patterns specify at what point in time a particular motor has to render the given shape. An example pattern is a series of sine wave shapes that run from the left wrist over the shoulder to the right wrist. Patterns thus define the spatial and temporal distribution of vibrations over the torso.

Finally, these patterns were played back on the emotion jacket at predetermined times. For the present study, most of the patterns were based tactile sensations that are linked to common sayings like having butterflies in your stomach, or having a shiver down one's spine.

3 User Study

So, starting with James' idea that each emotion has a distinct bodily component [16-18], we tried to recreate these bodily sensations to see whether they could be used to induce an emotion. We set up a user test in which participants had to view clips of movie content that was validated to elicit a certain emotional response. We created tactile emotion patterns for each of these clips, on the one hand, by reverting to common wisdoms like shivers down one's spine, a racing heartbeat, exploding with anger and, an arm around your shoulders, or a sigh of relief (etc.). On the other hand, we also created patterns that were specific to particular movies.

Fourteen participants (age range from 24 to 58, 4 females) viewed each clip in two versions: first the original version without tactile emotion patterns and on a second viewing the emotion patterns were projected onto their body. The participant(s) wore the vest during both viewings. The presentation order of the clips was randomized, although we made sure that each clip was first shown without tactile actuation.

Before and after each movie clip, the participants had to self-rate their emotional state using the Self-Assessment Manikin (SAM; [19]) which is a pictorial questionnaire that can be used to assess the positivity/negativity, level of arousal, and level of dominance/potency of an emotion. We also used a questionnaire that was employed in earlier work to determine immersion experiences in TV applications. This questionnaire included elements of emotional experience and immersion [20]. We expected that the after measurement would show indications of increased emotions (for instance, by a higher score on the arousal scale).

During the viewing, psychophysiological responses related to changes in emotional changes were recorded. Examples of these responses are the electrodermal response, heart rate, skin temperature, or respiration [21-25]. Again, we expected that these responses would shower larger deflections (i.e., indications of stronger emotional experiences) when the participants were viewing the movie clips with the emotion patterns present.

3.1 Questionnaire Data

The data of the SAM did not show that participants indicated higher positive states for positive clips when emotion patterns were present during viewing. Similarly, we did not find an effect of increased emotional arousal during movie viewing with emotion patterns present. Because we used a 5-point SAM, we presume that this absence of significant findings was due to the limited number of options to indicate changes. That is, the effects of the additional tactile emotion patterns appeared to be rather subtle and required a more detailed scale that enabled scoring of fine-grained differences.

For the immersion questionnaire, we obtained several significant effects that indicated that participants felt more involved in the movie viewing. Participants had more intense experiences, felt more drawn into the movie scenes, and felt that all senses were stimulated at the same time (see Table 1). Note that, interestingly, a question from the immersion questionnaire that explicitly taxed the experience of emotions did not reach the level of significance. We will return on this point in the discussion.

Table 1. List of relevant questions from the immersion questionnaire [20] with average scores (on a 5–point scale) for the actuation absent and actuation present conditions. * indicates significantly different from actuation absent at $p < 0.05$.

Item	Question	Absent	Present
7	My experience was intense.	2.68	3.04*
8	I had a sense of being in the movie scenes.	2.32	2.67*
12	I felt that all my senses were stimulated at the same time.	2.03	2.59*

3.2 Psychophysiological Responses

We observed significant effects on two of the three psychophysiological responses that we recorded. For skin-conductivity, we found higher levels of conductivity (indicative of higher arousal) in the viewing condition when the emotion patterns were present. Because this applied to all clips, it most likely reflected a generic increase in arousal due to the tactile stimulation itself (i.e. ignoring the emotional communication of the tactile stimulation).

More interesting was the statistically significant interaction showing that the increase in skin conductivity due to the presence of emotion patterns was different for the various movie clips. Figure 2 shows that for three clips, we observed stronger increases in skin conductivity than for the other clips. Interestingly, these clips were all intended to evoke negatively valenced emotions.

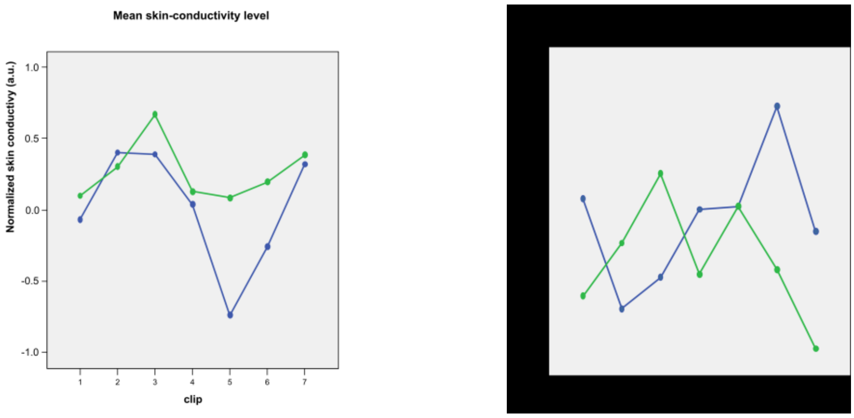

Figure 2. For all clips the effect of Actuation on skin–conductivity level (left panel) and heart rate (right panel). The blue line is for actuation absent; the green line is for actuation present. On the x-axis, the clips are numbered and ordered from left to right as: Braveheart, When Harry met Sally, Jurassic park III, The lion king, My bodyguard, Silence of the lambs, and Tom & Jerry. On the y–axis are the normalized physiological measurements.

For heart rate (Figure 2, right-hand panel), we observed a similar differentiation over movie clips of the effect of the added tactile emotion patterns. In this case a correlation with the valence of the clips for which we observed significant changes in heart rate, was not evident because one clip evoked positive emotions whereas another clip evoked negative emotions. Instead, it appeared as if the emotion patterns in these clips replicated increases in heart rate that the actors portrayed in the scenes.

4 Conclusions

The user study that we carried out showed that the addition to movie clips of emotions patterns that are projected on the torso using tactile actuators, results in a stronger or emotionally more immersive movie viewing experience. A few findings deserve some attention.

First, we did not observe a change in emotional state using the SAM of Bradley and colleagues [19]. We have already highlighted that this may have been due to a response scale that was not sufficiently granular to capture subtle changes in emotional state. However, a question from the immersion questionnaire that specifically asked for emotional experience also did not change in a statistically significant way. So did we change emotional state after all? In our view, we have, because a lot of subjective feedback from our participants that was not captured using formal questionnaires, indicated that they truly felt stronger immersion in the movie content when the emotion patterns were present. This is corroborated by the psychophysiological data. So, regarding subjective feedback, it is apparent that the questionnaires that we have chosen were not optimal in terms of level of granularity. The subtlety of our effects may require finer granularity than the questionnaires could deliver. In follow up studies it may even be necessary to select other questionnaires than we employed here.

The other point for discussion is that the psychophysiological recordings show that the emotion patterns, on average, have the strongest effects on negatively valenced clips. This may be a side effect of the type of tactile actuator that we chose. As we already mentioned, we were limited to vibrotactile stimulation due to our choice for eccentric rotating-mass motors. It is exactly this type of tactile stimulation that is often used for alerting functions, for instance, in cell phones. Therefore, it may be that these types of actuators are better suited for negatively valenced emotions because these emotions have an alerting and arousing function. It is clear that we need to study the efficacy of individual emotion patterns and that further detailed study of the design of the emotion patterns is needed to confirm or contradict this hypothesis.

On a higher level, we conclude that our findings highlight that it is possible to convey an emotional communication via relatively simple tactile technology. This corroborates the findings in the work of Hertenstein and colleagues on the code between tactile communication of emotional messages [2,15]. Our work shows that common wisdoms or sayings regarding bodily effects of emotions can actually be used to generate tactile stimuli that can trigger (or at least enhance) emotions. In our case, it still is the case that we require other content to set an emotional context. However, Hertenstein's findings that people can accurately decode a touch as an intended emotional communication shows that in principle our tactile stimuli could evoke emotions without a surrounding emotional context.

To conclude, by exploiting relatively simple and cheap tactile technology, we have been able to evoke strong psychological effects that show in increased emotional immersion during movie viewing. This finding shows that it is feasible to tactally enhance the emotional experiences of movie clips. This is a very useful addition to the enhancing of movie special effects using tactile stimulation that enables perceptually and emotionally rich movie-viewing experiences. Perhaps it is not without good reason that we say "I was touched" after having experienced a powerful emotion!

Acknowledgements. The author would like to thank two anonymous internal reviewers for their comments that helped to improve this manuscript.

References

[1] A. Haans and W.A. IJsselsteijn, "Mediated Social Touch: A Review of Current Research and Future Directions," *Virtual Reality*, vol. 9, pp. 149-159, 2006.

[2] M.J. Hertenstein, D. Keltner, B. App, B.A. Bulleit, and A.R. Jaskolka, "Touch Communicates Distinct Emotions," *Emotion*, vol. 6, pp. 528-533, 2006.

[3] M.J. Hertenstein, "Touch: Its Communicative Functions in Infancy," *Human Development*, vol. 45, pp. 70-94, 2002.

[4] J.B.F. Van Erp, "Tactile displays for navigation and orientation: perception and behavior," Ph.D.-thesis, Utrecht University, 2007.

[5] R.W. Lindeman, Y. Yanagida, H. Noma, and K. Hosaka, "Wearable Vibrotactile Systems for Virtual Contact and Information Display," *Virtual Reality*, vol. 9, pp. 203-213, 2006.

[6] K. Kaczmarek, J. Webster, P. Bach-y-Rita, and W. Tompkins, "Electrotactile and vibrotactile displays for sensory substitution systems," *IEEE Transactions on Biomedical Engineering*, vol. 38, pp. 1-16, 1991.

[7] "Tactile Gaming Vest - PC Gaming Hardware & Accessories with Physical 3D," http://tngames.com/products.

[8] A. Israr and I. Poupyrev, "Exploring Surround Haptics Displays," *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems - CHI EA '10*, Atlanta, Georgia, USA: , pp. 4171-4176, 2010.

[9] K. Tollmar, S. Junestrand, and O. Torgny, "Virtually living together," *Proceedings of the conference on Designing interactive systems processes, practices, methods, and techniques - DIS '00*, New York City, New York, United States: , pp. 83-91, 2000.

[10] A. Chang, S. O'Modhrain, R. Jacob, E. Gunther, and H. Ishii, "ComTouch," *Proceedings of the*

conference on Designing interactive systems processes, practices, methods, and techniques - DIS '02, London, England: , p. 312, 2002.
[11] K. Dobson, D. boyd, W. Ju, J. Donath, and H. Ishii, "Creating visceral personal and social interactions in mediated spaces," *CHI '01 extended abstracts on Human factors in computing systems - CHI '01*, Seattle, Washington: , p. 151, 2001.
[12] R. Hansson and T. Skog, "The LoveBomb: Encouraging the Communication of Emotions in Public Spaces," *CHI'01 extended abstracts on Human Factors in computing systems*, New York, NY, USA: ACM, , pp. 433-434, 2001.
[13] H.S.M. Cramer, "People's Responses To Autonomous And Adaptive Systems," Ph.D.-thesis, Universiteit van Amsterdam, 2010.
[14] M. Karam, C. Branje, G. Nespoli, N. Thompson, F.A. Russo, and D.I. Fels, "The Emoti-Chair: An Interactive Tactile Music Exhibit," *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems - CHI EA '10*, Atlanta, Georgia, USA: , pp. 3069-3074, 2010.
[15] M.J. Hertenstein, R. Holmes, M. McCullough, and D. Keltner, "The communication of emotion via touch," *Emotion*, vol. 9, pp. 566-573, 2009.
[16] W. James, "What is an Emotion?," *Mind*, vol. 9, pp. 188-205, 1884.
[17] J.J. Prinz, *Gut Reactions: A Perceptual Theory of Emotion*, Oxford University Press, , 2004.
[18] J.J. Prinz, "Are Emotions Feelings?," *Journal of Consciousness Studies*, vol. 12, pp. 9-25, 2005.
[19] M.M. Bradley and P.J. Lang, "Measuring emotion: the Self-Assessment Manikin and the Semantic Differential," *Journal of Behavior Therapy and Experimental Psychiatry*, vol. 25, pp. 49-59, Mar. 1994.
[20] R.J.E. Rajae-Joordens, "Measuring Experiences in Gaming and TV Applications - Investigating the Added Value of a Multi-View Autostereoscopic 3D Display," *Probing Experience - From Assessment of User Emotions and Behaviour to Development of Products*, J.W.D. Westerink, M. Ouwerkerk, T.J.M. Overbeek, W.F. Pasveer, and B. De Ruyter, eds., pp. 77-90, 2008.
[21] M.M. Bradley, "Emotion and Motivation," *Handbook of Psychophysiology*, J.T. Cacioppo, L.G. Tassinary, and Berntson, eds., Cambridge University Press, , pp. 602-642, 2000.
[22] J.T. Cacioppo, G.G. Berntson, D.J. Klein, and K.M. Poehlmann, "The Psychophysiology of Emotion Across the Lifespan," *Annual Review of Gerontology and Geriatrics*, vol. 17, pp. 27-74, 1997.
[23] P. Gomez, W. Stahel, and B. Danuser, "Respiratory Responses During Affective Picture Viewing," *Biological Psychology*, vol. 67, pp. 359-373, 2004.
[24] A. Kistler, C. Mariauzouls, and K. Von Berlepsch, "Fingertip Temperature as an Indicator For Sympathetic Responses," *International Journal of Psychophysiology*, vol. 29, pp. 35-41, 1998.
[25] Task Force of the European Society of Cardiology the North American Society of Pacing Electrophysiology, "Heart Rate Variability: Standards of Measurement, Physiological Interpretation, and Clinical Use," *Circulation*, vol. 96, pp. 1043-1065, 1996.

Multi-Haptics and Personalized Tactile Feedback on Interactive Surfaces

Hendrik Richter

University of Munich, Amalienstr. 17, 80333 Munich, Germany

Abstract

Tactile feedback on interactive surfaces such as touch-screens provides significant benefits in terms of reducing error-rates, enhancing interaction-speed and minimizing visual distraction [1]. However, current digital multitouch surfaces do not present any tactile information to the interacting user. Existing scientific approaches to provide tactile feedback on direct-touch-surfaces share the assumption that haptic feedback has to be given at the location of the interaction. We propose spatially disuniting the body-part of interaction (finger, hand) and the resulting tactile feedback. This approach is potentially beneficial for providing multi-haptic feedback on multi-touch surfaces and the communication of additional personalized information via the sense of touch. In the paper, we describe the potential benefits of our approach, present three first prototypes and discuss our current findings and intended future work.

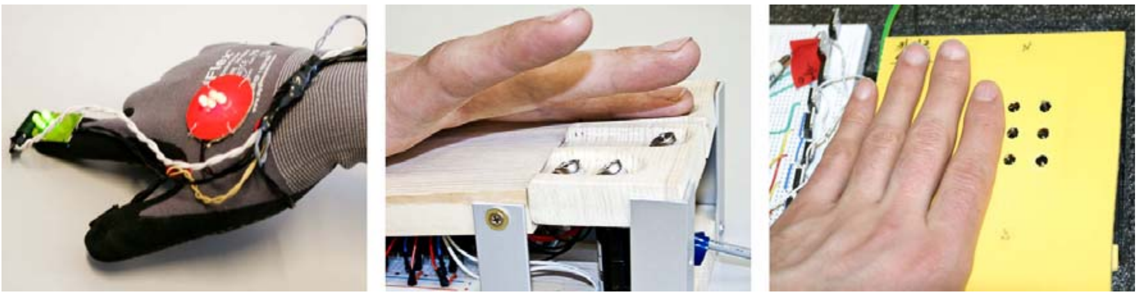

Figure 1: First prototypes (from left to right):
Tactile Thimble, Haptic Armrest, Edge Matrix

1 Introduction

Research about tactile feedback on direct-touch-surfaces can be categorized as follows: First, mobile actuator systems like the one used by Kaaresoja et al. [2] or Koskinen et al. [3] move the mobile device or the device's screen as a whole using motors or piezoelectric actuators. With this approach only a single touch-input can be augmented haptically. Second, shape displays, such as FEELEX [4] or Lumen [5] are based on the segmentation of the interactive surface into individually movable 'haptic pixels'. Currently, these systems only provide a small number of actuated points due to mechanical constraints.

The approaches mentioned above are based on the assumption that tactile feedback associated with an interaction should be applied on the interacting body part. Imagine touching a button on a tactile touchscreen with the tip of your index-finger. As a result, the whole screen would vibrate. This causes a repeated deformation of your finger's skin touching the screen. The body part of interaction (skin on your fingertip) is the location of tactile feedback. Only solitary user-inputs can be augmented haptically, the whole screen or device is moving as a single 'haptic pixel'. Or imagine touching an interactive shape display (like Relief [6] or Lumen mentioned above) with your hands. The moving pins stimulate the mechanoreceptors in the skin of your hand resting on the device. Again, the body part of interaction (skin on the palm of your hand) is the location of tactile feedback. The location of input and visual feedback on the device coincides with the location of haptic feedback. As a consequence, the visual output has to be mapped to the reduced resolution of the pin-array.

On the contrary, our approach is to disunite the body-part of interaction (finger, hand) and the resulting tactile feedback. In other words: while the user explores a virtual element on the interactive surface with his finger or hand, the resulting haptic stimuli are applied somewhere else on the body. Decisive for the position of application are human physiological conditions, the character of tactile information to be conveyed and the nature of the used haptic interface. We assume that our approach is beneficial for multi-haptic feedback on multi-touch surfaces, the accuracy of interaction as well as additional personalized information via the sense of touch.

In order to investigate the potential of our approach, we developed three first prototypes: the *Tactile Thimble*, the *Haptic Armrest* and the *Edge Matrix* (section 5). The prototypes differ in construction, body location of application and type of conveyed tactile information. Our three prototypes have one approach in common: tactile output (interactive surface → user) is spatially separated from manual input (user → interactive surface).

In the following, we describe the possible outcomes of our approach, present the three prototypes in detail and discuss our current findings and planned future work.

2 Tactile Feedback on Interactive Surfaces

Interactive surfaces exist in various types. From small, resistive single-touch screens (e.g. the one used in mobile devices like the Nintendo DS[1]) over capacitive screens (e.g. used in the Apple iPad[2]) to large interactive tables [7] or walls [8]. The term 'interactive surface' that we use in this paper refers to human-computer-interfaces with following characteristics [9]:

- *Combination of manual input and visual output:* In contrast to user interfaces like mice or touch-pads, the sensor area of input and the display area are congruent. Accordingly, positioning is absolute, the user's finger or hand is the tracking symbol.
- *Manipulation of interactive elements using fingers or hands*
- *Direct interaction with interactive elements*
- *Multiple points of sensed contact*: This way it is possible for multiple users to work on a shared interactive surface. Also gestures can be defined and interpreted as input.

During the interaction with most current touch-sensitive displays, touching is the same as activation. Tactile exploration on the device is not possible, tracking is done visually by the user while the finger is still "in the air". Constant visual attention is required during this targeting/ tracking phase. Tactile feedback during exploration and manipulation provides benefits. As stated above, tactile feedback during the interaction with touch-sensitive surfaces helps to reduce the errors made, enhances interaction-speed and reduces the required visual attention. This applies especially to situations or environments with increased cognitive or visual workload for the user [1] [10].

3 Related Work

Touch screens are used in a variety of mobile devices today. Mobile phones or portable music players often rely on vibrations to provide users with tactile feedback using built-in vibration motors. Vibrotactile feedback could help the user to know where he is on the screen and could support him in tasks like the selection of list items or during text entry. Poupyrev *et al.* used a single vibrotactile actuator to convey information. Different vibration patterns were used to communicate scrolling rate or position on the screen to the user. Selection of list items was 22% faster than when no tactile feedback was provided [11].

[1] http://www.nintendo.com/ds
[2] http://www.apple.com/ipad/

Mobile actuator systems like the one used by Kaaresoja *et al.* [2] or Koskinen *et al.* [3] move the mobile device or the device's screen as a whole using motors or piezoelectric actuators.

Another approach to provide the user with tactile information about interactive elements is the use of tactile displays. Existing electronic Braille displays for accessibility are limited to text-based information. Graphic tactile displays allow perceiving images by the sense of touch on a reusable surface and substitution of the visual/auditory sense. Tactile substitution can be used in augmenting accessibility for the blind or deaf in order to: (a) to enhance access to computer graphical user interfaces, (b) to enhance mobility in controlled environments [12]. Shape displays bring three-dimensionality to a table surface by using a matrix of individually movable elements. We already mentioned shape displays that combine input and output [5] [6] [4]. However, the resolution of haptic displays still is limited due to mechanical constraints [13].

A third possibility of bringing tactile feedback to the interactive surface is the application of tangible interfaces that are placed atop the table. On the one hand, this type of interface can be used as input device for the interaction with virtual elements on the surface. On the other hand, it works as output device to convey tactile information about the virtual objects underneath. An example for this approach is the Haptic Tabletop Puck by Marquardt *et al.* [13]. The prototype allows the user to explore vertical relief, malleability of materials, and horizontal friction of objects on the interactive surface. The tactile information is generated by a mechanically activated rod that moves vertically within a column to reach different heights above the table. Active tactile exploration of object characteristics and interaction with elements is enabled by a sensor on top of the device that detects the amount of pressure being applied to its top. The point of input is defined by a virtual arrow that starts under the device. The tracking symbol is the end of this arrow. This approach is really inspiring, but also has some drawbacks: the interaction using multiple fingers or even gestures is inhibited. The interaction isn't direct, the Haptic Tabletop Puck works as a mechanical separator between hand and interactive elements on the interactive surface.

We propose a different approach that tries to avoid these drawbacks.

4 Spatially Disuniting Interaction and Tactile Feedback

Our research is based on the approach to move the actively generated tactile feedback away from the interacting user's hand. The user is working with virtual elements on a multi-touch surface and is perceiving edges, areas or characteristics of these elements via tactile stimuli applied to his body. The interaction with virtual elements still is direct. In a multi-user environment like the collaborative work using a tabletop, additional tactile feedback could improve the accuracy during targeting tasks and the interaction with small objects [13]. Another possible goal we plan to examine is the exploration of distal objects on the interactive surface using tactile information. We are particularly investigating possible consequences such as Multi-Haptics (see section 4.1) and Personalized Tactile Feedback (see section 4.2).

4.1 Multi-Haptics

Using the stated approach, the first matter of our scientific interest is how to provide haptic feedback on multi-touch surfaces. In other words: the simultaneous transmission of tactile cues about a virtual object's geometrical, functional and semantic characteristics to more than one finger or hand. We refer to this goal as **Multi-Haptics**. Accordingly, the area of tactile stimulation is not limited to the size of a fingertip, haptically enlarging the tactile resolution of multiple contacts to the interactive surface becomes possible. Imagine the user is exploring an interactive surface using both hands. The location, form or surface characteristics of the virtual objects under the user's fingers are conveyed to the user's body (e.g. his underarm). The user is free to use both his hands and to directly interact with the surface and the depicted elements.

4.2 Personalized Tactile Feedback

In the next step, we plan to investigate the personalized transmission of tactile cues to multiple users of a shared interactive display using the aforementioned approach (**Personalized Tactile Feedback**). This means that one user is enabled to perceive tactile characteristics of an interactive element but another user of the shared touch-surface perceives different sensory cues (or none at all) during the tactile exploration of the same virtual object. One benefit would be the creation of an individual channel of tactile information to each user of a shared interactive surface. We plan to investigate how this channel can be used to carry private information. Imagine two people working on a shared interactive surface; one person is touching a virtual object that has a soft surface. When the other person is touching the same virtual object, the object could "feel" completely different, e.g. it could present the object's state using tactons [14]. The tactile feedback is personalized. One could also think of invisible objects on interactive surfaces that could be tactilely sensed by some users.

5 First Prototypes

In order to advance our research on the spatial separation of tactile feedback and body part of interaction, we designed and implemented three first prototypes of tactile interfaces (see Figure 1). All three prototypes stimulate the skin, i.e., are related to cutaneous touch. Our tactile interfaces are concerned with the mechanical deformation of the skin and leave out pain and temperature sensations. Pasquero [15] names basic engineering attributes for the coding of artificial perceptual information (e.g. amplitude, frequency, duration, resolution and signal waveform). An extra dimension is the locus of interaction. Our three prototypes provide the user with tactile signals that differ in the aforementioned attributes. Design of the prototypes and first results of our studies are presented in the following.

5.1 Tactile Thimble

This prototype applies tactile information to the user's interacting finger or the hand (see Figure 2). Using off- the-shelf components such as cylindrical vibration-motors, tactile location cues about close-by virtual objects are conveyed. A direct and private human-computer connection based on the perceived stimuli can be established.

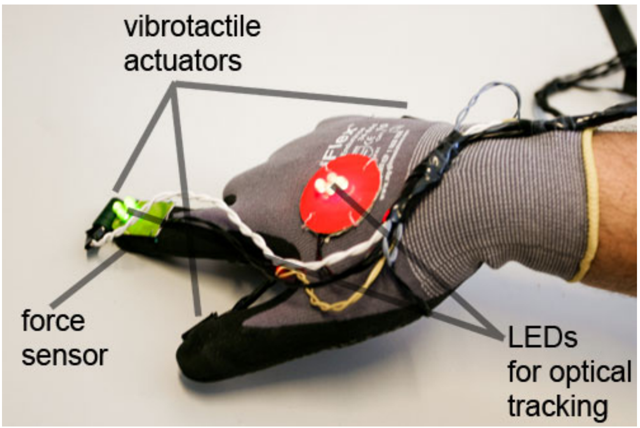

Figure 2: The Tactile Thimble

5.1.1 Design

The *Tactile Thimble* is a haptic interface that consists of a textile glove and three attached cylindrical vibration motors. The locations of the three actuators (forefront of fingernail on thumb, forefront of fingernail on index-finger and outside section of the palm) are based on physiological conditions of mechanoreceptors in the glabrous skin in the human hand [16] [17]. The glove's position and orientation is tracked optically using a camera. A force sensor is placed under the tip of the user's index finger, so the user is able to "press" a virtual button when he places his finger on it. The user who is wearing the glove may now be asked to "find" an invisible, virtual object on a 2D surface. This exploration is based on the tactile information that can be provided by the actuators. For example: if the virtual object is positioned to the right of the hand, the actuator on the right side of the glove may convey a signal. This concept is based on tactile way-finders such as [18].

5.1.2 Classification

The *Tactile Thimble* follows the principle of spatially separating body-location of interaction and location of tactile feedback. The tactile cues are applied to the same body-half, to the interacting hand, but not to the index-finger's skin that is in contact with the interactive surface. The *Tactile Thimble* is used to analyze the possibilities of supporting the process of tactile exploration using abstract tactile messages as described in [14]. In order to examine the potential of abstract cues for location information, we reduced the number of tactile actuators to three. Information like "move your hand backwards" is encoded using signal attributes such as duration or frequency instead of the location of the actuator.

5.1.3 Intended Research

With the *Tactile Thimble*, it is possible to perceive distant objects on an interactive surface using the sense of touch. In a pretest, we examined how our participants search for objects on a surface using only their sense of touch. We observed two search strategies. At the moment, we are trying to transfer these findings to the prototype in order to improve searching speed and object discrimination. Another focus of our ongoing studies lies on the design of the tactile signals that are conveyed to the user. The user has to be able to distinguish between states like "on the object", "object to the right" or "object too far away".

In the future, we plan to further examine the composition of vibrotactile signals to convey information about distance or location of an object. At the application level, we intend to use more than one device in order to study user-experience and potential benefits.

5.2 Haptic Armrest

While the user's dominant hand is interacting with virtual objects on an interactive surface, the non-dominant hand is placed on this prototype (see Figure 3). Again, the locations of exploration and the resulting sensory feedback are apart. Geometrical object-characteristics such as edges are conveyed using the movement of linear solenoid magnets. Additionally, tactile properties and functional or semantic characteristics of the virtual object are applied using multiple vibration motors. Based on these multiple motors, an additional carrier of abstract information is enabled by spatial encoding.

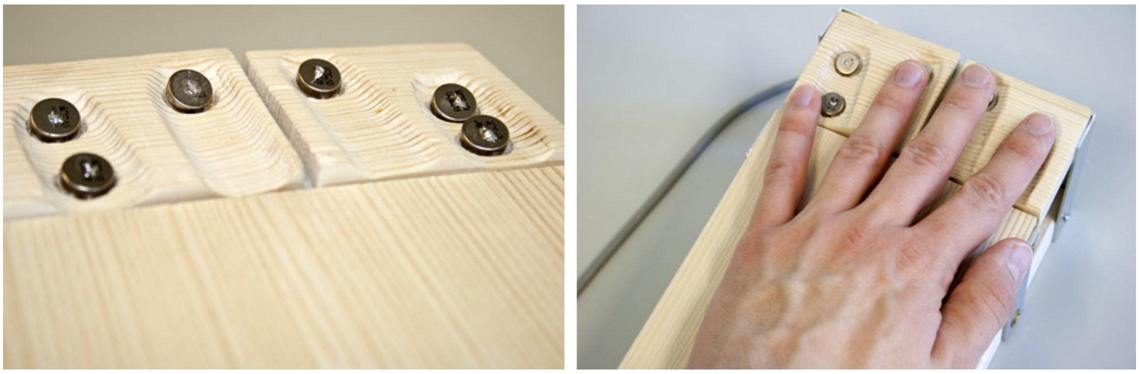

Figure 3: The Haptic Armrest. Left: two individually movable platforms and vibration actuators that are positioned under each fingertip. Right: the non-interacting hand is resting upon the actuators

5.2.1 Design

The *Haptic Armrest* is a tactile interface that consists of a wooden board that the user can rest his non-interacting hand upon. The interface contains two types of actuators: two solenoid magnets are used to move two individual platforms (~1cm stroke). Little finger & ring finger of the user's hand are resting on the first platform; middle finger and index finger are placed on the second platform. Vibration motors are attached to the two platforms in a way that they are placed under each fingertip. Following this approach, the *Haptic Armrest* combines properties of vibrotactile interfaces and shape displays. In other words, tactile information can be encoded using amplitude, frequency, duration or location of the signal, but also by actual displacement of the user's fingers.

5.2.2 Classification

The *Haptic Armrest* applies tactile information to the non-dominant hand of an interacting user (see Figure 4). The tactile feedback can be described as being "mirrored" to the other side of the user's body. Haptic information is applied to a body-area significantly larger than the body-area that is in contact with the interactive surface. Therefore, an enrichment of cutaneous experience is possible. Edges of virtual elements can be presented as pushes of the platforms. The characteristics of virtual objects can be conveyed using vibrotactile cues. Again, attributes of encoding are signal duration, amplitude, frequency and location.

5.2.3 Intended Research

This prototype is used to assess the importance of haptically perceivable edges and areas for object identification. An edge can be reduced to binary information (up/down) using the platforms. The orientation and course of an object's edge can not be perceived. At the moment, we are analyzing how the sense of touch is used to discriminate between virtual interactive elements that have the same visual appearance (see Figure 4). We are interested in the importance of object edges and the possibilities to describe these edges using actively generated tactile feedback.

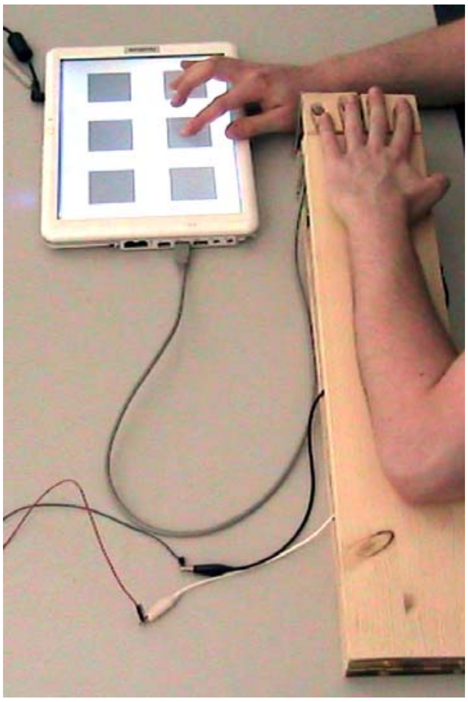

Figure 4: Usage of the Haptic Armrest in a first user-study (video still)

5.3 Edge Matrix

This prototype is basically a matrix of 3x3 linear solenoids (see Figure 5). Again, this haptic interface reacts to interactions of the user's hand on an interactive surface. A virtual object's orientation and rotation is conveyed to the non-interacting hand or arm that is rested on top of the matrix.

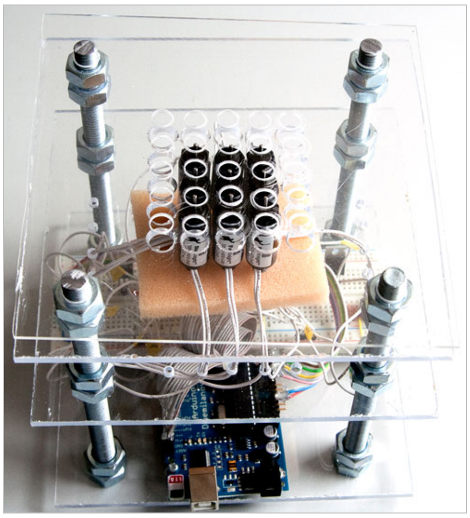

Figure 5: The Edge Matrix is a basic haptic display

5.3.1 Design

The *Edge Matrix* is an electromagnetic haptic display that can vibrate, for active feedback, as well as dynamically change its shape, for passive feedback. It consists of a matrix of electro-magnetic actuators, with individually controllable pins that can be pushed upwards. The solenoids can be activated individually. As a result, the 9 pins can be moved up and down (9 mm stroke). Additionally, it is possible to convey vibrotactile cues by continuously activating and deactivating single pins. We can simulate certain amounts of surface hardness by changing the applied voltage. With this prototype, it is possible to communicate shape information of objects on the interactive surface to other parts of the user's body. It is designed in a way that enlargement of the matrix is possible by adding solenoid actuators. The housing of the device enables the researcher to adapt the stroke of the pins as well as the force of pressure to the intended research question.

5.3.2 Classification

Again, this tactile interface dislocates tactile feedback to non-interacting body-parts. The body-location of application is not defined for this tactile interface. It may be positioned to convey tactile cues to the interacting arm (e.g. skin of the underarm). This approach would be similar to the *Tactile Thimble*. However, the tactile cues generated by this prototype may also be applied to the non-interacting hand. This approach would be similar to the one we used with *Haptic Armrest*.

The actuator technology combines attributes of vibrotactile interfaces and shape displays. Thus, information can be encoded by variations in frequency, duration or amplitude of vibrotactile information. Additionally, communication of shape information is possible by individually movable pins. As a result, the possible bandwidth of conveyed information is higher than the one provided by the other prototypes. E.g. the conveyance of edge orientation and course is possible by passive haptic feedback. An object's surface characteristics or abstract state is conveyable using active haptic feedback.

5.3.3 Intended Research

In the future, we plan to assess where to place the tactile display on the user's body in order to convey tactile object characteristics. Along with that, we are experimenting with size and form of the device. Similar to the *Tactile Thimble,* more than one tactile display could be placed on the user's body.

This prototype also has requirements on tracking technology. In order to analyze the relevance of haptically perceivable orientation and course of virtual elements, it is necessary to track the orientation and moving direction of the user's finger. We are currently working on solutions involving optical and sensor-based tracking technologies.

6 Conclusion and Future Work

This paper describes our approach to bring tactile feedback to an interactive surface. We analyze the potential of spatially separating body-part of interaction (finger, hand) and resulting tactile feedback. Three first prototypes of haptic interfaces are described in this paper: the *Tactile Thimble*, the *Haptic Armrest* and the *Edge Matrix*. The *Tactile Thimble* is a glove that communicates the proximity and position of distant virtual elements on a touch-sensitive surface. The information is encoded using vibrotactile signals. The *Haptic Armrest* is a haptic interface that is used to convey tactile information to the non-interacting hand. This prototype combines properties of vibrotactile interfaces and shape displays. Our third prototype, the *Edge Matrix* is a basic pin array. We plan to use this interface to convey tactile characteristics and spatial information (orientation, height) of virtual edges on touch-sensitive surfaces to the user.

Our research is based on the assumption that the separation of manual input and tactile stimuli provides a number of potential benefits. Using our approach of providing the user with tactile feedback, the interaction with virtual objects on the interactive surface remains direct. Multi-touch surfaces could be explored using multiple hands; every hand initiates a distinct transfer of information to the sense of touch. We refer to this goal as Multi-Haptics. On a shared interactive surface, our approach could result in personalized tactile feedback. The human sense of touch could be used as a channel carrying private information.

Our research is at an initial stage; our prototypes are evaluated in user-studies at the moment. We are improving our methods to measure the perception and causal connection of spatially separated input and feedback. In the future, we intend to improve our understanding of the perception of haptics on tabletops through more formal evaluations. We are constantly enhancing our tracking technologies in order to identify a finger or hand's orientation on the table. Another important part of our research is the design of the interfaces. We are experimenting how and where to apply the tactile cues on the user's body and how the form-factor of the interfaces can affect the user experience. At the application level, we assume that our approach is particularly beneficial in collaborative and dynamic scenarios.

7 Acknowledgment

The author would like to thank Hüsnü Güney, Doris Hausen, Kadri Januzaj and Sebastian Löhmann for technical support and valuable feedback.

References

[1] R. Leung, K. MacLean, M. Bertelsen, and M. Saubhasik, "Evaluation of haptically augmented touchscreen gui elements under cognitive load," *Proceedings of the 9th international conference on Multimodal interfaces*, ACM, 2007, p. 374–381.

[2] T. Kaaresoja, L. Brown, and J. Linjama, "Snap-Crackle-Pop: Tactile feedback for mobile touch screens," *Proc Eurohaptics*, Citeseer, 2006, p. 565–566.

[3] E. Koskinen, T. Kaaresoja, and P. Laitinen, "Feel-Good Touch: Finding the Most Pleasant Tactile Feedback for a Mobile Touch Screen Button," *Methodology*, 2008, pp. 297-304.

[4] H. Iwata, H. Yano, F. Nakaizumi, and R. Kawamura, "Project FEELEX: adding haptic surface to graphics," *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, 2001, p. 476.

[5] I. Poupyrev, T. Nashida, S. Maruyama, J. Rekimoto, and Y. Yamaji, "Lumen: interactive visual and shape display for calm computing," *ACM SIGGRAPH 2004 Emerging technologies*, 2004, p. 17.

[6] D. Leithinger and H. Ishii, *Relief*, New York, New York, USA: ACM Press, 2010.

[7] H. Benko, M. Morris, A. Brush, and A. Wilson, "Insights on Interactive Tabletops: A Survey of Researchers and Developers," *research.microsoft.com*, 2009.

[8] G. Morrison, "Interactive wall displays: interaction techniques and commercial applications," *SIGGRAPH '07: ACM SIGGRAPH 2007 courses*, ACM, 2007, pp. 54-64.

[9] K. Ryall, C. Forlines, M. Morris, and K. Everitt, *Experiences with and Observations of Direct-Touch Tabletops*, IEEE, 2003.

[10] S. Brewster, F. Chohan, and L. Brown, "Tactile feedback for mobile interactions," *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2007, p. 162.

[11] I. Poupyrev, S. Maruyama, and J. Rekimoto, "Ambient touch: designing tactile interfaces for handheld devices," *Proceedings of the 15th annual ACM symposium on User interface software and technology*, ACM, 2002, p. 60.

[12] V. Chouvardas, A. Miliou, and M. Hatalis, "Tactile displays: a short overview and recent developments," *ICTA ‘05: Proceedings of Fifth International Conference on Technology and Automation*, 2005, pp. 246-251.

[13] N. Marquardt, M. Nacenta, J. Young, S. Carpendale, S. Greenberg, and E. Sharlin, "The Haptic Tabletop Puck: Tactile Feedback for Interactive Tabletops," *ACM Interactive Tabletops and Surfaces - ITS'09. (Banff, Canada)*, 2009, pp. 93-100.

[14] S. Brewster and L. Brown, "Tactons: structured tactile messages for non-visual information display," *Proceedings of the fifth conference on Australasian user interface-Volume 28*, 2004, p. 15–23.

[15] J. Pasquero, "Survey on communication through touch," *Center for Intelligent Machines-McGill University, Tech. Rep. TR-CIM*, vol. 6, 2006.

[16] H. Oey and V. Mellert, "Vibration thresholds and equal vibration levels at the human fingertip and palm," *5th International Congress on Acoustics*, 2004.

[17] R.S. Johansson, "Tactile sensibility in the human hand: receptive field characteristics of mechanoreceptive units in the glabrous skin," *J. Phyeiol.*, vol. 281, 1978, pp. 101-123.

[18] W. Heuten, N. Henze, S. Boll, and M. Pielot, "Tactile wayfinder," *Proceedings of the 5th Nordic conference on Human-computer interaction building bridges - NordiCHI '08*, 2008, p. 172.

Assessing audiotactile interactions: Spatiotemporal factors and role of visual experience

Valeria Occelli[a]

[a] *Department of Cognitive and Education Science, University of Trento, Corso Bettini 31, Rovereto (Trento)*

Abstract

In the present paper, a brief overview of the results obtained in recent studies investigating audiotactile sensory interactions is provided. In particular, data emerging from studies on either monkeys or – brain-damaged and neurologically-intact – humans will be described, showing how the relative spatial position of the stimuli, the portion of space stimulated and the presence/absence of visual cues affect the sensory interplay between hearing and touch.

1 Introduction

Humans continuously interact with an environment providing a large amount of sensory information. The process by which the human nervous system tends to merge together the available pieces of information in unique events is commonly known as 'multisensory integration' [1].

The pioneering contribution to the understanding of the neural correlates of this process is owed to the studies in the superior colliculus performed by Stein and Meredith [2]. This midbrain structure is characterized by a high proportion of neurons responding to stimuli from more than a single sense (i.e., multisensory neurons). Multisensory integration is commonly assessed by considering the effectiveness of a crossmodal stimulus combination, in relation to that of its component stimuli, for evoking some responses from the organism. The crossmodal combination of stimuli evokes a number of impulses which is significantly higher than the number of impulses evoked by the most effective of these stimuli individually. Based on the study of such neurons' response properties, some principles for sensory integration have been formulated. The spatial principle is based on the evidence that only stimuli in spatial register (likely originating from the same external source), fall within the overlap between receptive fields of different sensory modalities, thus inducing an enhanced response; on the contrary, stimuli from disparate locations fall outside this area, failing to induce any enhancement or even causing a response depression. Moreover, only stimuli which occur close in time cause response enhancement, whereas stimuli separated in time just induce responses comparable to the ones evoked by unisensory stimuli (temporal rule). As it can be inferred, both spatial and temporal factors are of great importance for the assessment of multisensory interactions.

Unlike audiovisual and visuotactile sensory pairings, the interactions occurring at both neuronal and behavioural level between hearing and touch have been much less explored [3, 4]. This is somehow surprising, considering the wide range of everyday life situations in which we can experience – even though often in subtle and unconscious ways – the interplays occurring between these two sensory modalities [5]. Perceiving the buzzing and the itchy sensation of an insect approaching the rear surface of our neck; reaching a mobile phones ringing and vibrating from our trouser pocket. All these situations have in common the exclusive – or predominant – reliance on cue provided by senses other than vision. Besides these anecdotal reports, however, empirical evidence further support the existence of correlations between hearing and touch, thus justifying and corroborating additional investigations of this topic.

2 Experimental evidence

2.1 Studies on monkeys

Since the body is directly involved in the emergence of the tactile perceptual sensations, it follows that interactions between audition and touch are stronger within the space close to the body, the portion of space commonly known as ‘peripersonal space’ [6]. This has been shown in monkeys [7], patients [8] and neurologically intact humans [9].

For instance, Graziano and his colleagues [7, 10], have documented the existence of a population of trimodal neurons (i.e., neurons that respond to tactile, visual, and auditory stimuli) in the ventral premotor cortex (PMv, or ‘polysensory zone’, PZ). These neurons have receptive fields (RFs) that extend to a limited distance from the head, being able to respond to visual and auditory stimuli presented within roughly 30 cm from the tactile RFs. The RFs cover the contralateral side of the head as well as the space behind the head, thus providing a complete representation of the space surrounding the animal’s head. Interestingly, the gradient of firing of these neurons was found to vary not only as a function of the distance of the auditory stimuli from the animal’s head, but also as a function of their spectral complexity. Indeed, these neurons were found to preferentially respond to complex sounds (i.e., white noise bursts), with pure tones of different frequencies failing to elicit any significant response [7]. Given this evidence, these neurons would appear to make good candidates for the coding of multisensory characteristics of the space in close proximity to the body, i.e., within reach.

2.2 Studies on brain-damaged human patients

Evidence in support of audiotactile interactions in peripersonal space has been shown to have unique properties in human patients as well. In right brain-damaged patients suffering from left tactile extinction, the concurrent presentation of sounds on the right side of the head strongly interfered with the processing of the tactile stimuli on the left side of the neck (crossmodal auditory-tactile extinction) [8].

Interestingly, sounds interfered strongly with the processing of tactile inputs when they were delivered from close to the head (i.e., at a distance of 20 cm), but the effect was substantially reduced when they were presented far from the head (i.e., at a distance of 70 cm). Furthermore, this pattern of results primarily emerged with complex sounds, with pure tones inducing only a mild form of crossmodal extinction. Lastly, the portion of space from which the stimuli were presented, have proved to play a profound role in modulating audiotactile interactions. In particular, while white noise bursts exerted a stronger influence on tactile processing when the sounds were presented in the rear (vs. front) space, the – mild – effects induced by pure tones were selectively observed in rear space but not in frontal space [8], and furthermore were not modulated by the distance from which the sounds were presented.

2.3 Studies on neurologically-intact humans

The first study of audiotactile interactions taking place in the region of space surrounding the head, in particular, the space behind the head (i.e., in the part of space where visual cues are not available) of neurologically-intact people was conducted by Kitagawa and his coworkers [9]. In Kitagawa et al’s second experiment, a distractor interference task was used, with participants performing a tactile left/right discrimination task while auditory distractors (which were to be ignored) were presented simultaneously from the same or opposite side. In this task, the participants responded more slowly (and less accurately) when the auditory distractors were presented on the opposite side from the target tactile stimuli. Furthermore, the magnitude of this crossmodal interference effect varied significantly as a function of the distance and complexity of the auditory stimuli used. Whereas white noise bursts exerted a stronger crossmodal interference when they were presented from close to the participants’ head (i.e., within 20 cm) than when they were presented further from the head (i.e., 70 cm or more away), when the auditory stimuli consisted of pure tones, the overall effect was lower and was not modulated by the distance from which the sounds were presented.

Across the studies so far described, it has been demonstrated that crossmodal audiotactile interactions vary as a function of different parameters (e.g., distance between the stimuli and the body, spatial location of stimulation, and auditory complexity). First, stimuli presented in peripersonal space are distinctive and better able to attract attentional resources than stimuli presented in more distant regions. Second, the spatial arrangement of the stimuli has proved to play a significant role in modulating audiotactile interactions. Lastly, the interactions occurring between tactile and auditory complex stimuli are more pronounced than any interaction between tactile stimuli and pure tones.

Besides pointing to a high degree of similarity of the neural mechanisms coding for audiotactile interactions between primates and humans, the data emerging from these studies support the assumption that audiotactile spatial interactions are more prevalent in the region of space behind the head [8, 9].

Thus, the suggestion that has emerged from this kind of research is that the absence of vision (or visual information), as for stimulation occurring behind the head [8, 9] or as a result of blindness [11] seems to be related to the emergence of more prevalent spatial modulation of audiotactile interactions than those occurring in the frontal space [12].

This conjecture has received further support from a recent investigation that was designed to assess the potential existence of an audiotactile version of the so-called Colavita effect, as it has been termed from the name of its first investigator [13]. This term has been adopted to define the phenomenon by which neurologically-normal participants, performing speeded detection/discrimination response tasks, preferentially report the visual component of pairs of bimodal pairs of stimuli, sometimes neglecting to report the non-visual, auditory: [14, 15]; or tactile [16] component. As in a typical study of the Colavita effect, the participants in Occelli et al.'s study [13] had to make speeded detection responses to unimodal auditory, unimodal tactile, or bimodal audiotactile stimuli. The physical features of the auditory stimuli presented in frontal (Experiment 1) or rear space (Experiment 3), and the relative and absolute position of auditory and tactile stimuli in frontal (Experiment 2) and rear space (Experiment 3) were manipulated.

The most interesting results for the present theoretical context were observed in the third experiment, in which the tactile stimuli were presented to the side (right or left) of the back of the participant's neck, whereas the auditory stimuli – consisting of either white noise bursts or pure tones – were presented either in close spatial proximity to the body (i.e., over headphones) or far from it (i.e., from the loudspeakers located approximately 60 cm behind the participant's head).

The results demonstrated that sound complexity modulated the audiotactile Colavita effect, which was only observed when the auditory stimuli consisted of white noise bursts. If it is true that complex auditory stimuli are more likely to interact – and possibly to be integrated – with the tactile stimuli in rear space [7, 9], the auditory pure tone stimuli might have been more discriminable as compared to the white noise bursts [17]. This, in turn, could have facilitated the detection of both discrete sensory components of the bimodal trials in the present study.

Interestingly, the spectral complexity of the auditory stimuli had a differential effect as a function of the portion of space from which the stimuli were presented. Namely, when the frontal space was stimulated (Experiment 1), no audiotactile Colavita effect was observed for any kind of auditory stimuli (i.e., pure tones or white noise bursts), whereas when the stimuli were presented from the rear space (Experiment 3), the Colavita effect was selectively observed for those conditions in which the stimuli consisted of white noise bursts and not of pure tones. This data strengthen the hypothesis, already put forward in previous studies [8, 9], that in the rear space complex auditory stimuli (rather than pure tones) and tactile stimuli are more likely to interact.

The fact that a significant audiotactile Colavita effect was reported when the stimuli were presented from the same (vs. opposite) side only when presented from the back space and not when presented from the frontal space, can be explained by considering the fact that the spatial factors differentially affect audiotactile interactions as a function of the region of space in which they occur. As already mentioned, interactions involving auditory and tactile stimuli presented in frontal space tend to be less sensitive to spatial manipulations than are sensory interactions involving vision as one of the sensory components [18, 19, 20]. By contrast, the processing of the auditory and tactile spatial cues can be improved by presenting the stimuli in the portion of space where visual cues are typically not available [9]. The fact that this benefit was selectively observed for auditory white noise stimuli and not for pure tones is consistent with evidence showing that the spectrally-dense auditory stimuli (i.e., stimuli that contain a wide range of frequencies) induce noticeable benefits in auditory localization for the white noise bursts and not for pure tones [17].

Interestingly, the Colavita effect was exclusively observed when the auditory stimuli were presented via headphones, supporting the notion that the stimulation possibly involved a more pronounced integration of the signals, making the detection of the two discrete sensory components harder [15]. This result therefore suggests that when stimulation involves sounds originating from behind and close to the head it is somehow distinctive and gives rise to effects that cannot necessarily be detected when the stimulation is delivered far from the head [9] or in the peri-hand space [22].

Taken together, these results therefore suggest that the audiotactile Colavita effect is affected by both the relative and absolute locations from which the auditory and tactile stimuli are presented, as well as by the spectral complexity of the auditory stimuli. Of particular interest here is the fact that better performance in detecting/discriminating audiotactile stimuli presented from different locations (than from the same location) is selectively observed in those conditions where stimulation occurs in rear space. Moreover, the emergence of the Colavita effect was selective to those conditions in which the auditory stimuli consisted of white noise bursts and was never observed for pure tones. This pattern of results strengthens the hypothesis that when the portion of space behind (instead in front of) the head is stimulated, the interactions between auditory complex stimuli and tactile stimuli are more pronounced.

2.4 Studies on blind humans

The evidence that the spatial links between auditory and tactile stimuli are strengthened in the conditions where no visual cues are available has been further explored in another study, in which the potential effect of spatial factors on audiotactile temporal perception of both the sighted and blind was examined [11]. One of the most frequently used experimental designs to investigate temporal perception in humans has been the temporal order judgment (TOJ) paradigm [23]. In Occelli et al.'s study [11], participants were presented with pairs of stimuli at various different stimulus onset asynchronies (SOAs). Crucially, the stimuli were presented from either the same or different locations to the left and/or right of participants. The participants were required to make unspeeded responses regarding which sensory modality had been presented first. If the absence of vision is associated with a strengthening of spatial audiotactile interactions [9], then it can be hypothesized that the redundant spatial information provided by non-visual information in frontal space might exert a selective influence on the performance of blind participants in a audiotactile TOJ task, while the spatial arrangement of the stimuli would not be expected to modulate the performance of the sighted controls, i.e., based on the previous null effect of relative spatial position reported by Zampini et al. [12].

Studies of audiovisual audiovisual [19, 20] and visuotactile [18] TOJs have demonstrated that performance is modulated by the relative spatial position from which the stimuli are presented. In particular, participants are significantly more sensitive (i.e., the just noticeable difference, JND, is lower) when the stimuli are presented from different spatial positions rather than from the same position [18, 19, 20] The spatial modulation of TOJ performance may be attributable to the availability of redundant spatial information in the different-positions trials. As the TOJ task was presumably difficult for participants at the shorter SOAs, they may have utilized information concerning which position they perceived as having been stimulated first in order to facilitate their judgments concerning the correct order of presentation of the two modalities of stimuli. As these spatial cues are only available when the stimuli are presented from different positions, a selective facilitation of performance would only be expected to be observed in the different-positions (vs. same-position) condition.

An alternative explanation for the relative spatial position effect assumes that the performance of the participants may have been influenced by multisensory binding: stimuli from different sensory modalities that are presented from the same spatial position at approximately the same time tend to be bound together by the brain, thus giving rise to representations of unitary objects with multisensory properties [19].

Somewhat different findings have, however, been reported in the case of audiotactile TOJs [9, 12]. In particular, Zampini et al. [12] were unable to demonstrate any spatial modulation of audiotactile TOJ performance when the auditory and tactile stimuli were presented from in front of the participants on either the same or opposite sides. By contrast, Kitagawa et al. [9] observed a spatial modulation of audiotactile TOJs when the stimuli were presented from behind their participants' heads. Participants' performance was significantly better when the auditory and tactile stimuli were presented from different spatial positions rather than from the same position. Therefore, the comparison of these two audiotactile studies provides support for the claim that audiotactile spatial interactions may be more prevalent in the region behind the head (i.e., in the part of space where vision provides no direct information) than in front of it [3]. In their first experiment, the participants had to make a temporal order judgment (TOJ) regarding pairs of auditory and tactile stimuli presented from the left and/or right of participants at varying stimulus onset asynchronies (SOAs). In particular, the participants had to report the modality of the first stimulus that had been presented on each trial. The auditory stimuli were presented from loudspeakers situated 20 cm behind the participants' head, whereas the tactile stimuli were delivered via electrotactile stimulators attached to their earlobes. The results showed higher sensitivity (i.e., lower Just Noticeable Differences, JNDs, defined as the smallest temporal interval at which people can accurately discriminate the temporal order of the stimuli presented) for stimuli presented from different sides rather than from the same side.

The results of the TOJ task described in Occelli et al.'s study [11] showed that while the performance of the sighted participants was unaffected by the relative spatial position (same vs. different) from which the two stimuli (one auditory, the other tactile) were presented, thus replicating Zampini et al.'s earlier study [12], the performance of the blind participants was modulated by relative spatial position. In particular, the blind participants performed significantly more accurately when the two stimuli were presented from different spatial locations rather than from the same position. The fact that the performance of the blind – but not that of the sighted – participants was sensitive to the spatial separation between the auditory and tactile stimuli is consistent with the hypothesis that only the visually deprived group was influenced by spatial cues when performing the task. The fact that spatial incongruence (as compared to congruence) led to a selective improvement in the performance of our blind participants is consistent with previous data showing that visual deprivation results in an improved ability to process spatial cues in the spared sensory modalities, e.g. in touch [24] and in hearing [25]. Taken together, these results establish a similarity with the pattern of performance reported recently in sighted participants in both visuotactile [18] and audiovisual TOJ tasks [19, 20]. By contrast, the exclusive reliance on those sensory modalities that are typically considered adequate for conveying spatial information [26] failed to induce any advantage in terms of the performance of the blindfolded sighted participants [12].

3 Conclusions

Taken together, the studies here described show that the absence of visual information, as when stimulation occurs in the rear space [8, 9, 13], following blindness [11, 24], during blindfolding or while keeping the eyes closed, or in total environmental darkness [27] provides the most suitable conditions in which to reveal audiotactile interactions. More interestingly, the evidence reported here provides evidence that audiotactile interactions are significantly affected by spatial factors, especially when the stimulation occurs in the space close to the back of the participants' head [8, 9, 13] as compared to when the stimulation occurs in frontal space [12] or in the space behind the participants' backs but in the peri-hand space [28]. The processing of spatial audiotactile links, in turn, determines significant influences within another domains, such as the temporal perception [11] or the crossmodal congruency (e.g., correspondence between auditory frequency representation and tactile location) [27].

References

[1] Calvert, G. A., Spence, C., & Stein, B. E. (Eds.), *The handbook of multisensory processes*. Cambridge, MA: MIT Press.

[2] Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses.* Cambridge, MA: MIT Press.

[3] Kitagawa, N., & Spence, C. (2006). Audiotactile multisensory interactions in information processing. *Japanese Psychological Research*, **48,** 158–173.

[4] Soto-Faraco, S., & Deco, G. (2009). Multisensory contributions to the perception of vibrotactile events.

[5] Doel, K. van den, & Pai, D. K. (1998). The Sounds of Physical Shapes. *Presence: Teleoperators and Virtual Environments*, **7**, 382-395.

[6] Cardinali, L., Brozzoli, C., & Farné, A. (2009). Peripersonal space and body schema: Two labels for the same concept? *Brain Topography*, **21**, 252-260.

[7] Graziano, M. S., Reiss, L. A., & Gross, C. G. (1999). A neuronal representation of the location of nearby sounds. *Nature*, **397**, 428-430.

[8] Farnè, A., & Làdavas, E. (2002). Auditory peripersonal space in humans. *Journal of Cognitive Neuroscience*, **14**, 1030-1043.

[9] Kitagawa, N., Zampini, M., & Spence, C. (2005). Audiotactile interactions in near and far space. *Experimental Brain Research*, **166**, 528–537.

[10] Graziano, M.S., & Gandhi, S. (2000). Location of the polysensory zone in the precentral gyrus of anesthetized monkeys. *Experimental Brain Research*, **135**, 259-266.

[11] Occelli, V., Spence, C., & Zampini, M. (2008). Audiotactile temporal order judgments in sighted and blind individuals. *Neuropsychologia*, **46**, 2845-2850.

[12] Zampini, M., Brown, T., Shore, D. I., Maravita, A., Röder, B., & Spence, C. (2005). Audiotactile temporal order judgments. *Acta Psychologica*, **118**, 277–291.

[13] Occelli, V., Hartcher O'Brien, J., Spence, C., & Zampini, M. (in press). Assessing the audiotactile Colavita effect in near and rear space. *Experimental Brain Research.*

[14] Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics*, **16**, 409-412.

[15] Colavita, F. B. (1982). Visual dominance and attention in space. *Bulletin of the Psychonomic Society*, **19**, 261-262.

[16] Hartcher O'Brien, J., Gallace, A., Krings, B., Koppen, C., & Spence, C. (2008). When vision 'extinguishes' touch in neurologically-normal people: Extending the Colavita visual dominance effect. *Experimental Brain Research*, **186**, 643-658.

[17] Moore, B. C. J. (1997). *An introduction to the psychology of hearing* (4th Ed.). London: Academic Press.

[18] Spence, C., Baddeley, R., Zampini, M., James, R., & Shore, D. I. (2003). Multisensory temporal order judgments: When two locations are better than one. *Perception & Psychophysics*, **65**, 318–328.

[19] Zampini, M., Shore, D. I., & Spence, C. (2003a). Multisensory temporal order judgments: The role of hemispheric redundancy. *International Journal of Psychophysiology*, **50**, 165–180.

[20] Zampini, M., Shore, D. I., & Spence, C. (2003b). Audiovisual temporal order judgments. *Experimental Brain Research*, **152**, 198–210.

[21] Tajadura-Jiménez, A., Kitagawa, N., Väljamäe, A., Zampini, M., Murray, M. M., & Spence, C. (2009). Auditory-somatosensory multisensory interactions are spatially modulated by stimulated body surface and acoustic spectra. *Neuropsychologia*, **47**, 195-203.

[22] Menning, H., Ackermann, H., Hertrich, I., & Mathiak, K. (2005). Spatial auditory attention is modulated by tactile priming. *Experimental Brain Research*, 164, 41-47.

[23] Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *Journal of Experimental Psychology: General*, **130**, 799–832.

[24] Röder, B., Rösler, F., & Spence, C. (2004). Early vision impairs tactile perception in the blind. *Current Biology*, **14**, 121-124.

[25] Röder, B., Teder-Saläjärvi,W., Sterr, A., Rösler, F., Hillyard, S. A., & Neville, H. J. (1999). Improved auditory spatial tuning in blind humans. *Nature*, **400**, 162–166.

[26] Welch, R. B., &Warren, D. H. (1980). Immediate perceptual response to intersensory discrepancy. *Psychological Bulletin*, **88**, 638–667.

[27] Occelli, V., Spence, C., & Zampini, M. (2009). Compatibility effects between sound frequency and tactile elevation. *Neuroreport*, **20**, 793-797.

[28] Zampini, M., Torresan, D., Spence, C., & Murray, M. M. (2007). Auditory-somatosensory multisensory interactions in front and rear space. *Neuropsychologia,* **45***,* 1869-1877.

Combining Mediated Social Touch with Vision: From Self-Attribution to Telepresence?

Antal Haans   Wijnand A. IJsselsteijn

Eindhoven University of Technology, P.O. Box 513 5600MB Eindhoven

Abstract

Combining mediated social touch (i.e., interpersonal touch over a distance by means of a tactile display) with vision allows people to both see and feel their remote interaction partner's touches. This is expected to increase the user's sense of telepresence (i.e., the experience of "being there" in the same environment as one's remote interaction partner), thereby perhaps providing part of the immediacy that marks natural unmediated physical contact. Such a possible effect of combining touch with vision is expected to depend partly on the central nervous system's ability to categorize tools and technological artifacts as an actual part of one's body. Since such so-called self-attribution is facilitated by a morphological congruence between the artifact and the human body, we anticipated that the effect of combining touch with vision would be larger when people could see the mediated touches being performed on a morphologically congruent as compared to an incongruent input medium. In our experiment, we compared two input media: a sensor-equipped mannequin that enabled a one-to-one mapping between seen and felt touch, and a morphologically incongruent touch screen that displayed a set of buttons. When participants saw the touches being initiated on the morphologically congruent mannequin input medium, they, as expected, reported a higher sense of telepresence. In addition, they perceived of the mediated touches as more touch-like with the mannequin as compared to the touch screen input medium. These findings illustrate that visual feedback, especially when morphologically correct, can improve mediated social touch. Our experiment, however, suggests that other mechanisms than self-attribution might be involved.

1 Introduction

Mediated social touch allows geographically separated people to engage in physical contact by means of tactile or haptic displays. The notion of "touch over a distance", however, is somewhat contradictory, as physical contact and physical proximity are closely related. In contrast to visual and auditory interactions, which by their nature are less dependent on physical proximity, touching requires people to be in intimate distances of each other. In other words, part of what makes touch significant in our every day interactions, such as in the communication of intimacy or the display of power, can be explained by physical proximity or personal space violations. Despite the paradoxical nature of touch over a distance, several designers and researchers have developed prototypes that enable mediated social touch. Examples of such prototypes are the inTouch [1], and the FootIO [2] (for a recent overview, see [3]). Designers of these systems conjecture that the addition of a haptic or tactile communication channel will enrich mediated interactions. They refer, for example, to the symbolic and intrinsic (e.g., recovery from stress) functions of social touch, or to the supposed intimate nature of addressing the skin. In doing so, they implicitly assume that mediated social touch is perceived of in ways similar to real (i.e., unmediated) physical contact. However, there are several notable differences between a real touch and the tactile and haptic stimulation provided by the prototypes described earlier (also [4]). Simulating the sensation or "feel" of a human touch, for example, remains difficult and expensive, despite advancements in tactile and haptic display technologies (for an overview, see, e.g., [5,6]). As a result, current prototypes rely on simple electromechanical actuators, such as vibration motors, which are a poor substitute for real physical contact (i.e., in terms of the qualitative experience or "feel"). Such differences between mediated and real (i.e., unmediated) touch make questionable the assumption that mediated social touches are perceived of as similar to real touch. Attempts to test this assumption are scant and often inconclusive in their findings (e.g., [4,7,8]). These studies typically aim to replicate, under mediated conditions, existing socio-psychological research on interpersonal touch, thereby establishing whether people respond in ways similar to a mediated as to a real (i.e., unmediated) touch. Haans and IJsselsteijn [4], for example, aimed to replicate the well-known Midas touch effect under mediated conditions (i.e., with the touch act replaced by vibrating electromechanical actuators). In their experiment, participants engaged in a brief interaction with a confederate through an instant messaging service. During this interaction, half of the participants received a mediated touch from the confederate. Research on the Midas touch predicts that people who are touched by another person are more willing, than those not touched, to assist that person with picking up accidently dropped items. In the experiment by Haans and IJsselsteijn, only 45% of the participants that did not receive a mediated touch assisted the confederate with picking up coins she dropped “accidentally” out of her wallet. Consistent with the Midas touch effect, as much as 61% of the participants that did receive a mediated social touch assisted the confederate with picking up the coins. Although the Midas touch effect was not statistically significant, the 16% increase in helping behavior was found to be similar to the Midas touch effect in unmediated situations. Although such a finding is promising for the field of mediated social touch, one question remains: How might devices that allow for mediated social touch be improved?

To our knowledge, published work on prototypes that allow for mediated social touch have not yet considered to combine touch with vision, which would allow people to see, and at the same time feel, their interaction partner performing the touches (also [9]).[1] Combining touch with vision allows the central nervous system to extract real-time intermodal correlations, which is crucial in the design of transparent media technology [10]. Transparency is important, because when people “forget” about the interface, they are more likely to respond to the touches as if the mediating technology were not there [3]. They might thus perceive of the electromechanical stimulations as more natural or touch-like when combined with visual feedback. Moreover, by establishing reliable visuotactile correlations, a tactile sensation might be felt, not on the person’s own body, but in the location in which the stimulation is seen (so-called distal-attribution [11]). This, in turn, is expected to facilitate people in developing a sense of telepresence: the impression of being physically there in the same location as the person initiating the touch [12]. Relevant in this respect is a recent study by Lenggenhager and colleagues [13]. In their experiment, participants wore a head-mounted display through which they saw a virtual character standing with its back toward the participant. This virtual character was created by a real-time recording of a fake human figure (i.e., a mannequin). The experimenter stroked the backs of the participant and the mannequin in precise synchrony. The participant would thus feel the strokes on his or her body, while seeing the virtual character being stroked at the same time. After one minute of such visuotactile stimulations, many participants developed the vivid impression that the body they saw standing in front of them was actually their own, as if they were looking at themselves standing in the location of the virtual character (for another example of this illusion, see [14]). This telepresence-like experience is expected to result from the central nervous system’s ability to categorize the virtual character as a part of the participant’s body. This process, in which a discrimination is made between what is contained within and outside the boundaries of the body, is called self-attribution.[2] It has been studied mostly within the experimental paradigm of the rubber-hand illusion, in which self-attribution, rather than involving a complete fake body as the foreign object, is limited to a fake hand (e.g., [15]). To elicit the rubber-hand illusion, a person's concealed hand is stimulated in precise synchrony with a visible fake one. Similar to the illusion elicited by Lenggenhager and colleagues [13], which is referred to as a full-body analogue of the rubber-hand illusion, about a minute of such visuotactile stimulation is sufficient for the participant to experience the fake hand as an actual part of his or her own body. Research on the rubber-hand illusion suggests that self-attribution is largely dependent on sensorimotor integration, and thus on the capability of the central nervous system to extract correlations between the various sensory modalities. Research

[1] Interestingly, in the domain of internet-based adult toys (for which Ted Nelson coined the term teledildonics in the 1970s), several commercial systems are available that take advantage of combining tactile stimulation with visual feedback.

[2] Although this aspect of the phenomenon of telepresence has received relatively little attention, Held and Durlach already pointed toward the relation between self-attribution and telepresence [16] (For a more detailed discussion, see [17,18])

has, for example demonstrated that the rubber-hand illusion will diminish or break down when a delay of 500 ms or more is introduced between the stimulation of the fake hand and that of the participant's concealed hand (e.g., [19,20,21]). At the same, this sensorimotor integration is modulated top-down by a cognitive representation of what the human body is like [20]: The development of a vivid rubber-hand illusion is dependent on the extent to which the foreign object resembles, or is morphologically congruent to, a human hand (also [22]). A similar finding is reported by Lenggenhager and colleagues [13] with respect to the full-body analogue to the rubber-hand illusion: No such illusion was elicited when the virtual "character" did not resemble the human body. Based on these findings, one can expect that the positive effects of combining mediated social touch with vision depend on a morphological congruency between the human body and the input medium on which the touch acts are performed.

1.1 Research Aims

In this study, we investigate the effect of combining touch with morphologically correct visual feedback on people's responses to mediated social touch. We expect that, compared to an incongruent input medium, a morphologically congruent input medium, which resembles the human body and allows for a one-to-one mapping between seen and felt touch, will result in a more transparent interface. This expectation is translated into three related hypotheses. Our first hypothesis states that seeing the touch acts being initiated on a morphologically congruent input medium will increase the perceived naturalness of the felt touches. In other words, we expect that the vibrating electromechanical stimulation (i.e., the mediated social touches) will be experienced as more touch-like when the participants see the touches being initiated on a morphologically congruent as compared to a morphologically incongruent input medium. Our second hypothesis states that people will report a higher sense of telepresence (i.e., "being there") with a morphologically congruent input medium. Our third and final hypothesis states that the effects of morphological congruency on telepresence and the perceived naturalness of the touches can be explained by the central nervous system's ability to attribute the morphological congruent input medium to the self, as if it were an actual part of the participant's body.

2 Methods

2.1 Participants

Our sample was drawn from the participant database of the JF Schouten School at Eindhoven University of Technology, Eindhoven, the Netherlands. Twenty-two persons were invited to participate in the experiment. Two persons were excluded from the data set due to technical problems with the tactile display. Of the remaining 20 participants, the mean age was 35.0 (*SD* = 18.5; range 18 to 65 years); All participants were of Dutch nationality. All participants received a compensation of € 10.00.

2.2 Experimental Design

The experiment consisted of three sessions: two experimental and one control. In each session, the participants were remotely touched by another person (a confederate of the experimenter). At the same time, the participants could see the confederate performing the touches on an input medium through a television screen, which displayed a real-time recording of the confederate's actions. In the experimental "Mannequin" session, participants could see the confederate performing the touches on a human-like input medium (i.e., a mannequin; see Figure 1). This input medium was congruent with the morphology of the human body not only in appearance, but also in the mapping between seen and felt touch. In the experimental "Touch Screen" session, the human-like input medium was replaced by a set of buttons, each corresponding to a body location (displayed on a touch screen; see Figure 1). This input medium was incongruent with the human body both in appearance and in the mapping of felt and seen touch. The order of the Mannequin and Touch Screen sessions was counterbalanced across participants. In both experimental sessions (i.e., the Mannequin and Touch Screen session), participants received eight

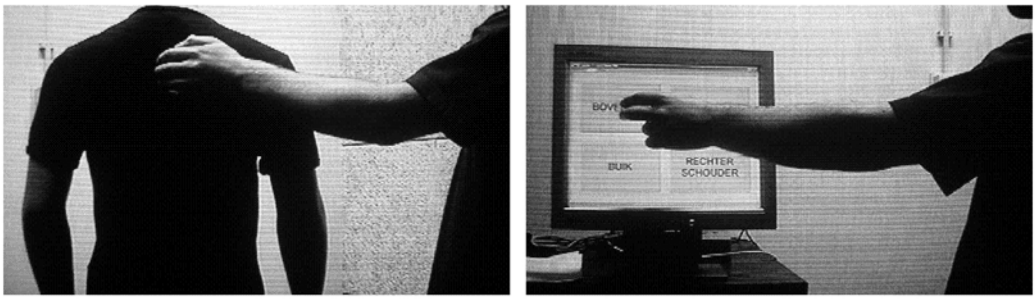

Figure 1: Visual stimuli for the morphologically congruent Mannequin (left) and incongruent Touch Screen session (right). The images are video stills converted to grayscale.

mediated touches on four locations on the body: stomach, upper back, lower back, and shoulder (each location was touched twice). The mediated social touches were presented in such an order that the same body location was not touched consecutively. The order of the touches was counterbalanced across participants.

Since we measured changes in participants' electrodermal activity in response to each mediated social touch (the result of which are reported elsewhere [23]), we were limited in the frequency with which the mediated touched could be delivered to the participant's body: Participants were touched only eight times, in only four specific ways, and with the time between two touches being relatively long (sometimes as much as 20 seconds). As a result, the visuotactile stimulation in the experimental sessions might have been too poor in information for the central nervous system to extract a sufficiently strong correlation between seen and felt touch. This in turn, is expected to impede the development of a vivid full-body analogue of the rubber-hand illusion [see 19, 21]. Therefore an additional control session, further referred to as the "Illusion Induction" session, was included in the experiment to assess the degree to which the illusion could be elicited with our particular experimental setup. In this Illusion Induction session, participants received a series of mediated touches, while watching the touches being performed on the morphologically congruent input medium (i.e., the mannequin). Again, the confederate touched the human-like input medium, and thus the participant's body, on the various body locations. In contrast to the experimental conditions, mediated touches were performed in a random fashion with the touches following shortly after each other, and thus with a high amount of information contained in the stimulation. Each participant was stimulated in such a manner for five minutes. If self-attribution occurs with our particular experimental setup, then not only morphological congruency, but also the amount of information in the visuotactile stimulation should affect the degree in which participants can develop a vivid full-body analogue of the rubber-hand. In other words, if self-attribution is involved, then a more vivid such illusion is expected to occur in the Illusion Induction control session as compared to the experimental Mannequin session.

2.3 Stimuli and Apparatus

Tactile stimulation was provided through a neoprene vest equipped with electromechanical actuators (i.e., vibration motors). This vest was similar to that used by Haans and colleagues [4,8]. Series of eight actuators were located at the stomach, 16 at the upper-back, and 12 at the lower back region. Two actuators were placed on the right shoulder. The morphologically congruent input medium (i.e., used in the experimental mannequin, and the control session) consisted of a male mannequin equipped with reed contacts (i.e., an electrical switch operated by applying a magnetic field; see Figure 1). The position of the reed contacts on the mannequin were matched to the position of the actuators in the vest, with each reed contact directly connected to one actuator. Small magnets were attached to the confederate's fingers. By bringing a magnet near a reed contact, the contact would close and the corresponding vibration motor was actuated. Mediated touches to the stomach and lower and upper back were given by stroking the magnet over the reed contacts on the mannequin. Mediated touches to the participant's shoulder were given by briefly touching the reed contacts on the shoulder of the mannequin, thereby resembling a tap rather than

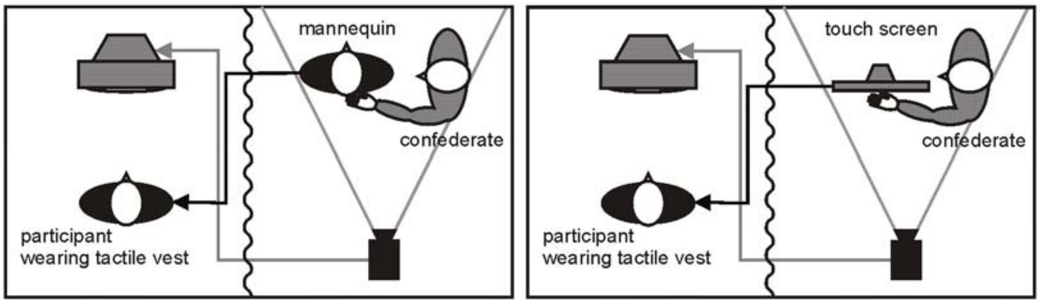

Figure 2: Experimental setup for the morphologically congruent Mannequin (left) and incongruent Touch Screen condition (right). The gray arrow depicts the visual channel, and the black arrow depicts the tactile channel.

a stroke. Whereas, touches to the stomach were done with the right hand, touches to the shoulder, the lower back, and upper back were done with the left hand. Since the reed contacts on the mannequin were matched to the position of the actuators in the vest, the mannequin input medium allowed for a one-to-one mapping between seen and felt touch for both body location and direction of touch. The morphologically incongruent input medium consisted of a touch screen on which four buttons were displayed. Each button corresponded to one of the four body locations (see Figure 1). Each button was labeled with the name of the body part it represented. In contrast to the mannequin interface, the touch screen did not provide a one-to-one mapping between seen and felt touch. Each time the confederate briefly pressed a button on the touch-screen, the software actuated the corresponding vibration motors to either produce a stroke (i.e., for the stomach, lower, and upper back) or a tap (i.e., for the shoulder). The hardware and software used for controlling the actuators with the touch screen was similar to that used by Rovers and Van Essen [2]. The buttons were always pressed with the left hand. The confederate was trained extensively to match the duration of the mediated touches in the Mannequin session with those in the Touch Screen session.

The room in which the experiment was conducted was divided in two sections by means of a black curtain (see Figure 2). One section housed the mannequin and touch-screen input media, as well as a digital camera that recorded the confederate's actions. In the other section of the room, a 21 inch television screen was placed at face height. During the sessions, the participant stood facing the television screen at a distance of approximately 1.50 meters. The television was connected to the camera standing in the other section of the room, allowing the participant to watch, in real-time, the confederate's actions on the input medium. To avoid that participants could see their own reflection in the television screen during the sessions, lights could be switched off in this section of the laboratory room.

2.4 Procedure

Participants were invited to the laboratory to evaluate a multimodal communication device together with another person. This other person was a male confederate of the experimenter. The confederate was casually dressed in a manner appropriate for a young man of his age (21 years). At their arrival, the experimenter assisted the participants with putting on the tactile vest. Next, the participant was asked to stand in front of the television screen, and the experimenter connected the tactile vest to either the Mannequin or the Touch Screen input medium. With the camera still off, the experimenter gave the participant one mediated touch to the upper back to familiarize him with the tactile stimulation.

At the start of each of the experimental sessions (one with the mannequin and one with the touch screen as the input medium), the participant was instructed to pay attention to the touches he saw and felt. Next, the experimenter would turn on the camera, enabling the participant to see both the input medium and the confederate through the television screen (see Figure 1). At that time, the confederate would initiate the eight mediated touches. After the last mediated touch, the participant completed a questionnaire. The second session, with the other input medium, proceeded in a similar manner as the first. After the experimental Mannequin and Touch Screen sessions, the experimenter asked the

Table 1: Items of the Naturalness of the Mediated Touches, and Telepresence self-report measures.

	Naturalness of the Mediated Touches
1	How natural, or unnatural, were the touches?
2	To what extent were the touches of a mechanical, or a human-like, nature?
3	How synthetic, or non-synthetic, were the touches?
4	To what extent did you perceive of the touches as realistic or unrealistic?
	Telepresence
1	Sometimes, it felt as if the other person was standing closely beside me.
2	Sometimes, I forgot that I was looking at a television screen.
3	Sometimes, it felt as if the other person was touching me directly on the skin.
4	Sometimes, it felt as if the other person was standing on my side of the room (within the curtained area).

participant to return to his place in front of the television screen for the final Illusion Induction control session. In contrast to the two experimental sessions, the various locations of the participant's body were touched in a random fashion (both tapping and stroking), for five minutes, with the touches following shortly after each other. After this session, the participant again completed a questionnaire.

2.5 Measures

2.5.1 Perceived Naturalness of Mediated Social Touches

After the Mannequin and the Touch Screen session, we assessed the extent to which the participants perceived the mediated touches as natural. Perceived naturalness was measured by means of four self-report items, such as "To what extent were the touches of a mechanical, or a human-like, nature" (see Table 1)? Participants could respond on a five-point scale with labels ranging from, for example, "mechanical" (coded with a 0), through "neutral" (coded with a 2), to "human-like" (coded with a 4). The mean score across these four items was used in the analyses. The reliability (Cronbach's alpha) of this aggregated Naturalness measure was α = .93 in the Mannequin, and α =.84 in the Touch Screen session. There were no missing responses.

2.5.2 Telepresence

After the Mannequin and the Touch Screen session, we assessed the degree to which the participants perceived a sense of telepresence. Telepresence was measured by means of four self-report items. These self-reports were written in the form of statements regarding the experience of "being there", such as “Sometimes, it felt as if the other person was standing closely beside me" (see Table 1). These statements were based on existing presence questionnaires (e.g., [24]). Participants could respond on a five-point scale with labels ranging from "disagree" (coded with a 0), through "neutral" (coded with a 2), to "agree" (coded with a 4). The mean score across these five items was used in the analyses. The reliability (Cronbach's alpha) of this aggregated Telepresence measure was α = .64 in the Mannequin, and α = .73 in the Touch Screen session. There were no missing responses.

2.5.3 Self-Attribution

After each of the three sessions (i.e., the experimental Touch Screen and Mannequin sessions, and the control Illusion Induction session), we assessed the degree to which people experienced a full-body analogue of the rubber-hand illusion by means of three self-reports, such as “Sometimes, it felt as if the mannequin was a part of my own body” (see Table 2). These self-reports were written in the form of statements, and were based on questionnaires commonly used to assess the vividness of the rubber-hand illusion (e.g., [15]) and its full-body counterpart [13,14]. Participants could respond on a five-point scale

Table 2: Items of the Self-Attribution measure, and percentage of affirmative responses in the Mannequin (M), Touch Screen (TS), and Illusion Induction (II) session.

		% Affirmative		
	Items	M	TS	II
1	Sometimes, it appeared as if I felt the touch on the location where I saw the input medium being touched, rather than on my own body.	60%	25%	65%
2	Sometimes, it felt as if the input medium was a part of my own body.	20%	10%	20%
3	Sometimes, I had the impression that I was looking at myself.	40%	15%	45%

The term input medium was replaced by the appropriate term for each session (i.e. "mannequin" or "touch screen"). Affirmative responses are either "slightly agree" or "agree"

with labels ranging from "disagree" (coded with a 0), through "neutral" (coded with a 2), to "agree" (coded with a 4). There were no missing responses. Since the reliability (Cronbach's alpha) of the Self-Attribution items was poor ($\alpha \geq .53$), we did not calculate aggregate scores, but used each item as an individual indicator. For this purpose, we recoded the three impression items into a dichotomous response format by collapsing "slightly agree" and "agree" into a single category "assert", and "disagree", "slightly disagree", and "neutral" into "refute".

3 Results

The average perceived naturalness of the mediated touches was $M = 1.2$ ($SD = 1.0$) in the Mannequin, and $M = 0.8$ ($SD = 0.9$) in the Touch Screen session. Since self-reported naturalness was found not to be normally distributed, the non-parametric Wilcoxon signed-ranks test was performed to investigate the effect of morphological congruence on perceived naturalness of the mediated touches. Although the touches were perceived of as relatively unnatural in both sessions, our participants reported, as expected, a higher perceived naturalness in the morphologically congruent (i.e., with the mannequin) as compared to the morphologically incongruent session (i.e., the with touch screen session), with $Z(N = 20) = 2$, and $p = .046$. Average self-reported telepresence was $M = 1.2$ ($SD = 1.0$) in the Mannequin session, and $M = 0.8$ ($SD = 0.9$) in the Touch Screen session. Although our participants' experience of telepresence was rather uncompelling, a paired-sample t-test demonstrated that participants, as expected, reported a higher degree of telepresence in the morphologically congruent (i.e., the Mannequin session) as compared to the morphologically incongruent session (i.e., the Touch Screen session), with $t(19) = 2.6$ and $p = .02$. With respect to self-attribution, more of our participants claimed to have encountered experiences related to the full-body analogue of the rubber-hand illusion in the Mannequin as compared to the Touch Screen session. In the Mannequin session, 20% of the participants affirmed to have encountered the impression that the input medium was a part of their body (Item 2 in Table 2). In contrast, only 10% of the participants reported to have encountered this impression in the Touch Screen session. This difference, however, was not found to be statistically significant, with $Z(N = 20) = 0.8$, and $p = .41$ (using the non-parametric Wilcoxon signed-ranks test). More participants claimed to have had the impression that they were looking at themselves through the television screen: 40% in the Mannequin session, but only 15% in the Touch Screen session (Item 3 in Table 2). This time, the difference between the Mannequin and Touch Screen session was statistically significant, with $Z(N = 20) = 2.2$, and $p = .03$. The most frequently encountered impression, however, entailed the experience that the touches were felt on the location where they were seen: 60% in the Mannequin session, but only 25% in the Touch Screen session (Item 1 in Table 2). Again, the difference between the two experimental sessions was statistically significant, with $Z(N = 20) = 2.3$, and $p = .02$. These data demonstrate that at least some of our participants reported to have encountered impressions related to a full-body analogue of the rubber-hand illusion, and that, as expected, the probability of developing these impressions was higher when they could see the touch act being performed on a morphologically congruent (i.e., the mannequin), as compared to an incongruent input medium (i.e., the touch screen).

In contrast to our expectations, however, increasing the amount of information contained within the visuotactile stimulation did not facilitate the participants in developing impressions related to a full-body analogue of the rubber-hand illusion. When comparing the percentages of participants who agreed with the three statements in the self-attribution questionnaire, we see that these percentages are rather similar for the Illusion Induction and the Mannequin session respectively: 65% and 60% for Item 1, 20% and 20% for Item 2, and 45% and 40% for Item 3 (see Table 2). Statistical analysis using a series of Wilcoxon signed-ranks test demonstrated that the Illusion Induction session did not differ significantly from the Mannequin session in eliciting these impressions, with $Z(N = 20) \leq 0.4$, and $p \geq .66$.

4 Discussion

With a sample of only 20 participants, we were able to support most of our hypotheses regarding the effects of combining mediated social touch with morphologically congruent visual feedback. As expected, our participants reported a higher sense of transparency and telepresence with the morphologically congruent (i.e., the mannequin) as compared the incongruent input medium (i.e., the touch screen). This finding illustrates the importance of morphologically congruent multisensory stimulation in the experience of transparency and telepresence (e.g., [10]). In addition, participants perceived the same electromechanical stimulation (by means of vibration motors) as more natural when they saw the mannequin being touched rather than the touch screen. This finding demonstrates that combining touch and vision can alleviate, at least partially, the technological limitations of current day tactile displays in accommodating natural interpersonal touch (i.e., in terms of "feel").

We hypothesized that these effects would result from the central nervous system’s ability to attribute the morphological congruent mannequin, but not the incongruent touch screen, to the self. As expected, more of our participants reported to have encountered impressions related to a full-body analogue of the rubber-hand illusion in the Mannequin as compared to the Touch Screen session (see Table 2). This is consistent with research on the rubber-hand illusion [20,22] and its full-body counterpart [13], which points towards a similar effect of morphological congruence. In contrast with research on the rubber-hand illusion [19,21], increasing the frequency and information richness of the visuotactile stimulation did not result in higher probabilities for developing impressions related to a full-body analogue of the rubber-hand illusion (see Table 2). One explanation for this unexpected finding is that our experimental setup and design might have constrained the development of a vivid illusion to such an extent that increasing the amount of information in the stimulation had no further facilitating effect. We used, for example, relatively simple mediating technology. Using a head-mounted display rather than a television screen, for example, would have provided a more immersive view on the remote site, thereby potentially increasing the extent to which self-attribution could occur. Similarly, replacing the vibrating electromechanical actuators with touches from a real human hand might have had a similar effect, as a discrepancy in the nature of expected and felt touch might constrain self-attribution [19,22]. Secondly, since the mannequin was positioned with its back toward the participant, the actual touches to the stomach could not be seen by the participants. This, in turn, might have negatively affected the development of a vivid full-body analogue of the rubber-hand illusion.

An alternative interpretation, however, is that we were unsuccessful in eliciting a full-body analogue of the rubber-hand illusion with our particular setup. It would, indeed, be rather surprising when eight touches which are spread over a relative long period of time can result in a similarly vivid illusion as five minutes of nonstop stimulation. If the observed effects of combining mediated touch with morphologically congruent visual feedback are not dependent on self-attribution, then what can explain the observed effects? Although more research is required to answer this question, there are at least two possible alternative explanations.

One alternative explanation for the observed effects is provided by research on the visual enhancement of touch. This research demonstrates that looking at the stimulated body part, even without seeing the stimulation itself, enhances a person’s tactile acuity (e.g., [25]). A similar facilitating effect is observed when the visual stimuli are presented on another person’s body, but not when they are presented on objects that do not resemble a human body (e.g., [26,27]). Recent research by Moseley, Parsons and Spence [28] demonstrates that the perception of pain is modulated by vision as well. In their study, people with chronic hand pain were asked to look at their affected hand through either a magnifying or a minifying glass. Whereas magnifying the affected hand increased pain, minifying the affected hand resulted in a significant decrease in pain. These studies indicate that combining mediated social touch with morphologically congruent visual feedback might in itself be a sufficient explanation for the effects observed in the present experiment.

A second alternative explanation for the observed effects is that participants did not perceive the mannequin as a part of their body, but as an object that represented them at the remote site as seen through the television screen (similar to an avatar being a representation of a person in a mediated environment). A similar argument has recently been made by Ehrsson and Petkova [29]. They argue that the full-body analogue of the rubber-hand illusion (as reported by [13,14]) is heavily dependent on the use of media technologies. Furthermore, they argue that the impressions reported by the participants in these studies might not result from self-attribution, but from people’s knowledge from using media technology (i.e., media schemata). In our experiment, 40 to 45 percent of the participants reported to have encountered the impression that they were looking at themselves when seeing the mannequin being touched on the television screen (i.e., Item 3 in Table 2). Such reports make sense both when people perceived the mannequin as a part of the body, and when they perceived the mannequin as a representation of themselves. The reported effects of combining mediated touch with morphological congruent visual feedback might thus result from psychological, rather than perceptual, means of self-identification. One psychological mechanism of self-identification that is worth investigating in this respect is synchronic identification. Synchronic identification entails the categorization of two objects, located at different locations, as being one and the same ([30]). It allows people to connect real objects with their mental or virtual counterparts ([31]). It thus allows a person to identify with, for example, his or her avatar in a mediated environment. Compared to self-attribution, synchronic identification does not rely, or at least not as heavily, on multisensory integration. It thus provides an explanation for the unexpected finding that increasing the amount of information in the stimulation did not facilitate the development of a more vivid full-body analogue of the rubber-hand illusion. Interesting future research might investigate how self-attribution and other mechanisms of self-identification interact, and how they can be empirically distinguished.

There were three notable limitations to the present study. First, there were many differences between the morphologically congruent (i.e., the mannequin) and the incongruent input medium (i.e., the touch screen), including differences in shape and color, the mapping of seen and felt touch, and the presence of text. Subsequent research is needed to determine more precisely which of these differences can explain the reported effects. Secondly, our experimental setup involved an actual interaction between two persons. Although this setup allowed for a more ecologically valid experimental situation, we could not ensure that the visuotactile stimuli were perfectly similar across sessions and participants. Our findings should therefore be confirmed with pre-recorded stimuli. A third limitation of the present experiment is that we have provided little evidence with respect to the role of self-attribution in media-related experiences, such as the phenomenon of telepresence. Interesting future research might, for example, determine whether the individual differences in people’s susceptibility for body image illusions are also reflected in their experience of telepresence.

Despite these limitations, our experiment demonstrates that combining touch and vision can alleviate, at least partially, the limitations of current mediated social touch devices. Visual feedback, if morphologically correct, renders the mediating technology more transparent. As a result, mediated touches by means of electromechanical actuators are perceived of as more natural and touch-like. Moreover, by increasing a person’s sense of telepresence, combining mediated touch with morphologically correct visual feedback can bridge, at least psychologically, the geographical distance between interaction partners. This, in turn, might aid in restoring the inevitable immediacy that marks natural unmediated physical contact.

Acknowledgements

The authors thank Dylan Schouten for assisting as a confederate, and Martin Boschman for assisting with the laboratory setup. Financial support from the EC through the PASION project (IST-6-27654) is gratefully acknowledged. Parts of this paper were presented at the 12th Annual International Workshop on Presence, November 11-13, 2009, Los Angeles, CA.

References

[1] S. Brave and A. Dahley. InTouch: A Medium for Haptic Interpersonal Communication. In S. Pemberton, editor, *Extended Abstracts of CHI 1997*, pages 363-364, 1997.

[2] A.F. Rovers and H.A. van Essen. Guidelines for Haptic Interpersonal Communication Applications: An Exploration of Foot Interaction Styles. *Virtual Reality,* 9: 177-191, 2006.

[3] A. Haans and W.A. IJsselsteijn. Mediated Social Touch: A Review of Current Research and Future Directions. *Virtual Reality,* 9: 149-159, 2006.

[4] A. Haans and W.A. IJsselsteijn. A Virtual Midas Touch: Helping Behavior after a Mediated Social Touch. *IEEE Transactions on Haptics,* 2: 136-140, 2009.

[5] G.C. Burdea. *Force and Touch Feedback for Virtual Reality.* John Wiley, New York, 1996.

[6] V. Hayward and K.E. MacLean. Do It Yourself Haptics, Part I. *IEEE Robotics and Automation Magazine,* 14: 88-104, 2007.

[7] S. Brave, C. Nass and E. Sirinian. Force-Feedback in Computer-Mediated Communication. In *Proceedings of HCI International, vol. 3,* pages 145-149, 2001.

[8] A. Haans, C. de Nood and W.A. IJsselsteijn. Investigating Response Similarities between Real and Mediated Social Touch: A First Test. In B. Begole, S. Payne, E. Churchill, R. St. Amant, D. Golmore and M.B. Rosson, editors, *Extended Abstracts of CHI 2007,* pages 2405-1410, 2007.

[9] A. Gallace and C. Spence. The Science of Interpersonal Touch: An Overview. *Neuroscience & Biobehavioral Reviews,* 34: 246-259, 2010.

[10] W.A. IJsselsteijn. Towards a Neuropsychological Basis of Presence. *Annual Review of CyberTherapy and Telemedicine,* 3: 25-30, 2005.

[11] J.M. Loomis. Distal Attribution and Presence. *Presence, Teleoperations and Virtual Environments,* 1: 113-119, 1992.

[12] T.B. Sheridan. Musings on Telepresence and Virtual Presence. *Presence, Teleoperations and Virtual Environments,* 1: 120-125, 1992.

[13] B. Lenggenhager, T. Tadi, T. Metzinger and O. Blanke. Video Ergo Sum: Manipulating Bodily Self-Consciousness. *Science,* 317: 1096-1099, 2007.

[14] H.H. Ehrsson. The Experimental Induction of Out-of-Body Experiences. *Science,* 317: 1048, 2007.

[15] M. Botvinick and J. Cohen. Rubber Hands ‘Feel’ Touch that Eyes See. *Nature,* 391: 756, 1998.

[16] R. Held and N. Durlach. Telepresence, Time Delay and Adaptation. In S.R. Ellis, editor, *Pictorial Communication in Virtual and Real Environments* (2nd ed.), pages 233-246, 1991.

[17] A. Haans and W.A. IJsselsteijn. Self-Attribution and Telepresence. In L. Moreno and Starlab Barcelona S.L., editors, *Proceedings of Presence 2007*, pages 51-58, 2007.

[18] A. Haans. *The Embodied User: Corporeal Awareness and Media Technology,* Eindhoven University of Technology, Eindhoven, The Netherlands, 2010.

[19] K.C. Armel and V.S. Ramachandran. Projecting Sensations to External Objects: Evidence from Skin Conductance Response. *Proceedings of the Royal Society of London B,* 270: 1499-1506, 2003.

[20] M. Tsakiris and P. Haggard. The Rubber Hand Illusion Revisited: Visuotactile Integration and Self-Attribution. *Journal of Experimental Psychology: Human Perception and Performance,* 31: 80-91, 2005.

[21] A. Haans, F.G. Kaiser, W.A. IJsselsteijn and D.G. Bouwhuis. *Vividness of the Rubber-Hand Illusion: A Function of Individual Susceptibility, Processing Demand, & Situational Impediments.* Submitted for publication.

[22] A. Haans, W.A. IJsselsteijn and Y.A.W. de Kort. The Effect of Similarities in Skin Texture and Hand Shape on Perceived Ownership of a Fake Limb. *Body Image,* 5: 389-394, 2008.

[23] A. Haans and W.A. IJsselsteijn. "I'm always Touched by Your Presence, Dear": Combining Mediated Social Touch with Morphologically Correct Visual Feedback. In *Proceedings of Presence 2009* [CD-Rom], 2009.

[24] J. Lessiter, J. Freeman, E. Keogh and J. Davidoff. A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence, Teleoperations and Virtual Environments,* 10: 282-297, 2001.

[25] S. Kennett, M. Taylor-Clarke and P. Haggard. Noninformative Vision Improves the Spatial Resolution of Touch in Humans. *Current Biology,* 11: 1188-1191, 2001.

[26] P. Haggard. Just Seeing You Makes Me Feel Better: Interpersonal Enhancement of Touch. *Social Neuroscience,* 1: 104-110, 2006.

[27] R. Thomas, C. Press and P. Haggard. Shared Representations in Body Perception. *Acta Psychologica,* 121: 317-330, 2006.

[28] G. Moseley, T. Parsons and C. Spence. Visual Distortion of a Limb Modulates the Pain and Swelling Evoked by Movement. *Current Biology,* 18: 1047-1048, 2008.

[29] H.H. Ehrsson and V.I. Petkova. E-letter response to: H.H. Ehrsson. The Experimental Induction of Out-of-Body Experiences. *Science,* 317: 1048, 2007.

[30] N. Bischof. *Das Rätsel Ödipus* [Oedipus's Riddle]. Piper Verlag, München, Germany, 1985.

[31] D. Bischof-Köhler. The Development of Empathy in Infants. In M.E. Lamb and H. Keller, editors, *Infant Development: Perspectives from German Speaking Countries,* pages 245-273, 1991.

Breathe with the Ocean: a System for Relaxation using Audio, Haptic and Visual Stimuli

Esko O. Dijk   Alina Weffers

Philips Research, High Tech Campus 34, Eindhoven, The Netherlands

Abstract

In this paper we present the "*Breathe with the Ocean*" system concept, which is a breathing guidance system that aims to help a user relax. It provides an immersive experience where the user is virtually present at an ocean shore. We describe the design and implementation of three embodiments of this concept and preliminary evaluations based on trial sessions. The feedback provided by the three systems to the user is in the form of audio, haptic (tactile) and visual (light) stimuli. Haptic stimuli are provided through a novel actuation device that we developed, the Touch Blanket. The three systems introduced are respectively a fixed-rate breathing guidance system, an adaptive breathing following system and an adaptive-rate breathing guidance system that maximizes heart-rate variability amplitude. We discuss the advantages and disadvantages of using open-loop versus closed-loop implementations of these types of systems, as well as our experiences so far in using multimodal stimuli for breathing guidance.

1 Introduction

Over the past decades numerous studies have been undertaken to determine the factors that cause relaxation at a psycho-physiologic level and the best methods to induce it. At the same time, a plethora of relaxation products (e.g. [1] [2] [3]) appeared on the market: some of these substantiated by scientific studies, others simply implementing one's idea of what might "feel good". We believe that in building a successful system for relaxation, one has to carefully choose the *stimuli* to be rendered, and the *policy* of rendering the stimuli.

Regarding the stimuli, several studies [4] have highlighted the potential positive effects of haptic (tactile) stimulation on the human body, emotions, mood, and health. Nevertheless tactile stimuli are not the only stimuli helpful for inducing relaxation. An interesting question is whether additional stimuli (i.e. sound or light) synchronized with haptic stimuli can emphasize their effect. A related question is whether a combined stimulation can induce relaxation more effectively than a single type of stimuli can do.

Other relevant questions are related to the activation time(s), intensity and duration of the stimulation itself for the purpose of relaxation. For example, should the stimuli be applied according to a predefined, fixed policy? Or should the stimulation be adaptive, responding to the current state and the personal needs of the user? And, in both cases, what should be the *policy* that determines the application of stimuli, in terms of modalities used, activation time(s), intensity and duration?

In this paper we aim to present and compare a number of system embodiments of our concept "*Breathe with the Ocean*", created for the purpose of relaxing a user. We explore three different system implementations, describing their advantages and disadvantages that were revealed during initial trials. The systems we present all implement a form of breathing guidance, a known technique to help people to relax [5] [6]. Breathing guidance means indicating to a user, during a session, in some way how the user should breathe in order to obtain beneficial effects for health or increase feelings of well-being. In our systems, breathing guidance is offered to the user by means of haptic, audio and light stimuli.

The main reason why we investigate a multi-sensory form of breathing guidance is our working hypothesis that involving more of the senses in a relaxation experience will increase both enjoyability and effectiveness of a relaxation session. For example, one aspect is that a guiding stimulus could be easier to follow if it is simultaneously presented through two or more modalities.

We organize our presentation as follows. Section 2 first presents the "Breathe with the Ocean" concept briefly, followed by a review of related work in Section 3 which also is intended to justify our choice of stimuli and choice for the technique of guided breathing. Section 4 presents the design and implementation of three embodiments of this concept, including the technologies employed and the chosen policies for rendering the stimuli. Section 5 describes our initial evaluation of the three embodiments, including our experiences so far in using multimodal stimuli for breathing guidance, followed by the conclusions section.

2 Concept

The “Breathe with the Ocean” concept entails a system that induces relaxation by providing a multi-sensory experience, where a user is virtually present at the shore of an ocean where he/she can hear and feel the waves of the ocean “washing” over the body. The concept does not rely heavily on visual stimuli and therefore it allows a user to close the eyes at any time, which may further help to increase relaxation.

We used the analogy with the ocean due to its natural association for most people with a relaxing setting, and because we identified the opportunity to integrate breathing guidance into the ocean setting of rhythmic ocean waves. Breathing guidance can have an additional relaxing effect. At the same time we recognize that this is not the only analogy we could have implemented, other settings from nature could be used as well, while the concept here represents one option out of many.

In our concept, the waves of the ocean are used to relax the user and to provide breathing guidance to the user, while the surrounding space is illuminated with a relaxing visual ambiance fitting the theme and purpose of the exercise. The visual ambiance is characterized by low-intensity, gentle tones of colored light. All three systems we present in this paper implement an imitation of ocean waves by combining

1. a "haptic wave” moving up and down the body, rendered using a matrix of vibration motors
2. ocean wave sounds rendered to audio headphones (or audio speakers) in synchrony with the haptic actuation.

In addition to the ocean wave sounds, a background relaxation music track can optionally also be played intermixed with the wave sounds.

The breathing guidance is linked to the haptic stimuli in the following way: when the haptic effect moves up from the feet towards the shoulders of the user, an activity of breathing in is implied. When the haptic effect moves down from the shoulders towards the feet of the user, an activity of breathing out is implied.

3 Related Work

To achieve the specific haptic stimulation which we refer to as "haptic waves", we make use of the *Touch Blanket* (see Figure 2 or [7]), which is a haptic actuation device developed within Philips that can provide personalized haptic effects to one or two users and can provide subtle haptic motion patterns, targeting specific body parts. The technology is the same as used in the work of Lemmens et al. [8].

Our rationale to use music is that various studies have shown that audio stimuli, in particular music can have a positive effect on relaxation. For example in [9] the authors show the influence of music on the level of nitric oxide and how this molecule is to a large degree responsible for physiological and psychological relaxing effects of listening to music. In [10] the authors show that although there are wide variations in individual preferences, music seems to directly affect the autonomic nervous system and as a consequence has further effects at a physiological level. Their tests have shown that music reduces anxiety and improves mood for patients in various situations such as while in intensive care units, while undergoing procedures or while receiving palliative care. In [11] the authors studied a “music enhanced therapy” which combines music and tactile stimulation embedded in a recliner, with traditional counseling techniques. The therapy was found to benefit musicians in that it reportedly reduced anxiety, improved mood, and reduced performance anxiety.

There is also evidence that haptics in the form of vibrations can induce relaxation. Wigram [12] reports an experiment with vibro-acoustics used on 60 healthy subjects (30 male, 30 female). Three conditions were evaluated: music + vibrations, music only, and the control condition (no stimulation).

The UWIST-MACL (MACL = Mood adjective check list) was used before and after the session to measure self-reported mood, and blood pressure and heart rate were measured. The self-reports show a significant decrease of *arousal* scores for the music + vibrations condition, compared to music alone or to the control group. In turn, the music-only condition shows a smaller but still significant decrease in arousal compared to the control condition. For blood pressure and heart rate, no significant results were found except that heart rate for participants in the two groups *music* and *music + vibration* combined was significantly lower than for participants in the control group.

Our work also builds on research conducted on the topic of guided breathing. Guided breathing exercises are linked to beneficial physiological effects related to relaxation such as decreased respiration rate (RSP rate) and increased Heart Rate Variability (HRV) amplitude. The connection between RSP rate and HRV amplitude, namely that guiding a user's breathing to progressively lower RSP rate leads to higher levels of HRV amplitude is documented in a number of literature studies [13] [14] [15] [16].

Many breathing guides available are quite straightforward, simply guiding the user at a fixed six respiration cycles per minute (c/min) according to a scientific finding [13] that around this rate *on an average person* the HRV amplitude is increased to a maximum. On the market, more notable products that offer breathing guidance for the purpose of relaxation are *emWave* developed by *HeartMath* [17] and *RESPeRATE* [18] developed by *Intercure.* Both products use a single sensor (respectively one for HRV and one for respiration) to adapt the breathing guidance rate based on the current state of the user.

Similar to *emWave* and *RESPeRATE* we also make use of HRV amplitude sensing and respiration sensing. The difference with existing systems is however that our breathing guide allows multimodal (haptic, audio and visual) stimulation, providing a whole experience aimed at inducing relaxation. The user in our case lies down on the Touch Blanket while hearing music and the sound of waves, while being surrounded by a visually relaxing ambiance. In contrast *emWave* is a small portable device which provides functional audio feedback and functional visual feedback by means of colored LEDs. *RESPeRATE* provides music feedback to guide respiration plus very simple indications in a LCD display. It is implemented as a portable device as well.

In comparison, the adaptive systems that will be presented by us in sections 4.3 and 4.4 provide multi-sensory stimulation and feedback, including tactile stimulation. Additionally the system described in section 4.3 takes a different approach than existing guides, by *following* (rather than just guiding) the natural breathing pattern of the user, while still inducing the user to progressively lower his/her respiration rate. In this case the system gives a form of bio-feedback, which allows the user to see/hear/feel whether their respiration lowers in rate and whether respiration is regular.

4 Design and Implementation

In this section we describe the design and implementation of three systems embodying the “Breathe with the Ocean” concept. Each system can be classified as either open-loop or closed-loop. We first describe the characteristics of each of these classes. Thereafter we describe the design and architecture of one open-loop and two closed-loop systems that all provide haptic, audio and visual feedback to the user.

4.1 Open-loop versus closed-loop

Open-loop systems execute according to a predefined strategy, and do not change their execution based on input(s) related to the state of the user. In other words, open loop systems are non-adaptive, usually a “one size fits all” type of system. In contrast, closed-loop systems are systems that adapt their execution according to input related to the state of the user. In particular, the closed-loop system in Section 4.3 includes the following components that make it closed-loop:

1. A measuring component which determines the current levels of different bio-signals or bio-signal features such as respiration rate, heart rate, or HRV amplitude;
2. An actuation component that renders the guided respiration rate by means of audio, haptic and/or visual feedback. The guided rate is based on the measurements.

In addition, the closed-loop system in Section 4.4 includes:

3. A component that interprets/assesses the relaxation state of the user, based on the measured and analyzed bio-signals;
4. A component that determines the user's personal and optimal respiration rate based on the data obtained by the previous component.

4.2 Embodiment 1 – fixed rate, open-loop breathing guide

The first embodiment of the “Breathe with the Ocean” concept that we built was an open-loop system. The usage scenario of this system is that while the user lies down on the Touch Blanket (see Figure 2), a haptic effect is rendered starting from the feet of the user towards the head of the user, and then back again. This cycle then repeats for the duration of the session. As mentioned before, these "haptic waves" are synchronized with audio samples of approaching and retreating ocean waves on a shoreline.

The intention of this system is that both the audio signal and the tactile stimulation provide breathing guidance to the user. Therefore, the audio/haptic waves occur at a frequency appropriate for a relaxed breathing pace. According to literature [13], a typical respiration rate at which human beings achieve maximum values in HRV amplitude – which is associated with relaxation – is at 6 breathing cycles per minute (c/min). Therefore, in this implementation the waves are rendered at the fixed, predefined rate of 6 c/min. A straightforward addition for this system could be a knob or slider to let the user easily adjust this rate.

The visual stimulus is provided by four Philips *LivingColors* [19] lamps set on a yellow/orange color intended to imitate the colors of sand and sun. Other color options such as blue to imitate the color of the sky or the ocean are possible as well - the current color is just one of many alternatives that we chose based on our own preference. The light intensity and color are fixed during the relaxation session.

The Touch Blanket exterior and interior are shown in Figure 2. On the outside, it looks like a conventional blanket and it can be conveniently spread on top of a seat or sofa. Inside, it contains 176 small, individually controllable small vibration motors (~1cm diameter), arranged in a 2D matrix. Each motor can be set to 25 different intensity levels with an update period of less than 50 milliseconds. The frequency and intensity of the motors are not separately controllable, because for this type of vibration motor higher intensity is directly linked to higher rotation speed and hence to a higher vibration frequency. Figure 3 shows how a single user can lie on the blanket, with (optionally) the top half of the blanket on top of his/her body to have haptic stimulation all around the body. Figure 4 shows a simplified system architecture diagram of the entire system.

Figure 1: Touch Blanket spread over a sofa. The blanket consists of two separate halves, one for the back area and one for the seat area.

To control the Touch Blanket, a C++ PC application was developed that generates tactile effects synchronized to the audio of approaching and retreating waves. This application uses effect scripts specifying what tactile effects are to be played when. In our experience so far and through user feedback we received, we learnt that what constitutes a good tactile experience, is a matter of personal taste. Especially, different intensities of haptic effects were preferred by different people. Therefore, in general a suitable UI to allow personalization of haptic effects is valuable to have. In our current system, only a basic slider control is available for users to adjust the intensity of effects between levels 0-25.

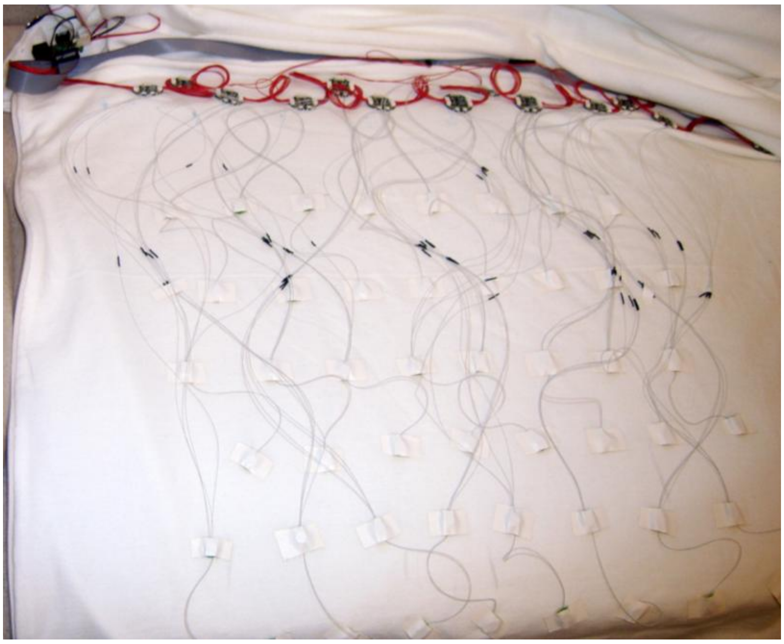

Figure 2: View inside the Touch Blanket, showing the wiring and part of the matrix of vibration motors. The motors are attached to the fabric with white adhesive tape.

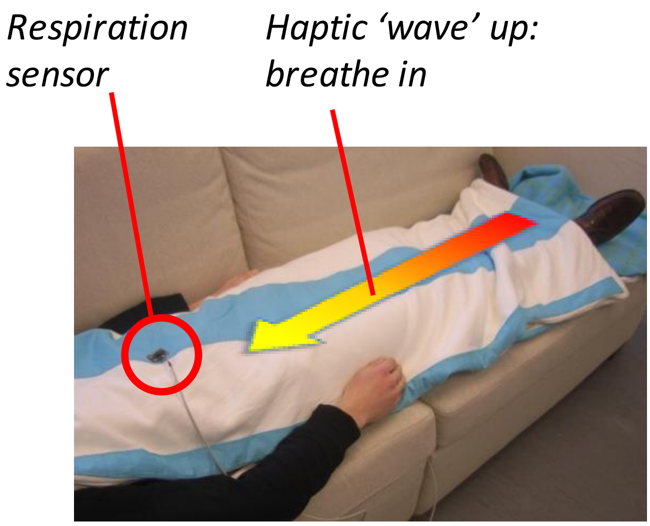

Figure 3: User lying on a sofa wrapped in the Touch Blanket, covering both back and front of the body. This provides haptic effects on both sides of the body. The respiration sensor (accelerometer) used in system embodiment 2 is placed at the abdomen. The arrow indicates a "haptic wave" effect travelling from the feet to the shoulders.

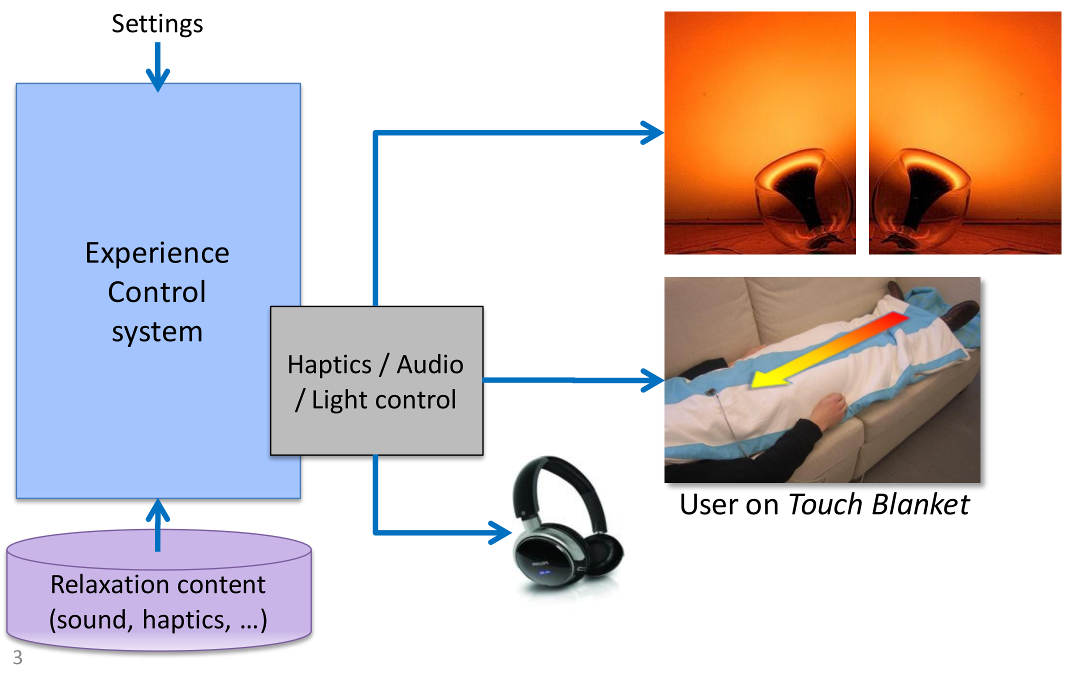

Figure 4: Simplified system diagram of embodiment 1, open-loop audio-haptic breathing guide

4.3 Embodiment 2 - Following the user's respiration pattern

Similar to the open-loop system described above, also in this embodiment the user lies down on the Touch Blanket, while haptic waves are rendered from the feet to the head of the user and vice-versa. Also here, the haptic waves are synchronized with audio content. Figure 5 shows a simplified system architecture diagram of the system.

In contrast to the previously described system, this implementation is a closed-loop system. It adjusts the stimulation based on the user’s current respiration rate and inhale/exhale pattern. Specifically, the haptic, audio and light effects are adjusted to *follow* the user's respiration. In this case the user is not forced to follow a fixed rhythm provided by a guide but can freely choose a natural inhale/exhale breathing pattern.

The idea here is that the system should not guide directly, but rather provide through the multi-sensory stimuli a form of *bio-feedback* to help the user become aware of how they breathe: whether they succeed to slow down their breathing during the exercise and whether they breathe regularly as intended by the exercise. For example, a user who would breathe quickly or irregularly would also perceive quick and irregular audio, haptic and light stimuli which we believe would be typically perceived as annoying by a user (based on our experience from a previous, yet unpublished study). This may provide a motivation to the user to breathe slower and more regular, thereby receiving the "reward" of calm and predictable stimuli from the system.

The feedback is provided as follows: a haptic wave is rendered from feet to head whenever the user inhales, and from head to feet whenever the user exhales. The time duration of the haptic wave rendering in either direction is adjusted to fully match the inhale/exhale respiratory cycle. Given the fact that the audio content is synchronized with the haptic effects, it follows that the sound of ocean waves matches the user's respiration rate and pattern as well.

Another difference compared to the previous system, is that this implementation provides feedback in the form of light, whereas previously the light was used purely for ambiental purposes. The visual feedback we provide here is given by varying the light intensity of the *LivingColors* lamps while the user performs the breathing exercise: the light intensity increases when the user inhales, and the intensity decreases when the user exhales.

As a result of the bio-feedback provided by the multimodal rendering we meant the user to be able to implicitly recognize the level of his/her performance while executing the breathing exercise. In response, the user should be learn over time to adjust his/her breathing accordingly. It is likely that spending more time on the breathing exercise will improve a user's skill in the technique. Once the user is able to breathe slowly and deeply with minimal effort, optimal relaxation may be achieved. The current implementation does not provide an explicit indication on the user's performance level, i.e. how close the respiration rate is to the optimal rate or how regular the breathing pattern is.

The LivingColors lamps that provide the light effect are wirelessly controlled using a proprietary RF protocol. This protocol allows a single controller to control the light intensity and color selection for one or more (up to four) lamps.

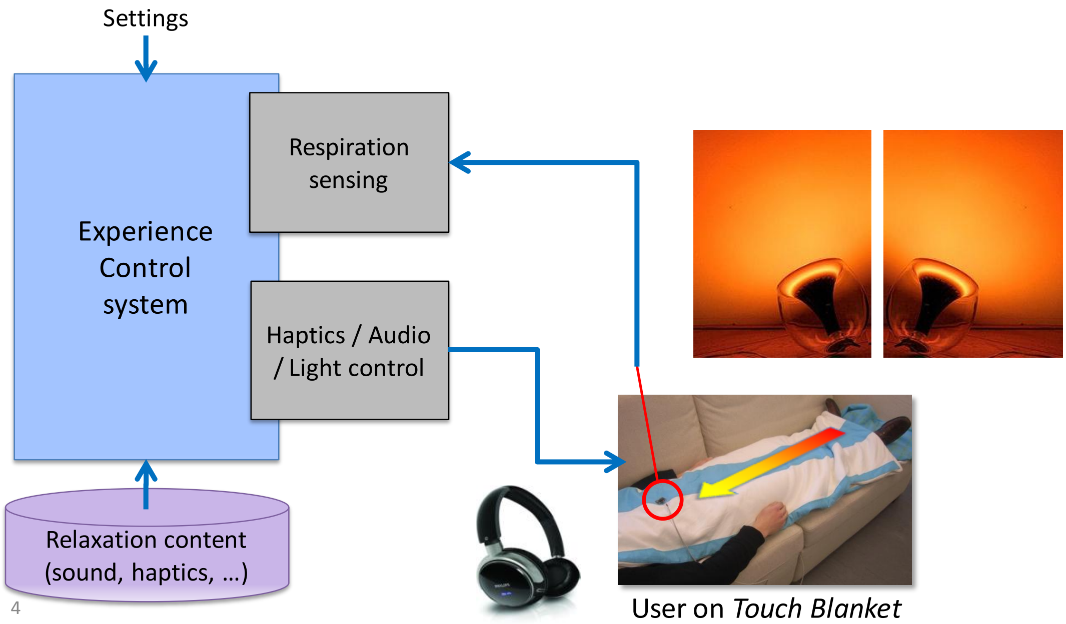

Figure 5: Simplified system architecture of closed-loop breathing following (bio-feedback) system

4.4 Embodiment 3 - adaptive breathing guide for maximizing heart-rate variability

Several publications ([13], [14], [15], [16]) show that guided breathing at steadily decreasing breathing guidance rates induces a lower respiration rate of the user and higher values of the HRV amplitude, which is linked to states of relaxation. This work also shows that for each user there exists a personal, optimal respiration rate which induces maximum HRV amplitude.

The disadvantage of fixed breathing guide systems (including our embodiment 1) is that these cannot maximize the HRV amplitude for each user individually, because they offer a single, fixed guidance rate which is only optimal on average for the population. Also, our embodiment 2 lets users freely choose their respiration rate which may not be close to the personal optimum at all. Given both arguments above, we decided to develop a fully adaptive closed-loop system that guides the user's breathing according to an adaptive policy intended to maximize HRV amplitude.

The solution we propose can detect in an adaptation phase, for each user, which respiration rate induces the highest HRV amplitude, and then offers that specific respiration rate during a guidance phase. In this system, again the user lies on the blanket as before in the two other embodiments. The system consists of a number of hardware and software components; see also Figure 6. The bio-signal values (RSP Rate, HRV) are measured using a portable Nexus-10 device with the following sensors attached: a photoplethysmograph (PPG) for measuring the heart signal (in the form of finger blood-volume pulse - BVP), and a respiration belt used to measure the respiration signal.

On a separate PC are running a data acquisition application, a Matlab application to analyze for real-time analysis of the bio-signals, and a C++ application for controlling the Touch Blanket. The Nexus sends its data to the PC via a Bluetooth connection. The Matlab application receives in real-time the raw bio-signal data from the data acquisition application, and then calculates in real-time the bio-signal features for respiration (such as mean respiration rate) and HRV. The features are calculated every 10 seconds, averaged over a window over the last 30 seconds. The Matlab application transmits every 10 seconds the feature values to the C++ client via a TCP/IP socket.

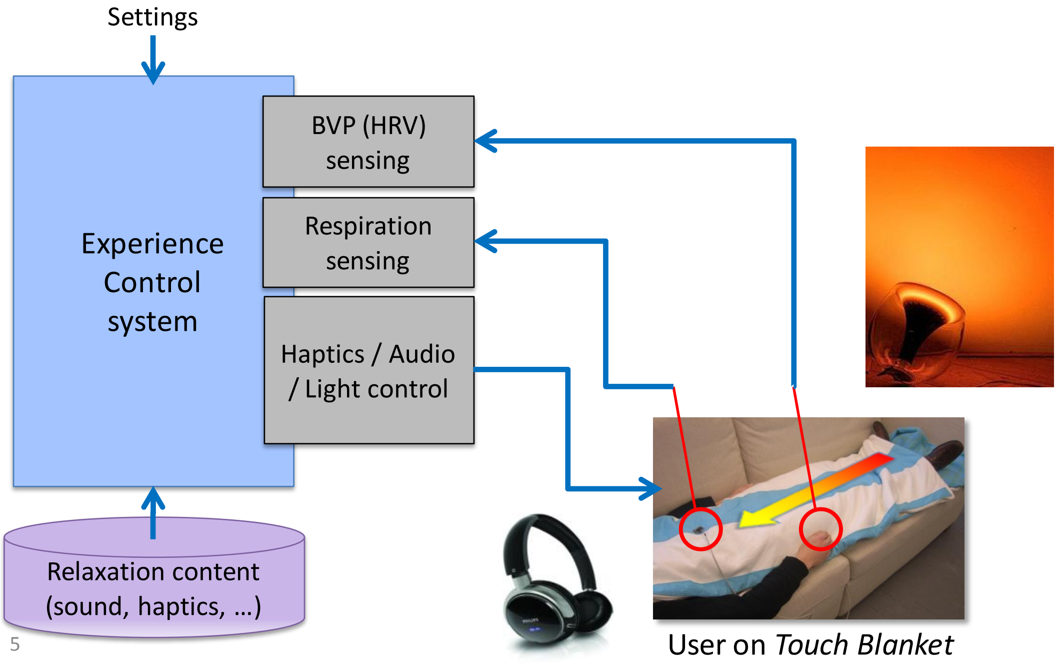

Figure 6 – Simplified system diagram of embodiment 3, a closed-loop adaptive breathing guide that uses respiration and HRV sensing to detect the optimal user respiration rate. The heart signal (BVP) is measured at the finger through a photoplethysmograph (PPG) clip-on sensor.

The C++ application receives in real-time the bio-signal data features from the Matlab application, calculates the optimal respiration rate based on the history of values of bio-signal features, and controls the Touch Blanket, light and audio which constitute the breathing guidance. Notable here again is that the current implementation does not provide explicit indications of the user's performance level, however such options could include pleasant sounds or vibrations when the user succeeds to follow the guide or when the exercise was successful.

Algorithm

The algorithm of the closed-loop breathing guide implements a strategy for guiding the user towards lower respiration rates, while constantly checking the effect of each progressively lower respiration rate on the HRV amplitude values which we intend to maximize. For each new respiration rate that the guide provides, the user is given in total at most 2 minutes time to adapt and follow the guide (initially 60 seconds, and 60 seconds for a retry only if necessary). The 60 seconds have been chosen as a threshold given that previous experiments with breathing guidance showed that 60 seconds is usually enough time to determine whether a user follows the guide, and in total 2 minutes are enough to determine if the user is able to follow the guide at the current respiration rate, or not. We allowed an error margin of 0.5

cycles/min given that usually users will seldom be able to follow the guide at the *exact* guided respiration rate. In our experience 0.5 cycles/min is an acceptable, realistic margin. Whenever the user is able to follow the guide at a specific respiration rate for 60 seconds, the system reduces the guided respiration rate with a step of 1 cycle/min. We used 1 cycles/min here to provide a smooth, gentle accommodation to a lower respiration rate, while also allowing an accurate enough identification of the respiration rate that renders the maximum HRV amplitude values. The algorithm can be summarized by the following steps:

1. **Adaptation phase:** Initially the system determines the respiration rate (RSP Rate) of the user *UBR* (user breathing rate) before the guided breathing exercise (i.e., the baseline) during a period of 30-60 seconds and sets the guide's breathing rate (*GBR*) to the initial user breathing rate UBR.
2. While GBR $\geq$ 5
    a. Over the following period of 60 seconds, while the user follows or tries to follow the guide with the rate GBR, the system collects 6 times the values of both RSP Rate and HRV Amplitude. Also the average UBR is again sampled.
    b. If the user can follow the guide (criterion: yes if UBR $\leq$ GBR + 0.5)
        - The GBR is decreased with a STEP value (measured in breathing cycles/min). Typically STEP = 1 c/min was used.
    c. If the user cannot follow the guide, allow a single retry by repeating step 2-a.
        - If retry successful, continue algorithm normally at step 2-b.
        - If retry *not* successful, abort the loop and continue at step 3.
3. Analyze all collected values so far for RSP Rate and corresponding HRV amplitude (if any), and set GBR to the RSP Rate which gave the maximum HRV amplitude.
4. **Guidance phase:** In the following 10 minutes continue a guided breathing exercise offering to the user the optimal rate GBR.

5 First evaluations

In this section we describe the results of the first evaluations we conducted for each of the previously described embodiments. Per embodiment, the insights we describe here have been inferred from discussions with users who tried the system (including the authors and their colleagues), mainly after informal trial practice sessions and during exhibition displays of these systems. We discuss the advantages and disadvantages of using open-loop versus closed-loop systems as well as our experiences so far in using multimodal stimuli for breathing guidance.

5.1 Embodiment 1 - fixed breathing guide

The first embodiment was tested with about 30-40 users, most of which visitors during a Philips-internal exhibition where the setting did not allow us to use the light stimulus. Our hypothesis for this evaluation was that users would appreciate the combined audio-haptic stimuli and that the breathing guidance (if followed properly) would increase their feelings of relaxation.

The open-loop system described in section 4.2 has the advantage of being easier to implement robustly due to its fixed, predefined execution and absence of on-body sensors. The disadvantage is the lack of automatic personalization. Indeed, some of the users found it difficult to follow the breathing guidance. Unfortunately the feedback related to this problem covered quite a range – from users who found the guide too fast, to those who found it excessively slow, to those who would like another balance in the relative times spent in inhaling/exhaling. This means that a quick fix by simply changing the breathing guidance frequency will not solve these problems.

Could however this issue be ignored? Unfortunately, not. The mismatch between the provided pace and the natural respiration rate of the user can lead to uncomfortable feelings caused by the difficulty to follow the guide, which ultimately diminishes the user experience and overall lowers the chance that the user will get relaxed. While some may argue that slow breathing is unnatural to people and is a skill that needs to be learned, an adaptive system could guide this learning process as well so that a user, in each session, can reach a slow respiration rate that they are still comfortable with. Adaptive guides can identify and take into account the user's personal lowest boundary at the time, whereas fixed guides cannot. Moreover, in the case of users who breathe too slow or too deep when following a fixed guide, the consequences can be even more unpleasant including potential hypoventilation, hyperventilation or dizziness. These considerations could indicate that open-loop breathing guide systems are not the best solutions for providing relaxation.

Considering the multi-sensory stimuli, the user comments revealed that most users found the synchronization between the haptic and the audio waves quite pleasing and useful in creating the right atmosphere as well. By letting some users try it with and without headphones, we found that this sub-group of users agreed that the combination of audio and haptic modalities provides a much more immersive experience than haptics alone.

On the haptic effect itself, it was described by users often as a "strange" and "new" sensation, and usually also "pleasant". The users to which we explained the concept of the haptic effect simulating the motion of ocean waves around the body, all understood the concept. However we did not test yet if users could find out about this analogy for themselves, nor how realistic users thought that the effect was compared to real ocean waves. A few users found it "tickling" and a number indicated it to be "like a very light massage". Only two users explicitly disliked it for being too mechanical and reminding them of a buzzing mobile phone, even after giving them the option to adjust the level of vibration.

The visual stimulus, although being perceived to be relevant for setting the ambiance, was not seen as important. One reason for this is that it did not change during a session and partly due to the tendency of users to close their eyes while breathing and attempting to relax.

5.2 Embodiment 2 - following the user's breathing pattern

The second embodiment (described in section 4.3) was tested with about 8 users. The comments of users regarding the haptic effects and the synchronization between the haptic and the audio waves for this system were similar to the previous embodiment, so we will only focus here on the noteworthy differences with respect to embodiment 1. Our hypothesis for this system was that all problems mentioned above, related to users not being able to correctly follow the guide, would disappear and that users would have an enjoyable and relaxing experience. However, we also expected that part of the users would still choose a too-high breathing rate in the absence of explicit guidance.

The advantages of this closed-loop implementation come from the fact that the system allows users to breathe at their own pace, and use their own natural breathing pattern. At the same time the system provides feedback about the performance of the user by the simple fact that users can feel and hear whether the rate of the waves becomes slower and regular in time or not. However, the current implementation does not include an explicit indication of the user's performance level. A disadvantage is that obviously a closed-loop system is more difficult to implement than an open-loop. We found that the user's respiration pattern is not easy to determine accurately in real-time due to inherent noise in the accelerometer measurements and artifacts related to motion or talking of the user. Also, we still have to investigate whether the vibrations of the Touch Blanket disturb the measurements in any way.

This implies that potentially for some users the feedback is not provided appropriately at all times. On the other hand, some users argued that it is better to guide users in breathing exercises rather than trust them to improve their breathing on their own as is the philosophy of the current breathing-following system. Given this argument, we decided to develop also a fully adaptive closed-loop system (Embodiment 3) that guides the user's breathing according to an adaptive policy.

In practice the breathing following algorithm did not always work for all users, because only a simple "first version" algorithm was used which was not yet optimized based on actual breathing recordings of a larger number of users. However, for users that breathed slowly (< 9 cycles/min), regularly and deeply and did not move nor talk during a session, the breathing following algorithm always worked.

Some users commented on a perceivable delay (or "lag") between their breathing and the haptic/audio effect. One user reported that this small delay had a strong influencing effect on him: it caused him to breathe progressively slower, as if the breathing guide indicated to the user to "hold back" in tempo. This type of effect of small delays is already described in earlier work [20]. Therefore, it may be possible to subtly guide a user towards lower respiration rate even while the system follows the breathing pattern of the user. This would be an interesting topic for further research.

The visual feedback was perceived as interesting, however many of the users who tried the system had a tendency to close their eyes while breathing and attempting to relax. In this situation the visual cues are not perceived. Also one could argue that too many feedback cues could be overwhelming and counterproductive for relaxation. For example, some users who did try to pay attention to all three feedback cues (haptic, audio and visual) found that it required quite a lot of concentration which made it rather hard to actually relax. Another possible cause for these reports could be slight (perceived) timing asynchronies between the various stimuli, but we did not further investigate this.

One last point is that choosing the right color to be rendered by the lamps may be very important. For example some suggested that to set a maritime ambiance blue light should be used, whereas studies in literature [21], [22], [23] indicate that blue light has a general energizing effect. Therefore blue light may be counter-productive to the purposes of a relaxation session. On the other hand, it may stimulate feeling awake and refreshed after a session and in this way add to a positive outcome of the relaxation session. Also personal preference plays an important role in what color of light a user can relax comfortably with. Other options are in simulating naturally occurring situations (e.g. sunset, dawn) or perhaps slowly changing the color to guide the user to a desired end state with associated end color.

5.3 Embodiment 3 - fully adaptive breathing guide for relaxation

The third embodiment (described in section 4.4) was tested with about 5 users, but not yet tested in a fully functional state, due to technical problems. Our hypothesis for this system, which we did not test yet, was that we could expect some problems with not being able to follow the guide during the first (adaptation) phase but that this would be compensated for in the guidance phase where the user would be guided with a personal optimal rate.

As with most closed-loop systems the advantage of this system comes from the fact that it offers a highly personalized solution to users. This is facilitated by the underlying algorithm which should lead to an increase of the HRV amplitude for those users that follow the guide. The current system supports only equal time durations for the inhale and exhale phases, which turned out to be not natural for some users and can become stressful when trying to precisely follow the breathing guide.

During initial evaluation of the system we also discovered a problem in the real-time estimation of the HRV amplitude values, which has not yet been solved to date. This problem caused estimated HRV amplitude values to be higher than the 'actual' HRV amplitude values (i.e. HRV amplitude estimates from a recorded BVP signal calculated afterwards off-line) during the first minutes of operation of the system. As a consequence, the selected breathing guidance rate based on the HRV amplitude was often chosen too high. Therefore we could not continue to test the system more extensively.

5.4 Other evaluation results

A general problem was found with the breathing guide embodiments 1 and 3, which both force a fixed breathing pattern on a user during a period of time. The problem occurs when a user for some reason gets "out of sync" i.e. "out of phase". This can be caused by an occasional deep sigh by the user, or an occasional urge to speed up or slow down a breathing cycle. A possible cause for this user behavior is the need of the body to regulate ventilation to optimal levels of oxygen inflow and carbon-dioxide outflow. In such cases the user's breathing starts to become misaligned to the guidance. The guidance algorithm, in the current implementations, does not detect such cases in order to adapt the guidance back to the user. This requires the user to "catch up" with the system by unnaturally quickening breathing pace, slowing it down, or by holding breath for a while. A fair amount of users find this situation stressful and not contributing to relaxation.

6 Conclusions

6.1 General

In this paper we have presented the "*Breathe with the Ocean*" concept, which envisions a breathing guidance system that induces relaxation by providing an immersive experience where the user is virtually present at an ocean shore. Haptic, audio and light effects are used to convey this experience and to guide the user's breathing. We have described the design and architecture of three embodiments of this concept, and gave a preliminary evaluation based on trial sessions. We discussed the advantages and disadvantages of using open-loop versus closed-loop implementations for these types of systems, as well as our experiences so far in using multimodal stimuli in this application. Although we have verified that multimodal stimuli can be used to create a compelling user experience, we did not compare our approach to other e.g. unimodal approaches in a systematic way.

6.2 Evaluations

Open-loop systems have the advantage of being easiest to realize due to the fact that they implement fixed policies of rendering stimuli and do not rely on bio-sensor input, which in practice adds a source of error to a system. The disadvantage is that these systems do not allow on-the-fly automatic personalization.

Our evaluations show that in the case of breathing guidance systems, a lack of personalization appears to be a significant drawback. This is due to the fact that different users have quite different inhale/exhale patterns and optimal respiration rates. Imposing respiration at a rate and pattern to which users cannot adapt comfortably can induce dizziness, hypoventilation or hyperventilation.

In contrast, closed-loop systems can provide personalization to a certain degree. For this reason the feedback from the users was more positive for the closed-loop implementations, in comparison to the open-loop system, when looking at the breathing guidance aspects. Naturally, as a result of having adaptive policies and requiring sensors, these systems are more complex to implement and test. The most critical and challenging part in providing a truly adaptive guidance is estimating the correct user breathing rate, phase and pattern over time. Pattern here includes breathing depth, inhale/exhale ratio and the durations of the pauses in the breathing cycle.

Regarding the sensory stimuli, the user evaluations revealed that most users found the haptic feedback pleasing, even more so if combined with the audio stimuli. Also they found the synchronization between the haptic effect and the audio ocean wave sounds quite pleasing and useful in setting the atmosphere. The visual stimulus, although being perceived to be relevant for setting the ambiance, did not seem to be as important for providing feedback. This was probably due to the tendency of users to close their eyes while breathing and attempting to relax.

In our view, adapting the rendering to the user in the correct manner is crucial to enabling a pleasant user experience. Failure to do so – even if only during a few moments within a session – can ruin the entire experience.

6.3 Future work

For future research we are considering a number of directions, such as

1. Investigating under which circumstances audio or visual stimuli can significantly increase the enjoyment and effectiveness of haptic stimuli, in the context of breathing guidance systems.
2. Comparing different strategies for breathing guidance in terms of effectiveness and enjoyability. E.g. do we need a system with "out of sync" recovery? Is it perhaps better to only guide the user during brief periods and then let a user breathe in his/her natural rhythm for a while? Are high values of HRV amplitude also linked to the most pleasurable experience, or not?
3. Performing formal user tests with the systems mentioned in this paper. To this end, the robustness and user experience of especially the closed-loop systems would still have to be improved. The initial findings in this paper provide a good lead on what needs to be improved.

Acknowledgements

This work was partially supported by the ITEA2 Metaverse1 (www.metaverse1.org) Project.

References

[1] http://www.inspiredliving.com/relaxation-products/

[2] http://www.photosonix.com/

[3] http://www.breathing.com/programs.htm

[4] E.O. Dijk et al., Audio-tactile stimuli to improve health and well-being – a preliminary position paper, *Proceedings of the EuroHaptics 2010 Special Symposium: Haptic and Audio-Visual Stimuli: Enhancing Experiences and Interaction*, Amsterdam, The Netherlands, July 7, 2010.

[5] R. Fried, Relaxation with biofeedback-assisted guided imagery: The importance of breathing rate as an index of hypoarousal, *Applied Psychophysiology and Biofeedback Journal,* **12**(4) / December, 1987, Springer Netherlands, 2005.

[6] R. Fried, Integrating music in breathing training and relaxation: II. Applications, *Applied Psychophysiology and Biofeedback Journal*, **15**(2) / June, 1990, Springer Netherlands, 2005.

[7] E.O. Dijk, A. Weffers-Albu, T. de Zeeuw, A tactile actuation blanket to intensify movie experiences with personalised tactile effects, *Demonstration papers Proc. 3rd International Conference on Intelligent Technologies for Interactive Entertainment*, Amsterdam, The Netherlands, June 22-24, 2009.

[8] P.M.C. Lemmens, F. Crompvoets, D. Brokken, J. Van den Eerenbeemd, G.J. De Vries. A Body-conforming Tactile Jacket to Enrich Movie Viewing. In *Proc. WorldHaptics 2009*, pp. 7-12, March 18-20, Salt Lake City, UT, USA, 2009.

[9] E. Salamon, M. Kim, J. Beaulieu, G.B. Stefano, Sound Therapy Induced Relaxation: Down Regulating Stress Processes and Pathologies, *Med Sci Monit.* 2003 May;**9**(5):RA96-RA101.

[10] K. J. Kemper, S.C. Danhauer, Music as Therapy, *Southern Medical Journal*, Vol. 98, Issue 3, p. 282-288, March 2005.

[11] W. Brodsky, Clinical Trial of a Music Generated Vibrotactile Therapeutic Environment for Musicians: Main Effects and Outcome Differences Between Therapy Subgroups, *Journal of Music Therapy*, XXXIV, p. 2-32, 1997.

[12] A.L.Wigram. *The effects of vibroacoustic therapy on clinical and non-clinical populations*, PhD thesis, St. George's Hospital Medical School, London University, 1996.

[13] E.G. Vaschillo et al., Characteristics of resonance in heart rate variability stimulated by biofeedback, *Applied Psychophysiology and Biofeedback* **31**(2), p. 129-142, 2006.

[14] C.-K. Peng et al., Heart rate dynamics during three forms of meditation, *International Journal of Cardiology* **95**(1), p. 19-27, 2004.

[15] R. McCraty et al., Heart rhythm coherence feedback: a new tool for stress reduction, rehabilitation, and performance enhancement, *Proceedings of the First Baltic Forum on Neuronal Regulation and Biofeedback*, Riga, Latvia, 2004.

[16] R. Patterson et al., Voluntary cardio-respiratory synchronization, *IEEE Engineering in Medicine and Biology Magazine* Nov/Dec 2004, p. 52-56.

[17] http://www.heartmathstore.com/item/6300/emwave-personal-stress-reliever

[18] http://www.resperate.com/us/discover/resperatedemo.aspx

[19] http://www.lighting.philips.com/microsite/living_colors/

[20] B.M. Saperston, *Method for influencing physiological processes through physiologically interactive stimuli*, US patent 5,267,942, filed April 20, 1992.

[21] S.W. Lockley, E.E Evans, F.A.J.L. Scheer et al. Short-wavelength sensitivity for the direct effects of light on alertness, vigilance, and the waking electroencephalogram in humans. *SLEEP* 2006;**29**(2): 161-168.

[22] http://www.sciencedaily.com/releases/2008/10/081029105807.htm

[23] http://sleep.med.harvard.edu/news/28/Blue+Light+May+Fight+Fatigue

Identification of virtual grounds using virtual reality haptic shoes and sound synthesis

Stefania Serafin, Luca Turchet, Rolf Nordahl and Smilen Dimitrov
Medialogy, Aalborg University Copenhagen
Lautrupvang 15, 2750 Ballerup, Denmark
sts,tur,rn,sd@media.aau.dk

Amir Berrezag and Vincent Hayward
UPMC Univ Paris 06, Institut des Systèmes Intelligents et de Robotique,
4 place Jussieu,75005 Paris, France
amir.berrezag,vincent.hayward@isir.upmc.fr

Abstract

We describe a system which simulates in real-time the auditory and haptic sensation of walking on different surfaces. The system is based on physical models, that drive both the haptic and audio synthesizers, and a pair of shoes enhanced with sensors and actuators. In a discrimination experiment, subjects were asked to recognize the different simulated surfaces they were exposed to with uni-modal (auditory or haptic) and bi-modal (auditory and haptic) cues. Results show that subjects perform the recognition task better when using auditory feedback versus haptic feedback. The combination of auditory and haptic feedback only in some conditions significantly enhances recognition.

Keywords: walking simulation, physical models, audio-haptic interaction

1 Introduction

Research on multimodality has been focused on the interaction between vision and other modalities. However, lately the interest in investigating the interaction between other senses, such as touch and audition, has grown. This has been facilitated by the rapid development of haptic devices, together with efficient and accurate simulation algorithms. Several studies have indeed investigated audio-tactile phenomena such as the multimodal recognition of textures [11] and stiffness, both using physical stimuli [5] and simulations based on physical models [1].

The cited studies have focused on the interaction between touch and audition in hand-based interactions. To our knowledge, the interaction of auditory and haptic feedback in foot-based devices is still an unexplored topic. A notable exception is the work of Giordano et al., who showed that the feet were also effective at probing the world with discriminative touch, with and without access to auditory information. Their results suggested that integration of foot-haptic and auditory information does follow simple integration rules [9].

In previous research, we described a system able to recreate the auditory and haptic sensation of walking on different materials and presented the results of a preliminary surface recognition experiment [7]. This experiment was conducted under three different conditions: auditory feedback, haptic feedback, and both.

By presenting the stimuli to the participants passively sitting in a chair, we introduced a high degree of control on the stimulation. However, this method of delivery is highly contrived since it eliminates the tight sensorimotor coupling that is natural during walking and foot interaction. It is true for the auditory channel, but even more so for the haptic channel. In spite of these drastically constrained conditions, performance was surprisingly good.

In this paper, we extend such work by allowing subjects to walk in a controlled laboratory, where their steps are tracked and used to drive the simulation. We investigate whether introducing a higher level of interactivity will significantly enhance the recognition rates as well as the perceived quality and realism of the simulation.

2 Simulation Software

We developed a physically based synthesis engine able to simulate the auditory and haptic sensation of walking on different surfaces. Acoustic and vibrational signatures of locomotion are the result of more elementary physical interactions, including impacts, friction, or fracture events, between objects with certain material properties (hardness, density, etc.) and shapes. The decomposition of complex everyday sound phenomena in terms of more elementary ones has been an organizing idea in auditory display research during recent decades [8]. In our simulations, we draw a primary distinction between solid and aggregate ground surfaces, the latter being assumed to possess a granular structure, such as that of gravel.

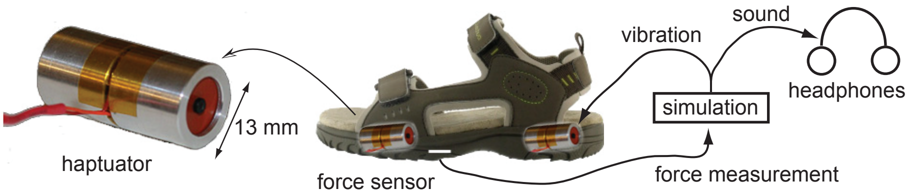

Figure 1: System (one shoe shown). Left: recoil-type actuation from Tactile Labs Inc. The moving parts are protected by an alumimum enclosure able to bear the weight of a person. Middle: approximate location of the actuators in the sandal. Right: system diagram showing the interconnections.

Figure 2: A picture of one pressure sensor and two actuators embedded in the shoes.

The impact model and its discretization are described elsewhere in detail [3]. The model has been recently adapted to the audio simulation of footsteps [12]. Here, we used the same model to drive the haptic and the audio synthesis. It is briefly recalled below.

A footstep sound may be considered to cause multiple micro-impacts between a sole, i.e., an *exciter*, and a floor, i.e., a *resonator*. Such interaction can be either discrete, as in the case of walking on a solid surface, or continuous, as in the case of a foot sliding across the floor.

In the simulation of discrete impacts, the excitation is brief and has an unbiased frequency response. The interaction is modelled by a Hunt-Crossley-type interaction where the force, f, between two bodies, combines hardening elasticity and a dissipation term [10]. Let x represent contact interpenetration and $\alpha > 1$ be a coefficient used to shape the nonlinear hardening, the special model form we used is

$$f(x, \dot{x}) = -kx^{\alpha} - \lambda x^{\alpha}\dot{x} \quad \text{if } x > 0, \quad 0 \text{ otherwise.}$$

The model described was discretized as proposed in [2].

If the interaction called for slip, we adopted a model where the relationship between relative velocity v of the bodies in contact and friction force f is governed by a differential equation rather than a static map [6]. Considering that friction results from a large number of microscopic damped elastic bonds with an average deflection z, a viscous term, $\sigma_2 v$, and a noise term, $\sigma_3 w$, to represent roughness, we have

$$f(z, \dot{z}, v, w) = \sigma_0 z + \sigma_1 \dot{z} + \sigma_2 v + \sigma_3 w.$$

The force specified by these models is applied to a virtual mass which produces a displacement signal that is then processed by a linear shaping filter intended to represent the resonator.

Stochastic parameterization is employed to simulate particle interactions thereby avoiding to model each of many particles explicitely. Instead, the particles are assigned a probability to create an acoustic waveform. In the case of many particles, the interaction can be represented using a simple Poisson distribution, where the sound probability is constant at each time step, giving rise to an exponential probability weighting time between events.

We used this approach to model both solid and aggregate surfaces. A solid surface is represented by an impact and a slide. The impact model alone was used to recreate the sound and the feel produced when walking on wood. The friction model was tuned to simulate walking on creaking wood. To simulate walking on aggregate grounds, we used a physically informed sonic models (PhiSM) algorithm [4]. The synthesis was tuned to simulate snow and forest underbrush.

These algorithms were implemented as an extension to the Max/MSP platform[1] to drive both the auditory and haptic feedback.

3 Simulation Hardware

In order to provide both audio and haptic feedback, haptic shoes enhanced with pressure sensors have been developed. A pair of light-weight sandals was procured (Model Arpenaz-50, Decathlon, Villeneuve d'Ascq, France). This particular model has light, stiff foam soles that are easy to gouge and fashion. Four cavities were made in the tickness of the sole to accommodate four vibrotactile actuators (Haptuator, Tactile Labs Inc., Deux-Montagnes, Qc, Canada). These electromagnetic recoil-type actuators have an operational, linear bandwidth of 50–500 Hz and can provide up to 3 G of acceleration when connected to light loads. As indicated in Fig. 1 and Fig. 2, two actuators were placed under the heel of the wearer and the other two under the ball of the foot. There were bonded in place to ensure good transmission of the vibrations inside the soles. When activated, vibrations propagated far in the light, stiff foam. In the present configuration, the four actuators were driven by the same signal but could be activated separately to emphasize, for instance, the front or back activation, to strick a balance, or to realize other effects such as modulating different, back-front signals during heel-toe movements.

The sole has two FSR pressure sensors[2] whose aim was to detect the pressure force of the feet during the locomotion of a subject wearing the shoes. The two sensors were placed in correspondence to the heel and toe respectively in each shoe. The analogue values of each of these sensors were digitized by means of an Arduino Diecimila board[3] and were used to drive the audio and haptic synthesis.

A cable came out from each shoe, with the function of transporting the signals of the pressure sensors and of the actuators. Such cables were about 5 meters long, and they were connected through DB9 connectors to two 4TP (twisted pair) cables: one 4TP cable carried the sensor signals to a breakout board-which contained trimmers, that formed voltage dividers with the FSRs, which then interfaced to an Arduino board. The other 4TP cable carried the actuator signals from a pair of Pyle Pro PCA1[4] mini 2X15 W stereo amplifiers, driven by outputs from a FireFace 800 soundcard.[5] Each stereo amplifier handled 4 actuators found on a single shoe, and each output channel of the amplifier drove two actuators connected in parallel. The PC handled the Arduino through a USB connection, and the FireFace soundcard through a FireWire connection.

[1] www.cycling74.com
[2] I.E.E. SS-U-N-S-00039
[3] http://arduino.cc/
[4] http://www.pyleaudio.com/manuals/PCA1.pdf
[5] http://www.rme-audio.com/english/firewire/ff800.htm

4 CONTROLLING THE SYNTHESIS ENGINE

The audio-haptic synthesis engine is controlled by a ground reaction force (GRF), i.e., a force exerted by the ground on a body in contact with it. We first created a database of GRFs, by extracting the amplitude envelopes from different audio files of recorded footsteps, chosen among those available on the Hollywood Edge sound effects library.[6] Figure 3 shows the waveform of one of the footstep sounds used and its corresponding extracted GRF.

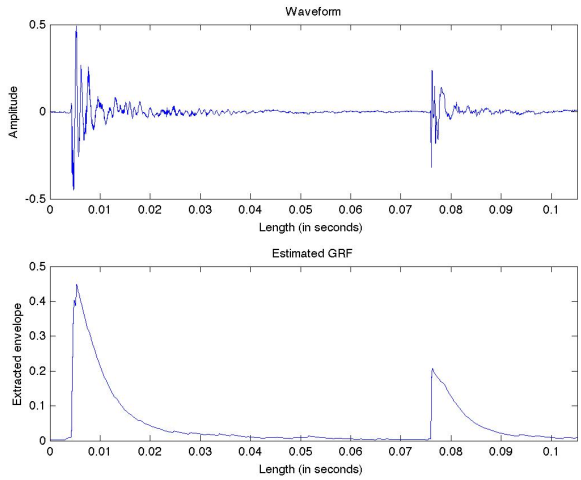

Figure 3: Time domain waveform of a footsteps sound (top) and its corresponding extracted GRF (bottom).

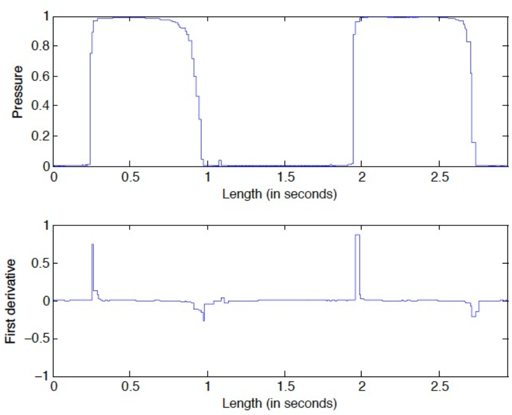

Figure 4: Waveform of a characteristic signal captured by the pressure sensors (top), and its derivative (bottom).

We then used the data captured by the pressure sensors and calculated their first time derivatives, which are related to the intensity with which the foot hits the ground. An example of data captured from the pressure sensors and its corresponding derivative is shown in Figure 4. The values of the first derivative were used as control for triggering the footsteps synthesizer, some GRFs corresponding to heel or toe

[6]www.hollywoodedge.com/

according to the activated sensor. Five types of heel and toes audio files were used for the purpose, and randomly chosen at the moment of triggering, giving rise to 25 possible combinations.

Each time the value of the first derivative became bigger than a threshold, the GRF corresponding to the activated sensor was triggered into the engine. More precisely we checked only positive changes in the derivative value, since we were interested in the generation of the sound when the step hits – and not when it leaves – the ground. Other thresholds, both on the signals and on their first derivatives, were used in order to handle some boundary conditions, like the standing of the subject, with the aim of controlling the generation of sound. Such thresholds were set in a phase of calibration of the system, which had to take into account the different weights of the subjects wearing the shoes, in order to have an average value suitable for all the possible cases.

5 Evaluating the system

We conducted a between-subjects experiment whose goal was to investigate the ability of subjects to recognize the different simulated auditory and haptic stimuli they were exposed to. The experiment consisted of three conditions: in the first condition subjects were exposed to auditory feedback only, in the second condition to haptic feedback only, and in the third condition to a combination of audio and haptic feedback.

The sounds provided in the first condition were synthesized sounds generated in real time while subjects were walking using the interactive system described in the previous section.
The same stimuli were provided in the second condition by using the haptic shoes which provided haptic feedback in real-time. In the third condition, a combination of auditory and haptic stimuli was provided.

5.1 Setup

All experiments were carried out in an acoustically isolated laboratory. The walking area was approximately 18 square meters, delimited by the walls of the laboratory. The setup for the first and third condition of the experiment consisted of the pair of sandals described in section 2, an Arduino board, a Fireface soundcard, a laptop and a set of headphones (Sennheiser HD 650, http://www.sennheiser.com).

In the first condition with audio only, the haptic actuators were not used, and the pressure sensors were driving the synthesis engine only to provide auditory feedback.
In the second condition (haptics only), participants were asked to wear earplugs and sound protection headsets instead of headphones, in order to block any audible feedback produced by the actuators. Indeed such sounds were not in the same quality range of the sounds designed to be conveyed through the headphones, and they could have biased the judgments of the participants.

In order to facilitate the navigation of the subjects, the wires coming out from the shoes in all setups, as well as the wires connecting the headphones to the soundcard, were linked to a bumbag or to snaplinks attached to trousers (see Fig. 5).

5.2 Participants

Thirty participants were divided in three groups (n = 10) to perform the between-subjects experiment. The three groups were composed respectively of 7 men and 3 women, aged between 20 and 35 (average age=24.6, sd=4.67), 9 men and 1 woman, aged between 20 and 31 (average age=23.4, sd=3.23), and 7 men and 3 women, aged between 21 and 25 (average age=22.7, sd=1.07). All participants reported normal hearing conditions. All participants were naive with respect to the experimental setup and to the purpose of the experiment.

The participants took in average about 10, 13 and 11 minutes for condition 1, 2 and 3 respectively.

5.3 Task

During condition 1 and 3 the participants were asked to wear the pair of sandals and the headphones described in sections 2 and 5.1, and to walk in the laboratory. During condition 2 they were asked to wear the pair of sandals, earplugs and sound protection headsets, and to walk in the laboratory.

Participants were exposed to 16 trials in each condition. During the experiment, 8 stimuli were presented twice in randomized order. The stimuli consisted of audio and haptic simulations of footsteps sounds

Figure 5: A subject performing the experiment in the audio-haptic condition.

on the following surfaces: beach sand, gravel, snow (in particular deep snow), forest underbrush (a forest floor composed by dirt, leaves and branches breaking), dry leaves, wood, creaking wood and metal.

Participants were given a list of different surfaces to be held in one hand, presented as non-forced alternate choice. Such list included a range of materials wider than those presented in experiments. During the act of walking they listened simultaneously to footsteps sounds and/or vibrations on a different surface according to the stimulus presented. The task consisted of answering by voice the following three questions after the presentation of the stimulus:

1. Which surface do you think you are walking on? For each stimulus choose an answer in the following list: 1) beach sand, 2) gravel, 3) dirt plus pebbles, 4) snow, 5) high grass, 6) forest underbrush, 7) dry leaves, 8) wood, 9) creaking wood, 10) metal, 11) carpet, 12) concrete, 13) frozen snow, 14) puddles, 5) water, 16) I don't know.

2. How close to real life is the sound in comparison with the surface you think it is? Evaluate the degree of realism on a scale from 1 to 7 (1=low realism, 7=high realism).

3. Evaluate the quality of the sound on a scale from 1 to 7 (1=low quality, 7=high quality).

The participants were informed that they could choose the same material more than one time and that they were not forced to choose all the materials in the list. In addition, they could use the interactive system as much as they wanted before giving an answer. When passed to the next stimulus they could not change the answer to the previous stimuli.

6 Results and discussions

The collected answers were analyzed and compared between the three conditions. Results are summarized in the confusion matrices reported in Table 1,2 and 3. Results confirm that auditory modality is dominant on the haptic modality and that the haptic task was more difficult than the other two.

Table 1: Confusion matrices with audio condition

	BS	GL	SW	UB	DL	WD	CW	MT	FS	CC	HG	DR	PD	WT	CP	—
BS	6	4	3					3			3				1	
GL		8		5	2						1	1	1	1		1
SW	1		15						3		1					
UB	1	4		6		1			6							2
DL	1	5		5	5				1		1					2
WD						10	2	1	1	2						4
CW							19									1
MT								13					1			6

Legend:								
	WD	wood	CW	creaking wood	SW	snow	UB	underbrush
	—	don't know	FS	Frozen snow	BS	beach sand	GL	Gravel
	MT	metal	HG	High grass	DL	dry leaves	CC	concrete
	DR	dirt	PD	puddles	WT	Water	CP	carpet

Table 2: Confusion matrices with haptic condition

	BS	GL	SW	UB	DL	WD	CW	MT	FS	CC	HG	DR	PD	WT	CP	—
BS	3	2	5	2			1		1			3				3
GL	5	1	3	1	1		1				2	2	1		1	2
SW	4		7		1		2		5							1
UB	1	3	5			2	1		1			3				4
DL	1	2	1		3				1			2			4	6
WD		2	2			6	3	1		1						5
CW			2			1	7		3	1	1			2		3
MT					1		5	7				2			1	4

Legend:								
	WD	wood	CW	creaking wood	SW	snow	UB	underbrush
	—	don't know	FS	Frozen snow	BS	beach sand	GL	Gravel
	MT	metal	HG	High grass	DL	dry leaves	CC	concrete
	DR	dirt	PD	puddles	WT	Water	CP	carpet

Table 3: Confusion matrices with audio-haptic condition

	BS	GL	SW	UB	DL	WD	CW	MT	FS	CC	HG	DR	PD	WT	CP	—
BS	8	4	4		1				2		1					
GL		18		2												
SW			13						6							1
UB		1	2	8	1				6			2				
DL	2	5		3	5		1		1		1	2				
WD						18				1						1
CW							20									
MT								14								6

Legend:								
	WD	wood	CW	creaking wood	SW	snow	UB	underbrush
	—	don't know	FS	Frozen snow	BS	beach sand	GL	Gravel
	MT	metal	HG	High grass	DL	dry leaves	CC	concrete
	DR	dirt	PD	puddles	WT	Water	CP	carpet

Table 4: Average realism scores from a seven-point Likert scale.

	BS	GL	SW	UB	DL	WD	CW	MT
audio	5	5.1875	5.5	3.6667	5.1667	4.2	4.4211	3.1538
haptics	5	5	6	-	3.6667	4	4.1429	4.8571
combined	4.125	5.2778	5.4615	5.25	3.6	2.8333	3.95	3.0714

Table 5: Average quality scores from a seven-point Likert scale.

	BS	GL	SW	UB	DL	WD	CW	MT
audio	5.05	4.9722	5.275	4.7222	5.2222	4.5625	4.4211	4.5
haptics	3.8824	4.2222	4.7895	3.75	3.8571	3.8667	3.6875	3.9375
combined	4.5	5.15	5.6316	4.7	4.85	3.3684	3.7	4.0714

The addition of the haptic to the audio modality seemed to help the recognition only in few cases. Indeed percentages are on average slightly higher in condition 3 rather than condition 1, although an in-depth analysis shows significant difference only for gravel (χ^2 = 8.9011, df = 1, p-value = 0.00285) and wood (χ^2 = 5.8333, df = 1, p-value = 0.01573).

In the audio-haptic condition subjects recognized quite well most of the modeled surfaces. Very high percentages were found for gravel, wood, creaking wood.

An analysis performed on the wrong answers of conditions 1 and 3 reveals that on average subjects tended to classify erroneously a surface as another belonging to a same category (e.g., wood-creaking wood-concrete, snow-frozen snow, dry leaves-forest underbrush-dirt) rather than to different categories (e.g., wood-water, wood-gravel, metal-dry leaves). In particular considering the category snow-frozen snow, snow was recognized with a percentage of 90% in the audio condition and of 95% in the audio-haptic condition.
Moreover, what emerges from these results is the ability of the subjects in distinguishing materials in the same category for solid surfaces, and their difficulties in the recognition of aggregate surfaces (aspect also confirmed by the comments of the participants). The same tendencies were found in another experiment we conducted using the same physical models controlled by another system [12] as well as in a previous experiment with audio-haptic stimuli, where subjects were receiving the stimuli passively [7].

Subjects were able to distinguish at haptic level solid surfaces with a percentage of 53%, and were able to distinguish aggregate surfaces with a percentage of 74%.

As regards the percentage of the "I don't know" answers, it was higher for condition 2 (17.5%), rather condition 1 (10 %) and lower in condition 3 (5%) and this is an indication of the difficulty of the task proposed. Condition 2 took longer (13 minutes) rather than condition 1 (9 minutes and 30 seconds) and 3 (11 minutes). This is also an indication of the difficulty of the task proposed.

Table 4 and 5 report results on the perceived quality and realism of the simulation. The degree of realism was calculated by looking only at that data from correct answers, i.e., when the surfaces were correctly recognized. As far as the quality judgement is concerned, the data was based on all the answers different from "I don't know". The results are similar to those presented in [7], even if in the previous studies a lower number of actuators was present (two for each shoe instead of four) and the experiment was performed in passive conditions. Moreover, results show that the combination of auditory and haptic stimuli does not always enhance realism and perceived quality of the simulation.

7 Conclusion

In this paper, we introduced a real-time footsteps synthesizer able to provide audio and haptic feedback, and which is controlled by the user during the act of walking by means of shoes embedded with sensors and actuators. The system was tested in a between-subjects experiment.

Results show that subjects are able to recognize most of the synthesized surfaces using the interactive system with high accuracy. Similar accuracy can be noticed in the recognition of real recorded footsteps sounds, which is an indication of the success of the proposed algorithms and their control.

The developed system is ready to be integrated in computer games and interactive installations where a user can navigate.

In future work, we indeed plan to utilize the system in multimodal environments, and include visual feedback, to understand the role of the different sensorial modalities to enhance sense of immersion and presence in scenarios where walking plays an important role.

Acknowledgment

The research leading to these results has received funding from the European Community's Seventh Framework Programme under FET-Open grant agreement 222107 NIW - Natural Interactive Walking.[7]

[7]www.niwproject.eu

REFERENCES

[1] Avanzini, F. and Crosato, P. (2006). Haptic-auditory rendering and perception of contact stiffness. *Lecture Notes in Computer Science*, 4129:24.

[2] Avanzini, F. and Rocchesso, D. (2001). Modeling collision sounds: Non-linear contact force. In *Proc. COST-G6 Conf. Digital Audio Effects (DAFx-01)*, pages 61–66.

[3] Avanzini, F., Serafin, S., and Rocchesso, D. (2005). Interactive simulation of rigid body interaction with friction-induced sound generation. *IEEE transactions on speech and audio processing*, 13(5 Part 2):1073–1081.

[4] Cook, P. (1997). Physically Informed Sonic Modeling (PhISM): Synthesis of Percussive Sounds. *Computer Music Journal*, 21(3):38–49.

[5] DiFranco, D., Beauregard, G. L., and Srinivasan, M. (1997). The effect of auditory cues on the haptic perception of stiffness in virtual environments. *Proceedings of the ASME Dynamic Systems and Control Division.*

[6] Dupont, P., Hayward, V., Armstrong, B., and Altpeter, F. (2002). Single state elastoplastic friction models. *IEEE Transactions on Automatic Control*, 47(5):787–792.

[7] et al., R. N. (2010). Preliminary experiment combining virtual reality haptic shoes and audio synthesis. In *Proc. Eurohaptics*.

[8] Gaver, W. (1993). What in the world do we hear?: An ecological approach to auditory event perception. *Ecological Psychology*, 5(1):1–29.

[9] Giordano, B. L., Mcadams, S., Visell, Y., Cooperstock, J., Yao, H. Y., and Hayward, V. (2008). Non-visual identification of walking grounds. *Journal of the Acoustical Society of America*, 123(5):3412–3412.

[10] Hunt, K. H. and Crossley, F. R. E. (1975). Coefficient of restitution interpreted as damping in vibroimpact. *ASME Journal of Applied Mechanics*, 42(2):440–445.

[11] Lederman, S. (1979). Auditory texture perception. *Perception*.

[12] Nordahl, R., Serafin, S., and Turchet, L. (2010). Sound synthesis and evaluation of interactive footsteps for virtual reality applications. In *Proc. IEEE VR 2010*.

Design for the Periphery

Saskia Bakker, Elise van den Hoven, and Berry Eggen

Industrial Design Department, Eindhoven University of Technology
P.O.Box 513, 5600MB Eindhoven, the Netherlands
{s.bakker, e.v.d.hoven, j.h.eggen}@tue.nl

Abstract

In everyday life, we are able to perform various activities simultaneously without consciously paying attention to them. In line with Weiser and Brown s [25] vision of calm technology, we see major opportunities to leverage these skills in interaction with technology by designing interactions that can take place in the periphery of our attention. In order to design such interactions however, a detailed understanding of human attention skills is important. This paper therefore provides an extensive theoretical background on attention theory and links this to the design of interactive systems. The aim is to lay a basis for design-research on interaction design for the periphery.

1 Introduction

With the upcoming of pervasive technologies, the computer is becoming ubiquitous in everyday life. As a result, intelligence and thus information can be everywhere nowadays. However, with the traditional methods of human computer interaction (screens, beeps, keyboards, mouses and the like), we are at risk being overburdened with information. Emails, reminders, advertisement, the news, recommendations, all are trying to attract our attention. When we look at the physical world, a lot of information is present here too; the weather, the time of the day, street-signs, the activity of people around us. This information however, does not overburden us in any way. We can monitor it in the periphery of our attention, but also attune to it in the center of our attention if we wish. This allows us to be aware of what s going on around us without specifically attending to it. This human skill is frequently used in everyday life, but not often leveraged in interaction with technology (exceptions are [8, 15]). Therefore, we see major opportunities for this skill to be leveraged in order to avoid information overload. This direction is closely related to the vision of calm technology, which engages both the center and the periphery of our attention, and in fact moves back and forth between the two [25, p. 79]. We think this is a valuable approach to fitting new technologies into our lives, by leveraging human interaction skills gained in the real world.

From the second half of the twentieth century on, several theories of human attention skills have been developed. Though developed to gain insight in human capabilities from a psychological or neuro-scientific point of view, these theories also provide valuable insights for designers of interactive systems that leverage these skills in interaction. Currently only few studies in the area of calm technology are grounded in theory on human attention skills. However, in order to effectively design systems that can be perceived in the periphery as well as in the center of the attention, a detailed understanding of such skills may be valuable. This paper therefore aims at laying a basis for design-research on interaction design for the periphery, by linking an extensive theoretical background on attention theory to the area of calm technology and the combination of audio and physical interaction in particular.

2 Calm Technology

Most current computing technology is designed to be in the center of the attention. Particularly audio is frequently used for alerts, alarms and reminders [23] which are meant to attract the attention of the user. Weiser however, envisioned the interaction with the computers of the future vanishing into the background, so that we are freed to use them without thinking and so to focus beyond them on new goals [24, p. 3]. In other words, when we can perceive and interact with information from the computer in the periphery of our attention, computing technology can fit into our everyday life the way everyday information sources fit into our lives. This is what is meant with the term Calm Technology [25], which enables users to monitor information without specifically paying attention to it, while at the same time facilitating them to focus on it if desired. Weiser and Brown [25] use the inner office window, which connects the office to the hallway, as an everyday example of calm technology. This window allows all kinds of small informative clues: a lot of motion in hallway subtly informs you of an upcoming event and a light shining into the hallway late at night says that someone is working late. These clues are part of the ambience of the environment and are usually in the background but may be focused on if desired. Similarly, the aim of calm technology is to form a part of the ambience so that presented information can be perceived in the periphery. Some other terms have also been used to describe similar concepts. For example, Matthews et al. [12] use the term Peripheral Displays to describe computing applications that allow a person to be aware of information without being overburdened by it (p. 247). Ishii [7] uses the term Ambient Media to describe a class of interfaces that is designed to smooth the transition of the users' focus of attention between background and foreground (p. xx).

An early example of a calm technology design is the Dangling string [25], a plastic spaghetti string that hangs from the ceiling in an office context. The string is connected to a motor that will spin based on the information sent through the Ethernet cable. If the network is busy, the motor spins fast and if the network is not busy it will spin slowly. Therefore a visual and sonic representation of the business of the network is provided. Motion Monitor [12] is an example of calm technology that uses only visual information to engage the periphery; it is a ball that lights up in different colors resembling the amount of activity at a remote location, providing users with a sense of connectedness to friends and family. Other calm technology systems primarily use the auditory modality, for example audio aura [15], a system that informs office workers of information such as the availability of co-workers through background auditory cues. Birds Whispering [6] uses bird-sounds to subtly reveal information about the activity in the office. AmbientROOM [8] explored background monitoring of information through light, sound and movement in an office. The activity level of people on a distant location is displayed through shadows of rippling water on the ceiling or through light effects. Other information (e.g. the number of unread emails or the value of a stock portfolio) is displayed through a subtle soundscape of bird and rain sounds.

3 Attention Theory

In the 19th century, James [9] stated that everyone knows what attention is (p. 403), referring to the many different ways the word attention is used in everyday situations. Attention can be devoted to stimuli that we perceive through our senses (*sensorial attention*), but also to cognitive processes or thoughts (*intellectual attention*) [9]. The world around us is constantly full of stimuli that we can potentially attend to. Furthermore, our memory and reasoning capabilities add a large number of potential cognitive processes to undertake. As we cannot fully appreciate all that takes place at any one time [16, p. 6], a process of selective allocation of attention is needed to make sense of the world. As a result of this process, attended stimuli or thoughts will always be just a small fraction of all available streams [28].

Throughout the past decades, several models of this process of attention management have been developed in the areas of cognitive psychology as well as neuropsychology. Although various different functions of attention are distinguished in literature, the two most important functions described are *selective attention* and *divided attention* [17, 19, 27]. Selective attention is the process of selectively focusing the attention on one stimulus while intentionally ignoring others [19]. Models of selective attention therefore describe attention by analogy with a mental filter [27] that selects certain stimuli to attend to and rejects others. Such models are therefore often referred to as filter-models. Divided attention is the process in which we carefully divide our attentional resources over multiple attentional tasks at once [19], e.g. when multitasking [27]. Models of divided attention therefore see attention as a finite mental resource that can be divided over several sensorial or intellectual processes. Although these two functions of attention may not be mutually exclusive, we will first separately review literature on both functions. At the end of this section, we will present a model that captures our current understanding of the theory and that can be used to inform design-research in the area of calm technology.

3.1 Selective Attention

Selective attention theories usually only concern sensorial attention [17] and are often grounded in research on speech perception (e.g. [2, 14, 21]). In almost any given situation, several stimuli of multiple modalities will reach our senses simultaneously. These stimuli however, must be processed before they can be perceived. This processing, for example, enables us to distinguish our friend s voice from the voice of a passerby. In psychoacoustics this is called *auditory scene analysis* [1], which takes place when incoming signals are grouped into different streams based on the likelihood of them coming from the same source. Stimuli in other modalities are likely processed in a similar manner [1]. It is generally accepted that this form of processing is performed at an unconscious level of awareness, and happens before the process of attention starts to take place [28]. The attention theories we will describe in this paper therefore assume that incoming stimuli have been grouped into streams that can be attended to.

Selective attention theories describe attention as a mental filter. Influential early work by Cherry [4] forms the basis of a series of theories of selective attention [16]. Cherry s experiments [4] address the problem of selective attention, which he called the cocktail party problem referring to the situation at a cocktail party where one stands in a room full of sounds and is able to focus attention on a single conversation [19]. Cherry presented subjects with two spoken messages simultaneously and instructed them to attend to one message by directly repeating, or *shadowing* it. The other message had to be ignored, or *rejected*. As a result, Cherry found that subjects were able to almost entirely separate the two messages. Furthermore, the subjects did not remember any words mentioned in the rejected message. They could only recall that another message was present. Many researchers performed experiments similar to the ones performed by Cherry [4], which lead to a series of models of selective attention.

3.1.1 Early Selection Theory

Early selection theories suggest a limited capacity channel in the perceptual process that is only capable of handling one perceptual stream at a time [17]. The first well-known early selection theory was proposed by Broadbent [2], who suggested a selective filter in the brain that allows only one channel of information to pass and rejects others. One stream is selected based on the subjective attributes of the sound (e.g. pitch, volume). As became clear from Cherry s experiments, words in the rejected messages are not remembered. Broadbent s model [2] therefore assumes that the meaning of words is extracted after one stream has been selected [16]. The selection thus takes place early in the process, see Figure 1.

Late Selection Theory

In shadowing experiments similar to the ones performed by Cherry, Moray [14] showed that words that are important to the person in question (e.g. one s name) are consciously noticed when present in a rejected stream [16]. Counter to early-selection theory, this shows that the meaning of at least some words may be extracted before one channel is selected. Based on such evidence, an alternative theory was proposed by Deutsch and Deutsch [5], who suggest that the selection of attention is located later in the process, at a moment in which the meaning of all incoming streams has been identified [16, 19], see Figure 1. This late selection theory proposes that the identification process happens involuntarily and below the level of awareness.

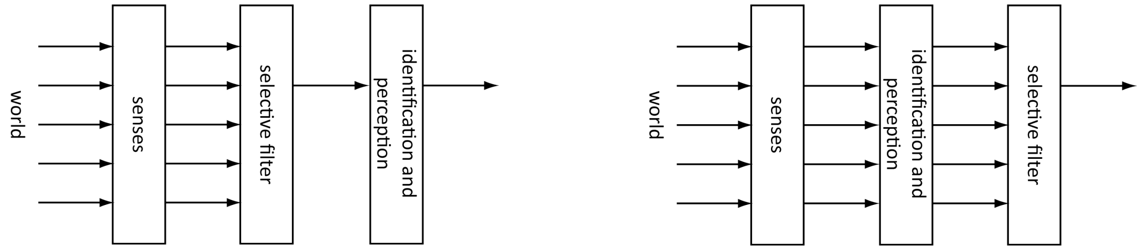

Fig. 1. Simplified illustration of an early selection model (left), adapted from [19, p. 94], and of a late selection model (right), adapted from [19, p. 96].

3.1.2 Attenuation Theory

Around the same time Deutsch and Deutsch [5] proposed their late-selection theory, Treisman [21] found that when words in the rejected channel are relevant to the information in the attended channel, they consciously or unconsciously influence the perception of the information in the attended channel. For example, listeners did notice words of the message in the rejected channel when the information was related to that in the attended channel [16]. Given the results of her experiments, Treisman [21] suggests that the selection process is among other things influenced by the relevance of the information in the incoming channels. She suggests that this happens through activations of the detector units for related concepts. This process is called *priming*. Based on the content of the information in the attended channel, related concepts are primed and therefore the threshold for identifying them is lowered. For example when having a conversation about a concert, related words such as the location of the concert, or the name of the performer may be primed. When a passerby says any of these words, one is more likely to recognize them and attend to this channel than when words with a less relevant meaning are heard.

Different from late-selection theories, Treisman proposes that the meaning of words in rejected channels is not identified before reaching the selective filter. Counter to early-selection theories however, Treisman [21] proposes a filter that attenuates unattended channels rather than completely blocking them. This way, a weakened version of these unattended channels is passed on [19]. Given the lower threshold for detecting primed stimuli, even an attenuated version of them is enough to be recognized and attended to. The idea of priming could also explain Moray s finding[14] showing that listeners tend to notice their own name in rejected channels; detectors for one s own name are primed, see Figure 2.

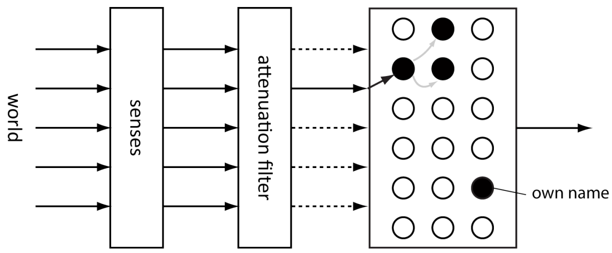

Fig. 2. Illustration of an attenuation model, adapted from [20, p. 25]. Representations of words in memory are illustrated by circles. Primed words, that have a lower threshold for being recognized, are illustrated by black circles.

The process of priming was already noticed by James [9], who referred to it as preperception. He illustrated this phenomenon by giving the following example. If I have received an insult, I may not be actively thinking of it all the time, yet the thought of it is in such a state of heightened irritability, that the place where I received it or the man who inflicted it cannot be mentioned in my hearing without my attention bounding, as it were, in that direction, as the imagination of the whole transaction revives [9, p. 449]. In other words, we may not only be primed for topics related to our current attentional focus or for stimuli that are generally relevant to us (such as our own name), we may also be primed for topics that are in the back of our minds for other reasons. Similar to the concept of priming is a process that Shallice and Burgess [18] call the supervisory attention system , when describing the selection of action and thought processes rather than perceptual stimuli. The supervisory attention system is in charge of top-down activation of schematic representations of actions or thoughts which increases the probability of them being selected.

Knudsen [11], who recently investigated attention in terms of neurobiological components, mentions top-down sensitivity control as one of the component processes fundamental to attention. Though describing a neurobiological phenomenon, this is rather similar to Treisman s idea of priming. The process of top-down sensitivity control increases the sensitivity of neural representations of attended information, which increases the chances of high signal strength at these representations. Knudsen [11] furthermore describes a bottom-up salience filter . Stimuli that are highly salient to us, such as a loud noise or a sudden flash of light, will pass this filter and can be attended.

Most psychology literature that summarizes attention theory address both early and late selection theories (e.g. [17, 20]). However, early selection theories cannot explain why relevant words in rejected channels are recognized. The amount of processing required to identify all incoming information in late selection theory on the other hand, does also not seem likely as most of the information is never used. Therefore, attenuation theory is often seen as the most plausible explanation of results of shadowing experiments ([17, 20]).

3.2 Divided Attention

As mentioned before, theories of selective attention usually only concern sensorial attention. Furthermore, most previously mentioned theories are grounded in research on the auditory modality only. Taking a broader approach, divided attention theories explain how we can perform multiple attentional tasks at once. These tasks may involve sensorial and/or intellectual attention and may include multiple modalities. Divided attention theories describe attention as the division of a limited amount of mental resources over different activities [19].

According to divided attention theory, the extent to which we can perform multiple tasks at once depends on the mental effort required for each task. The amount of mental effort needed decreases with practice and experience [27]. For example walking; when learning how to walk this activity requires a lot of mental effort and cannot be performed simultaneously with other activities. However, when we get more experiences, less effort is required and more can be done simultaneously. Highly trained processes such as walking require hardly any resources and are therefore called *automatic processes*. Such processes need little mental effort, involve no conscious control and many of these processes can be performed in parallel. On the contrary, *controlled processes* are processes that do require conscious control (such as reading a book). Only one controlled task can be performed at once.

A theory of divided attention proposed by Kahneman [10], suggests that resources can be allocated to any of the possible activities that one can perform as a result of information input (sensorial or intellectual). The amount of required resources may vary based on various aspects of the activity such as difficulty or automation. The distribution of resources over activities depends on a person s own intentions, as well as on so called *enduring dispositions*. This refers to relevant words such as our own name, or salient sensorial stimuli that attract our attention immediately such as a loud noise.

Furthermore, Kahneman [10] suggests a link between arousal and attention. The more aroused we are, the narrower our focus of attention will be. This means that as arousal increases, the attention to controlled processes increases, but also implies that the attention to automated processed decreases. For example, when we are highly engaged while reading a book, we may not notice any other streams of information, even when highly relevant information is present such as our own name. Another example: normally we are able to read (controlled process) and walk (automated process) at the same time, but when the reading task requires deep thought or causes high levels of arousal, we will often stop walking to focus on the reading [16]. This indicates that automated processes may require some attentional resources, though much less than controlled processes.

Though Kahneman [10] proposes only one kind of resource available for all tasks, Wickens and McCarley [27] describe multiple attentional resources, which may explain why certain tasks are more easily performed simultaneously than others. For example, driving while reading a book is more difficult than driving while listening to the same book being read to you.

3.3 Overview

The purpose of this theoretical overview is to gain a better understanding of human attentional processes and abilities in order to inform design for the periphery . Such designs are interactive systems that leverage these abilities so that information can be perceived without specifically focusing your attention on it. For this purpose, we have created an illustrative overview of the parts of the attentional process that we think are important in everyday life situations, see Figure 3. This overview is primarily meant to structure our own understanding of the process and is based on the above-mentioned literature.

We have seen that one may attend to sensorial stimuli (sensorial attention) or cognitive processes (intellectual attention). In many everyday situations however, an attentional activity may involve both sensorial and intellectual attention. For example when in a conversation, one will attend sensorial streams (e.g. listening to conversational partners, looking at their facial expressions) and intellectual processes (e.g. speaking, recalling information from memory). We will therefore not refer to streams that can be attended, but to *potential activities* which may consist of several kinds of tasks as well as involve multiple modalities. At any given moment in time, there will be multiple potential activities that we can attend to. These potential activities emerge from sensorial stimuli (e.g. hearing music makes the activity of listening to this music available or hearing a passerby speaking about politics may elicit potential thought processes about this subject) or from intellectual processes (e.g. planning to go call a friend or to take a shower). These potential activities form the center of the overview in Figure 3 and are represented by vertical bars of varying height and opacity. To keep the overview clear, Figure 3 only illustrates ten potential activities, however at any moment many more will be possible.

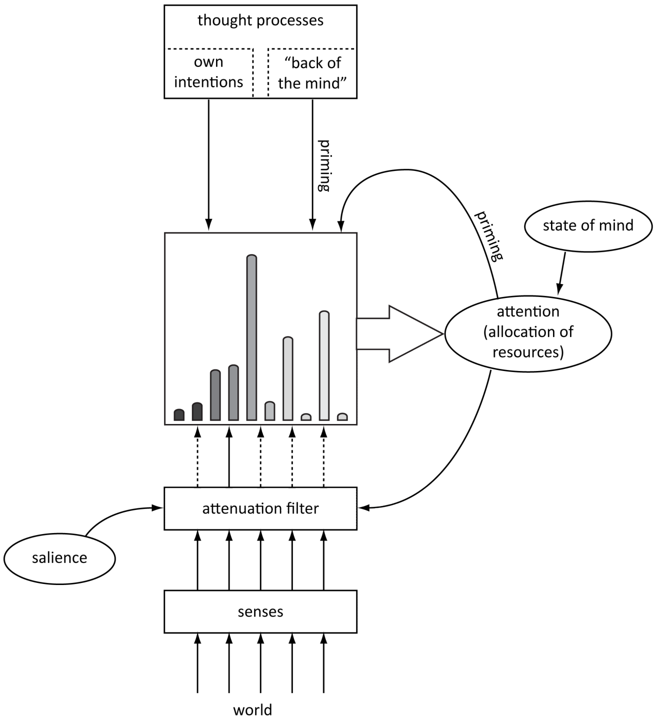

Fig 3. Illustrative overview of the attentional process, aimed to structure our understanding of this process to inform and support design for the periphery .

Obviously not all these potential activities can be performed at the same time. As suggested by divided attention theory, attentional resources have to be divided over these potential activities and only the ones to which resources are allocated will be performed. The height of the different bars indicates the resource demand of the activity, which is the amount of resources needed to perform the task. The opacity of the bars indicates the likelihood of recourses being allocated to that specific activity. The darker the bar, the more likely the activity is to receive resources. This likelihood is influenced by a person s own intentions (consciously deciding which activity to undertake) as well as by (unconscious top-down) priming which makes activities more likely to receive attentional resources. Primed activities are related to things that are generally relevant to us (e.g. our own name), things that are the current focus of attention or things that are in the back of our minds .

We discussed theories of both selective and divided attention proposed in literature. The difference between these two functions of attention is not straight forward. As a matter of fact, both selectivity and resource allocation characterize the attentional process [17]. Selectivity primarily seems to play a role in sensorial attention, whereas both sensorial and intellectual attention may involve resource allocation. We therefore suggest that the attentional process does involve a selective filter, but we define attention as the allocation of resources to one or more potential activities. In line with Treisman s attenuation theory, the filter illustrated in Figure 3 attenuates unattended incoming stimuli. This filter is influenced by the salience of incoming stimuli, for example a loud noise will pass the filter without attenuation.

The division of resources further depends on the state of mind of the individual, such as the level of arousal. Figure 4 illustrates the division of resources, and thus the attention, in two different situations. In these overviews, attentional resources are illustrated by white circles. In a situation is which one is highly engaged in reading a book, all attentional resources are allocated to this activity and no other activities can be performed (Figure 4a). A more frequently occurring everyday situation is illustrated in Figure 4b. In this situation, one s main activity is preparing dinner, but at the same time he is listening to the radio, looking at his watch and chewing chewing-gum. The other listed potential activities do not receive resources and are thus not performed. As illustrated in Figure 4b, not all recourses are used in this situation. At any moment one may (consciously) decide to call a friend or (unconsciously) be attracted to pay more attention to the radio as one s name is suddenly heard in that stream. This would change the recourse demand of some activities (e.g. carefully listening to the radio) as well as the priming of certain activities (e.g. when calling a friend while preparing diner, the relevance of monitoring the pans on the stove may become more relevant). This illustrates that the process of dividing resources over potential activities is highly dynamic and may in fact be at no moment a static overview.

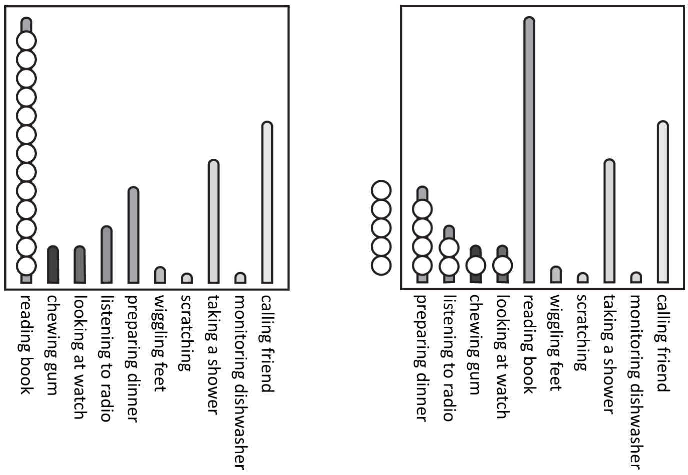

Fig. 4. Division of resources over different possible activities in two situations. **4a**: High attentional task reading a book (left). **4b**: A combination of low attentional tasks (right).

4 Center and Periphery of the Attention

In the previous section, we have presented an overview of our understanding of the attentional process, which may be useful to inform design for the periphery . This means, the design of systems that can be perceived or interacted with in the periphery of our attention, but also be focused on in the center of the attention. However, for such design processes to take place, it is essential to define what we mean by *center* and *periphery* of the attention.

In psychology literature reporting on attention [26, 27], the word periphery, which literally means external boundary or outward boundary [13], is generally used in the context of visual perception. In that area, the concept of peripheral vision refers to the parts of vision that occur outside the center of the visual field [26]. Vision in the center of the visual field is referred to as central or foveal vision [26]. Although visual displays are frequently used in calm technology designs, for example peripheral displays [12], authors in the area of ubiquitous computing often use the term periphery in a broader context. Brown and Duguid [3] describe how media contain peripheral cues that subtly direct users along particular interpretive paths by invoking social and cultural understandings (p. 131). These cues help us shape our expectations of the content. For example; by the cover, paper and typeface of a book we can determine if it is a novel or a study book. A little broader, Weiser and Brown [25] use the word periphery to name what we are attuned to without attending to explicitly (p. 79). Although this definition involves multiple modalities it is not yet very explicit, which we think is important to inform design processes.

As we describe attention as the division of resources over potential activities, we will explain the center and periphery of the attention in the same context, which is illustrated in Figure 5. What we see as the center of the attention is the one activity that most resources are allocated to. In the situation illustrated in Figure 5a, the center of the attention is the activity of reading a book, while Figure 5b illustrates preparing dinner as the center of the attention. The periphery of the attention consists of all potential activities that are not in the center, regardless of the number of resources being allocated to them. For example, in both situations illustrated in Figure 5, the activity of listening to the radio is in the periphery of the attention, while in situation 5a it is not performed and in situation 5b it is performed, be it with only a low amount of resources. However, the activity of listening to the radio is closer to the center in Figure 5b compared to Figure 5a, which means that it is more likely to shift to the center of the attention. As mentioned before, the attentional process is highly dynamic; the resource demand, priming and proximity to the center of each potential activity (represented by respectively height, opacity and location of each vertical bar) are subject to constant change. For example when driving a car and having a conversation with a passenger at the same time, both the activity of driving and the conversation will constantly move between the center and the periphery of the attention.

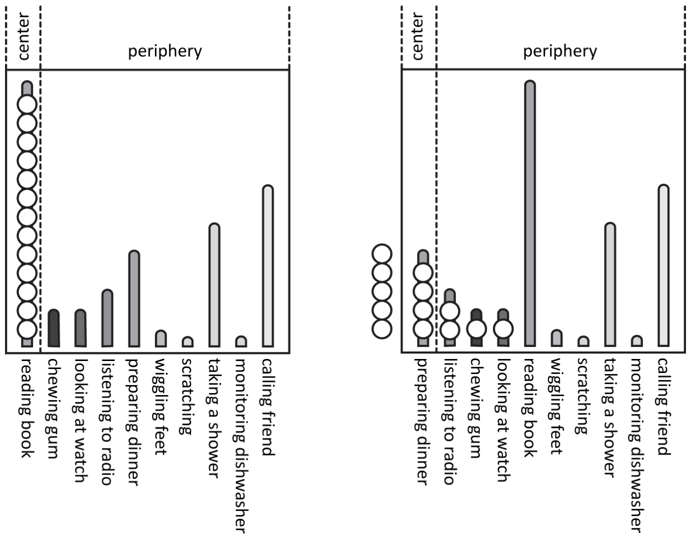

Fig. 5. Illustration of the center and periphery of the attention in two situations. **5a**: High attentional task reading a book (left). **5b**: A combination of low attentional tasks, in which not all available resources are used (right).

5 Design for the Periphery

As a result of the theoretical overview provided in this paper, we have defined attention as the division of attentional resources over potential activities. Furthermore we have described the *center* of the attention as the one activity to which most resources are allocated and the *periphery* of the attention as all remaining potential activities. Most current interaction with technology takes place in the center of the attention: screens, beeps, reminders, ringtones and keyboards are all designed to attract the attention. As information provided by the physical world is often perceived and interacted with in a much less obtrusive way, we see major opportunities for leveraging above describe human attention abilities in interaction with technology. This is what we refer to as design for the periphery .

The value of design for the periphery primarily lays in the idea that potential activities (related to interactive technology) can reside in the periphery of the attention, where they hardly require resources. However, when such an activity becomes relevant to the user, it may shift to the center of the attention and intentionally be performed. The weather is an everyday life example of this phenomenon. Information about the current weather will in most situations be available in the periphery, but when it becomes relevant (e.g. when one is about to go for a walk) or extraordinarily noticeable (e.g. a sudden storm), it will shift to the center of the attention. When designing the display of information that may shift between the periphery and center of the attention, it is important to think about how these intended shifts may be facilitated. A principle that may be drawn upon is that of salience, which ensures that extraordinary stimuli (e.g. loud sounds, sudden movement) are immediately noticed (see Figure 4).

Drawing upon salience however would not exactly contribute to technology being calm and unobtrusive. A more interesting principle to draw upon would be the idea of priming, which lowers the threshold for perception of stimuli that are relevant based on the current center of attention, things that are in the back of one s mind or intrinsically relevant stimuli such as one s own name. It would therefore be interesting for design-research in this area to address the design of stimuli for which priming may naturally occur or that may facilitate a learning process that could lead to priming after some experience has been gained.

Apart from stimuli that can shift from the periphery to the center of the attention, certain stimuli may never shift to the center even though they may influence the activities in the center of the attention. To take the example of the weather information again, one may never consciously attend to weather related information provided by the world around him, but still know that the sun is shining. This knowledge may influence his conscious activities such as deciding to go out for lunch. When designing this kind of peripheral stimuli, the principle of priming will not play a role, but the clarity of the designed stimuli will be of crucial importance as these stimuli likely need to be learned to reside in the periphery. An iterative design process could help in successfully designing such stimuli.

Since we have taken a broad definition of the term periphery, designing for the periphery may head into different directions. One of these directions could address the design of information displays that are intended to be perceived in the periphery of the attention. This direction relates to the area of peripheral display [12], which primarily uses the visual modality for information display, and in the area of calm technology [25], in which also examples that use audible information are known (e.g. [8, 15]). Furthermore, the tactile modality may be interesting to explore in the context of design for the periphery. Apart from perceiving sensorial stimuli, potential activities in the periphery may also be elicited by cognitive processes, such as tying your shoelaces in the periphery while watching TV in the center of the attention. An alternative direction may therefore address the design of interactions that can be performed in the periphery of the attention. As we are very skilled in manipulating objects with our hands, tangible interaction [22] seems to be a relevant interaction style for this purpose.

6 Conclusion and Future Work

In this paper, we have reviewed literature on calm technology and given an overview of attention theory. We have described our understanding of the attentional process, as well as of the center and periphery of the attention. Furthermore, potential directions for design for the periphery , which we see as a valuable approach for fitting new technologies into everyday life, have been discussed. We now have a detailed theoretical understanding of human attention abilities, which lays a basis for future design-research work in this area. Our next steps will therefore involve iteratively exploring the design space of the proposed directions.

References

[1] A.S. Bregman. Auditory Scene Analysis. *The Perceptual Organization of Sound*. MIT Press, Cambridge, MA, 1990.

[2] D.E. Broadbent. *Perception and Communication*. Pergamon Press, London, UK, 1958.

[3] J.S. Brown and P. Duguid. Keeping It Simple: Investigating Resources in the Periphery. *T. Winograd (Eds.) Bringing Design to Software*. ACM Press/Addison-Wesley, New York, NY: 129-150, 1996.

[4] E.C. Cherry. Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America*, 25(5): 975-979, 1953.

[5] J. Deutsch an D. Deutsch. Attention: Some theoretical considerations. *Psychological Review* 70: 80-90, 1963.

[6] B. Eggen and K. van Mensvoort. Making Sense of What Is Going on Around : Designing Environmental Awareness Information Displays. *P. Markopoulos, B. de Ruyter and W. Mackay*

(Eds.) Awareness Systems: Advances in Theory, Methodology and Design. Springer London, UK: 99-124, 2009.

[7] H. Ishii. Tangible bits: beyond pixels. *Proceedings of the 2nd international conference on Tangible and embedded interaction (TEI 08)* page xv-xxv, 2008.

[8] H. Ishii, C. Wisneski, S. Brave, A. Dahley, M. Gorbet, B. Ullmer and P. Yarin. ambientROOM: integrating ambient media with architectural space. *CHI 98 conference summary on Human factors in computing systems*, page 173-174, 1998.

[9] W. James. *The Principles of Psychology*. Dover Publications, Inc., New York, NY, 1890.

[10] D. Kahneman. *Attention and effort*. Prentice-Hall, Englewood Cliffs, NJ, 1973.

[11] E.I. Knudsen. Fundamental components of attention. *Annual Review of Neuroscience* 30: 57-78, 2007.

[12] T. Matthews, A.K. Dey, J. Mankoff, S. Carter and T. Rattenbury. A toolkit for managing user attention in peripheral displays. *Proceedings of the 17th annual ACM symposium on User interface software and technology*, page 247-256, 2004.

[13] Merriam-Webster Online Dictionary. Retrieved from http://www.merriam-webster.com/.

[14] N. Moray. Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology* 11(1): 56-60, 1959.

[15] E.D. Mynatt, M. Back, R. Want, M. Baer and J.B. Ellis. Designing audio aura. *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI 98)* page 566-573, 1998.

[16] D.A. Norman. *Memory and Attention. An introduction to human information processing*. John Wiley & Sons, Inc., New York, NY, 1976.

[17] H.E. Pashler. *The Psychology of Attention*. MIT Press, Cambridge, MA, 1998.

[18] T. Shallice and P. Burgess. Supervisory control of action and thought selection. *A.D. Baddeley and L. Weiskrantz (Eds.) Attention: selection, awareness, and control: a tribute to Donald Broadbent*. Oxford University Press, Oxford, UK, 1993.

[19] R.J. Sternberg. *Cognitive Psychology*. Harcourt College Publishers, Orlando, FL, 1999.

[20] E.A. Styles. *The Psychology of Attention*. Psychology Press Ltd., East Sussex, UK, 1997.

[21] A.M. Treisman. Verbal Cues, Language, and Meaning in Selective Attention. *The American Journal of Psychology* 77(2): 206-219, 1964.

[22] B. Ullmer and H. Ishii. Emerging frameworks for tangible user interfaces. *IBM Systems Journal* 39(3-4): 915-931, 2000.

[23] B.N. Walker and M.A. Nees. Theory of Sonification. *Principles of Sonification: An Introduction to Auditory Display and Sonification*. Retrieved from http://sonify.psych.gatech.edu/publications/.

[24] M. Weiser. The computer for the 21st century. *ACM SIGMOBILE Mobile Computing and Communications Review* 3(3): 3-11, 1999.

[25] M. Weiser and J.S. Brown. The Coming Age of Calm Technology. *P.J. Denning and R.M. Metcalfe (Eds.) Beyond Calculation: the next fifty years of computing*. Springer-Verlag, New York, NY: 75-85, 1997.

[26] C.D. Wickens and J.G. Hollands. *Engineering Psychology and Human Performance*. Prentice-Hall Inc., NJ, 2000.

[27] C.D. Wickens and J.S. McCarley. *Applied Attention Theory*. CRC Press, Boca Raton, FL, 2008.

[28] S.N. Wrigley. *A theory and computational model of auditory selective attention*. PhD Thesis, Department of Computer Science, University of Sheffield, UK, 2002.

List of authors