Ian Oakley
Stephen Brewster (Eds.)

LNCS 7989

Haptic and Audio Interaction Design

8th International Workshop, HAID 2013
Daejeon, Korea, April 2013
Revised Selected Papers

Springer

Lecture Notes in Computer Science 7989

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Ian Oakley Stephen Brewster (Eds.)

Haptic and Audio Interaction Design

8th International Workshop, HAID 2013
Daejeon, Korea, April 18-19, 2013
Revised Selected Papers

Springer

Volume Editors

Ian Oakley
Ulsan National Institute of Science and Technology (UNIST)
School of Design and Human Engineering
UNIST-gil 50, Ulsan 789-798, Korea
E-mail: ian.r.oakley@gmail.com

Stephen Brewster
University of Glasgow, School of Computing Science
Sir Alwyn Williams Building, Lilybank Gardens, Glasgow G12 8QQ, UK
E-mail: stephen.brewster@glasgow.ac.uk

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-41067-3 e-ISBN 978-3-642-41068-0
DOI 10.1007/978-3-642-41068-0
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013948104

CR Subject Classification (1998): H.5.1-2, H.3-4, H.1.2, H.5.5, K.3.1, J.3

LNCS Sublibrary: SL 3 – Information Systems and Application,
incl. Internet/Web and HCI

© Springer-Verlag Berlin Heidelberg 2013

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Introduction

HAID 2013, the 8th International Workshop on Haptic and Audio Interaction Design was held during April 18–19, 2013, in Daejeon, Korea. The workshop took place immediately after the IEEE World Haptics Conference 2013 (WHC 2013) and with the generous support of this event's organizing team. We are particularly thankful for the strong support of Dong-Soo Kwon (KAIST), the General Chair of WHC 2013. Collocating the two events was also a fruitful merger. Indeed, HAID, with its focus on how computer systems can be best designed to leverage and appeal to the senses of touch and hearing matches naturally with WHC's focus on haptic systems in the broadest terms – from mechanical design of hardware to psychophysical studies of perception. The events both benefited from co-attendance and the different but related foci led to high levels of synergy.

A total of 14 papers were accepted to HAID 2013, each containing novel work merging technological and human concerns. A panel of experts spanning industry and academia reviewed and vetted each paper. We extend our sincerest thanks to all our reviewers, who provided invaluable commentary, advice, and feedback both promptly and respectfully. Their work was essential, not only as input to the authors, but also to establishing the high quality of the final program. HAID 2013 also benefitted from two outstanding invited keynote speakers. Jinsil (Hwaryoung) Seo from Texas A&M University opened the workshop with an inspiring overview of her visceral, evocative, and multisensory digital art. Seungmoon Choi from POSTECH (Pohang University of Science and Technology) closed the event with a discussion of a series of research projects and prototypes tackling the challenging question of how to effectively and easily author rich and expressive haptic, and particularly vibro-tactile, content. These very different speakers both bracketed the workshop temporally and framed it conceptually. They covered a broad range of perspectives from exploratory art through to practical engineering that was mirrored in the paper presentations of the main program.

The papers were organized into five thematic sessions, briefly introduced below.

Non-intrusive and Thermal Haptics

The field of haptic interaction is rapidly maturing and diversifying, trends that were highly evident in the papers presented in this session. Rather than focusing on traditional mechanical haptic feedback systems (such as force feedback or vibrating actuators), this work explored either how changes in the temperature of a touched surface could be used as feedback in computer interfaces (Wilson et al.) or to convey emotions (Salminen et al.). The remaining paper in this session looked even further afield, at the potential of fully non-contact haptic systems – those using remote actuation techniques based on the sensations evoked by focused jets of air or sound (Tsalamlal et al.). Overall, the work presented in this session highlighted new areas in which the field of haptics is developing both in terms of the interaction scenarios it outlines and the actuation technologies that are emerging.

New Interfaces and Interactions

This session collected contributions that apply haptic and audio cues to existing areas of human–computer interaction, expanding the scope of these feedback modalities to novel domains. In the first paper in this session, L'Orsa et al. detail a new take on how multi-modal cues can enhance remote surgery systems through providing contextual warnings during critical tasks that offer the potential to reduce unintentional tissue damage. Girard et al., on the other hand, expand the literature on haptic collaboration by exploring whether haptic cues can objectively increase the efficiency of tightly coupled physical manipulation tasks. Finally, Mori et al. present a study exploring the potential of haptic cues to form the basis of advanced brain–computer interfaces for tasks such as robotic tele-operation. The focus of all these papers on remote or collaborative interaction highlights this as a promising area for future studies.

Emotion and Affect

Touch is an important affective sense, relating to deep and powerful expressions of emotion. This session included two papers investigating different aspects of this relationship. Fontana revisited a classic experiment that explores how individuals map spatial stimuli to particular sound forms. His work shows that haptic stimuli are associated with sounds in ways that are broadly similar to the way this is achieved with visual cues, a finding which has the potential to cast light on the central cortical processing of shape information. Gaffary et al. also focused on prior work by highlighting the methodological limitations of existing studies linking haptic cues to emotions. To better understand these problems, they contrast the explanatory outcomes achieved through three different statistical procedures, concluding with a concise and valuable summary of the advantages and disadvantages of each. Taken together, these papers highlight how novel empirical methods can advance the field of haptic and audio interaction design, particularly in complex non-traditional application areas.

Music

The session on music dealt with the intimate relationship between using an instrument to produce sound and the highly physical, haptic, skill that is required to do so. This strong focus on merging the senses makes musical contributions an excellent fit for HAID and the three contributions presented in this session showed the richness and diverse scope of work in this area. Ziat et al. presented a complex psychophysical study looking at the effects of plucking virtual haptic strings of different stiffness values on audio perception of the notes this produced. Giordano and Wanderly presented a complementary but more practically focused paper that surveyed the literature on how haptic cues can be used to augment or enhance music making practices. The paper culminates with a valuable taxonomy of the possibilities in this space. The final paper, by Esteves et al., was entirely application focused – it described the design of a tabletop system that enabled musical novices to collaborate with live performances through manipulating tangible blocks on a tabletop computer.

Mobile Devices and Applications

The final session at HAID 2013 tackled mobile interaction, a recurring and rapidly developing theme in the field. Kim et al. presented fundamental work on the design and evaluation of a novel miniature and expressive tactile actuator designed expressly for embedding within mobile devices. Panëels et al. discussed a more applied HCI study looking at how cues delivered by tactile actuators mounted around the wrist can be understood in real-world tasks and scenarios where orientation of the arm may varying considerably. Finally, Rassmus-Gröhn et al. described a field study of a mobile application that integrated audio cues into a navigation system aimed at pedestrian tourists in order to support both inclusion and lower the attentional demands of the software. Spanning a broad range of perspectives from hardware through empirical experiment to fieldwork, this session truly demonstrated the interdisciplinary nature of the HAID workshop.

In summary, HAID 2013, the 8th International Workshop on Haptic and Audio Interaction Design, showcased work in new sensory interaction modalities and novel HCI application domains. It explored the links between haptic cues and emotions and investigated the rich relationships between touch and sound in the domain of music, a natural fit for such studies. Finally, it expanded on the growing body of research applying non-visual interfaces to mobile devices and scenarios where users are busy and fully engaged with the world, scenarios where cues and information that can be heard and felt stand to convey key advantages.

Organization

The 8th International Workshop on Haptic and Audio Interaction Design was organized in collaboration with IEEE World Haptics 2013.

General Chair

Stephen Brewster	University of Glasgow, UK

Paper Co-chairs

Hong Tan	Microsoft Research Asia, Beijing, China
Ian Oakley	UNIST, Ulsan, Korea
Ki-Uk Kyung	ETRI, Daejeon, Korea

Poster Co-chairs

Andrea Bianchi	Sungkyunkwan University, Suwon, Korea
Darren Edge	Microsoft Research Asia, Beijing, China

Publicity Chairs

Kwangsu Cho	Sungkyunkwan University, Seoul, Korea
Federico Fontana	Università degli Studi di Udine, Italy

SV Chair

Dongbum Pyo	KAIST, Daejeon, Korea

Local Arrangements

Junseok Park	ETRI, Daejeon, Korea

Program Committee

ErcanAltinsoy	Technical University Dresden, Germany
Seungmoon Choi	Pohang University of Science and Technology, Korea
Andrew Crossan	Glasgow Caledonian University, UK
Antonio Frisoli	Scuola Superiore Sant'Anna, Italy

Inwook Hwang	Pohang University of Science and Technology, Korea
Hiroyuki Kajimoto	The University of Electro-Communications, Tokyo, Japan
Johan Kidal	Nokia Research, Finland
Ravi Kuber	University of Maryland, Baltimore County, USA
David McGookin	Glasgow University, UK
Sile O'Modhrain	University of Michigan, USA
Charlottle Magnusson	Lund University, Sweden
Valentina Nisi	University of Madeira, Portugal
Antti Pirhonen	University of Jyväskylä, Finland
Evalotta Sallnäs	KTH, Sweden
Jongman Seo	Pohang University of Science and Technology, Korea
Steven Strachan	Commissariat à l'énergie atomique et aux énergies alternatives, France
Gerhard Weber	Dresden University of Technology, Germany
Gi-Hun Yang	Korea Institute of Industrial Technology, Korea

Table of Contents

Non-intrusive and Thermal Haptics

New Interfaces and Interactions

Emotion and Affect

Music

Mobile Devices and Applications

Non-intrusive Haptic Interfaces: State-of-the Art Survey

Mohamed Yacine Tsalamlal[1], Nizar Ouarti[2], and Mehdi Ammi[1]

[1] CNRS/LIMSI, University of Paris-Sud, 91403 Orsay, France
`{yacine.tsalamlal,ammi}@limsi.fr`
[2] CNRS/ISIR, University of Pierre and Marie Curie, 75005 Paris, France
`ouarti@isir.upmc.fr`

Abstract. Haptic rendering technologies are becoming a strategic component of the new Human-Machines Interfaces. However, many existing devices generally operate with intrusive mechanical structures that limit rendering and transparency of haptic interaction. Several studies have addressed these constraints with different stimulation technologies. According to the nature of contacts between the device and the user, three main strategies were identified. This paper proposes to detail them and to highlight their advantages and drawbacks.

Keywords: haptic interface, intrusive device, workspace, stimulation strategies.

1 Introduction

Today, haptic technologies are involved in different fields. For instance, in the Product Lifecycle Management (PLM), where haptic based VR approaches have increased both speed and accuracy of human-computer interactions for the edition and the assembly of the 3D CAD models [Bordegoni, 2009]. In the field of the training and rehabilitation, haptics plays an important role for the training of sensory motor skills and to alleviate the motor system impairments [Broeren, et al.,]. Education, games, and entertainment can be also quoted as disciplines where several studies have shown the role of haptics to improve the learning and the interactivity though serious games approach [Sourina, 2011]. However, the deployment of haptic devices in public and industrial applications is still limited. This is due to the intrusiveness and the limit of some performance factors of existing haptic devices.

Astley and Hayward have listed a number of measures relating the performance of haptic devices [Hayward and Astley, 1996]. Among these, we can find the number of degrees of freedom, the nature of the contact between the user and the device, the amplitude of movement, the instantaneous and continuous forces, the peak acceleration, the inertia and damping of the device, the sensor resolution, the accuracy of the generated force, the stability of interaction and rendering, the environmental factors (such as noise), and the weight of the device.

The human haptic perception is extended in space and sensitivity range. Therefore, the haptic device design has to respond to workspace constraints and dynamics of sensory feedback. Moreover, all aspects of risk management related to the use of haptic devices have to be taken into account.

I. Oakley and S. Brewster (Eds.): HAID 2013, LNCS 7989, pp. 1–9, 2013.

The workspace of most actual haptic interfaces is often restricted even to making simple gestures. Moreover, haptic devices based on articulated robotic structures like exoskeletons [Nakai et al., 1998], or cable systems [Sato, 1992] have several limitations. In fact, these strategies adopt devices that must be physically connected to the user through mechanical systems to interact with the virtual object. These systems are often intrusive, limiting the comfort and the transparency of interaction with the environment. This survey aims at highlighting studies and new actuation technologies that address workspace and intrusiveness constraints.

2 State of the Art

According to the nature of contacts between the device and the user, we identified three main classes of strategies for non-intrusive haptic stimulation. The first approach corresponds to attached or wearable haptic devices. The second approach provides a limited contact only when tactile feedback is required. The last approach does not require direct contacts with the material device. We propose to detail these strategies in the following sections.

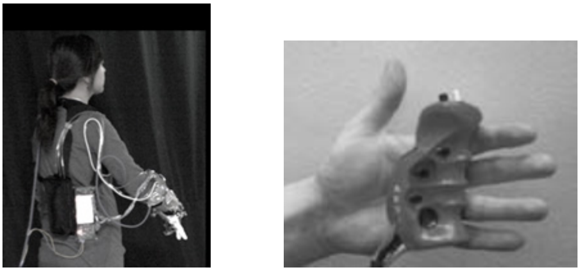
2.1 Attached Haptic Devices

The first strategies consisted in attaching tactile devices to the user's fingers or palms. For example, the CyberTouch device [CyberTouch, 2010] involves wearing a data glove equipped with vibrators that individually stimulate the fingers and palms (see Fig 1.a). Dave Anderson et al., [Anderson et al., 1999] developed a wearable tactile interface for virtual reality (Fig 1.b). This device consists in a pin-matrix of 2 × 3 electromagnetic actuators, mounted on a frame, and fixed on the finger of the user. Each actuator operates in the range of frequency of 8–100Hz, is at a maximum pressure of 1.2N/cm².

(a) (b)

Fig. 1. (a) Vibrotactile Glove device (CyberTouch). (b) Sandia National Laboratories Tactile.

Kim et al. [Kim, 2008] proposed a wearable, small, and lightweight tactile display system composed of piezoelectric ultrasonic actuator arrays (see Fig 2.a). In these strategies, the skin and the device are always in contact, which leads to undesired touch feelings. The Salford university group developed The ARRL, a pneumatic haptic interface (Fig 2b). It integrates three pneumatic actuators, in order to reproduce simple tactile feedback, when the hand (or the user virtual cursor) enters in contact with a virtual object [Stone, 2001]. However, even though data gloves allow more freedom of movement, they provide low spatial and force resolutions.

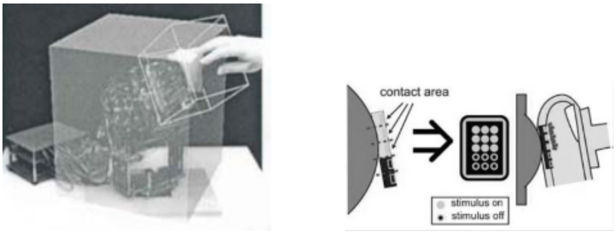


(a) (b)

Fig. 2. (a) Piezoelectric ultrasonic tactile display. (b) ARRL Interface.

2.2 Haptic Devices with Limited Contact

The second strategy consisted in controlling the positions of tactile devices so that they make contact with the skin only when tactile feedback is required. Hirota, and Hirose [Hirota, and Hirose, 1995] investigated the implementation of a haptic feedback for universal surface display (see Fig 3.a). The interface consists of 4 × 4 pins, on a square surface with a dimension of 20mm². Each pin has a stroke of 50mm. The different combinations of the pins' amplitudes produce different 3D surfaces that can be explored. Sato et al. [Sato, 2007] have proposed a multi-fingered master-slave robotic system featuring electrotactile displays on each finger of the master hand (Fig 3.b). The position of the electrotactile display is controlled so that it is in contact with the user's finger only when the slave robot grasps or touches objects. Major drawbacks of such systems are that they require bulky robot arms and a complicated control method. Drif et al. [Drif, 2004] proposed an alternative approach through a multi-level display concept. The kinesthetic feedback is based on moving and orienting a mobile surface on the contact point when a collision between the real finger and the virtual surface is detected. This approach makes it possible to touch an object without handling or wearing an intrusive mechanical structure. Moreover, it provides a good haptic interaction transparency. The main drawback of this strategy concerns the reactivity of the system and the limit of working space.

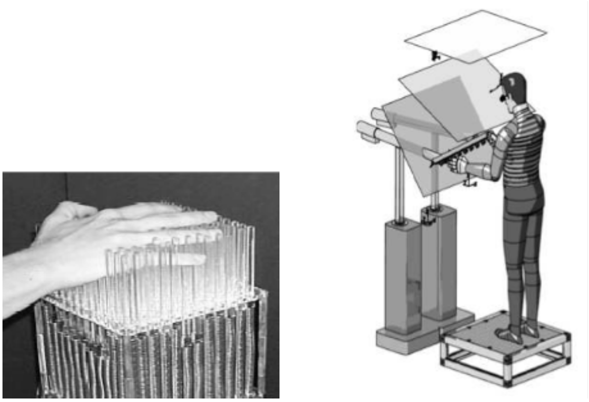


(a) (b)

Fig. 3. (a) Tactile interface based on returned of form. (b) Conceptual representation of the electrotactile system.

Bordegoni et al. [Bordegoni, 2010] proposed the use of an active tangible interface for the evaluation of virtual shapes of aesthetic products (SATIN). The system consists of two MOOG haptic devices which position and rotate a robotic spline. The user can modify the shape of the spline by applying local pressures (see Fig 4.b). The main advantage of this approach is the realism of the interaction with the 3D models. However, it cannot display an arbitrary surface.

The MATRIX (Fig 4.a) developed by Overholt et al. [Overholt et al., 2001], is a device that offers real-time control of a deformable surface, enabling the manipulation of a wide range of audiovisual effects. Interface consists of 144 rods which move vertically in a 12 by 12 grid, and are held up by springs at rest. The device uses this bed of rods to provide a large number of continuous control points, thereby improving the communication bandwidth with the computer. But, this device provides a limited workspace.



(a) (b)

Fig. 4. (a) MATRIX (A Multipurpose Array of Tactile Rods for Interactive eXpression). (b) Conceptual representation of the SATIN system.

2.3 Haptic Devices without Direct Contact

The last strategy consists in providing tactile feedback from a distance without any direct contact. Two approaches have been explored: air-jet-based strategies and acoustic-radiation-based strategies.

Air Jet Based Strategies. Several groups have explored the use of air jets as a means of conveying kinesthetic or tactile information. In the following sections we summarize the main strategies.

Haptic Feedback with Direct Contact with the Air Jet. Bianchi et al. [Bianchi, 2011] studied an air jet approach for generating a lump percept. This work aims to develop a tactile feedback system to enhance palpation using robot-assisted minimally invasive surgery (RMIS) (see Fig 5.a). The proposed approach consists of directing a thin stream of air through an aperture directly on the finger pad, which indents the skin in a hemispherical manner, producing a compelling lump percept. This work investigated the relationship between aperture size and air supply pressure by means of tactile sensor measurements and psychophysical experiments to understand how they affect the perceived pressure on the finger pad. The results suggested that the JND of air pressure on the finger pad is constant, regardless of aperture size. Moreover, this study clearly shows the effectiveness of well-prototyped air jets for tactile feedback. The main limit of this work concerns the stimulation distance, which does not exceed 2 cm. Recently, in the context of the design of a novel non-contact haptic device providing an important work space, Tsalamlal et al. [Tsalamlal et al., 2013] have proposed to use the air jet for direct tactile stimulation with greater distances (>10 cm). This configuration provides an important workspace for a better freedom of movement (Fig 5.b). It can also provide a greater stimulation area with a good force resolution. In this work, authors carried out psychophysical experiments in order to characterize the human perception of the air jet tactile stimulation.

Haptic Feedback through a Flexible Intermediate Surface. Inoue et al. [Inoue, 2009] presented a haptic device using a flexible sheet and air jet. This approach presents the haptic sensation of virtual lumps under the skin to the user's finger. A tensioned flexible sheet is regarded as a virtual skin. The user touches the sheet directly with his or her finger. Then he or she feels the softness of normal skin as the sheet compliance. A nozzle fires a thin beam of an air jet onto the back of the sheet when the user touches the sheet at the location of a virtual lump. The main drawback of this strategy concerns the limit of working space, since the exploration of complex surfaces requires the translation/orientation of the flexible sheet.

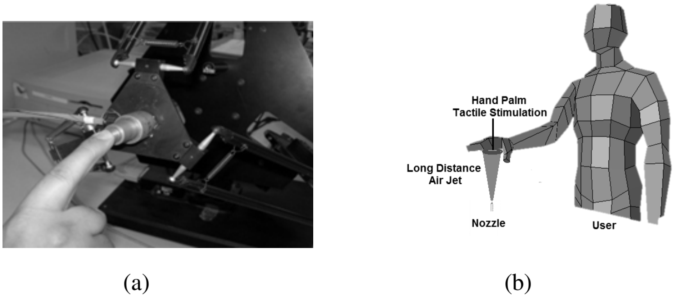
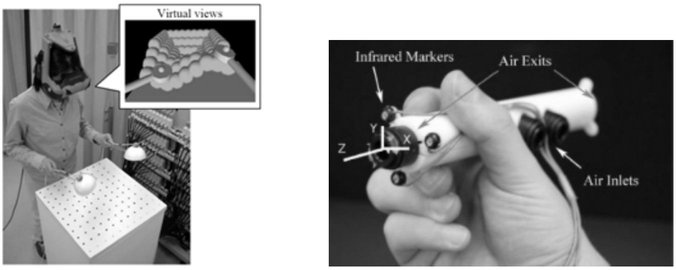


Fig. 5. (a) Air lump display for RIMS. (b) Conceptual representation of long distance air jet stimulation.

Haptic Feedback through a Rigid End-Effector. Suzuki [Suzuki, 2005] proposed a haptic device consisting of multiple air jets arranged in a matrix and air receivers held by the user (Fig .a). The air jets from the nozzles hit the air receiver and apply forces on the user's hand. The purpose of the receiver is to act as an interface for receiving the air jets. This haptic device enables interaction with static or dynamic surfaces. This strategy presents different drawbacks, such as the limit of the force and position resolutions. Moreover, the user needs to handle an end-effector.

Haptic Feedback with Portable Jets. Xu et al. [Xu, 1991] developed a system where a single air jet was mounted on the wrist of a user, and short force perturbations were applied to the user's arm by quickly opening and closing the air jet valve. The system generates a binary force sequence with a steady state thrust of 4 N. The signal frequency varies between 75 Hz and 150 Hz. This system was mainly used for the study and identification of the mechanical properties of the human arm joint. Romano and Kuchenbecker [Romano, 2009] presented a portable haptic interface (AirWand) with one degree of freedom (DoF), which provides a large workspace (15 m3 instead 0.006 m3 with standard devices). The system is based on a 1-DoF tool that has two air jets aligned along the longitudinal axis of the tool, indicated as air exits (see Fig 6.b). These two jets are used to create forces along the longitudinal axis in both the positive and the negative Z direction. The maximum peak force experienced during trials was around 7.58 N and the maximum continuous force was F = 3.16 N. Gurocak et al. [Gurocak, 2003] proposed a 3-DoF version of the same concept. The system consists of six jets attached around the wrist and oriented along three orthogonal axes. By controlling the air flow through each valve, the system commands three degrees of force output at the user's wrist. The main advantage of this strategy concerns the generation of a real kinesthetic feedback with an important workspace. However, the user needs to handle an end-effector, which limits the transparency of the interaction.



(a) (b)

Fig. 6. (a) Air jets matrix array for virtual reality. (b) AirWand device.

Acoustic-Radiation-Based Strategies. The second candidate for providing haptic feedback in the free space without direct contact with a mechanical structure is based on the acoustic radiation pressure generated by ultrasound transducers (see Fig 7). Takayuki et al. [Takayuki, 2008] proposed the use of acoustic-radiation pressure to provide tactile feedback in 3D environment.

Fig. 7. Airborn ultrasound tactile display

The concept is based on controlling the phase delays of acoustic waves to generate a focal point. The prototype consists of 91 airborne ultrasound transducers controlled through a 12-channel driving circuit. The measured total output force within the focal region is 0.8 mN, the spatial resolution is 20 mm, and the prototype produces vibrations of up to 1 kHz. A recent version of this interface significantly increases the working space and provides an output force of 16 mN at the focal point [Takayuki, 2010]. The main drawback of this haptic technology concerns the limit of force intensity. Moreover, the authors highlight some medical risks of interactions with sensitive regions (e.g., head and face). Currently, the interface is used with the hand and arm only. A medical study is in progress.

3 Conclusion

Haptic interfaces are relatively recent devices. These interfaces physically stimulate a user through tactile and kinesthetic perception, which provides a better presence and immersion in virtual environments. For the user, the issues focus on comfort and fidelity of the interaction and perception.

Current haptic devices operate essentially with articulated robotic structures, or cable systems. Most of these systems are intrusive and are not practical in many fields such as games and desktop applications. A number of systems, using direct (e.g., data gloves) or limited contacts with the user, allow more freedom of movement but provide low spatial and force resolutions. Currently, other works investigate non-contact stimulation technologies. Two main strategies were investigated: the acoustic radiations tactile stimulation, and the air jets tactile or kinesthetic stimulations.

The use of acoustic radiations does not allow high intensity stimulations and could present a significant risk for the user. The air jet approaches seem more promising. But, existing works exploit either intermediate object for interaction with air jet (i.e., air receiver), or exploit very short stimulation distances, restricting the user's workspace. However, more recent works are investigating actuation technologies based on air jet stimulation with longer distances, and larger workspace.

References

[Bordegoni, 2009] Bordegoni, M., Cugini, U., Belluco, P., Aliverti, M.: Evolution of a haptic-based interaction system for virtual manual assembly. In: Proceedings of the 3rd International Conference on Virtual and Mixed Reality, Berlin, Heidelberg, pp. 303–312 (2009)

[Broeren, 2008] Broeren, J., Georgsson, M., Rydmark, M., Sunnerhagen, K.: Virtual reality in stroke rehabilitation with the assistance of haptics and telemedicine. In: Proc. 7th ICDVRAT with ArtAbilitation, Maia, Portugal (2008)

[Sourina, 2011] Sourina, O., Wang, Q., Nguyen, M.K.: EEG-based Serious Games and Monitoring Tools for Pain Management. In: Proc. MMVR 18, Newport Beach, California, vol. 8(12), pp. 606–610 (February 2011)

[Bai, 2011] Bai, M.R., Tsai, Y.K.: Impact localization combined with haptic feedback for touch panel applications based on the time-reversal approach. J. Acoust. Soc. Am. 129(3), 1297–1305 (2011)

[Senseg, 2011] Senseg, electrostimulation (2011), `http://senseg.com/technology/`

[iPhone, 2010] iPhone Haptics, Google Code Page, University of Glasgow, `http://code.google.com/p/iphonehaptics/` (accessed August 29, 2010)

[Hayward and Astley, 1996] Hayward, V., Astley, O.R.: Performance measures for haptic interfaces. In: Robotics Research: The 7th International Symposium, pp. 195–207. Springer, Heidelberg (1996)

[Hayward et al., 2004] Hayward, V., Astley, O.R., Cruz-Hernandez, M., Grant, D., Robles-De-La-Torre, G.: Haptic interfaces and devices. Sensor Review 24(14), 16–29 (2004)

[CyberTouch, 2010] CyberTouch (2010), `http://www.estkl.com/products/datagloves/cyberglovesystems/cybertouch.html`

[Anderson, 1999] Anderson, T., Breckenridge, A., Davidson, G.: FGB: A Graphical and Haptic User Interface For Creating Graphical, Haptic User Interfaces, Sandia National Laboratories (1999)

[Kim, 2008] Kim, S.C., Kim, C.H., Yang, T.H., Yang, G.-H., Kang, S.C., Kwon, D.S.: SaLT: Small and Lightweight Tactile Display Using Ultrasonic Actuators. In: Proc. 17th IEEE Int'l Symp. Robot and Human Interactive Comm. (RO-MAN 2008), pp. 430–435 (2008)

[Stone, 2001] Stone, R.J.: Haptic Feedback: A Potted History, From Telepresence to Virtual Reality, MUSE Virtual Presence, Chester House, UK (2001)

[Hirota, and Hirose, 1995] Hirota, K., Hirose, M.: Surface Display: Concept and Implementation Approaches, ICAT/VRST. In: Int. Conf. on Artificial Reality and Tele-Existance, Japan, pp. 185–192 (1995)

[Sato, 2007] Sato, K., Minamizawa, K., Kawakami, N., Tachi, S.: Haptic telexistence. In: ACM SIGGRAPH 2007 Emerging Technologies, New York, Article 10 (2007)

[Drif, 2004] Drif, A., Citerin, J., Kheddar, A.: A multilevel haptic display design. In: 2004 IEEERSJ International Conference on Intelligent Robots and Systems IROS IEEE Cat No04CH37566, vol. 4, pp. 3595–3600 (2004)

[Bordegoni, 2010] Bordegoni, M., Ferrise, F., Covarrubias, M., Antolini, M.: Haptic and sound interface for shape rendering. In: Presence: Teleoperators and Virtual Environments, vol. 19(4), pp. 341–363. The MIT Press (August 2010)

[Overholt et al., 2001] Overholt, D., Pasztor, E., Mazalek, A.: A Multipurpose Array of Tactile Rods for Interactive sXpression, Technical Application, SIGGRAPH (2001)

[Bianchi, 2011] Bianchi, M., Gwilliam, J.C., Degirmenci, A., Okamura, A.M.: Characterization of an Air Jet Haptic Lump Display. In: 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (2011)

[Tsalamlal et al., 2013] Tsalamlal, M.Y., Ouarti, N., Ammi, M.: Psychophysical study of air jet based tactile stimulation. In: IEEE World Haptics Conference (accepted, 2013)

[Inoue, 2009] Inoue, K., Kato, F., Lee, S.: Haptic device using flexible sheet and air jet for presenting virtual lumps under skin. In: Proc. IEEE/RSJ Intl. Conference on Intelligent Robots and Systems, pp. 1749–1754 (2009)

[Takayuki, 2008] Iwamoto, T., Tatezono, M., Shinoda, H.: Non-Contact Method for Producing Tactile Sensation Using Airborne Ultrasound. In: Ferre, M. (ed.) EuroHaptics 2008. LNCS, vol. 5024, pp. 504–513. Springer, Heidelberg (2008)

[Suzuki and Kobayashi, 2005] Suzuki, Y., Kobayashi, M.: Air jet driven force feedback in virtual reality. IEEE Computer Graphics and Applications, 44–47 (2005)

[Xu, 1991] Xu, Y., Hunter, I.W., Hollerbach, J.M., Bennett, D.J.: An air jet actuator system for identification of the human arm joint mechanical properties. IEEE Transactions on Biomedical Engineering 38, 1111–1122 (1991)

[Romano, 2009] Romano, J.M., Kuchenbecker, K.J.: The AirWand: Design and Characterization of a Large-Workspace Haptic Device. In: Proceedings, IEEE International Conference on Robotics and Automation, pp. 1461–1466 (May 2009)

[Gurocak, 2003] Gurocak, H., Jayaram, S., Parrish, B., and Jayaram, U.: Weight Sensation in Virtual Environments Using a Haptic Device With Air Jets. Presented at J. Comput. Inf. Sci. Eng., 130–135 (2003)

[Takayuki, 2010] Hoshi, T., Takahashi, M., Iwamato, T., Shinoda, M.: Noncontact Tactile Display Based on Radiation Pressure of Airborne Ultrasound. IEEE Transactions on Haptics 3(3), 155–165 (2010)

Thermal Feedback Identification in a Mobile Environment

Graham Wilson[1], Stephen Brewster[1], Martin Halvey[2], and Stephen Hughes[3]

[1] School of Computing Science, University of Glasgow, G12 8QQ, UK
[2] Glasgow Caledonian University, Glasgow, G4 0BA, UK
`first.last@{glasgow.ac.uk, gcu.ac.uk}, stephenahughes@gmail.com`

Abstract. Audio and vibrotactile feedback are not always suitable or desirable, as noise and/or movement may mask them, and so thermal feedback may provide a salient alternative. In this paper, the identification of 'thermal icons' (structured thermal feedback) was tested as a means of conveying information when users were sitting and walking in an outdoor location. Overall identification rate for thermal icons was 64.6%, but identification of individual parameters was promising, at 94% accuracy for *direction of thermal change* (warming/cooling) and 73.1% accuracy for *subjective intensity* (moderate/strong). Results showed that walking outdoors did not significantly worsen icon identification compared to sitting outdoors, but the environmental temperature had a strong influence. Recommendations are given on how better to design and adapt thermal feedback for use in outdoor mobile scenarios.

Keywords: Thermal feedback, mobile interaction, non-visual feedback.

1 Introduction

Conveying information non-visually is important for mobile interaction, so that visual attention can be paid to the environment. Both Earcons [1] and Tactons [2] have been shown to be effective in conveying information in mobile scenarios. However, there are mobile environments in which audio and vibrotactile feedback may not be suitable, such as very loud (e.g., rock concerts) or very quiet places (e.g., libraries or religious buildings) for audio, or simultaneously loud and bumpy environments, such as public transport, that are unsuitable for both [7]. Thermal feedback is entirely silent and so may be suitable for quiet environments. It could also be more salient in bumpy environments. Further, user preference for when a feedback modality is desired varies by location and situation [7] and thermal feedback may provide a third alternative to audio and vibrotactile feedback. Thermal feedback is an under-studied aspect of touch and so warrants further investigation into potential uses and benefits.

HCI research has measured 'yes/no' detection and subjective comfort/intensity ratings of thermal stimuli in indoor [20] and outdoor [5] environments. However, this research only looked at whether any changes were felt, and not absolute identification of the unique form of those changes (as has been done with Earcons [1] and Tactons

I. Oakley and S. Brewster (Eds.): HAID 2013, LNCS 7989, pp. 10–19, 2013.

[2]). There is an important difference between simply acknowledging a change in stimulation and being able to uniquely identify specific forms and decode the meaning of them. For thermal feedback to be a viable information source, users must be able to identify unique forms of thermal stimulation. Previously [19], we developed two-dimensional structured 'thermal icons', which could be identified with 83% accuracy, however, the participants in this study were sat indoors. Both walking [20] and environmental temperatures [6, 15] significantly influence thermal perception, so it was necessary to test identification of thermal icons when the user is sitting and walking outdoors, to judge the feasibility of thermal feedback for more realistic mobile interaction. As such, this paper reports an experiment that tested identification of two-parameter thermal icons presented to the palm of the hand from a mobile device when the user was both sitting and walking in an open-air, outdoor environment. Ambient temperature and humidity were measured to study potential environmental influences.

2 Related Work

A distinction is made here between perception and identification of thermal stimuli. *Perception* of thermal changes – moving from no sensation to the production of a sensation – is well understood and involves simply acknowledging a change has occurred. However, *identification* of thermal changes – classifying unique forms of encoded stimuli and using those to convey information – is not well understood.

Thermal *perception* can be highly precise, with experts able to detect changes of < 0.2°C from skin temperature in ideal laboratory conditions [14]. The skin naturally rests at a neutral temperature of between 26°C and 36°C in moderate environmental temperatures [9, 13]. Detection of changes within this range is dependent more on the rate of change (ROC) of the stimulus than the actual extent of the change itself [13]. Faster changes feel stronger and are felt sooner than slow changes. Cold perception is generally faster [11, 20] and more precise [10] than warmth perception. Cold or warm environments have the effect of cooling or warming the skin, respectively [6, 15] and we become more sensitive to changes that move further away from neutrality towards the pain thresholds [13]. The thermal sense is not a good "thermometer": it is not good at identifying specific temperatures, as it is based on changes in the overall magnitude of sensation, which translates into a subjective appraisal of the intensity.

Wettach *et al.* [17] trained users to uniquely *identify* five different degrees of warmth at up to 75% accuracy after several days of training. They also report the use of five temperatures to indicate the correct direction of travel in an outdoor navigation task, but few details about the hardware, experimental design or results are given. Exactly what temperatures were used, and so how different they were from each other, was not reported. Other research has attempted to communicate affective information thermally. Suhonen *et al.* [16] studied how thermal feedback was used to convey emotions during remote communication and found that warmth was used to represent or reinforce agreement/positivity, while cold represented disagreement/negativity. However, they did not examine user responses to varying extents of warmth/cold, or their identification of those extents. Iwasaki *et al.* [8] suggested conveying emotional state using warmth on a mobile device, but did not test feedback perception. Emotional responses to thermal stimuli have also been measured by Salminen *et al.* [12] and Halvey *et al.* [4], but only subjective perception of stimuli was measured, not identification. Only Wettach *et al.* [17] had users in a mobile setting, and the paper provides no details on how well the stimuli could be differentiated.

Wilson *et al.* [20] measured *perception* and subjective comfort/intensity ratings of various thermal stimuli for use in HCI when the user was sitting and walking indoors. They found that walking significantly reduced the number of stimuli detected. Halvey *et al.* [5] found that outdoor environmental temperatures also influenced perception of stimuli, with particularly low and high temperatures leading to poorer perception/detection, but only tested perception when sitting. Based on these results, we previously designed and tested *identification* of two-dimensional 'thermal icons' which could convey information during mobile interaction [19]. This structured thermal icon design has the advantage of being capable of conveying two pieces of information. We used two thermal parameters: direction of change (warming and cooling) and subjective intensity of change (moderate and strong) to create four icons conveying the "Source" (Personal or Work) and "Importance" (Standard and Important) of a received text message. Users in this study were able to uniquely identify the two pieces of information with an accuracy of 82.8%, but did so sitting indoors.

Thermal feedback is promising, as it is a truly private feedback method, while vibrations can still be heard or felt by those nearby (for example, sitting on the same bench). It provides unique sensations and is also inherently hedonic [13]. However, if thermal feedback is to be considered a useful alternative means of conveying information in mobile interaction, absolute identification of unique, coded forms of thermal stimulation must be tested in realistic outdoor environments. Currently only *perception* of thermal changes has been tested outdoors and only when the individual was sitting. Therefore, we ran a study testing identification of two-dimensional thermal icons using compact hardware when the participants sat on a bench and walked a route in an open-air outdoor environment.

Fig. 1. Peltier modules used to produce thermal icons (left); attached to back of mobile (right)

3 Evaluation

The apparatus was built by SAMH Engineering and consisted of two Peltier modules, each attached to a heat sink for heat dissipation (see Figure 1, left). The Peltiers were controlled (with ~0.1°C accuracy) by a small microcontroller, powered by four AA batteries. Custom software on a Nexus One Android device (see Figure 1, right) communicated with the microcontroller over Bluetooth. The Peltiers and heat sinks were attached to the back of the Nexus One, in a position to make contact with the palm of the left hand, which held the device. Both the microcontroller and the battery pack were placed in a small shoulder bag that the participant carried (Figure 2). The apparatus was entirely silent: the Peltiers made no audible sounds when in operation.

3.1 Thermal Icons

The thermal icons were designed to convey two pieces of information: the “Source” and “Importance” of a hypothetical text message. The Source could be either “Personal” or “Work” and the Importance could be either “Standard” or “Important”. This gave four different message types: Standard Personal, Important Personal, Standard Work and Important Work. The thermal icons were created in our previous research [19], based on thermal perception when sitting and walking indoors [20] and sitting outdoors [5]. Two salient parameters of thermal stimulation were used to create the icons: *direction of thermal change* and *subjective intensity of change*. Each of these had two levels: Warming and Cooling for *direction of change* and Moderate and Strong for *subjective intensity*, giving four thermal icons: Moderate Warmth, Strong Warmth, Moderate Cooling and Strong Cooling.

A starting neutral skin temperature of 32°C was chosen, as it sits within the skin’s resting thermal range [9] and stimuli warmed and cooled from there. Warmth represented Personal messages, as there is evidence of an innate association between physical warmth and interpersonal warmth or trust [18]. Work messages are an alternative to personal messages and so were mapped to cool changes. More important messages were mapped to subjectively stronger changes as they are more attention-grabbing [13]. Both the extent of thermal change (Δ temperature from skin temperature) and the rate of temperature changes (ROC) influence the perceived magnitude of sensation [20]. Therefore, the two *subjective intensity* levels of ‘Moderate’ (Standard) and ‘Strong’ (Important) were created by mixing both Δ temperature change and ROC. Changing temperature by 3°C at 1°C/sec produced the ‘Moderate’ intensity and changing by 6°C at 3°C/sec produced the ‘Strong’ intensity [20]. These Δ and ROC values were chosen based on stimuli that produced detectable sensations in previous research [20], as smaller Δ values were less likely to be detected outdoors [5]. These changes were in both *directions*, starting from 32°C, giving thermal icons of:

- Strong Cooling 6°C @ 3°C/sec, to 26°C: Important Work message
- Moderate Cooling 3°C @ 1°C/sec, to 29°C: Standard Work message
- Moderate Warmth 3°C @ 1°C/sec, to 35°C: Standard Personal message
- Strong Warmth 6°C @ 3°C/sec, to 38°C: Important Personal message

3.2 Design and Procedure

Thirteen participants took part (2 female), aged from 22 to 31 (mean 25.6), and were paid £6 for participation. The evaluation had a within-subjects design, with Mobility (sitting, walking) as a factor. The experiment was conducted in an enclosed courtyard adjacent to a university building. There were benches to test icon identification when sitting outdoors, and large, flat concrete paths to test identification when walking outdoors. The area was quiet, away from road traffic, but there was a degree of footfall from students, staff and tourists. A nearby indoor area was used for instruction and training, so that icons could be learnt in a thermally stable environment.

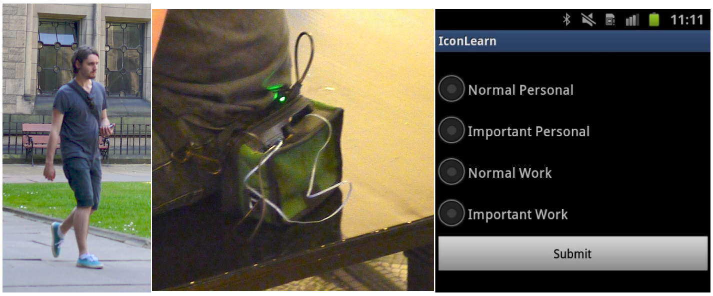


Fig. 2. Participant walking in the courtyard (left), carry bag (centre) and task GUI (right)

The procedure was the same for the sitting and walking conditions and participants took part in both, in a counterbalanced order. The task started with a 10-minute training session seated in the indoor location. First, the mapping of feedback parameters to message types was explained. The training session started with 60 seconds of skin "adaptation", where the Peltiers were set to the neutral starting temperature of 32°C and held in the hand, to equalize skin and Peltier temperatures. The participants were then given 10 minutes to feel each thermal icon as many times as desired and learn the mapping of icon to message type. Training software on the Nexus One showed four radio buttons, each labeled with one of the message types (Figure 2, right). After training, all participants expressed confidence in having memorized the mappings. The participants then took part in the first mobility condition, followed by the second.

In both conditions, the Nexus One was held in the left hand and input was given by the right. The palm of the hand was chosen for stimulation as it is the most sensitive area [14, 20], but other locations may also be suitable such as the wrist/forearm and upper arm [20], where a watch or exercise arm band, augmented with thermal elements, could be worn. During the sitting condition, participants simply sat on the bench. During the walking conditions, participants were asked to walk in a simple square route around the courtyard at their normal walking pace (Figure 2). During both conditions, each icon was presented four times in a random order, with a 30 second gap between subsequent presentations, where the Peltiers were returned to 32°C. The device screen showed the same radio buttons as during training. Participants were instructed, whenever they identified an icon, to press on the radio button corresponding to the interpreted message type, and press "Submit". The Peltiers were immediately set back to 32°C, the user response was recorded and another icon was presented at random after 30 seconds. If the system received no input within 20 seconds of an icon being presented, a "missed" event was logged, and a different icon was presented.

The Independent Variables were *Mobility* (sitting and walking) and *Icon* (four icons). The Dependent Variables were: *Accuracy* (whether the right message type was identified) and *Identification Time* (the time between the start of an Icon presentation and when a response was recorded by the "Submit" button). For overall Accuracy, both pieces of information had to be correctly identified. Accuracy rates for both parameters individually were also recorded and "missed" events counted as an error in all three Accuracy measures. Environmental temperature and humidity were both recorded throughout the study using a thermometer.

4 Results

Influence of Ambient Temperature & Humidity. The mean temperature across all conditions was 20.27°C (SD = 3.99; min = 12.7°C, max = 27.4°C). The potential relationship between environmental Temperature and Accuracy was investigated using Pearson's product-moment correlation coefficient. A significant negative correlation was found between Temperature and Accuracy (r (23) = -.562, $p<0.01$), with Accuracy decreasing as Temperature increased. Humidity had a significant negative correlation with Temperature (r (22) = -.872, $p<0.01$). Humidity ranged from 46.6% to 87.1%, with a mean of 64.22% (SD = 9.29). There was a positive but non-significant relationship between Humidity and Accuracy (r (23) = .381, $p>0.05$).

Table 1. Thermal Icon confusion matrix, showing each icon presented and the number of each icon they were perceived as

	Perceived As				
	Mod Warm	**Strong Warm**	**Mod Cool**	**Strong Cool**	**Missed**
Mod Warm	70	10	3	1	10
Strong Warm	31	57	1	3	2
Mod Cool	0	3	54	33	4
Strong Cool	2	1	27	62	2

Accuracy. Pearson's correlation coefficient found no relationship between trial number and Accuracy (r (30) = .149, $p>0.05$), suggesting performance was similar throughout the study. The overall identification rate for the two-parameter thermal icons was 64.6% (SD = 47.75). The mean Accuracy for each individual thermal parameter was 96.3% for direction of change and 73.1% for subjective intensity, with 18 missed thermal icons (4.79% of all icons). The confusion matrix for thermal icons is shown in Table 1. Accuracy data was not normally distributed, so non-parametric analyses were used. A Wilcoxon T test found no effect of Mobility on Accuracy, as identification rates were similar when walking (mean = 61%, SD = 49.0) to when sitting (mean = 69%, SD = 46.5). A Friedman's test found no effect of Icon on identification rate either, with mean Accuracy of 74% (SD = 43.8), 61% (SD = 49.1), 57% (SD = 49.7) and 66% (SD = 47.6) for the moderate warm, strong warm, moderate cool and strong cool icons respectively.

Identification Time. Identification time (IDT) correlated positively and significantly with Trial number (*r* (30) = .423, $p<0.05$), with IDT increasing as the number of completed trials increased: identification became more time-consuming over time. IDT data was also not normally distributed, so non-parametric analyses were used. A Wilcoxon *T* test found no effect of Mobility on IDT, with mean times of 9.20s (SD = 3.88s) when sitting and 8.57s (SD = 3.78s) when walking. A Friedman's test found no effect of Icon on IDT, with means of 8.33s (SD = 3.27), 8.83s (SD = 3.77), 9.00s (SD = 3.80) and 9.48s (SD = 4.36) for Moderate Warm, Strong Warm, Moderate Cool and Strong Cool icons respectively. Overall mean IDT across all icons was 8.91s.

5 Discussion

There were some encouraging results from the study, however, some significant issues were encountered which have major implications for the use of thermal feedback for conveying information in mobile, outdoor interaction. Therefore, both positive recommendations and potential obstacles are discussed, which others might draw upon and use to advance the design of thermal feedback.

5.1 Saliency of Direction of Change

The results show that *direction of change* was extremely well identified, at 96%. Therefore, basic warming and cooling thermal stimulation is highly salient, even when walking outdoors, and both warm and cold stimuli can be felt using simple, compact apparatus. The low-bandwidth feedback designs that simply warm or cool to provide information are therefore likely to be useful even when walking outdoors. Also, thermal *direction of change* may be a suitable replacement for problematic Tacton parameters (roughness or spatial location) for mobile interaction [2].

Recommendations. *Direction of change* is a useful parameter for thermal feedback in mobile environments. If only a single piece of information, with two alternatives, is to be conveyed, then thermal feedback *direction of change* is a suitable means. Based on results here and elsewhere [20], a change of at least 3°C is recommended.

5.2 Walking Does Not Significantly Impair Identification

Walking outdoors did not significantly affect identification Accuracy (61%), compared to sitting outdoors (69%) in this study. Given the negative effects on perception of thermal changes from walking [20] and environmental temperatures [5] individually, a more pronounced drop in identification when the two influences acted together might be expected. While the Accuracy for sitting and walking outside is quite low, it is encouraging that there appears to be only a small interaction cost when walking.

Recommendations. This result suggests that thermal feedback may be as suitable for use when walking outdoors as when sitting outdoors, however, future research should test identification in a wider range of realistic mobile scenarios, such as on transport.

5.3 Environmental Influences

Outdoor environmental temperatures significantly impacted identification of thermal icons as, even within the small range of temperatures recorded during the experiment (13-27°C), thermal icons became significantly less identifiable as temperature increased, with long Identification Times of 8-9 seconds. Participants also took longer to identify icons as time went on. No significant correlation was found between trial number and Accuracy, however, so it seems that extra time was taken to maintain Accuracy. Environmental temperature influences skin temperature, which, in turn, influences thermal perception [6, 13, 15]. Warmer environments (and walking) may have elevated the temperature of skin surrounding the Peltiers, potentially leading to domination or referral [3] of warmth and erroneous interpretation of greater warmth at the stimulation site. This could then mask warm changes, as the Δ between skin and stimulus is smaller; and also enhance cooling changes, as Δ is then larger.

Average identification accuracy for both bits of information was 64%, ranging from 87% (during 14.9°C outdoor temperature) down to 33% (at 25-26°C outdoor temperature). The overall value is markedly lower than the 83% accuracy we found for the same thermal icons when sitting indoors [19], although the high value of 87% is slightly higher. The individual differences are worthy of note, however, as one participant managed only 62% at 13.5°C and another managed 83% at 23°C.

Recommendations. From the results, we hypothesize that feedback designs may have to adapt to the environment and adjust the starting temperature and/or the extent/rate of thermal change (see Issues with Subjective Intensity, below) to make the feedback more salient. An example might be to match the starting temperature to current skin temperature. Future research should examine dynamically adjustable feedback.

5.4 Issues with Subjective Intensity

The main source of error in the study came from the subjective intensity (SI) parameter, as 73.1% were identified correctly (similar to Wettach *et al.* [17]). Analysing the confusion matrix shows that more cold SI were confused than warm SI (60 vs. 41 respectively). This is unexpected, as we are generally more sensitive to cold stimuli [10], however cold stimuli may also feel less intense than warm stimuli [20], which may mean that they were more difficult to tell apart. The Strong Warm icon was closer to the heat pain threshold than the Strong Cold icon was to the cold pain threshold. This may have given the Strong Warm a unique, more intense, quality, making it easier to tell apart from the moderate warmth, a distinction possibly lacking in the two cold icons. Given the performance of subjective intensity, the range of thermal stimuli that can be used to convey information in mobile HCI may be limited. The icon design was based on research that suggests faster and larger changes feel subjectively more intense than slower, smaller changes [11, 15, 20]. The ROCs and temperature Δ values used here may simply not have been fast or large enough to reliably tell apart.

Recommendations. To use subjective intensity as a way of conveying information in mobile HCI, the temperature Δ values should be larger than the 3°C used here, and/or the rates at which the temperature is changed should be more different than 1°C/sec vs. 3°C/sec. Only two Δ values and two ROCs were used in the design of thermal icons. A more thorough examination of the different possibilities could yield stimuli that are more reliably perceivable and perceptually distinct, compared to those used here. Alternatively, our previous work suggested that thermal and vibrotactile feedback can be combined to produce salient “intramodal” icons [19].

5.5 Feedback Alternatives

If subjective intensity remains an unreliable parameter in icon design, a replacement would need to be sought. Area of stimulation and spatial location are example candidate parameters and research should be conducted to test their suitability. Spatial location may be the more suitable parameter of the two, and has been used successfully in Tactons [2]. Varying the area of stimulation also varies the subjective intensity of the sensation [13], which this research has shown to be a problematic means of conveying information outdoors. However, using a larger stimulator may make feedback more salient [13]: only two $2cm^2$ Peltiers were used in our research.

6 Conclusions

This experiment has been the first to test absolute identification of encoded, two-dimensional thermal icons presented from a mobile phone while participants sat and walked outdoors. In this way, two pieces of information could be conveyed. Identification of both bits of information was lower than expected, at 64.6%, but identification of each individual thermal parameter was promising, particularly direction of thermal change (warming/cooling). Walking outdoors also had no significant impact on identification compared to sitting outdoors. Environmental temperature significantly affected information transmission, however, so our findings have led to several recommendations about how thermal feedback may be better designed to suit mobile interaction, and so improve thermal icon design.

References

1. Brewster, S.: Overcoming the lack of screen space on mobile computers. Personal and Ubiquitous Computing 6, 188–205 (2002)
2. Brown, L., Brewster, S., Purchase, H.: Multidimensional Tactons for Non-Visual Information Presentation in Mobile Devices. In: Proc. MobileHCI 2006, pp. 231–238 (2006)
3. Green, B.: Localization of Thermal Sensation - Illusion and Synthetic Heat. Perception & Psychophysics 22(4), 331–337 (1977)

4. Halvey, M., Henderson, M., Brewster, S., Wilson, G., Hughes, S.: Augmenting Media with Thermal Stimulation. In: Magnusson, C., Szymczak, D., Brewster, S. (eds.) HAID 2012. LNCS, vol. 7468, pp. 91–100. Springer, Heidelberg (2012)
5. Halvey, M., Wilson, G., Brewster, S., Hughes, S.: “Baby It’s Cold Outside”: The Influence of Ambient Temperature and Humidity on Thermal Feedback. In: Proc. CHI 2011, pp. 715–724 (2011)
6. Hirosawa, I., Dodo, H., Hosokawa, M., Watanabe, S., Nishiyama, K., Fukuichi, Y.: Physiological Variations of Warm and Cool Sense with Shift of Environmental-Temperature. Int. J. Neuroscience 24(3-4), 281–288 (1984)
7. Hoggan, E., Crossan, A., Brewster, S.: Audio or tactile feedback: which modality when? In: Proc. CHI 2009, pp. 2253–2256 (2009)
8. Iwasaki, K., Miyaki, T., Rekimoto, J.: AffectPhone: A Handset Device to Present User’s Emotional State with Warmth/Coolness. In: BIOSTEC 2010 (2010)
9. Jones, L.A., Berris, M.: The Psychophysics of Temperature Perception and Thermal-Interface Design. In: Proc. HAPTICS 2002, pp. 137–142 (2002)
10. Kenshalo, D., Holmes, C., Wood, P.B.: Warm and Cool Thresholds as a Function of Temperature Change. Perception & Psychophysics 3(2A), 81–84 (1968)
11. Pertovaara, A., Kojo, I.: Influence of the rate of temperature change on thermal thresholds in man. Experimental Neurology 87(1), 439–445 (1985)
12. Salminen, K., Surakka, V., Raisamo, J., Lylykangas, Pystynen, J., Raisamo, R., Makela, K., Ahmaniemi, T.: Emotional Responses to Thermal Stimuli. In: Proc. ICMI 2011, pp. 193–196 (2011)
13. Stevens, J.C.: Thermal Sensibility. In: Heller, M.A., Schiff, W. (eds.) The Psychology of Touch. Lawrence Erlbaum, New Jersey (1991)
14. Stevens, J.C., Choo, K.: Temperature sensitivity of the body surface over the life span. Somatosensory & Motor Research 15(1), 13–28 (1998)
15. Strigo, I., Carli, F., Bushnell, M.: Effect of ambient temperature on human pain and temperature perception. Anesthesiology 92(3), 699–707 (2000)
16. Suhonen, K., Muller, S., Rantala, J., Vaananen-Vainio-Mattila, K., Rasiamo, R., Lantz, V.: Haptically Augmented Remote Speech Communication: A Study of User Practices and Experiences. In: Proc. NordiCHI 2012, pp. 361–369 (2012)
17. Wettach, R., Behrens, C., Danielsson, A., Ness, T.: A thermal information display for mobile applications. In: Proc. MobileHCI 2007, pp. 182–185 (2007)
18. Williams, L., Bargh, J.: Experiencing physical warmth promotes interpersonal warmth. Science 322, 606–607 (2008)
19. Wilson, G., Brewster, S., Halvey, M., Hughes, S.: Thermal Icons: Evaluating Structured Thermal Feedback for Mobile Interaction. In: Proc. MobileHCI 2012, pp. 309–312 (2012)
20. Wilson, G., Halvey, M., Brewster, S., Hughes, S.: Some Like it Hot? Thermal Feedback for Mobile Devices. In: Proc. CHI 2011, pp. 2555–2564 (2011)

Cold or Hot? How Thermal Stimuli Are Related to Human Emotional System?

Katri Salminen[1], Veikko Surakka[1], Jukka Raisamo[2], Jani Lylykangas[1],
Roope Raisamo[2], Kalle Mäkelä[3], and Teemu Ahmaniemi[4]

[1,2] Tampere Unit for Computer-Human Interaction, Kanslerinrinne 1,
33014, University of Tampere, Finland
[1] Research Group for Emotions, Sociality, and Computing
[2] Multimodal Interaction Research Group
`{katri.salminen,veikko.surakka,jukka.raisamo,`
`jani.lylykangas,roope.raisamo}@sis.uta.fi`
[3] Tampere University of Technology, Department of Electronics,
P.O. Box 692, 33101 Tampere
`kalle.makela@tut.fi`
[4] Nokia Research Center, Itämerenkatu 11 – 133, 00180 Helsinki, Finland
`teemu.ahmaniemi@nokia.com`

Abstract. The aim was to study emotional responses to thermal stimulation. Stimuli were varied by increasing or decreasing temperature by 2, 4 or 6°C in respect to the participants' hand temperature. The stimuli were either dynamic (i.e. heated or cooled while touching) or pre-adjusted (i.e. heated or cooled to the target temperature before touching). The results showed, for example, that 6°C change in temperature was rated as unpleasant, arousing, dominant, and avoidable especially when the stimulus was warm. 4°C increase was rated as arousing, dominant, and pleasant. In addition, pre-adjusted 6°C increase elevated the physiological arousal in terms of skin conductance response.

Keywords: Affective haptics, Thermal stimulation, Human emotions.

1 Introduction

We often use thermal attributes while describing meaningful events in our lives. For example, over a lunch hour conversation we can describe someone as a warm person or as cold as ice. Meeting an old friend by coincidence can be heartwarming while losing a dear pet can make one feel cold inside. Therefore, it is not surprising that temperature is frequently argued to be connected with sociality and emotions. Already in the 1950's Harlow [1] showed that touch and warmth of a caregiver was essential for normal social and emotional development of monkey infants. Temperature has also short term effects on social behavior. For example, holding a cup containing warm coffee was found to make one characterize other people in more positive manner than holding a cup containing cold coffee [2]. In general, it seems that the experience of pleasantness is affected by thermal stimulation so that warm temperatures are perceived as comfortable and pleasant [3] while cold temperatures are perceived as uncomfortable [4].

I. Oakley and S. Brewster (Eds.): HAID 2013, LNCS 7989, pp. 20–29, 2013.

In human-technology interaction (HTI) thermal stimulation has been studied in a lesser extent. Most of the previous studies have concentrated, for example, on providing information on the temperature of objects via thermal sense [5]. The results have shown that people are able to recognize object surface materials (e.g. foam or copper) based on only thermal information. In these studies participants' hands have been preheated to improve thermal sensitivity.

A recent study [6] has mapped how easily people can recognize small changes in temperature when the stimulation is provided, for instance, to arm area with thermal actuators. In this study also the subjective experience of thermal comfort was taken into account by asking the participants whether the stimulus was uncomfortable or comfortable. The results showed that people can detect even one degree temperature changes, and that intense temperature changes (e.g. 6°C) are experienced as more uncomfortable than less intense changes (e.g. 1°C).

Systematic studies mapping emotional responses to thermal stimuli are virtually non-existent in the field of HTI. When studying ratings of emotional responses to haptic stimulation [7], the use of bipolar rating scales has been functional. These rating scales are based on the dimensional theory of emotions which suggests that there are three basic dimensions (pleasantness, arousal, and dominance) covering the dimensional emotional space when rating different types of stimuli [8]. As these three dimensions have been associated to a fourth dimension (i.e. ones' approach – withdrawal behavior) this motivational tendency has recently been measured as well [9]. In addition, in emotion studies psychophysiological responses to stimuli are often detected. Especially, the level of emotional arousal can be analyzed using the skin conductance response (SCR) measurement. SCR is argued to be faster and higher in magnitude to arousing stimuli and slower and smaller in magnitude to calming stimuli [10].

Studying thermal stimulation can be challenging. The temperature of the skin is dependent on the current body temperature, and it fluctuates based on environmental conditions. In addition, when the user operates with a device equipped with thermal actuator, the temperature of the thermal actuator is roughly the same as the user's hand. Thus, the skin temperatures should be defined (i.e. measured) before stimulus presentation in order to be able to stimulate the participants hand exactly with certain temperatures. Therefore, we have chosen an approach where we first measure the participants hand temperature and then present the change in temperature (i.e., stimulus) in respect to that temperature.

In a pilot study [11] we tested how thermal stimulation might evoke emotions. Two target temperatures (4°C increase and decrease in respect to the hand temperature) were presented to the participants with two presentation methods (i.e. pre-adjusted and dynamic). The task was to rate the stimuli with previously mentioned emotion-related scales while SCR was measured. A 4°C increase was rated as arousing and dominant when compared to 4°C decrease suggesting that warm stimuli are activating the dimensions of arousal and dominance while cold stimuli do that in a lesser degree. The SCR was higher in magnitude when the stimulus was warm than when the stimulus was cold confirming the results of subjective ratings of arousal. Interestingly, the presentation method affected the speed of the SCR. The warm stimuli activated the SCR faster when the stimulus was pre-adjusted and cold when the stimulus was dynamic. Together, both obtained SCR results suggest that SCR is responding mostly to pre-adjusted stimuli. However, it the temperature is cold, dynamic presentation method may be better to evoke a rise in SCR.

Even though the results of the pilot study indicated that thermal stimuli can be related to the functioning of human emotional system, the study had several limitations. First, we had only 8 participants. Second, only two target temperatures were used. Third, the algorithm used to create the stimuli was sensitive to the contact with the skin so that, for example, when the participant was touching the thermal actuator when it was cold the temperature was slightly shifted towards to the hand temperature from the aimed target temperature. Therefore, the results were only indicative of the way temperature variations could affect human emotions. In the current study we aimed to gain further insight about temperature and emotions. For this purpose, we recruited 24 participants and presented them six target temperatures with two presentation methods. The algorithm for stimulus presentation was improved. Participants rated the stimuli with four emotion-related rating scales and one asking their ratings about the temperature. Physiological responses were analyzed by SCR measurements.

2 Methods

2.1 Participants

24 voluntary participants (12 female and 12 male with a mean age of 25.4 years, range 18–46) took part in the study. All had normal sense of touch by their own report. Their mean hand temperature was 30.3°C (range 23.9–33.5°C). The experiment was approved by the Department of Science Center of the Pirkanmaa hospital district.

2.2 Apparatus

The thermal stimuli were provided with a Peltier thermal actuator. The actuator contained a Multicomp PF-031-10-25 (15mm x 15mm x 4.8mm) thermoelectric module attached to an aluminum heatsink with double-sided heat-conducting tape. The actuator also had an integrated Labfacility DM-310 PT1000 temperature sensor attached to the surface of the actuator to monitor the temperature of the actuator. The actuator was driven with a driver module that was hosted by an Atmel ATmega324P microcontroller to receive drive instructions via serial connection and convert the instructions into PWM signal to drive the actuator accordingly. E-Prime 2.0 Professional® application suite [12] was used for running the experiment and for collecting the user input. An Arduino prototyping platform with Atmel ATMega328P was used both for controlling the driver module via serial connection and for listening to another serial connection where E-Prime sent the ID of the next stimulus. After receiving the stimulus ID Arduino queried the temperature of the hand placed over a Melexis MLX90614 IR temperature sensor. Then, Arduino used a simple proportional–integral-derivative (PID) control loop to adapt the Peltier temperature close to that of the hand to meet the starting condition. Third, Arduino used the PID control loop to meet the parameters defined by the current stimulus with the user's hand placed over the Peltier actuator. When running the control loop the driver module constantly sent temperature data from the actuator that was used to refine the driving instructions. All the temperature data received from the actuator as well as the initial hand and room temperature measurements were written to a log for later inspection. The control loop was programmed to monitor the surface temperature of the Peltier during the stimulus presentation. If the temperature of the Peltier actuator went below 17°C or above 43°C, the experiment was automatically terminated. This was done avoid the possibility of the actuator to damage the skin (e.g., cause skin burn). The SCR was measured with Nexus-10 platform and Biotrace+ software version 2010A. A Bluetooth connection was used to transfer SCR data to a laptop PC during the measurement.

2.3 Design and Stimuli

The experiment was a within-subject repeated measures design. Six target temperatures were created by decreasing or increasing the stimulus temperature by 2, 4 or 6°C in respect to the participants hand temperature measured before the stimulus presentation. The stimuli were divided in two groups based on whether the presentation method was dynamic or pre-adjusted. When the presentation method was dynamic, the participants sensed the cooling or heating towards the target temperature during the stimulus presentation. When the presentation method was pre-adjusted the stimulus was heated to the target temperature before the participant was allowed to touch the stimulus, and then the temperature was kept constant during stimulus presentation. To compare the possible effects of thermal variation to a neutral reference point, a stimulus where the temperature was kept constant at the participant’s hand temperature (i.e. neutral stimulus) was presented in both stimulus groups. Therefore, we had a total of 14 stimuli (i.e. dynamic increase or decrease of 2, 4 and 6°C, pre-adjusted increase or decrease of 2, 4 and 6°C, and two neutral stimuli). Because the hand temperature of the participants varied, the participant with lowest hand temperature received stimuli at the range of 17.9 - 29.9°C, and the participant with highest hand temperature at the range of 27.5 - 39.5°C.

2.4 Procedure

After arriving to the laboratory the participant was seated in a chair, sensors to measure SCR throughout the experiment were attached to the participants’ nondominant hand’s index finger and middle finger. The experiment was divided in two experimental blocks. In the first block a stimulus presentation trial proceeded as follows. The participant put the dominant hand’s palm on the infrared sensor for 10 s to measure the current hand temperature with a sampling rate of 10 Hz. The average of these measurements was used as the current hand temperature. Then, the participant put the same palm on the Peltier actuator to feel a thermal stimulus. The presentation time depended on the stimulus so that when the presentation method was pre-adjusted the stimulus was always presented for 10 s. However, when the presentation method was dynamic, the temperature change was roughly 0.5°C/s so that the presentation time was either 4 s (2°C change), 8 s (4°C change) or 12 s (6°C change). The neutral stimulus was presented in both stimulus groups for 10 s. After the stimulus presentation the participant was instructed to rate the stimulus with four nine-point scales for pleasantness (varying from "the stimulus felt unpleasant" to "the stimulus felt pleasant"), approachability (varying from "the stimulus felt avoidable" to "the stimulus felt approachable"), arousal (varying from "I felt calm during stimulus presentation" to "I felt aroused during stimulus presentation"), and dominance (varying from "I was in control during stimulus presentation" to "the stimulus was in control during presentation"). The mid-point of each scale represented a neutral experience (i.e., the stimulus felt neither unpleasant nor pleasant). The participant was instructed to rate the immediate impression of the stimulus. The next trial was initialized after the participant had rated one stimulus with all the four scales. The second block proceeded similarly to the first block except that instead of emotional ratings the participants rated their experience of the stimulus temperature with a nine-point scale varying from cold to hot. The mid-point of the scale represented a neutral experience (i.e. the stimulus felt neither cold nor hot). The order of both the stimulus groups and rating scales was fully counterbalanced. In addition, the stimulus presentation was fully randomized within a group. The total amount of stimulus presentation trials was 28. Conducting the experiment took approximately 45 minutes.

2.5 Data Analysis

The subjective rating data was first analyzed with Friedman tests in order to test whether varying the temperature of dynamic stimuli affected the ratings. Then, Friedman test was used to test whether varying the temperature of pre-adjusted stimuli affected the ratings. If statistically significant differences were found, Wilcoxon signed-ranks test were used for pairwise comparisons. Wilcoxon signed-ranks test was also used to test whether the presentation method affected the ratings (e.g. dynamic and pre-adjusted 6°C increases were compared). The rise time (i.e. the time from the beginning of the response until the highest peak) and the magnitude (i.e. the highest peak) of the SCR were analyzed with the within-subject repeated measures analysis of variance (ANOVA). The SCR reactions were calculated from the data collected in the first experimental block during 5000 ms time frame from the stimulus onset (i.e. 1000 to 6000 ms). Three participants were excluded from the SCR analysis due to technical problems. Bonferroni corrected pairwise comparisons were used as post hoc tests.

3 Results

Friedman tests showed that the temperature affected the ratings of pleasantness when the stimuli were dynamic $\chi = 28.1$, $p < 0.001$. However, when the pre-adjusted stimuli were compared no statistically significant differences were found $\chi = 11.5$, $p > 0.05$. Tables 1-3 show the results of the Wilcoxon signed-ranks tests when the presentation method was dynamic. Because the pre-adjusted method showed no statistically significant findings, the results are not included in the table for the ratings of pleasantness. Wilcoxon signed-ranks test showed also that the presentation method affected the pleasantness ratings. When the temperature was -6°C, pre-adjusted stimulus was rated as more pleasant than dynamic stimulus $T = 2.8$, $p < 0.01$. The pre-adjusted stimulus was also rated as more pleasant than dynamic stimulus when the temperature change was +4°C, $T = 2.3$, $p < 0.05$, and +6°C $T = 2.4$, $p < 0.05$.

Friedman tests showed that the temperature affected the ratings of approachability when the presentation method was dynamic $\chi = 32.4$, $p < 0.001$ and when it was pre-adjusted $\chi = 17.6$, $p < 0.01$. Tables 1-3 show the results of the Wilcoxon signed-ranks tests. Wilcoxon signed-ranks test showed also that the presentation method affected the approachability ratings but only when the temperature was increased. Pre-adjusted stimuli were rated as more approachable than dynamic stimuli when the temperature change was +2°C $T = 2.5$, $p < 0.05$, +4°C $T = 2.4$, $p < 0.05$, and +6°C $T = 2.4$, $p < 0.05$.

Friedman tests showed that the temperature affected the ratings of arousal when the presentation method was dynamic $\chi = 64.7$, $p < 0.001$ and when it was pre-adjusted $\chi = 50.5$, $p < 0.001$. Tables 1-3 show the results of the Wilcoxon signed-ranks tests. Wilcoxon signed-ranks test did not show any statistical significant differences between temperatures for the ratings of arousal when the presentation method was varied.

Friedman tests showed that the temperature affected the ratings of dominance when the presentation method was dynamic $\chi = 78.9$, $\mathrm{p} < 0.001$ and when it was pre-adjusted $\chi = 70.9$, $\mathrm{p} < 0.001$. Tables 1-3 show the results of the Wilcoxon signed-ranks tests. Wilcoxon signed-ranks test showed also that the presentation method affected the dominance ratings. Dynamic stimuli were rated as more dominant than pre-adjusted stimuli when the temperature change was -6°C $T = 2.1$, $p < 0.05$, -2°C $T = 2.1$, $p < 0.05$, and +6°C $T = 2.3$, $p < 0.05$.

Friedman tests showed that the temperature affected the ratings of temperature when the presentation method was dynamic $\chi = 119.6$, $p < 0.001$ and it was pre-adjusted $\chi = 127.1$, $p < 0.001$. Tables 1-3 show the results of the Wilcoxon signed-ranks tests. Wilcoxon signed-ranks test showed also that the presentation method affected the temperature ratings. Pre-adjusted stimuli were rated as hotter than dynamic stimuli when the temperature change was -6°C $T = 2.1$, $p < 0.05$, or -4°C $T = 2.5$, $p < 0.05$. However, when the temperature change was +6°C, dynamic stimulus was rated as hotter than pre-adjusted stimulus $T = 3.8$, $p < 0.001$.

Tables 1-3. Pairwise comparisons between the subjective ratings. The temperatures being compared are on the top of the table, and the rating scale on the left of the table. Each cell is divided in two parts. On the white area of the cell there are pairwise comparisons for the stimuli with dynamic presentation method and on the blue area for the stimuli with pre-adjusted presentation method. In each cell there is Wilcoxon's T value, the temperature change rated as, for example, more pleasant or hotter and indication of P value (p < 0.05*, p < 0.01**, and p < 0.001***) for both presentation methods. Ns stands for non-significant.

Table 1.

	-6 vs -4	**-6 vs -2**	**-6 vs 0**	**-6 vs +2**	**-6 vs +4**	**-6 vs +6**	**-4 vs -2**
pleasantness dyn	*T* = 2.0 [-4] *	*ns*	*ns*	*ns*	*ns*	*T* = 2.2 [-6] *	*ns*
approachal dyn	*T* = 2.5 [-4] *	*ns*	*ns*	*ns*	*ns*	*T* = 2.3 [-6] *	*ns*
approachal pre	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*	*ns*
arousal dyn	*T* = 2.4 [-6] *	*T* = 2.3 [-6] *	*T* = 2.6 [-6] **	*ns*	*ns*	*T* = 3.4 [+6] ***	*ns*
arousal pre	*ns*	*ns*	*ns*	*ns*	*T* = 2.5 [+4] *	*T* = 2.9 [+6] **	*ns*
dominance dyn	*T* = 3.2 [-6] **	*T* = 3.3 [-6] ***	*T* = 3.5 [-6] ***	*T* = 2.1 [-6] *	*ns*	*T* = 2.9 [+6] **	*ns*
dominance pre	*T* = 2.5 [-6] *	*T* = 3.3 [-6] ***	*ns*	*ns*	*T* = 2.7 [+4] **	*T* = 3.5 [+6] **	*ns*
temperature dyn	*ns*	*T* = 3.7 [-2] ***	*T* = 4.0 [0] ***	*T* = 4.2 [+2] ***	*T* = 4.2 [+4] ***	*T* = 4.3 [+6] **	*T* = 4.1 [-2] ***
temperature pre	*T* = 2.3 [-4] *	*T* = 3.5 [-2] ***	*T* = 3.9 [0] ***	*T* = 4.3 [+2] ***	*T* = 4.2 [+4] ***	*T* = 4.3 [+6] **	*T* = 3.6 [-2] ***

Table 2.

	-4 vs 0	**-4 vs +2**	**-4 vs +4**	**-4 vs +6**	**-2 vs 0**	**-2 vs +2**	**-2 vs +4**
pleasantness dyn	*ns*	*ns*	*ns*	*T* = 3.1 [-4] **	*ns*	*ns*	*ns*
approachal dyn	*ns*	*ns*	*ns*	*T* = 3.2 [-4] ***	*ns*	*ns*	*ns*
approachal pre	*ns*	*ns*	*ns*	*T* = 2.6 [-4] **	*ns*	*ns*	*ns*
arousal dyn	*ns*	*ns*	*T* = 2.8 [+4] **	*T* = 4.1 [+6] ***	*ns*	*T* = 2.3 [+2] *	*T* = 3.3 [+4] ***
arousal pre	*ns*	*ns*	*T* = 3.4 [+4] **	*T* = 3.6 [+6] ***	*ns*	*ns*	*T* = 3.9 [+4] ***
dominance dyn	*ns*	*ns*	*T* = 3.1 [+4] **	*T* = 4.1 [+6] ***	*ns*	*ns*	*T* = 3.5 [+4] **
dominance pre	*ns*	*ns*	*T* = 3.8 [+4] ***	*T* = 4.1 [+6] ***	*ns*	*T* = 2.4 [+2] *	*T* = 3.9 [+4] ***
temperature dyn	*T* = 4.2 [0] ***	*T* = 4.3 [+2] ***	*T* = 3.7 [+4] ***	*T* = 4.3 [+6] ***	*ns*	*T* = 3.9 [+2] ***	*T* = 3.6 [+4] ***
temperature pre	*T* = 3.7 [0] ***	*T* = 4.3 [+2] ***	*T* = 4.3 [+4] ***	*T* = 4.3 [+6] ***	*T* = 2.8 [0] *	*T* = 4.1 [+2] ***	*T* = 4.4 [+4] ***

Table 3.

	-2 vs +6	**0 vs +2**	**0 vs +4**	**0 vs +6**	**+2 vs +4**	**+2 vs +6**	**+4 vs +6**
pleasantness dyn	*T* = 2.9 [-2] **	*ns*	*T* = 2.0 [0] *	*T* = 3.9 [0] ***	*T* = 2.0 [+2] *	*T* = 3.6 [+2] ***	*T* = 3.5 [+4] ***
approachal dyn	*T* = 2.7 [-2] **	*ns*	*T* = 2.5 [0] *	*T* = 4.0 [0] ***	*ns*	*T* = 3.2 [+2] ***	*T* = 3.5 [+4] ***
approachal pre	*T* = 2.3 [-2] *	*ns*	*ns*	*T* = 3.3 [0] ***	*ns*	*T* = 3.1 [+2] **	*T* = 3.0 [+4] **
arousal dyn	*T* = 3.7 [+6] ***	*T* = 2.0 [+2] *	*T* = 3.5 [+4] ***	*T* = 4.0 [+6] ***	*T* = 2.7 [+4] **	*T* = 3.9 [+6] ***	*T* = 3.9 [+6] ***
arousal pre	*T* = 3.7 [+6] ***	*ns*	*T* = 3.7 [+4] ***	*T* = 3.9 [+6] ***	*T* = 2.6 [+4] **	*T* = 3.1 [+6] **	*ns*
dominance dyn	*T* = 4.1 [+6] ***	*T* = 2.0 [+2] *	*T* = 4.0 [+4] ***	*T* = 4.3 [+6] ***	*T* = 3.4 [+4] ***	*T* = 4.3 [+6] ***	*T* = 3.6 [+6] ***
dominance pre	*T* = 4.0 [+6] ***	*ns*	*T* = 3.7 [+4] ***	*T* = 4.0 [+6] ***	*T* = 3.0 [+4] **	*T* = 3.7 [+6] ***	*ns*
temperature dyn	*T* = 4.3 [+6] ***	*T* = 3.6 [+2] ***	*T* = 3.6 [+4] ***	*T* = 4.4 [+6] ***	*T* = 2.9 [+4] **	*T* = 4.3 [+6] ***	*T* = 4.2 [+6] ***
temperature pre	*T* = 4.3 [+6] ***	*T* = 3.5 [+2] ***	*T* = 4.2 [+4] ***	*T* = 4.3 [+6] ***	*T* = 3.4 [+4] **	*T* = 4.4 [+6] ***	*T* = 3.6 [+6] ***

For the rise time of the SCR a 2 × 7 (presentation method × temperature change) ANOVA showed no statistically significant effects of the stimuli. For the magnitude of the SCR a 2 × 7 (presentation method × target temperature) ANOVA showed a statistically significant interaction of the main effects of presentation method and target temperature $F(1,6) = 6.21$, $p < 0.001$. To analyze the interaction of the main effects, two one-way ANOVA's were performed to test whether the used target temperatures elevated the SCR differently when the presentation method was dynamic than when it was pre-adjusted. One-way ANOVA's showed that when the presentation method was dynamic, there were no statistically significant differences in the magnitude of the SCR between different target temperatures $F(1,6) = 2.2$, $p > 0.05$. However, when the presentation method was pre-adjusted there were statistically significant differences in the magnitude of the SCR between different target temperatures $F(1,6) = 10.3$, $p < 0.001$. Bonferroni corrected pairwise comparisons showed that the statistically significant result was due to the fact that pre-adjusted 6°C increase in temperature elevated the magnitude of SCR more than 6°C decrease $MD = 0.29$, $p < 0.01$, 4°C decrease $MD = 0.29$, $p < 0.01$, 2°C decrease $MD = 0.35$, $p < 0.01$, 0°C $MD = 0.23$, $p < 0.01$, or 2°C increase $MD = 0.22$, $p < 0.01$.

4 Discussion

The participants rated 6°C increase as less pleasant than any other stimuli but only when the presentation method was dynamic. A 6°C increase in both dynamic and pre-adjusted method was rated as less approachable than any of the other stimuli. Warm stimuli were in general rated as arousing and dominant so that the higher the intensity, the more arousing and the more dominant rating (e.g. 6°C increase was rated as more arousing and dominant than 2°C increase or decrease). High intensity cold stimulus (i.e. 6°C decrease) was rated as more arousing and dominant than the other cold stimuli when the presentation method was dynamic. It should be noted that even the high intensity cold stimuli were always rated as less arousing and less dominant than any of the warm stimuli. In addition, the dynamic stimuli were rated as less pleasant and approachable but more dominant than pre-adjusted stimuli especially when stimulus temperature was warm.

The participants also rated the stimuli adequately in respect to the experienced temperature despite of the presentation method. This result suggesting that 2°C changes are sufficient for creating distinguishable thermal stimuli for haptic user interfaces is in line with earlier findings [6, 11]. However, also the presentation method affected the ratings so that, for example, 6°C increase was rated as hotter when the presentation method was dynamic than when it was pre-adjusted. This result may be due to the reason that when the method is pre-adjusted, the participant is able to feel the stimulus temperature immediately after touching the actuator. But when the stimulus presentation is dynamic the participant may become cautious about the final limit of heating.

SCR measurements showed that only pre-adjusted 6°C increase significantly elevated the SCR. It seems possible that when the temperature changes during touching the actuator the autonomic nervous system (ANS) adapts to the change in the temperature but when the shift in stimulus temperature is abrupt (i.e., presentation method is pre-adjusted) there is no time to adapt and therefore a stronger ANS response is triggered.

In general, the use of wider set of temperature changes than in a pilot study [11] clearly gave new fine grained insight. Unlike in the pilot study, now the ratings of pleasantness and approachability were affected by temperature changes. 6°C increase was rated as unpleasant and avoidable while 2°C increase was rated as rather pleasant and approachable. 4°C shift in temperature in respect to the hand temperature was rated as rather neutral. So, it seems that in respect to experiences of pleasantness and approachability, smaller (e.g. 2°C) or larger (e.g. 6°C) shifts in temperature (in contrast to the 4°C shifts used earlier) are more effective.

In a previous study [6] there was a tendency to rate cold stimuli as more comfortable than warm stimuli. Also the stimulus intensity affected the experience of thermal comfort so that the warm stimuli with higher intensity were in general rated as less comfortable than the stimuli with lower intensity. Our current results seem to contradict this finding at some degree. As in the previous study warm stimuli with high intensity were rated as rather unpleasant. However, this result was evident only with dynamic presentation method. Therefore, it seems that the experience of comfort or pleasantness is not tied to the stimulus intensity or temperature alone, but the way the temperature is presented to the user. This may indicate that instead of temperature as such, the change in temperature is the factor making the experience of thermal stimulation as pleasant or unpleasant.

The dimensional theory of emotions [8] suggests that the elevated level of arousal represents the activation of the motivational tendency related to approaching or avoiding the stimulation. The pre-adjusted warm stimuli were rated as arousing yet pleasant. Then, dynamic stimuli were in general rated as more dominant but less pleasant and approachable than pre-adjusted stimuli. This suggests that the pre-adjusted stimuli can be used to make the stimuli arousing and approachable while dynamic presentation method makes the stimuli avoidable.

In a practical scenario a thermal actuator would be attached to the back of a mobile phone to provide stimuli to the user's palm. A pre-adjusted stimulus is provided if the mobile phone is, for instance, on the table and the user touches the phone after stimulus has reached the target temperature. A dynamic stimulus is provided, for example, when the user is holding the mobile phone during a phone call. Intuitively, stimulus rated as arousing and dominant could be used to efficiently catch the users' attention. In addition, the stimulus should be pleasant. A 4°C increase is suitable for this purpose with both presentation methods.

The current results could be tested with interfaces utilizing auditory feedback. At this point, the results can only be speculated. There is some evidence that vibrotactile feedback in conjunction with speech works mostly as emphasizing emotional content of the speech (e.g. making it more arousing) [13]. Based on the current results seems likely that thermal stimulation could be used to communicate information related both to pleasantness and arousal in conjunction with speech. This assumption could be supported by the fact that thermal stimulation is processed differently than vibrotactile stimulation in skin and at the central nervous system level.

In summary, the results suggest that warm stimuli work better than cold and neutral stimuli in activating human emotional system when measured with both subjective rating scales and physiological responses. If one wants to elevate the level of arousal and dominance, at least 4°C increase in temperature is needed. A pre-adjusted presentation method is more suitable for this purpose than dynamic presentation method as in general pre-adjusted stimuli were rated as more pleasant and approachable. The stimuli with pre-adjusted presentation method also affected the physiological responses more efficiently than the stimuli with dynamic presentation method suggesting that ANS is affected more when the stimulus is pre-adjusted than when it is dynamic. These results can be seen as a step forward in the knowledge needed in creating devices that are capable of using thermal feedback as a haptic interaction method.

Acknowledgments. This research was a part of HEI project funded by the Finnish Funding Agency for Technology and Innovation (Tekes). In addition, the research was supported by Tampere Doctoral Programme in Information Science and Engineering (TISE), and the HapticAuto project (Tekes decision number 40289/11).

References

1. Harlow, H.F.: The nature of love. American Psychologist 13(2), 673–685 (1958)
2. Williams, L.E., Bargh, J.A.: Experiencing physical warmth promotes interpersonal warmth. Science 322, 606–607 (2008)
3. Kanosue, K., Sadato, N., Okada, T., Yoada, T., Nakai, S., Yoshida, K., Hosono, T., Nagashima, K., Yagishita, T., Inoue, O., Kobyashi, K., Yonekua, Y.: Brain activation during whole body cooling in humans studied with functional magnetic resonance imaging. Neuroscience Letters 329, 157–160 (2002)
4. Sung, E.-J., Yoo, S.-S., Yoon, H.W., Han, S., Park, H.W.: Brain activation related to affective dimension during thermal stimulation in humans: A functional magnetic resonance imaging study. International Journal of Neuroscience 117, 1011–1027 (2007)
5. Jones, L.A., Ho, H.: Warm or cool, large or small? The challenge of thermal displays. IEEE Transactions on Haptics 1, 53–70 (2008)
6. Wilson, G., Halvey, M., Brewster, S.A., Hughes, S.A.: Some like it hot: thermal feedback for mobile devices. In: CHI 2011, pp. 2555–2564 (2011)
7. Salminen, K., Surakka, V., Lylykangas, J., Raisamo, J., Saarinen, R., Rantala, J., Raisamo, R., Evreinov, G.: Emotional and behavioral responses to haptic stimulation. In: CHI 2008, pp. 1555–1562 (2008)
8. Bradley, M.M., Lang, P.J.: Measuring emotion: the self-assessment manikin and the semantic differential. Journal of Behavior Therapy and Experimental Psychiatry 25, 49–59 (1994)
9. Davidson, R.J.: Cerebral asymmetry, emotion and affective style. In: Davidson, R.J., Hugdahl, K. (eds.) Brain Asymmetry, pp. 361–387. The MIT Press, Cambridge (1995)
10. Witvliet, C.V.O., Vrana, S.R.: Psychophysiological responses as indices of affective dimensions. Psychophysiology 32, 436–443 (1995)
11. Salminen, K., Surakka, V., Raisamo, J., Lylykangas, J., Pystynen, J., Raisamo, R., Mäkelä, K., Ahmaniemi, T.: Emotional responses to thermal stimuli. In: ICMI 2011, pp. 193–196 (2011)
12. Schneider, W., Eschman, A., Zuccolotto, A.: E-Prime User's Guide. Psychology Software Tools, Inc., Pittsburgh (2002)
13. Salminen, K., Surakka, V., Lylykangas, J., Rantala, J., Ahmaniemi, T., Raisamo, R., Trendafilov, D., Kildal, J.: Tactile Modulation of Emotional Speech Samples. In: Adv. Human-Computer Interaction 2012 (2012)

Potential Tissue Puncture Notification during Telesurgery

Rachael L'Orsa[1], Kourosh Zareinia[2], Liu Shi Gan[2,3],
Chris Macnab[1], and Garnette Sutherland[2]

[1] Department of Electrical and Computer Engineering,
[2] Project neuroArm, Faculty of Medicine,
University of Calgary, Canada
{ralorsa,cmacnab,kzareini,garnette}@ucalgary.ca
[3] IMRIS Inc., Winnipeg MB, Canada
lgan@imris.com

Abstract. This paper proposes the use of vibrotactile feedback during telesurgery to notify surgeons of potential tissue puncture. Puncture trials using an experimental telesurgical apparatus were performed on an artificial membrane to characterize general force ranges at which punctures occur. The average force threshold during puncture was established, and human operators then attempted to apply a maximum force to the membrane without causing a puncture via the telesurgical apparatus. As the surgical tool-tip approached the pre-established force threshold, a wrist-mounted haptuator worn by the operators vibrated a warning. Warnings via different sensory modalities (auditory and tactile) were compared both with and without force feedback. Results show that the use of a warning via either sensory modality decreases the maximum force applied by the operator, thereby decreasing the occurrence of unintentional punctures. The inclusion of force feedback achieved similar results, though task completion times were significantly increased.

Keywords: haptics, haptuator, keyhole surgery, teleoperation, telesurgery, tissue puncture.

1 Introduction

The 'minimally invasive' or 'keyhole' paradigm for surgery or microneurosurgery offers many advantages to patients over traditional surgical approaches, such as decreased postoperative hospitalization times, improved cosmesis, and reduced surgical site infections [1], [2]. However, it decouples surgeons from their sense of touch, reduces their dexterity within the patient's body, and has a steep learning curve [3]. The use of telesurgical robots can alleviate many of these drawbacks [4], [5], [6], [7], but currently no commercially-available systems restore the sense of touch to surgeons, i.e. incorporate haptics. This lack of sensation can lead to difficulties ascertaining tissue condition and/or properties [8], which can in turn lead to the unintentional perforation of blood vessels and other critical

I. Oakley and S. Brewster (Eds.): HAID 2013, LNCS 7989, pp. 30–39, 2013.

structures. These unintentional punctures constitute one of the major sources of complications for minimally invasive surgical procedures, particularly during initial entry into the patient's body [9]. A telesurgical system that could assist surgeons in identifying tissue condition and properties using alternative sensory channels and also provide notifications of high applied forces could potentially ease surgeon workload during surgery while reducing surgical complications.

This paper seeks to introduce a new method for avoiding unintentional tissue puncture by providing surgeons with additional sensory (particularly haptic) feedback during telesurgical procedures. To our knowledge, this is the first application of haptic feedback for force threshold notifications in the surgical field. Some research has been done comparing the use of visual, auditory, and tactile feedback for notifying drivers of impending collision events, where it was found that haptic feedback was less invasive [10] and more alerting [11] than the other two modes alone. Some studies have found that multisensory notifications rendered synchronously can have an additive effect, such as when auditory notifications paired with their tactile counterparts appeared to be louder [12].

While these studies report mixed findings for the efficacy of single channel versus multisensory feedback for vehicle operators, there are no comparable results in the literature for telesurgery and no indications that findings from these studies would generalize to activities such as tele-microneurosurgery. Most experimental haptics research to date for telesurgery has focused on the accurate reproduction of force feedback for surgeons (and to a much lesser extent, vibrotactile feedback [13]), haptic displays [14] and sensors [15] to allow for tissue property determination via telepalpation or surgical simulation [16], and haptic fixtures [17] to provide fixed position-based safety controls during surgery. While haptic fixtures are the most commonly applied haptic safety control for telesurgical systems, such so-called forbidden regions generally neither allow for modulation of applied forces nor permit surgeons the autonomy to decide to proceed if it is safe or necessary to do so, with very few exceptions [18]. Thus this study represents a first foray into vibrotactile notifications to reduce unintended tissue puncture during telesurgery.

Furthermore, the importance of haptic (particularly force) feedback to surgical tasks remains a controversial topic, with very little data available in the literature to either support or contradict its inclusion in telesurgical systems. We seek to add to this body of knowledge by investigating the impact of force feedback on our puncture notification trials.

The remainder of this paper is divided as follows: Section 2 introduces the experimental design and apparatus, Section 3 describes the results, and Section 4 provides conclusions and future directions.

2 Experimental Design

Experiments utilize a custom prototype 7 degree of freedom (DOF) telesurgical system, though it remains in Z-lock with movement restricted solely to the Z-axis. A square metal tube (1.5" x 1.5" x 3") forms a simulated membrane enclosing an air-filled cavity, achieved by mounting it on a wooden base for stability and fitting it with an adhesive strip of translucent plastic. This flat and taut artificial membrane, while dissimilar from real tissue, allows for a relatively static contact geometry during tool-tip interactions. An automated, increasing orthogonal force repeatedly disturbs the artificial membrane until the system's end-effector tool-tip punctures through it. These trials establish an average puncture force, which in turn establishes a range of threshold forces at which warnings should occur during teleoperation.

2.1 Apparatus

Figure 1 depicts a basic schematic diagram of the system setup. A SensAble Phantom Premium 3.0L/6DOF haptic interface is used as he master, controlling a Kuka KR-6 slave robot for all experiments. A custom 7th DOF in the form of an actuated fine gripper microsurgical tool coupled with an ATI Gamma 6DOF force/torque sensor attaches to the robot's end-effector. The tool allows for precision gripping and rolling, and while neither function is utilized for these experiments, the tool's actuators remain powered throughout to maintain its grip and roll at constant values. The Premium 3.0 reproduces forces from the Gamma with some scaling (a force gain of 0.8) but without smoothing when its amplifiers are activated.

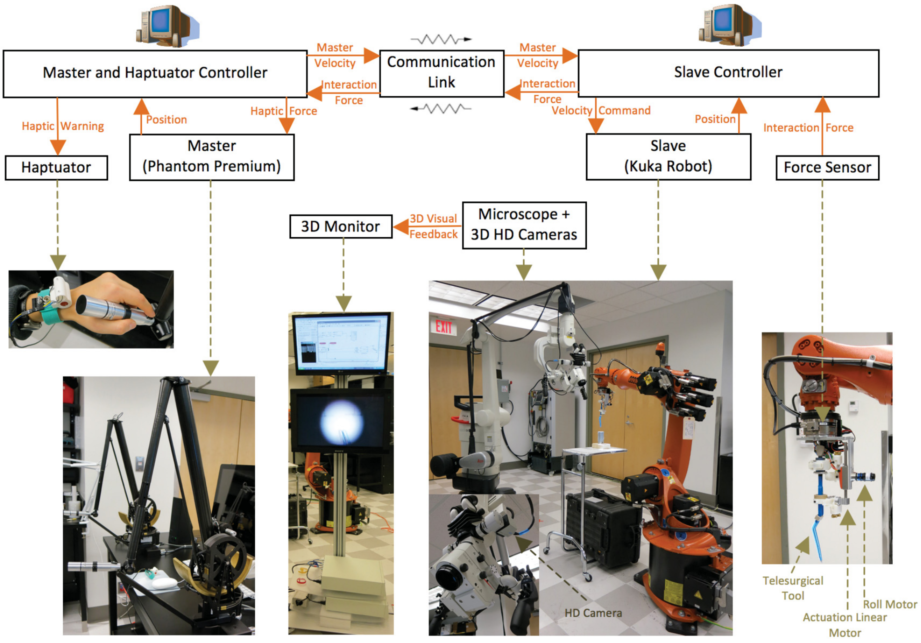


Fig. 1. System diagram

A Dell Dimension 4700 with a 3.0GHz Pentium 4 processor running 32-bit Windows XP Professional SP3 with 4GB of RAM processes information for the master system, connecting the Premium 3.0, an external speaker, and a haptuator. The haptuator provides high-bandwidth, iron-free, recoil-based electromagnetic vibrotactile actuation at different frequencies and amplitudes [19]. Microsoft Visual C++ 2010 interfaces between the haptuator, the external speaker, and the master PC via the UDP protocol.

The slave system uses a custom PC with a 3.3GHz Intel CORE i5 processor running 32-bit Windows 7 Professional SP1 with 4GB of RAM. The slave PC connects to the Kuka workstation and a Quanser Q2-USB data acquisition board (DAQ). The Quanser DAQ interfaces with the precision grip/roll motor, the force sensor and its accompanying hardware (a National Instruments DAQ and signal conditioning box provided by ATI), and a linear Faulhaber minimotor with accompanying controller for actuating the custom microsurgical tool. Matlab/Simulink R2011a with Quanser QUARC 2.2 blocks provides a real-time interface between hardware and the master and slave PCs respectively. Master and slave PCs communicate over a LAN via the TCP/IP protocol.

Also, a Leica M525 OH4 surgical microscope coupled with two Ikegami HDL 20D microscope camera systems, a Sony LMD2451MD LCD monitor, and RealD 3D glasses provides a magnified real-time 3D video feed of the tool-tip's interaction with the artificial membrane.

2.2 Characterization of Average Puncture Force

The slave robot starts at a home position where the angles for joints 1 through 6 are [0 -90 76 0 103 0] degrees respectively. As previously mentioned, the linear actuator for the custom microsurgical tool maintains a 'closed' position. The artificial membrane sits on a free-standing tray centered approximately one inch below the slave robot end-effector's tool-tip. A Matlab/Simulink program controls the end-effector's tool-tip as it advances in the negative direction along the Z-axis (i.e. down, into the membrane) with a constant velocity of 0.01 m/s, until stopped by an experimenter when membrane puncture is visually and audibly confirmed. The slave robot then reinitializes to its home position, and a new artificial membrane is fitted for the next experiment.

Twenty such individual trials are performed using the automated program to help maintain experimental conditions between trials. The maximum force for each trial defines the puncture force. As the variance in puncture force is low, the average puncture force from all twenty trials defines the force threshold (hereafter referred to as the 'puncture-force threshold') for the real-time potential puncture notification trials that follow.

2.3 Real-Time Potential Puncture Notification

Defining a force margin allows operators time to react to a potential puncture notification before they apply forces that exceed the predefined puncture-force threshold. All notification trials use a roughly 50% force margin, though future experiments should investigate the impact of this factor on operator response. Depending on the trial, notifications are generated by either the haptuator (for a vibrotactile warning), the external speaker (for an auditory warning), or both. As identified through preliminary trials, the two major sources of variance between experiments are the operator and the type of feedback (or lack thereof) applied. Each of the eight sets of trials has its own code, as shown in Table 1, where a given set either serves as a control (visual feedback only, no warnings) or defines the type of feedback applied.

Table 1. Set codes

FF	V	V + H	V + A	V + H + A
no	S1	S2	S3	S4
yes	F1	F2	F3	F4

The top four sets (S1 through S4) are visual feedback only (V), visual and haptic (vibrotactile) feedback together (V + H), visual and auditory feedback together (V + A), and all three types of feedback together (V + H + A). The bottom four sets (F1 through F4) are the same, but include force feedback (FF).

For the first sets (S1 and F1), a real-time 3D video feed provides visual feedback of the tool-tip's interaction with the artificial membrane, and operators must gauge their applied force based on the minimal (∼2mm) membrane deformation alone. The second sets (S2 and F2) provide vibrotactile haptic feedback via the wrist-mounted haptuator. The haptuator is mounted on the operator's non-dominant wrist, i.e. not on the arm with which they operate the haptic interface. The characteristics of the haptic and auditory feedback (a 60 dB, 0.02 second 'beep' repeated at a frequency of 50 Hz) is kept constant throughout all trials for all operators, as is the position of the external speaker.

Five successful trials are performed by each operator for each of the eight sets of experiments, where a successful trial is one in which no puncture occurs. Thus any given data set may contain more than five trials if punctures do occur, but the data for each operator will contain exactly five trials for which the operator caused no punctures. For each consecutive trial, a pseudo-random integer generator in Matlab chooses which set to apply. Operators are unaware of this choice and cannot anticipate either notifications or force feedback. The forty successful trials constitute one block, which is then repeated by different operators. Two operators, one with no previous teleoperation experience and one with a multi-case history of neuroArm [20] operation, perform the full experiment. Both are neurosurgeons, but neither is trained on the experimental apparatus.

The operators familiarize themselves with the system via 3-5 initial trials. They then apply as much force as possible to the artificial membrane using the telesurgical setup, without causing a puncture and without decreasing the applied force, until they are satisfied that they have achieved a maximum.

3 Results

Since the variance in puncture force during characterization trials was 0.14 N, the 13.4 N average for all twenty force characterization trials became the puncture-force threshold for the real-time potential puncture notification trials. Given the roughly 50% force margin (rounded down to the nearest integer), this means that notifications occurred when operators applied a force in excess of 6.0 N to the membrane.

Figure 2 compares the force data for all five successful trials of one operator's S4 and F4 sets. Force feedback trials show lower forces yet higher task completion times. The force fluctuations in the F4 sets is typical of operators unaccustomed to force feedback, and highlights the need for further investigation into ideal amounts of training.

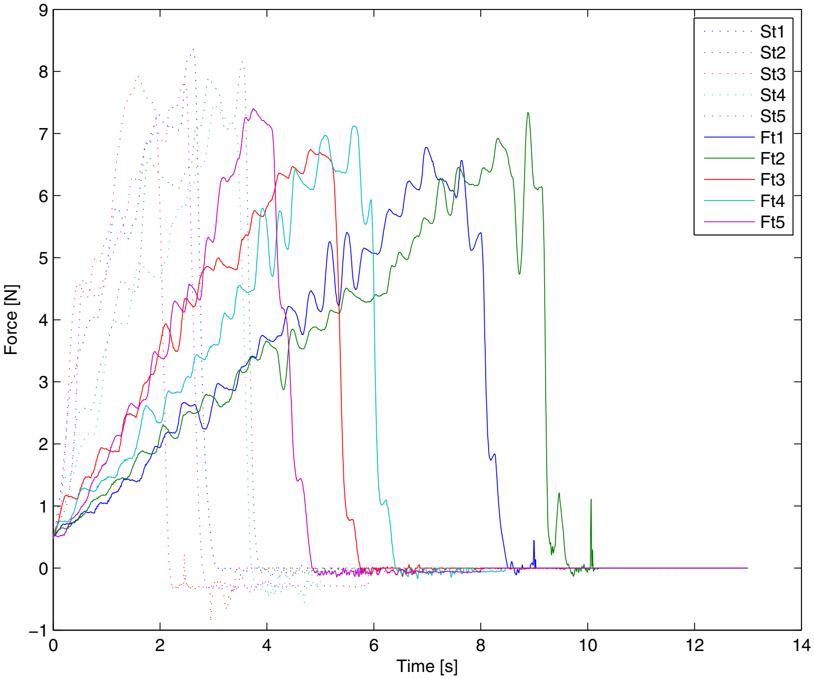


Fig. 2. Representative comparison of data with and without force feedback

The maximum force, the time to achieve maximum force, and a Boolean denoting puncture occurrence are stored as metrics for each trial. Within each of the eight sets, maximums, minimums, and means are calculated that reveal several general insights:

1. On average, the use of notifications decreased the amount of applied force and the task completion time regardless of sensory modality.
2. There were no discernible trends between modalities.
3. On average, the use of force feedback decreased the amount of applied force yet increased the task completion time regardless of sensory modality.

Multiple ANOVAs performed in Excel on the experimental data using a value of $\alpha = 0.05$ support these findings. Table 2 shows the number of punctures caused by each operator over all sets of trials. By inspection, the number of punctures differs between operators, though neither caused any punctures when audio or the combination of haptic and audio notifications were applied.

Table 2. Puncture results

Operator	$S1$	$S2$	$S3$	$S4$	$F1$	$F2$	$F3$	$F4$
1	4	0	0	0	2	0	0	0
2	2	1	0	0	0	0	0	0

Table 3 shows a summary of the total punctures caused by both operators. In general, fewer punctures occurred when any type of notification was applied. Furthermore, the inclusion of force feedback decreased the number of punctures that occurred both with and without notifications.

Table 3. Summary of warning impact on total punctures

Set	No Warnings	Warnings
S	6	1
F	2	0
Both	8	1

Applied forces and task completion times for all successful trials for both operators were compared with respect to the use of force feedback. Table 4 shows that on average, 8.16% less force was applied by the operators when force feedback was used, though it took them 95.62% longer to complete the task. These effects were confirmed as being significant via ANOVAs, by an $F = 4.71 > F_{\text{critical}} = 3.97$ with $p = 0.03$ for the applied forces and an $F = 29.17 >> F_{\text{critical}} = 3.97$ with $p < 0.001$ for the task completion times. No significant variation was attributed to the differences between operators. That the use of force feedback produces fewer punctures is likely due to this decrease in applied force.

Table 4. Comparison of results with and without force feedback

FF	Ave. Force [N]	Ave. Time [s]
no	8.91	2.82
yes	8.18	5.52
% change	-8.16	95.62

Applied forces and task completion times for all successful trials for both operators were also compared with respect to the use of notifications. Table 5 highlights the percent increase or decrease in the average applied force and task completion times when no notification was used (set codes S1 and F1), versus when only haptic notifications (S2 and F2), only audio notifications (S3 and F3), both haptic and audio notifications (S4 and F4), or any kind of notifications (S2 through S4 and F2 through F4) were used. In general, decreases of 22.57% in average applied force ($F = 97.11 >> F_{\text{critical}} = 3.92$ with $p < 0.001$) and 34.61% in average task completion time ($F = 17.55 >> F_{\text{critical}} = 3.92$ with $p < 0.001$) were achieved with the use of any type of notification. Again, no significant variation was attributed to the differences between operators.

Table 5. Comparison of results with and without notifications

Type of warning	Ave. Force [N]	% Change	Ave. Time [s]	% Change
none	10.29	N/A	5.64	N/A
haptic	8.25	-19.78	3.16	-43.98
audio	7.72	-24.91	3.38	-39.97
both	7.92	-23.03	4.52	-19.88
any	7.96	-22.57	3.69	-34.61

Though it appears from Table 5 that purely audio feedback and purely haptic feedback most successfully decreased the applied force and task completion times respectively, the ANOVAs indicated no significant variation could be attributed to the effect of the type of notification on either metric. However, the differing reactions of the two operators to the various types of notifications were significant for applied force ($F = 5.49 > F_{\text{critical}} = 4.02$ with $p = 0.02$).

4 Conclusion

Based on this preliminary investigation, it appears that giving operators a scale (i.e. warnings) with which to gauge their proximity to a given force threshold allows them to moderate forces applied during telesurgery and thus to avoid unintentional tissue punctures. Of the sensory channels tested for this paper, it was not possible to identify whether auditory, haptic, or a combination of the two improved operator performance the most. However, it was found that both the inclusion of force feedback and the provision of notifications through either modality decreased applied forces, though at the cost of increased task completion times for force feedback. Follow-up experiments should be performed in a clinical setting and expanded to a group of operators with a larger sample size. A characterization of operator preferences regarding feedback modalities for notifications during surgery should be performed, and the impact of varied force margins for notifications should be explored. An investigation of reaction times based on notification modality would be appropriate, and a comparison of operator performance between fully random trials (as reported in this paper) and set-random trials (where the sets are performed in random order, but all five trials within a set are performed consecutively) could give insight into learning curves for the operation of such systems. Furthermore, a comparison of operator performance with and without feedback anticipation (when the operators are expecting to receive a notification and thus rely less on visual information) could help determine a more clear relationship between sensory feedback channels and surgical performance.

References

1. Varela, J.E., Wilson, S.E., Nguyen, N.T.: Laparoscopic surgery significantly reduces surgical-site infections compared with open surgery. Surg. End. 24(2), 270–276 (2010)
2. Reisch, R., Stadie, A., Kockro, R.A., Hopf, N.: The Keyhole Concept in Neurosurgery. World Neurosurgery (2012), `http://www.sciencedirect.com/science/article/pii/S187887501200157X`
3. Meireles, O., Horgan, S.: Applications of surgical robotics in general surgery. In: Rosen, J., Hannaford, B., Satava, R.M. (eds.) Surgical Robotics: Systems Applications and Visions, pp. 791–812. Springer (2010)
4. Ahlering, T.E., Skarecky, D., Lee, D., Clayman, R.V.: Successful Transfer of Open Surgical Skills to a Laparoscopic Environment Using a Robotic Interface: Initial Experience With Laparoscopic Radical Prostatectomy. J. Urology. 170(5), 1738–1741 (2003)
5. Ballantyne, G.: Telerobotic Gastrointestinal Surgery: Phase 2Safety and Efficacy. Surg. End. 21(7), 1054–1062 (2007)
6. Carlsson, S., Nilsson, A.E., Schumacher, M.C., Jonsson, M.N., Volz, D.S., Steineck, G., Wiklund, P.N.: Surgery-related Complications in 1253 Robot-assisted and 485 Open Retropubic Radical Prostatectomies at the Karolinska University Hospital, Sweden. Urology. 75(5), 1092–1097 (2010)
7. Grantcharov, T.P., Kristiansen, V.B., Bendix, J., Bardram, L., Rosenberg, J., Funch-Jensen, P.: Randomized clinical trial of virtual reality simulation for laparoscopic skills training. Brit. J. Surg. 91(2), 146–150 (2004)
8. Xin, H., Zelek, J.S., Carnahan, H.: Laparoscopic surgery, perceptual limitations and force: A review. In: First Canadian Student Conference on Biomedical Computing, Kingston (2006)
9. Cuss, A., Abbott, J.: Complications of laparoscopic surgery. Obstetrics, Gynaecology and Reproductive Medicine 22(3), 59–62 (2012)
10. Lee, J.D., Hoffman, J.D., Hayes, E.: Collision warning design to mitigate driver distraction. In: SIGCHI Conference on Human Factors in Computing Systems, Vienna, pp. 65–72 (2004)

11. Spence, C., Ho, C.: Multisensory warning signals for event perception and safe driving. Theoretical Issues in Ergonomics Science 9(9), 523–554 (2008)
12. Okazaki, R., Kajimoto, H., Hayward, V.: Vibrotactile stimulation can affect auditory loudness: A pilot study. In: Isokoski, P., Springare, J. (eds.) EuroHaptics 2012, Part II. LNCS, vol. 7283, pp. 103–108. Springer, Heidelberg (2012)
13. Bark, K., McMahan, W., Remington, A., Gewirtz, J., Wedmid, A., Lee, D.I., Kuchenbecker, K.J.: In Vivo Validation of a System for Haptic Feedback of Tool Vibrations in Robotic Surgery. Surg. End., 1–9 (2012)
14. Bianchi, M., Gwilliam, J.C., Degirmenci, A., Okamura, A.M.: Characterization of an air jet haptic lump display. IEEE Eng. Med. Biol. Soc., 3467–3470 (2011)
15. Gwilliam, J.C., Pezzementi, Z., Jantho, E., Okamura, A.M., Hsiao, S.: Human vs. robotic tactile sensing: Detecting lumps in soft tissue. In: IEEE Haptics Symposium, Waltham, pp. 21–28 (2010)
16. Okamura, A.M., Webster III, R.J., Nolin, J.T., Johnson, K.W., Jafry, H.: The Haptic Scissors: Cutting in Virtual Environments. In: International Conference on Robotics and Automation, Taipei, pp. 828–833 (2003)
17. Park, J.W., Choi, J., Park, Y., Sun, K.: Haptic Virtual Fixture for Robotic Cardiac Catheter Navigation. Artificial Organs. 35(11), 1127–1131 (2011)
18. Gibo, T.L., Verner, L.N., Yuh, D.D., Okamura, A.M.: Design Considerations and Human-Machine Performance of Moving Virtual Fixtures. In: International Conference on Robotics and Automation, Kobe, pp. 671–676 (2009)
19. Yao, H.Y., Hayward, V.: Design and analysis of a recoil-type vibrotactile transducer. J. Acous. Soc. Am. 128(2), 619–627 (2010)
20. Lang, M.J., Greer, A.D., Sutherland, G.R.: Intra-operative Robotics: NeuroArm. In: Intraoperative Imaging Acta Neurochirurgica Supplementum, pp. 231–236. Springer-Verlag (2011)

Visuo-haptic Tool for Collaborative Adjustment of Selections

Adrien Girard, Yacine Bellik, Malika Auvray, and Mehdi Ammi

CNRS-LIMSI, Univ. Paris Sud,
Rue John Von Neumann, 91403 Orsay, France
`{adrien.girard,yacine.bellik,malika.auvray,mehdi.ammi}@limsi.fr`
`http://www.limsi.fr`

Abstract. Mutual awareness between users working in collaborative virtual environments is an important factor for efficient collaborations. Several studies have reported that haptic feedback improves performance in collaborative tasks. However, few researches have tried to evaluate the influence of haptic feedback on mutual awareness, and to link the corresponding measures with the performance and efficiency factors. In the context of collaborative 3D polygonal modelling, we present a collaborative interaction method that dynamically adjust the selection area of the different involved partners. The aim of this interaction method is to improve the efficiency of collaborative working by improving the partners' mutual awareness. The experimental evaluation compares the proposed collaborative method of selection with a standard and individual method of selection used in most polygonal modelling software. The experimental results show an improvement in working efficiency and a better work distribution between the partners. Moreover, the analysis of awareness measure shows that the proposed approach balances self awareness and mutual awareness.

Keywords: Haptic feedback, polygonal modelling, shared situation awareness, selection adjustment.

1 Introduction

Distant collaboration is a promising solution to expedite team working and to associate experts with different skills in complex projects. However, this working approach presents a real challenge for the coordination of actions between the remote partners. In the case of the Collaborative Virtual Environments (CVE), the communication between partners is limited by the network latency, by issues related to the world consistency, and by the limits of communication systems which do not support some components such as gestural and facial expressions. All these issues limit the shared situation awareness (SSA), and may lead to a reduction of performance and efficiency of the group. Shared situation awareness (SSA) is the degree of similarity between each member's perception of a same given situation [1]. This component plays an important role in actions' coordination between partners and thus provides a better efficiency for collaborative tasks.

During the edition of the geometry of a shared object, a lack of SSA might produce unwanted deformations. For instance, when a first user manipulates a large selection and

I. Oakley and S. Brewster (Eds.): HAID 2013, LNCS 7989, pp. 40–49, 2013.

at the same time a second user manipulates a small selection overlapping the selection of the first user. If the first user didn't notice this second selection, it can conduct to unexpected deformations.

In fact, to complete reliable and effective collaborative tasks, the user needs to coordinate his actions and selections with the partner's actions and workspace. Thus, SSA plays an important role for collaboration in CVE.

Another issue for collaborative tasks concerns unbalanced workload between partners. This issue is due to social loafing and to the difference in skill level between partners. Social loafing corresponds to the tendency of the group members to do less than what they are capable of as individuals, the experiments of Blaskovich [2] suggest that social loafing is stronger in distant and virtually supported collaboration than in collocated collaboration. Unbalanced workload leads to a reduction in performance and efficiency during collaborative tasks.

In order to improve SSA during collaborative tasks, and to limit the unbalanced workload, we propose a collaborative method of selection which would force the user to pay attention to the partner's actions and workspace. This approach increases the user's attention to relevant information about the partner instead of providing additional information which may be useless [1]. Moreover, the proposed method integrates a haptic communication component that effectively communicate, through a visuo-haptic guidance, the required area to manipulate to the partner. The metaphor, named Collaborative Selection Adjustment (CSA), was experimented in the context of collaborative modelling. The investigated task concerns the manipulation of geometry of 3D polygonal models at different scales to produce new shapes.

2 Background

Several researches have investigated the role of the haptic channel in CVE. Sallnäs et al.[3], Basdogan et al.[4] and Groten et al. [5] have investigated the implicit haptic communication (i.e., feed-through) between partners during collaborative skill games applications. Their results have shown that haptic feedback significantly improves gestural accuracy and task performance. Groten et al. [5] have studied the influence of haptic feedback on efficiency (i.e., ratio between performance and physical effort). The results have shown that haptic feedback reduces working efficiency but improves performances. These studies highlight the importance of haptic communication during virtual collaboration. However, they focus on a standard haptic feed-through mechanism, which corresponds to the natural force interaction, and they do not investigate advanced communicative features of the haptic channel.

Oakley et al. [6] were among the first to use haptic feedback to create metaphoric communication. They have proposed different mechanisms of haptic communication to enhance interactions between partners during the collaborative edition of 2D diagrams. The cursors can attract or be attracted by each other, therefore the attraction mechanisms can indicate information about position. The subjective results have revealed that users find the haptic communication engaging and helpful. Girard et al.[7] have proposed an attraction mechanisms dedicated to 3D molecular modelling applications and they showed that the attraction mechanisms can be used to improve selection performances.

Ullah [8] has proposed several haptic interactions to improve the coordination during collaborative manipulations. The haptic functions enabled communication of the partner's actions. Moreover, they assisted the partners by keeping them in contact with the shared object. The different results have shown, through subjective measures, how haptic functions improved SSA and gestural coordination. The use of haptic channel to support metaphorical communication through artificial forces is well accepted by users. The proposed method was inspired by these metaphorical haptic communications to support team coordination.

Beyond haptic-centred approaches, Nova et al.[9] have investigated the link between the communicated information and SSA in the context of 3D video games. A visual awareness tool was proposed in order to provide different information to the partner. The results have shown that the awareness tool failed to improve SSA, but it increased performance. Endsley and Garland [1] noticed that providing more data is different from providing more information. Users can be overwhelmed by an excessive amount of data which may hide the useful information.

As introduced in a previous work [10], these two last works lead us to propose an interaction method which increases the user's attention on relevant information instead of providing additional information. The impact of the proposed collaborative interaction method on the SSA was measured by a Situation Awareness Global Assessment Technique (SAGAT) [11].

3 Proposed Approach

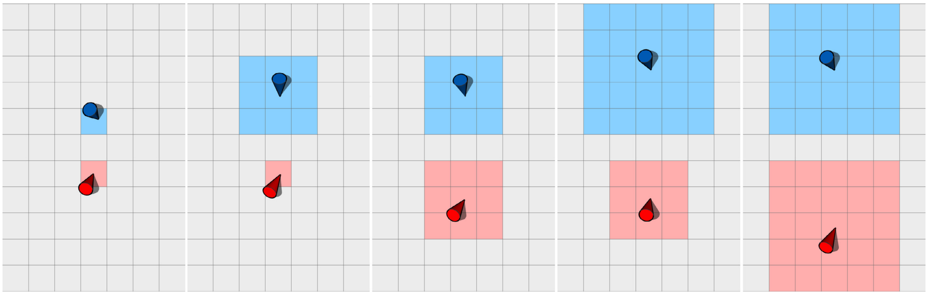
3.1 Context

In several fields, such as molecular modelling [12], virtual sculpting [13], and computer aided design [14], users need to have an access to different scales of manipulation. For example, during the molecular assembly, biologists need to manipulate the molecule from atoms (i.e., elementary structure), residues (i.e., intermediate structure), or fragments (i.e., important region of the molecule) in order to deform some local regions or the overall shape of the molecule for a good geometric matching with the second molecule during the assembly.

The study reported here investigates this issue in the context of multilevel manipulation of 3D polygon meshes. In fact, these geometric structures require manipulations at different scales such as a single face, a small group of faces, or an important part of the mesh in order to control the overall shape and the different levels of details of the designed object. The existing approaches to adjust the scale of manipulation consists in defining explicitly the required selection with a graphic user interface, for instance, with buttons or sliders to increase or decrease the size of the selection. However, this approach is not adapted for synchronous collaborative manipulations where users need to manipulate the shared object at the same scale. We propose to adapt this single user function to a collaborative function that involves two partners simultaneously.

3.2 Dynamic Adjustment of Selection

The proposed approach for collaborative adjustment of selection consists in dynamically adapting the areas selected by each user. More precisely, it increases or reduces


(a) (b) (c) (d) (e)

(a) One single face separate the two users, they both have a selection of one face. (b) Users are separated by 2 faces. (c) Users are separated by 3 faces. (d) Users are separated by 4 faces. (e) Users are separated by 5 faces.

Fig. 1. Size of the selection according the distance between the cursors

the selected areas (i.e., the number of selected faces) according to the distance (i.e., the number of faces) separating the two partners. The size of each user's selection area is determined by a propagation mechanism. From the faces touched by each user (i.e., faces in contact with the cursors), the selection areas are extended to adjacent faces which are connected by edges or vertices. In order to avoid overlapping between the two selection areas, which can lead to conflicting actions, the propagation is stopped before the two areas overlap. Thus, there is always at least one face between the two selected areas.

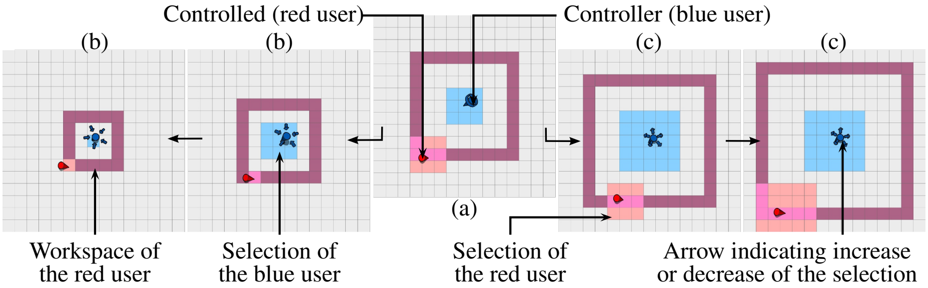
Figure 1 presents some cases of selections according to the distance separating the two cursors. On a flat surface with a uniform distribution of square faces (see Figure 1), the number of faces selected by each partner is defined by:
- $F_1 = (N+1-N(mod\ 2))^2$
- $F_2 = (N-1+N(mod\ 2)))^2$

Where N is the number of faces separating the users, F_i is the number of faces selected by the user i and $a(mod\ b)$ is the modulo operation of a by b.

With this selection method, partners just have to move closer or to move away in order to access to different scales of selections and manipulations. Thus, they can easily switch from a local manipulation to the manipulation of the overall mesh. This method generates two selection areas with similar sizes, thereby the partners always manipulate the mesh at the same scale. It should be noted that this function requires that both partners are simultaneously in contact with the mesh to adjust the selections, if only one user touch the mesh his selection is extended to the entire mesh. Once the selection completed, the two partners can manipulate independently their selections.

This approach is designed for collaborative tasks involving two partners. However, it can be extended to more users, for example, by taking into account the geometric distance with the nearest partner. Finally, the evaluation of this approach was limited in our study to the manipulation of planar shapes for an easier perception of the selection areas, and to avoid constraints related to the multiple points of view which are external to the current issue. However, the proposed approach can be applied to more complex shapes without any modification.

3.3 Control of the Selection Size



(a) Initial situation. The selection size is controlled by the blue user (controller user). The purple faces are the only faces where the controlled user (red cursor) can move. The controller user can not move away from the face he selects.

(b) The controller user reduces his selection size by bringing the controlled partner closer.

(c) The controller user increases his selection size by pushing the controlled partner away.

Fig. 2. Example of selection size control

The dynamic adjustment of selections allows the two partners to manipulate the shape at different scales. However, the partners need to coordinate their relative positions to obtain the correct selections and thereby perform the correct deformation. More precisely, once the first user selects a given face, the second partner have to select the correct relative position in order to get the correct size and position of selections to his partner. This involves the communication of the relative distance and orientation of the region to select to the partner. However, the designation of targets in 3D CVE is a complex and difficult task due to the limits of depth perception in 3D environments. To address this issue, we propose to provide the first user that starts the selection (named the controller user) with a communication tool that limits the workspace of the his partner (named controlled user). This workspace corresponds to the faces that respect the required distance between the two active selection. On a flat shape constituted by a homogeneous distribution of faces the limited workspace describes a square shape around the controller user. The movements of the controlled partner are haptically constrained on the limited workspace (Figure 2). The controller user controls the relative distance separating him from the controlled partner (i.e., the number of the faces separating the two cursors), and thus he can control the size of his selection. Inside this limited workspace, the controlled partner can move freely and thus can adapt his relative orientation according to the verbal indications of his partner. The controlled user can break the workspace limitation if he considers that the proposition of selection of his partner is incorrect or he can accept it and start the mesh deformation.

To increase (Figure 2.c) or reduce (Figure 2.b) the size of the active workspace (i.e., the radius of the faces separating the two cursors), the controller pushes or pulls on the face he is touching in order to push away or pull closer the active workspace. A haptic feedback imitates a press button effect to provide the perception of transitions.

4 Experimental Evaluation

4.1 Methods

The aim of this experiment is to evaluate the impact of the collaborative selection adjustment (CSA) on efficiency, workload balance and shared situation awareness. Based on these goals, we address the following hypotheses:
-**H1** CSA improves collaborative work's efficiency.
-**H2** CSA improves workload balance between users.
-**H3** CSA balances self situation awareness and shared situation awareness.

Hardware and Software Setup: The experimental platform was based on a standard desktop station. Two PHANToM Omnis from SensAble and two 23 inch screens were connected to the same computer in order to avoid network latency issues. The two working spaces were separated with a curtain to avoid visual contact between the partners. However, the participants could communicate verbally.

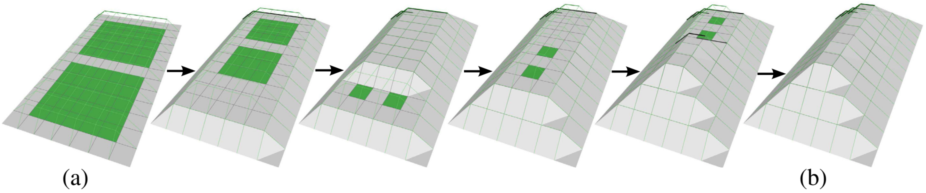


Fig. 3. Example of deformation task.(a)Initial 3d object: a plane. The two sets of green faces represent the faces the users have to select and translate. Only one user can see the green faces and knows the required selection. The green grid over the mesh represents the shape to match after the first pair of deformations.(b)Final 3d object after 5 sequential pairs of deformations correctly performed.

Conditions: Two conditions were presented to participants:
-**CSA:** collaborative selection adjustment
-**ISA:** individual selection adjustment
The CSA condition is based on the approach presented in this paper (cf. section 3.2). The ISA condition is based on the standard method of adjustment of the size selection used in computer graphics softwares. The selection size is controlled by a manual and individual method without interaction with the partner. The control of the selection size is based on the same interaction mechanism as the CSA approach. The user pushes or pulls the selected face to increase or decrease the size of his selection area (cf. section 3.3). Unlike the CSA method, the ISA condition does not constrain the position of the partner on an active workspace.

Participants: 24 participants (12 pairs, 21 men and 3 women) recruited in the LIMSI lab and University of Paris-Sud, aged between 22 and 34 years old, completed the experiment.

Procedure: The objective of the task is to deform a flat shape in order to turn it into new defined shapes: a pyramid (Figure 3), stairs and a canyon shape. The expected deformation is displayed with a green wire mesh (Figure 3.a) and five pairs of deformation are required to succeed. The correct selection to perform the required deformation is displayed with two sets of green faces (Figure 3.a). The required deformation (green wires) is displayed for both users, but the required selections (set of green faces) are only displayed for one single user. This additional information, provided to one user, is important to create an unbalanced knowledge and to force communication between the partners. Every selection indication (green faces) is presented in pairs, in order to encourage parallel work.

The subjects follow a progressive tutorial before performing the evaluated task. Furthermore, in order to simplify deformation tasks, the deformations are constrained according to the normal vector of the manipulated face.

Measures: Several objective measures were collected for both conditions. The following measures concern the evaluation of performance and efficiency.

- **M1** Completion score (%): mean percentage of successful deformations.
- **M2** Completion time (s): mean time to perform all the required deformations.
- **M3** Travelled distance (dm): average distance travelled by both users during the task.
- **M4** Deformation distance (dm): average distance travelled by both users during the deformations process.

The following measures concern the analysis of the work distribution within pairs:

- **M5** Distance difference (dm): difference between the distance travelled by each of the two users (| M3 of user 1 - M3 of user 2|).
- **M6** Deformation difference (dm): difference between the distance travelled by each of the two users during the deformations (|M4 of user 1 - M4 of user 2|).

Finally, the following measures are proposed to study the evolution of situation awareness (SA) and shared situation awareness (SSA). During each modelling task, the interactions with the 3D object are frozen at a random instant. The cursor of the partner is removed from the visual display. The indications of selection (green areas) are also hidden. Then, we ask participants to indicate the following information:

- **M7** Next selection (SA): What will be your next selection?
- **M8** Partner's next selection (SSA): What will be your partner's next selection?

For these two measures the participants have to indicate the planned next selection by touching the corresponding face. This measure provides a binary value; 1 means the user guesses correctly the next selection performed, 0 means the user did not find the next selection.

Subjective Measures: A questionnaire composed of 3 statements was presented to each participant after each condition. The subjects indicate on a Likert-scale if they agree or disagree with the statements. Scores ranging from 1 to 5 mean strongly disagree and strongly agree respectively.

- **S1** The tutorial was not necessary.
- **S2** I have done the same number of deformations as my partner.
- **S3** The work was evenly distributed between me and my partner.

4.2 Results

As our population does not respect a normal distribution, we used a non-parametric statistical test named Wilcoxon signed-rank test to determine if the results are significant.

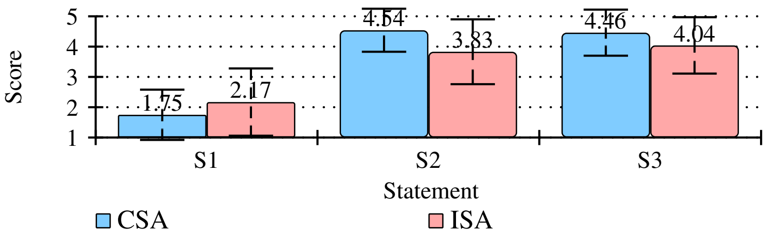


Fig. 4. Subjective results and standard deviation

Table 1. Results for the efficiency measures and for the distribution measures

	M1(%) Completion score	M2 (s) Completion time	M3 (dm) Distance travelled	M4 (dm) Deformation distance	M5(dm) Distance difference	M6 (dm) Deformation difference
ISA condition	69.9	125	372	21.4	55.55	7.99
CSA condition	77.4	113	284	19.1	51.32	3.78
Reduction (%)	-10.7	9.5	23.7	10.6	7.61	52.7
Wilcoxon (p-value)	0.295	0.063	0.001	0.658	0.278	0.035

Work Efficiency: Table 1 summarizes the results related to the completion score (M1) and completion time (M2). The mean completion score presents a non significant improvement of 10.7% from ISA to CSA ($p-value > 0.05$). However, the completion time is reduced by 9.5% with a borderline p-value ($p-value < 0.1$). The improvement in performance under the CSA condition is arguable. We can consider that the performances under the two conditions are similar. The score of **S1** (Figure 4) suggests that the ISA condition is easier to understand than the CSA condition. Users are indeed faster with the ISA condition which may improve the performance at the beginning of the experiment.

The analysis of the travelled distance (M3) shows a significant reduction of 23.7% between the ISA and CSA conditions ($p-value < 0.05$). Under the CSA condition, users are more careful about their movements since they influence the selection of their partner. This leads to a reduction of the travelled distance. The efficiency is defined as the capacity to produce outcome with minimum effort, so we can consider that CSA improves working efficiency since the effort is reduced without affecting performance. Thus, hypothesis **H1** is validated. By contrast, the deformation distance measure (M4) shows no-significant difference ($p-value > 0.05$) between the ISA and CSA conditions. CSA reduces the effort involved during the overall task but not the effort involved during the deformation process.

Work Distribution: Table 1 presents the results related to the difference of travelled distance (M5) and deformation distance (M6). These results show a non significant improvement in the difference of travelled distance of 7.61 % ($p-value > 0.05$) between the ISA and CSA conditions. The two partners travelled similar distances.

By contrast, the deformation difference (M6) measure shows a significant reduction of 52.7 % ($p-value < 0.05$) between the ISA and CSA conditions. Indeed, some users are more efficient than their partners. They may have better skills in 3D environments and haptic interaction. CSA reduces this difference because the two partners have to work simultaneously in order to deform the mesh. Under the ISA condition, the two partners can work independently, which can lead to an unbalanced activity. If one participant is far more efficient than his partner, he can individually perform all the tasks. The statements **S2** and **S3** of Figure 4 concern the perception of the work distribution. The scores of **S2** and **S3** (Figure 4), corresponding to the perception of the work distribution, are more important for the CSA condition (0.71 and 0.42 respectively) than for the ISA condition. This subjective result suggests that users perceived a better balance of work distribution under the CSA condition. Thus, hypothesis **H2** is validated.

SA and SSA: Table 2 compares the M7 measure (SA) with the M8 measure (SSA) according to the ISA and CSA conditions. Under the ISA condition, the M7 measure is significantly more important ($p - value < 0.05$) than the M8 measure. This result shows that users are more accurate and aware of the selection that they plans to do (SA), than of the selections planned by their partner (SSA). This difference between shared and self awareness is not observed under the CSA condition. Compared to the ISA condition, CSA does not improve self situation awareness (SA), however CSA improves the balance between situation awareness (SA) and shared situation awareness (SSA), the hypothesis **H3** is thus validated. The user's awareness is better shared between his activity and the activity of his partner. This results suggest that awareness capacities are limited, and that in this application context awareness can not be increase but they are rather redistributed.

Table 2. Comparison of results between M7 and M8 measures

	ISA condition	CSA condtion
M7 Next selection	0.69	0.61
M8 Partner next selection	0.44	0.55
Reduction (%)	36	9.1
Wilcoxon (p-value)	0.014	0.53

5 Discussion and Conclusion

The experimental results confirm that haptic feedback improves working performance during collaborative tasks [4] [3] [5]. Moreover, in contrast to the results of Groten et al.[5], the proposed approach improves working efficiency. The differences in context (i.e., 2D skill games) and metrics used (i.e., force and velocity) may explain this difference in the obtained results. The collaborative method of selection balances the situation awareness and the shared situation awareness which can be useful for a more efficient supervision of the collaborative process. We experimented our approach with simple modelling tasks which were especially designed to be suitable for collaborative work. It would be interesting to evaluate this same approach with a more generic and complex modelling tasks. Moreover, the investigation of new contexts like virtual sculpting [13] could highlight new constraints requiring some adaptation such as new criterion to adjust the selection areas (e.g., euclidean distance between cursors instead of a number of face).

References

1. Endsley, M., Garland, D.J.: Theoretical underpinnings of situation awareness: A critical review. In: Situation Awareness Analysis and Measurement, pp. 3–32. Lawrence Erlbaum Associates, Mahwah (2000)
2. Blaskovich, J.L.: Exploring the effect of distance: An experimental investigation of virtual collaboration, social loafing, and group decisions. Journal of Information Systems 22(1), 27–46 (2008)
3. Sallnäs, E.L., Rassmus-Gröhn, K., Sjöström, C.: Supporting presence in collaborative environments by haptic force feedback. ACM Transactions on Computer-Human Interaction 7(4), 461–476 (2000)
4. Basdogan, C., Ho, C.H., Srinivasan, M.A., Slater, M.: An experimental study on the role of touch in shared virtual environments. ACM Transactions on Computer-Human Interaction 7(4), 443–460 (2000)
5. Groten, R., Feth, D., Peer, A., Buss, M., Klatzky, R.: Efficiency analysis in a collaborative task with reciprocal haptic feedback. In: Proceedings of the 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 461–466. IEEE, St. Louis (2009)
6. Oakley, I., Brewster, S., Gray, P.: Can You Feel the Force? An Investigation of Haptic Collaboration in Shared Editors. In: The Proceedings of EuroHaptics, pp. 54–59 (2001)
7. Girard, A., Ammi, M., Simard, J., Auvray, M.: Improvement of collaborative selection in 3D complex environments. In: 2012 IEEE Haptics Symposium, Vancouver, British Columbia, Canada, pp. 281–288 (2012)
8. Ullah, S.: Multi-modal Assistance for Collaborative 3D Interaction: Study and analysis of performance in collaborative work. PhD thesis, Université d'Evry-Val d'Essonne (2011)
9. Nova, N., Wehrle, T., Goslin, J., Bourquin, Y., Dillenbourg, P.: Collaboration in a multi-user game: impacts of an awareness tool on mutual modeling. Multimedia Tools and Applications 32(2), 161–183 (2006)
10. Girard, A., Bellik, Y., Auvray, M., Ammi, M.: Collaborative approach for dynamic adjustment of selection areas in polygonal modelling. In: Proceedings of the 18th ACM Symposium on Virtual Reality Software and Technology, VRST 2012, pp. 205–206. ACM, New York (2012)
11. Bolstad, C., Cuevas, H., Gonzalez, C., Schneider, M.: Modeling shared situation awareness. In: Proceedings of the 14th Conference on Behavior Representation in Modeling and Simulation (BRIMS), Los Angeles (2005)
12. Chastine, J.W., Brooks, J.C., Zhu, Y., Owen, G.S., Harrison, R.W., Weber, I.T.: AMMP-Vis: A Collaborative Virtual Environment for Molecular Modeling. In: ACM Symposium on Virtual Reality Software and Technology, pp. 8–15 (2005)
13. Gunn, C.: Collaborative virtual sculpting with haptic feedback. Virtual Reality 10(2), 73–83 (2006)
14. Nam, T.J., Wright, D.: The development and evaluation of syco3d: a real-time collaborative 3d cad system. Design Studies 22(6), 557–582 (2001)

Multi-command Tactile Brain Computer Interface: A Feasibility Study

Hiromu Mori[1], Yoshihiro Matsumoto[1], Victor Kryssanov[2], Eric Cooper[2],
Hitoshi Ogawa[2], Shoji Makino[1], Zbigniew R. Struzik[3,4],
and Tomasz M. Rutkowski[1,4,⋆]

[1] Life Science Center of TARA, University of Tsukuba, Tsukuba, Japan
[2] Ritsumeikan University, Shiga, Japan
[3] University of Tokyo, Tokyo, Japan
[4] RIKEN Brain Science Institute, Wako-shi, Japan
`tomek@tara.tsukuba.ac.jp`
`http://about.bci-lab.info/`

Abstract. The study presented explores the extent to which tactile stimuli delivered to the ten digits of a BCI-naive subject can serve as a platform for a brain computer interface (BCI) that could be used in an interactive application such as robotic vehicle operation. The ten fingertips are used to evoke somatosensory brain responses, thus defining a tactile brain computer interface (tBCI). Experimental results on subjects performing online (real-time) tBCI, using stimuli with a moderately fast inter-stimulus-interval (ISI), provide a validation of the tBCI prototype, while the feasibility of the concept is illuminated through information-transfer rates obtained through the case study.

Keywords: tactile BCI, P300, robotic vehicle interface.

1 Introduction

Contemporary BCIs are typically based on mental visual and motor imagery prototypes, which require extensive user training and non-impaired vision of the users [13]. Recently alternative solutions have been proposed to make use of spatial auditory [1,10] or tactile (somatosensory) modalities [6,5,12] to enhance brain-computer interface comfort and increase the information transfer rate (ITR) achieved by users. The concept proposed in this paper uses the brain somatosensory (tactile) channel to allow targeting of the tactile sensory domain for the operation of robotic equipment such as personal vehicles, life support systems, etc. The rationale behind the use of the tactile channel is that it is normally far less loaded than visual or even auditory channels in such applications.

One of the first reports [6] of the successful employment of steady-state somatosensory responses to create a BCI targeted a low frequency vibrotactile stimulus in the range of $20 - 31$ Hz to evoke the subjects' attentional modulation, which was then used to define interfacing commands. A recent report [12]

⋆ Corresponding author.

I. Oakley and S. Brewster (Eds.): HAID 2013, LNCS 7989, pp. 50–59, 2013.

proposed using a Braille stimulator with 100 ms static push stimuli delivered to each of six fingers to evoke a somatosensory response potential (SEP) related P300. The P300 response is a positive electroencephalogram event-related potential (ERP) deflection starting at around 300 ms and lasting for $200 - 300$ ms after an expected stimulus in a random series of distractors (the so-called oddball EEG experimental paradigm) [7]. P300 responses are commonly used in BCI approaches and are considered to be the most reliable ERPs [9,13] with untrained subjects. The results indicated that the experiments achieved information transfer rates of 7.8 bit/min on average and 27 bit/min for the best subject.

This paper proposes a novel tactile brain-computer interface based on P300 responses evoked by tactile stimuli delivered via vibrotactile exciters attached to the ten fingertips of the subject's hands.

The rest of the paper is organized as follows. The next section introduces the materials and methods used in the study and also outlines the experiments conducted. The results obtained in psychophysical and electroencephalogram experiments with eleven BCI-naive subjects are then discussed. Finally, conclusions are formulated and directions for future research are outlined.

2 Materials and Methods

Eleven paid BCI-naive subjects (ten males and one female) participated in the experiments. The subjects' mean age was 21.82, with a standard deviation of 0.87. All the experiments were performed at the Life Science Center of TARA, University of Tsukuba, Japan. The psychophysical and online (real-time) EEG tBCI prototype experiments were conducted in accordance with the *WMA Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects.*

2.1 Tactile Stimuli

The tactile stimuli were delivered as sinusoidal waves generated by a portable computer with MAX/MSP software [2]. The stimuli were generated via ten channel outputs (one for each fingertip of the subject) of an external *digital-to-analog* signal converter MOTU UltraLite-mk3 Hybrid coupled with three YAMAHA P4050 power amplifiers (four acoustic frequency channels each).

The stimuli were delivered to the subjects' fingertips via the tactile exciters HiWave HIAX25C10-8/HS working in the range of $100 - 20,000$ Hz, as depicted in Figure 1. Each exciter in the experiments was set to emit a sinusoidal wave at 200 Hz to match the exciter's resonance frequency and to stimulate the *Pacini endings* (fast-adapting type II afferent type tactile sensory hand innervation receptors) [3].

The subjects placed their fingertips on the exciters (see Figure 1) and attended only to the instructed locations (with a button-press response in the

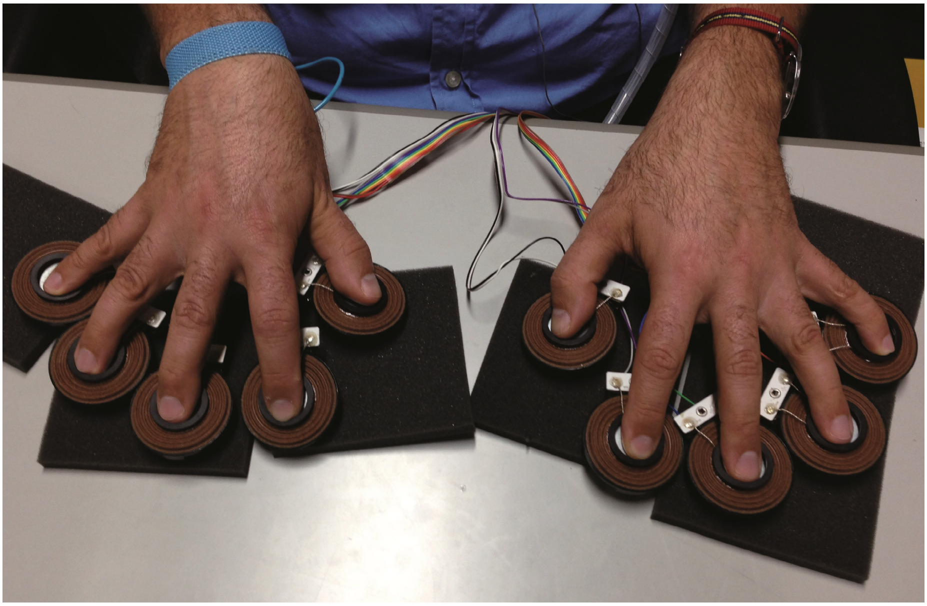


Fig. 1. The experimental set-up. Tactile stimuli are delivered to the fingertips by ten HIAX25C10-8 exciters. The set-up was used in both the psychophysical and the EEG experiments.

psychophysical experiments, or a mental count of the targets in the EEG experiments). The training instructions were presented visually by means of the MAX/MSP program, as depicted in Figure 2.

2.2 Psychophysical Experiment Protocol

The psychophysical experiment was conducted to investigate the stimulus carrier frequency influence on the subjects' response time and accuracy. The behavioral responses were collected using a small numeric keypad and the MAX/MSP program. Each subject was instructed which stimulus to attend to by a cross above the target fingertip ($TARGET$), shown by the program (Figure 2). Then the subject pressed the response button with the dominant foot. In the experiment, each subject was presented with 50 $TARGETS$ and 450 $non-TARGETS$ as stimuli.

Each trial was composed of a randomized order of 100 ms tactile bursts delivered to each fingertip separately with an inter-stimulus-interval (ISI) of 600 ms. Every random sequence thus contained a single $TARGET$ and nine $non-TARGETS$. A single session included five trials for each $TARGET$ fingertip (resulting in 50 $TARGETS$ and 450 $non-TARGETS$). The choice of the relatively long ISI is justified by slow behavioral responses in comparison to EEG evoked potentials [7], as described in the next section.

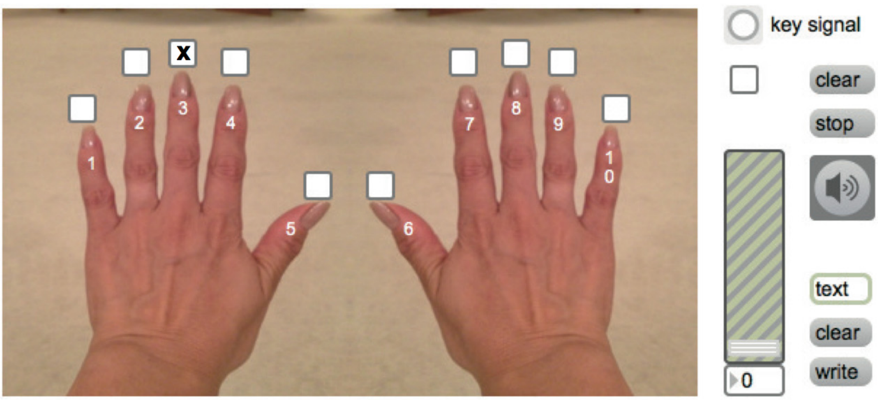


Fig. 2. The visual instruction screen presented to the subjects during the psychophysical experiment. Each fingertip was assigned a number encoding a command in the interactive application. The controls on the right side were used by the subject to adjust the stimulus intensity (with a fader), and also by the experimenter to start the experiment (the button with the speaker launches the *digital-to-analog* signal converter MOTU UltraLite-mk3 Hybrid) and to save/clear the results.

The response time delays were registered with the same MAX/MSP program, also used for the stimulus generation and instruction presentation.

2.3 EEG tBCI Experiment

The exciters were attached to the fingertips in the same manner (see Figure 1). During the experiment, EEG signals were captured with a portable wireless EEG amplifier system g.MOBIlab+ and g.SAHARA by g.tec, using eight dry electrodes. The electrodes were attached to the head locations: Cz, CPz, P3, P4, C3, C4, CP5, and CP6, as in the 10/10 extended international system [4] (see the topographic plot in the top panel of Figure 3). The ground and reference electrodes were attached behind the left and right ears respectively. In order to limit electromagnetic interference, the subjects' hands were additionally grounded with armbands connected to the amplifier ground. No electromagnetic interference was observed from the exciters. Details of the EEG experimental protocol are summarized in Table 1.

The recorded EEG signals were processed by a BCI2000-based application [9], using a stepwise linear discriminant analysis (SWLDA) classifier [8] with features drawn from the $0 - 700$ ms ERP interval. The sampling rate was set at 256 Hz, the high pass filter at 0.1 Hz, and the low pass filter at 40 Hz. The ISI was 400 ms, and each stimulus duration was 100 ms. The subjects were instructed to spell out the number sequences (corresponding to the interactive robotic application commands shown in Table 2) communicated by the exciters in each session. Each $TARGET$ was presented five times in a single command trial, and the averages of the five ERPs were later used for the classification. Each subject performed three experimental sessions (randomized 50 $TARGETS$ and 450 $non-TARGETS$ each), which were later averaged as discussed in Section 3.

3 Results

This section discusses the results obtained in the psychophysical experiment and in the EEG experiment.

3.1 Psychophysical Experiment Results

The psychophysical experiment results are summarized in Figure 4, where the median response time and the range are depicted for each fingertip as boxplots (see also Figure 2 for the fingertip numbering).

The Wilcoxon rank sum tests for pairwise comparisons revealed no differences (at the 0.05 level) among the median values for all the fingertip pairs of each subject. This result confirms the stimulus similarity since the behavioral responses for all the fingers were basically the same. This finding validates the design of the EEG experiment.

3.2 EEG Experiment Results

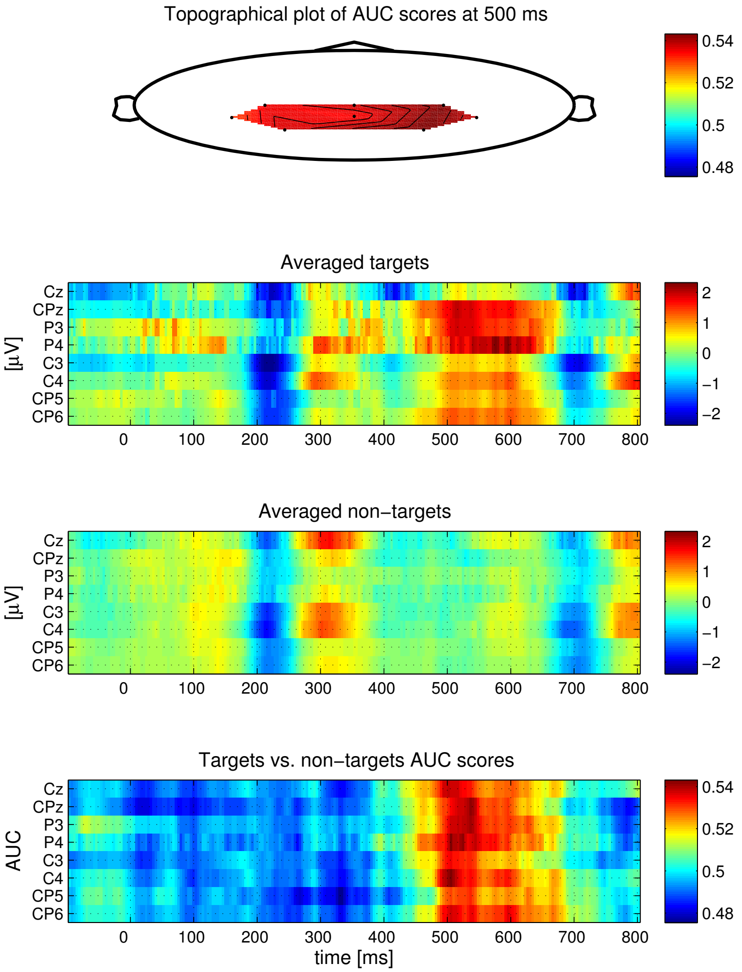
The results of the EEG experiment are summarized in Table 3 and Figure 3. All eleven BCI-naive subjects scored well above the chance level of 10%, reaching an ITR in the range from 0.19 bit/min to 4.46 bit/min, which may be considered to be a good outcome for experiments with beginners (naive subjects). The ITR was calculated as follows [10]:

$$ITR = V \cdot R \tag{1}$$

$$R = log_2 N + P \cdot log_2 P + (1-P) \cdot log_2 \left(\frac{1-P}{N-1} \right), \tag{2}$$

Table 1. Conditions of the EEG experiment

Number of subjects	11
Tactile stimulus length	100 ms
Stimulus frequency	200 Hz
Inter-stimulus-interval (ISI)	400 ms
EEG recording system	g.SAHARA & g.MOBIlab+ active dry EEG electrodes system.
Number of the EEG channels	8
EEG electrode positions	Cz, CPz, P3, P4, C3, C4, CP5, CP6.
Reference and ground electrodes	Behind both of the subject's ears
Stimulus generation	10 HIAX25C10-8 exciters
Number of trials for each subject	5



Topographical plot of AUC scores at 500 ms

Averaged targets

Averaged non−targets

Targets vs. non−targets AUC scores

Fig. 3. Grand mean averaged results of the fingertip stimulation EEG experiment for all 11 subjects. The top panel presents the head topographic plot of the $TARGET$ versus $non - TARGET$ area under the curve (AUC), a measure commonly used in machine learning intra-class discriminative analysis. ($AUC > 0.5$ is usually assumed to be confirmation of feature separability [11]). The top panel presents the largest difference as obtained from the data displayed in the bottom panel. The topographic plot also depicts the electrode positions. The fact that all the electrodes received similar AUC values (red) supports the initial electrode placement. The second panel from the top presents averaged SEP responses to the $TARGET$ stimuli (note the clear P300 response in the range of $450-700$ ms). The third panel presents averaged SEP responses to the $non-TARGET$ stimuli (no P300 observed). Finally, the bottom panel presents the AUC of $TARGET$ versus $non - TARGET$ responses (again, P300 could easily be identified).

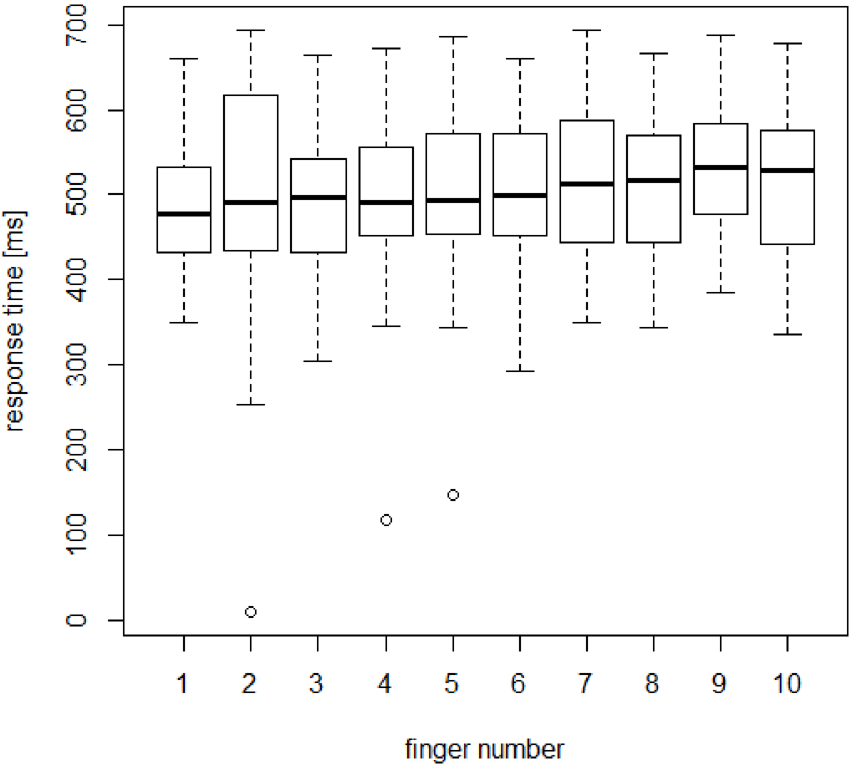


Fig. 4. Results in boxplots (note the overlapping quartiles visualizing no significant differences between medians) of the grand mean averages (11 subjects) of the psychophysical experiment response time delays. Each number represents a finger, as depicted in Figure 1. The dots represent outliers.

Table 2. Interactive application commands encoded with the finger numbers (see also Figure 1)

Finger number	Command
1	low speed
2	medium speed
3	high speed
4	stop
5	slow speed reverse
6	go left ($-90°$)
7	go straight–left ($-45°$)
8	go straight ($0°$)
9	go straight–right ($45°$)
10	go left ($90°$)

Table 3. The fingertip stimulation EEG experiment accuracy and ITR scores. The theoretical chance level was 10%. For the classifier, features were derived from the averages of the five ERPs of all the subjects.

Subject number	Maximum accuracy	ITR
1	20%	0.19 bit/min
2	40%	1.34 bit/min
3	50%	2.21 bit/min
4	30%	0.66 bit/min
5	30%	0.66 bit/min
6	40%	1.34 bit/min
7	70%	4.46 bit/min
8	40%	1.34 bit/min
9	20%	0.19 bit/min
10	30%	0.66 bit/min
11	20%	0.19 bit/min

where R stands for the number of bits/selection; N is the number of classes (10 in this study); P is the classifier accuracy (see Table 3); and V is the classification speed in selections/minute (3 selections/minute for this study). The maximum ITR it was possible for the BCI-naive subjects to achieve in the settings presented was 9.96 bit/min.

4 Conclusions

This paper reports results obtained with a novel ten-command tBCI prototype developed and used in experiments with eleven BCI-naive subjects. The proposed interface could be used for real-time operation of robotic vehicles. The experiment results obtained in this study confirmed the general validity of the tBCI for interactive applications.

In the psychophysical experiment, it is shown that all the tested fingertip zones are equally sensitive to the stimuli and all can be used with the tBCI prototype. The EEG experiment with the prototype has confirmed that tactile stimuli can be used to operate robotic devices with up to ten commands and with the command interfacing rate ranging from 0.19 bit/min to 4.46 bit/min for untrained users.

The results presented offer a step forward in the development of neurotechnology applications. Due to the still not very practical interfacing rate achieved, allowing for only about three commands per minute, the current prototype would obviously need improvements and modifications. These needs determine the major lines of study for future research. However, even in its current form, the proposed tBCI can be regarded as a practical solution for totally-locked-in (TLS) patients, who cannot use vision or auditory based interfaces due to sensory or other disabilities.

Acknowledgments. This research was supported in part by the Strategic Information and Communications R&D Promotion Program no. 121803027 of The Ministry of Internal Affairs and Communication in Japan, and by KAKENHI, the Japan Society for the Promotion of Science, grant no. 12010738. We also acknowledge the technical support of YAMAHA Sound & IT Development Division in Hamamatsu, Japan.

References

1. Halder, S., Rea, M., Andreoni, R., Nijboer, F., Hammer, E., Kleih, S., Birbaumer, N., Kübler, A.: An auditory oddball brain-computer interface for binary choices. Clinical Neurophysiology 121(4), 516–523 (2010)
2. `http://cycling74.com/`: Max 6 (2012)
3. Johansson, R.S., Flanagan, J.R.: Coding and use of tactile signals from the fingertips in object manipulation tasks. Nature Reviews Neuroscience 10(5), 345–359 (2009)
4. Jurcak, V., Tsuzuki, D., Dan, I.: 10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems. NeuroImage 34(4), 1600–1611 (2007)
5. Mori, H., Matsumoto, Y., Makino, S., Kryssanov, V., Rutkowski, T.M.: Vibrotactile stimulus frequency optimization for the haptic BCI prototype. In: Proceedings of The 6th International Conference on Soft Computing and Intelligent Systems, and The 13th International Symposium on Advanced Intelligent Systems, Kobe, Japan, November 20-24, pp. 2150–2153 (2012)
6. Muller-Putz, G., Scherer, R., Neuper, C., Pfurtscheller, G.: Steady-state somatosensory evoked potentials: suitable brain signals for brain-computer interfaces? IEEE Transactions on Neural Systems and Rehabilitation Engineering 14(1), 30–37 (2006)
7. Niedermeyer, E., Da Silva, F.L. (eds.): Electroencephalography: Basic Principles, Clinical Applications, and Related Fields, 5th edn. Lippincott Williams & Wilkins (2004)
8. Potes, C.M.: P300 Classifier (2009)

9. Schalk, G., Mellinger, J.: A Practical Guide to Brain-Computer Interfacing with BCI2000. Springer-Verlag London Limited (2010)
10. Schreuder, M., Blankertz, B., Tangermann, M.: A new auditory multi-class brain-computer interface paradigm: Spatial hearing as an informative cue. PLoS ONE 5(4), e9813 (2010)
11. Theodoridis, S., Koutroumbas, K.: Pattern Recognition, 4th edn. Academic Press (2009)
12. van der Waal, M., Severens, M., Geuze, J., Desain, P.: Introducing the tactile speller: an ERP-based brain-computer interface for communication. Journal of Neural Engineering 9(4), 045002 (2012)
13. Wolpaw, J.R., Wolpaw, E.W. (eds.): Brain-Computer Interfaces: Principles and Practice. Oxford University Press (2012)

Association of Haptic Trajectories to Takete and Maluma

Federico Fontana

Università di Udine, Department of Mathematics and Computer Science
via delle Scienze, 206 – 33100 Udine, Italy
`federico.fontana@uniud.it`

Abstract. An experiment has been made, in which participants grasping the stylus of a robotic arm were physically guided along a jagged or rounded trajectory, and then were asked to associate either trajectory to the word "takete" or "maluma". A significant preference (nine out of eleven participants) for associating the jagged trajectory to "takete" and the rounded trajectory to "maluma" has been found, indicating the existence of a connectivity between haptic trajectories and words. This result suggests to interaction designers to avoid the association of counter-intuitive labels or verbal meanings to (yet rarely used) structured synthetic kinesthetic messages ("haptons") that are perceived as jagged or rounded. The experiment complements existing research on cross-modal associations between stimuli belonging to other sensory channels, such as vision or taste, and words having demonstrated verbal equivalence to "takete" and "maluma". Furthermore, it raises interest on currently unanswered questions about the perceptual importance of temporal aspects in the haptic recognition of shapes by rectilinear or curvilinear contour patterns, and their higher-level decoding and connectivity at cortical level.

Keywords: takete and maluma, bouba and kiki, kinesthetic-verbal associations.

1 Introduction

In 1929, Wolfgang Köhler asked a group of Spanish speakers to make an association between the words "takete" or "maluma" and the images of two shapes, one jagged and the other rounded, like those in Figure 1. His results showed a significant preference of the speakers for associating "takete" with the jagged, and "maluma" with the rounded shape. The experiment has been repeated by several psychologists using different pairs of words, in particular "kiki" and "bouba", as well as involving speakers from different languages and levels of literacy. Apart from some specific exceptions reported for a population of Papua New Guinea, these experiments have all shown a general tendency of speakers, including young children aged 2.5 years old [1], to map rounded shapes on words containing the vowels "o" and "u", and, conversely, jagged shapes on words containing "e" and "i".

I. Oakley and S. Brewster (Eds.): HAID 2013, LNCS 7989, pp. 60–68, 2013.

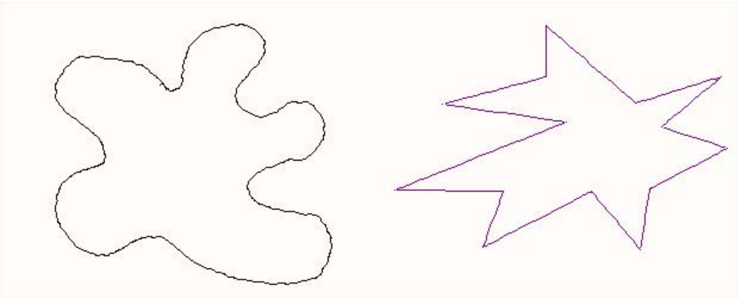


Fig. 1. Images similar to those used by Köhler in his experiments

Taken together, these results provide evidence of a powerful cross-modal effect, linking visual shapes to sounds of words. The existence of this effect in pre-literate children supports hypotheses on the existence of active connections among contiguous cortical areas, making possible for humans to link characteristic geometrical shape contours to similar geometries assumed by the speaker's lips. According to Ramachandran and Hubbard [2] such connections exist before language, hence they represent a general invariant influencing and constraining its development. Furthermore Ramachandran and Hubbard speculate that synesthesia, a phenomenon causing unusual, vivid cross-modal perceptions in a small percentage of the population, represents an amplified manifestation of the simultaneous activity occurring in connected cortical areas.

This body of research has been expanded by Spence and colleagues [3,4], who investigated on the existence of cross-modal associations between words and tastes. In evaluations where they asked subjects to associate perception of food and liquid tastes to takete/maluma or bouba/kiki, significant correlations emerged suggesting at least that counter-intuitive branding and labeling may be detrimental to the success of a food product. Besides its commercial implications, this investigation has demonstrated that the cross-modal associations of takete/maluma are not limited to visual shapes.

In this paper, a simple extension to the domain of non-visual shapes is presented. In particular, visual stimuli of shape contours such as those in Figure 1 have been substituted by kinesthetic feedback, providing cues of piece-wise rectilinear or curvilinear closed trajectories to blindfolded subjects. In the broader context of shape perception, haptic guidance has been mainly researched as a non-visual conveyor (often in connection with sound) of geometric and calligraphic information for training and teaching purposes, especially for enabling blind children to draw, to handwrite, and to perform collaborative explorations together with sighted users [5,6,7].

We avoided to investigate on the association of takete/maluma to tactile percepts of 3D shapes: this investigation would have represented a direct counterpart to the Köhler experiment once exposing two solids to the subjects' manipulation, respectively with rounded and piece-wise flat surfaces meanwhile having identical weight and texture; on the other hand, the use of 3D haptic trajectories as stimuli for the test restricts the displayed information to proprioceptive cues of contour, hence maintaining a clear parallelism with the original experiment. Furthermore, as we will see in the discussion, the employment of a simple robotic arm for the generation of kinesthetic feedback allows to link the experimental results directly to the design of haptic computer interfaces and virtual environments, especially concerning the realization of abstract haptic messages or "haptons" [8], as we may call them in analogy to tactons [9] to underline their kinesthetic nature, and in spite of the fact that their use in human-computer interaction has not been widely adopted yet [10].

2 Method

The experiment took place in a quiet room of the Department of Computer Science of the University of Verona, belonging to the Video, Image Processing and Sound Laboratory.

2.1 Participants

Eleven graduating students (six male and five female) participated to the test. All were Italian speakers, reporting normal proprioception and sight as well as no episodes of synesthesia. One subject had previous knowledge of the takete/maluma effect based on visual shapes, while the others were new about it. They were not economically rewarded, however upon individual request their participation to the experiment was acknowledged to the graduation course council of their Faculty.

2.2 Apparatus and Stimuli

A widely used robotic arm (Phantom Omni by SensAble Technologies) was programmed to draw a piece-wise rectilinear or curvilinear trajectory. This specific device has the advantage to expose a pen-like termination (called stylus) to the user's hand: this characteristic facilitated the simulation of an active pen, autonomously drawing rounded or jagged contours meanwhile being grasped by the subjects. Concerning the kinesthetic feedback,

- rectilinear movements reproduced a trajectory, whose projection on the subject's visual plane is illustrated in Figure 2 (below). This trajectory was framed within an ideal box sized approximately 10×10×4 centimeters. The entire drawing was covered at approximately piece-wise constant speed in about two seconds, in absence of a resistive grasping of the robotic arm. Reactive forces were exerted by the device to keep the robotic arm on the right trajectory during the task;

- curvilinear movements were obtained by driving the robotic device with sinusoidal forces acting on the 3D spatial domain (x, y, z). With respect to the x direction, the force F_x in Newton at time t was given by

  $$F_x(t) = 5 \cdot \sin\big[\pi S_x\big(x(t) + \mu_x(t)\big)\big],$$

  with $S_x = 10\ \mathrm{m}^{-1}$, and μ_x, included to avoid stops of the robotic arm caused by incidental absence of force along the three directions simultaneously, randomly assuming values between $0.5 \cdot 10^{-3}$ and $5.5 \cdot 10^{-3}$ m across time. An identical relation governed the forces acting respectively on the y and z direction. The arm termination was initially set at the position $(0, 0, 0)$.

  The net result was a rounded three-dimensional trajectory, again occupying a box sized approximately 10×10×4 centimeters. Figure 2 (above) illustrates a possible projection on the xy (corresponding to the visual) plane of these trajectories, whose final realization in any case depended on the force exerted by the user's hand: in particular, the smaller extension along the z direction was consequence of the larger inertia of subjects in following the movement of the robotic arm along depth[1]. Similarly to the rectilinear case, the quasi-periodic cycle was completed in approximately two seconds in presence of a gentle grasping of the robotic arm.

The psychophysics of tactile perception of 3D shapes essentially relies on studies of active exploration and manipulation, and has been extensively covered in the scientific literature. Conversely, the *haptic guidance* of subjects during passive motor tasks has been sparsely studied by psychologists. Feygin and colleagues [11] have measured the performance of subjects in manually reproducing, either with or without use of their vision, a 3D trajectory that had previously been learned under three different conditions: haptic, visual, haptic and visual. Results showed that visual training has an important role for learning the spatial development of a trajectory, while haptic training imprints the memory about its temporal development.

Besides the subjective performance in learning, and then reproducing a 3D trajectory through different modalities — an issue which is relatively important in the context of our experiment — two observations that have been reported as anecdotal by Feygin and colleagues are conversely worth mentioning here: i) among various learning strategies, subjects used verbalization and singing, and ii) under the visual learning condition, at least 7 out of 36 subjects mimicked the displayed trajectory with their hands. Together, these observations reveal the existence of a cross-modal association between the visual perception of a trajectory and the subjective production of movement, as well as words and structured (i.e., musical) sounds.

---

[1] The force feedback provided by the Phantom Omni amounts in any case to few Newtons, and can be easily overwhelmed by an intentional reaction of the user.

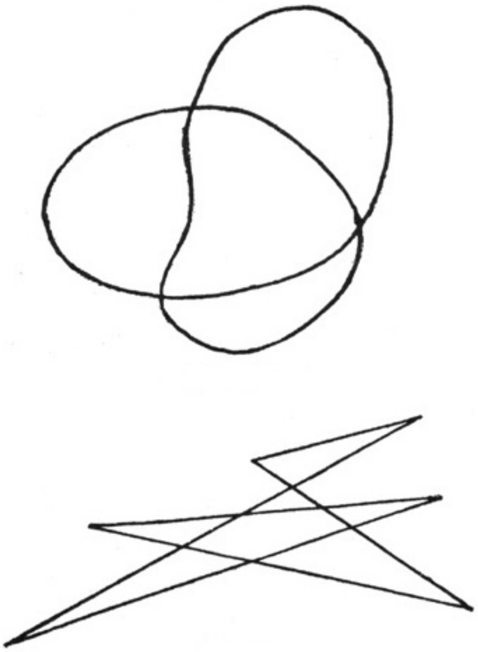


Fig. 2. Trajectories used in the experiment

2.3 Procedure

Subjects were asked to sit in front of the Phantom Omni, and to grasp the stylus with their preferred hand as if it were a normal pen. Then, they were suggested to keep their arm and wrist relaxed while holding it. At this point they were made aware of the gray and black buttons existing at finger reach on the stylus of the device, and were instructed to use them respectively to trigger the beginning of a motion pattern and to switch between two different trajectories of the robotic arm when the device was idle. Finally, subjects were blindfolded and the test started:

1. during the training phase, under the supervision of the experimenter subjects got accustomed with either trajectory by freely triggering as well as swapping them through the buttons. This phase took approximately two minutes;
2. once the training was over, before the evaluation task the experimenter asked them to remember the words "takete" and "maluma".[2] Then, to autonomously repeat the previous phase and try to associate either word to the corresponding trajectory;

---

[2] In Italian, the vowel "e" sounds like in the English word "fence"; "maluma" sounds almost the same as in English.

3. during the evaluation task, subjects were left free to trigger and swap between as many instances of the two trajectories as they wished, until they came up with their decision. The decision concluded the test.

The time to perform the test (training plus evaluation) amounted to about five minutes.

2.4 Results

Table 1 lists the decisions, subject by subject. The polarization of the evaluation is evident. A two-tailed z-test of a single population proportion X, with $X \in \mathcal{N}(0.5, \sqrt{\frac{0.5 \cdot 0.5}{11}})$ the percentage of subjects associating the jagged trajectory to "takete", yields a p-value equal to 0.035.

Table 1. Subjects' decision

Subject no.	Sex	Previous knowledge	Decision on jagged trajectory
1	male	no	maluma
2	female	no	takete
3	female	no	maluma
4	male	no	takete
5	male	yes	takete
6	male	no	takete
7	male	no	takete
8	male	no	takete
9	female	no	takete
10	female	no	takete
11	female	no	takete

3 Discussion

The percentage of subjects associating the jagged trajectory to "takete" confirms the existence of a cross-modal effect between the haptic task and the words used in the experiment. However, its strength is less pronounced than the effect found by Köhler, or by Ramachandran and Hubbard using "kiki" and "bouba" in connection with visual shapes.

The decreased strength may be due to the minor sharpness of the angles displayed between two rectilinear movements by the haptic device at hand, due to its performance limits. These limits, in other words, may result in smoothed reproductions of the jagged trajectory compared to its geometrically ideal visual counterpart which, for instance, can be drawn on a piece of paper. As we mentioned before, it is easy to distort the haptic feedback provided by the Phantom Omni, just by forcing its arm to travel on a different trajectory. Conversely, the variability of the rounded trajectories should have no impact in the results, provided also the existence in the literature of successful studies making use of different types of rounded shapes [1].

Now, one may think to resort to more powerful robotic devices: on the one hand they guarantee increased robustness in reproducing a jagged trajectory; on the other hand, their motors are almost inevitably noisy. Specifically, their noise characterizes completely either trajectory: stationary when they reproduce rounded patterns and, conversely, spiky every time a piece of rectilinear trajectory is switched to another while reproducing a jagged pattern. The Phantom Omni in this sense is almost completely silent, and its residual noises should not have interfered with the task. Future experiments may opt for the use of more powerful robotic arms, once guaranteeing that subjects are not biased by the auditory feedback coming from the haptic device.

Alternatively to a robotic device, subjects during the task may have been guided by another person, or may have been asked to follow a contour by autonomously navigating around a path. By maximizing the potential of the temporal description of a shape, both such different user experiences might have led to stronger association effects in subjects: these and other alternatives to the proposed implementation are worth addressing in future experiments.

It would also be interesting to know if providing the words as printed, rather than verbally, would lead to different results. Such an experimental design, however, would significantly differ from the present one only if preliminary making sure that subjects are able to recall the words visually, and not verbally. Furthermore, the respective design should let the blindfolded subjects have a look to the printed words each time they need to recall them during the test, hence introducing an occasional visual component potentially lowering the level of control in the experiment.

The use of haptic instead of visual feedback opens an interesting question, which has been left aside by the previous discussions about the takete/maluma experiment based on vision: do subjects associate shapes or trajectories to words? In other words, do subjects reconduct visual contour patterns to the corresponding shapes that they delimit, or rather to an image of the motor activity producing the contours themselves? The latter hypothesis, which may be supported by the results of the proposed experiment and, at least partially, by the observations of Feygin and colleagues reported in Sec. 2.2, would require to extend the speculations of Ramachandran and Hubbard, who connect shapes to words mainly through a spatial analogy while giving less emphasis to the temporal aspects.

Recent studies reporting about fMRI-based analyses of the activity occurring in specific areas of the human brain, are progressively shedding light on the functional roles of the lateral occipital complex and the intraparietal sulcus of the brain, in decoding visual and haptic 3D shape information incoming from the periphery of the nervous system, and in the integration of this information with object representations stored in the frontoparietal regions [12]. In particular, spatial properties of objects such as size, shape, and relative position, seem to be encoded into the lateral occipital complex in a modality-independent format, as opposed to pictorial properties such as color and texture. The existence of a common representational system would find confirmation in the individual ability to explore objects either visually or haptically, with comparable spatial learning performances.

Specifically, the lateral occipito-temporal sulcus has been supposed to contain a spatial representation of objects, that is accessed from both the periphery and the memory. The former access occurs independently of the sensory modality. However, the process is modulated by the object familiarity: particularly during the haptic exploration of unfamiliar objects, the activity in the lateral occipito-temporal sulcus is largely disconnected from the frontoparietal regions, and the recognition is then mostly supported by local spatial imagery processes: in other words, the haptic recognition of unfamiliar objects would receive little or no support from visual-haptic analogies existing in the memory, until some spatial imagery is consolidated in the lateral occipital complex. Interestingly for our discussion around takete/maluma, these conclusions have been drawn in experiments where participants listened to pairs of words, and decided whether the objects designated by such words had similar shape or not; then, in a separate session the same participants performed a haptic discrimination task of familiar and unfamiliar objects [12].

Unfortunately, these functional experiments could not measure any temporal aspect of the decoding process taking place in the lateral occipito-temporal sulcus during the haptic recognition: especially in the case of the exploration of unfamiliar objects, any measure of this kind could have informed our discussion with substantial arguments. Precious information could have additionally been drawn from our experiment as well, had we recorded quantitative data about the temporal process (such as deviations from the desired trajectory and jitter on task completion times caused by reactive forces exerted by subjects) during the haptic recognition task. Especially for this reason, we hope to collect such data in a future instance of the experiment, as well as to repeat the task with congenitally blind participants: for these subjects, in fact, any mental imagery of spatial features of objects results exclusively by a haptic experience of 3D shapes.

As a final consideration, the proposed experiment is likely to target a subset of the existing connections between spatial ability and haptic feedback, and hence have limited scope. Words like “takete” and “maluma” may in fact be significantly associated to temporal cues going beyond their instantiation into simple geometrical shapes. Furthermore, as for any forced choice, the two words can probably be associated to different pairs of qualities using whatever modality: this fact calls for an even more cautious use of the results.

All this said, the outcome of the proposed experiment suggests the existence of a robust decoding of unfamiliar haptically-explored trajectories from users gifted with normal proprioception and sight. In absence of this decoding, any connectivity with words such as “takete” and “maluma” should in fact be absent. This conclusion implies that it is important, for interaction designers, to avoid counter-intuitive associations between haptically-displayed shapes and their labels: similarly to what has been recommended in the case of cross-modal associations between tastes and words [3], this concern may prevent from the design of "haptons" with an inherently odd verbal semantics and related meaning.

Acknowledgments. This research has been hosted and funded by the Department of Computer Science of the University of Verona. The author acknowledges the software implementation work of Davide Conigliaro, Marco Schivi and Pietro Lovato, attending the course of Non-visual interaction in the same Department at the time when this research was conducted. One anonymous manuscript reviewer has provided especially constructive criticism, that helped the author discuss better the limits of the hypothesis and consequent experimental design in the related section.

References

1. Maurer, D., Pathman, T., Mondloch, C.: The shape of boubas: sound-shape correspondences in toddlers and adults. Developmental Science 9, 316–322 (2006)
2. Ramachandran, V., Hubbard, E.: Synaesthesia: A window into perception, thought and language. J. of Counsciousness Studies 8, 3–34 (2001)
3. Crisinel, A.S., Jones, S., Spence, C.: 'The sweet taste of maluma': Crossmodal associations between tastes and words. Chemical Perception 5, 266–273 (2012)
4. Gallace, A., Boschin, E., Spence, C.: On the taste of "bouba" and "kiki": An exploration of word-food associations in neurologically normal participants. Cognitive Neuroscience 2, 34–46 (2011)
5. Crossan, A., Brewster, S.: Multimodal trajectory playback for teaching shape information and trajectories to visually impaired computer users. ACM Trans. Access. Comput. 1, 12:1–12:34 (2008)
6. Rassmus-Grohn, K., Magnusson, C., Eftring, H.: Ahead - audio-haptic drawing editor and explorer for education. In: IEEE International Workshop on Haptic, Audio and Visual Environments and Games, HAVE 2007, pp. 62–66 (2007)
7. Tanhua-Piiroinen, E., Pasto, V., Raisamo, R., Sallnäs, E.L.: Supporting collaboration between visually impaired and sighted children in a multimodal learning environment. In: Pirhonen, A., Brewster, S. (eds.) HAID 2008. LNCS, vol. 5270, pp. 11–20. Springer, Heidelberg (2008)
8. Hancock, D.: An interactive computer graphics system with force feedback display. Technical report, Hewlett Packard Labs., San Mateo, CA (1996), `http://dennishancock.com/haptics/hpfinal.html`
9. Brewster, S.A., Brown, L.M.: Non-visual information display using tactons. In: CHI 2004 Extended Abstracts on Human Factors in Computing Systems. CHI EA 2004, pp. 787–788. ACM, New York (2004)
10. MacLean, K., Enriquez, M.: Perceptual design of haptic icons. In: Proc. of Euro-Haptics, Dublin, UK (2003)
11. Feygin, D., Keehner, M., Tendick, F.: Haptic guidance: Experimental evaluation of a haptic training method for a perceptual motor skill. In: Proc. 10th Symposium On Haptic Interfaces For Virtual Environments and Teleoperator Systems (HAPTICS 2002). IEEE (2002)
12. Lacey, S., Tal, N., Amedi, A., Sathian, K.: A putative model of multisensory object representation. Brain Topography 21, 269–274 (2009)

Comparison of Statistical Methods for the Analysis of Affective Haptic Expressions

Yoren Gaffary[1], Victoria Eyharabide[2], Mohamed Yacine Tsalamlal[1],
Jean-Claude Martin[1], and Mehdi Ammi[1]

[1] CNRS/LIMSI, Bâtiment 508, Université de Paris-Sud, 91403 Orsay Cedex, France
`firstname.lastname@limsi.fr`
[2] STIH, Université Paris-Sorbonne, 28 rue Serpente, 75006 Paris, France
`maria-victoria.eyharabide@paris-sorbonne.fr`

Abstract. Several studies were conducted to show the relevance of haptics for conveying emotions to users. These studies usually cover recognition rate of emotions from haptic expressions. Surprisingly, the analysis of features of these haptic expressions has been in counterpart often limited to a classical analysis of variance. This method is limited since it can neither highlight multiple possible expressions of a given emotion nor compare several emotions or features simultaneously. This paper presents a methodological approach for collecting and analyzing haptic expressions of emotions. We compare three statistical methods, namely analysis of variance, principal component analysis, and clustering. Over this study we will highlight the advantages and drawbacks of each method for the analysis of haptic expressions of emotions.

Keywords: Emotion, Haptics, Experimental Study, Statistics.

1 Introduction

Emotions play an important role in human-human communication [12]. The capabilities of some modalities as facial expressions to express emotions during human-computer interactions are addressed in multiple studies [13].

Recently, several works have investigated the role of haptics to improve the recognition and discrimination of some emotions expressed with facial expressions [3,2,14,7]. These works were based on the identificati on of discriminative features in haptic expressions for each investigated emotion. These studies have exploited the analysis of variance (ANOVA) as a mainstream statistical method to exhibit these discriminative features. However, classical ANOVA suffers from three main limitations. First, emotions are compared pairwise. It is thus not possible to compare simultaneously more than two emotions. Second, features of haptic expressions are considered one at a time and it is not possible to identify the correlations between these features. Third, these studies do not focus on multiple ways to express a given emotion, while studies in other modalities have observed and suggest multiple and different expressions of the same emotion [5].

I. Oakley and S. Brewster (Eds.): HAID 2013, LNCS 7989, pp. 69–78, 2013.

To overcome these limitations of ANOVA, we have explored its complementarity with two other statistical methods. The first method is based on the Principal Components Analysis (PCA) [10] which allows the highlighting of correlations between features. The second method is based on the clustering analysis (using the Expectation-Maximization (EM) algorithm [6]) to analyze multiple groups of emotions in a given set of haptic expressions.

This paper starts with the description of a corpus of haptic expressions that we collected. The next section introduces the results of the three statistical methods applied to the collected data. Finally, we summarize the advantages and drawbacks of each method, and highlight the complementarity between the three methods.

2 HAPTEMO: A Corpus of Affective Haptic Expressions

The first step of this study concerns the creation of a corpus of haptic expressions corresponding to different emotions. This implies i) the selection of a set of emotions, and ii) the definition of an experimental protocol to collect corresponding haptic expressions. We propose to study pairs of close emotions in order to identify haptic features that enable an efficient recognition and discrimination.

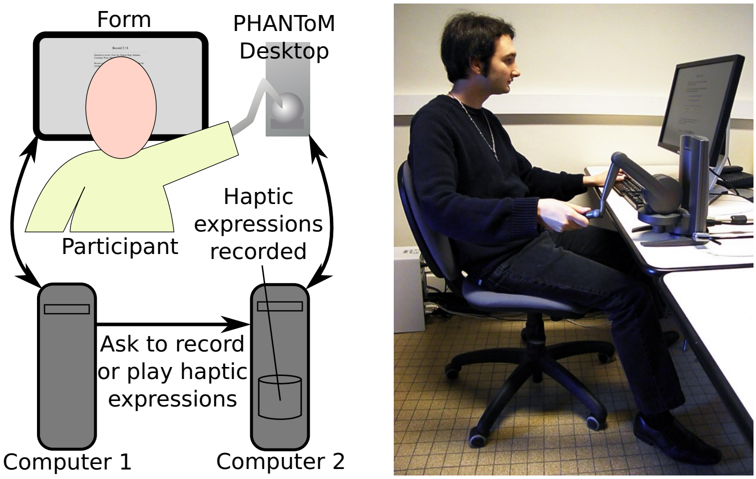
The dimensional representation of emotions suggests that emotions can be represented using three continuous and orthogonal dimensions : Pleasure (degree of well-being), Arousal (degree of mental or physical activity) and Dominance (degree of control of a situation) [11].

This dimensional approach enables to compute a distance between two emotions and compare features expressing emotions that are either close or far from each other. Thus, we have selected the following emotions (according to the PAD axes as observed in a previous study [11]): "Joy", "Elation", "Disgust", "Contempt", "Anxiety","Fear", "Irritation" and "Rage".

Multiple protocols, using realistic acted or spontaneous expressions, have been defined for collecting expressions of emotions in several modalities. We propose to use the acted approach since we want to analyze haptic expressions, and their related features, that conveys without ambiguity a single emotion to users.

2.1 Experimental Platform

The experimental platform is based on a PHANToM Desktop haptic arm. This device enables recording and rendering of 3D Kinesthetic expressions. The platform includes two computers (see Fig. 1). The main computer displays instructions to users with a screen and process keyboard's inputs. It also includes the UDP protocol to support the communication with the second computer which records and render haptic expressions. This configuration ensures a better stability for haptic rendering.



(a) experimental set-up.

(b) a participant interacting with the platform.

Fig. 1. Experimental platform to collect haptic expressions of emotions

2.2 Participants

Forty subjects (eight women, thirty-two men), aged between twenty and fifty-three, thirty-one average age (SD = 8) participated in this experiment. We did not analyze the influence of gender, handedness or education on the collected haptic expressions. This is due to the high majority of European, right-handed (thirty-three subjects were right-handed, and thirty-five received a European education) and males.

2.3 Procedure

The experimental procedure includes the three following steps :

Step 1. Participants have filled a form asking for their age, gender, dominant hand, if they have already used an haptic device, and their cultural education.

Step 2. They had five minutes of training session, during which participants were asked to explore the workspace of the haptic device and to express an emotion that was not in the investigated set (”Surprise”).

Step 3. Once this training session was finished, participants were asked to express each of the eight emotions (”Joy”, ”Elation”, ”Disgust”, ”Contempt”, ”Anxiety”, ”Fear”, ”Irritation” and ”Rage”). The order of presentation of the emotional labels was counterbalanced across subjects. A textual description of a relevant emotional situation, selected from the MindReading database [8], was displayed with each emotional label in order to ensure a common understanding of the meaning of the emotional labels.

Participants had ten seconds to express the requested emotion using the haptic device. We have asked them to hold the stylus of the haptic device as if it was the hand of somebody else. Subjects freely moved their hand, by holding the stylus, inside the workspace of the haptic device. They had only one trial for each emotion in order to collect spontaneous haptic expressions. Afterwards, the recorded haptic expression was rendered to the participant with the same haptic device. Then, the participant had to assess, via a seven point Likert scale, his level of confidence about the expressed emotion.

Forty haptic expressions corresponding to forty subjects were collected for each of the eight emotions. The HAPTEMO corpus is thus made of $40 \times 8 = 320$ haptic expressions of emotions.

2.4 Measures

For each collected haptic expression, we have computed several measures derived from studies investigating haptic and gestural affective expressions [1,4]:

- **M1** *Distance*: overall traveled distance by the participant's hand (end-effector of the haptic device) between the beginning and the end of the haptic expression.
- **M2** *Mean speed*: average speed of the participant's hand.
- **M3** *Fluidity*: degree of suppleness of the expression given by the following equation: $\frac{\sum_{t=0}^{duration} |a(t+1)-a(t)|)}{duration}$. Where $a(t)$ corresponds to the acceleration at t ($(t+1)-t = \Delta t \approx 1$ ms), and *duration* corresponds to the overall duration of the expression. Notice that a low level for this measure corresponds to movements with a high fluidity.
- **M4** *Amplitude*: distance between the two farthest corners of the bounding box containing the haptic expression.
- **M5** *Expansion Index*: degree of expansion of the haptic expression given by the following equation: $\frac{\sum_{t=0}^{duration} d(p(t),\text{isobar})}{duration}$. Where $d(p_1, p_2)$ corresponds to the distance between positions p_1 and p_2, $p(t)$ corresponds to the position of the end-effector at t ($(t+1)-t = \Delta t \approx 1$ ms), *isobar* corresponds to the isobarycenter of the expression, and *duration* corresponds to the duration of the expression.
- **M6** *Duration*: overall duration of the haptic expression.
- **M7** *Major Axis (X coordinate, Y coordinate, Z coordinate)*: major axis of the gesture, computed with a Singular Value Decomposition (SVD) [9].
- **M8** *Weight of Major Axis*: prevalence of the major axis on the movement (based on SVD).
- **M9** *Weight of Second Major Axis*: prevalence of the second major axis on the movement (based on SVD).
- **M10** *Repetitivity*: estimation of the repetitive phases of the expression. This is obtained by calculating the barycenter of the haptic expression and the major axis. We use the projection of each point of the trajectory on this axis and the barycenter and count 1 repetition each time the barycenter is crossed two times by the projection.

Measures M1 to M5 are also computed using a single axis of movement (left-right, up-down and depth). This leads to a total of 27 measures.

In addition to these objective measures, we propose a evaluate the level of confidence reported by the participant, for each haptic expression (a 7-points Likert scale). It indicates if the participant thinks his/her expression expresses well the targeted emotion.

3 Analyzes of the Haptic Expressions

The goal of this study is to highlight the similarities and differences in expressions of different emotions. We have investigated three statistical methods: an ANOVA, a PCA and an EM-cluster analysis. For each method, we have highlighted the main advantages and drawbacks.

Before the analysis, we filtered the corpus to keep only haptic expressions which present a positive level of confidence (minimum 5/7). For the rest of this paper, the term "haptic expressions" refers to the 194 expressions out of a total 320 expressions (equal to 60% of the total, corresponding between 47% and 75% of expressions collected for each emotion) that fulfilled this criterion.

3.1 Analysis of Variance

The commonly used ANOVA method enables to identify differences between two emotions according to a given measure. The previous filtering step provides sets with different numbers of haptic expressions for each emotion. Thus, we used a Wilcoxon test that enables the comparison of populations with different sizes.

The Wilcoxon test was applied to each quantitative measure of two compared emotions. Table 1 summarizes the number of measures presenting a significant difference ($p < 0.05$) between two emotions. This table shows that some pairs of emotions (bold values) present more differences than the mean number of differences (mean of 9.4). This means that those pairs of emotions are statistically more different using the identified measures. For instance, "Elation" and "Disgust" present many significant differences for the following measures: M1 (component of movement along the up-down axis), M2 (along up-down and depth axes), M3 (along up-down and depth axes), M4 (along up-down axis) and M7 (along all three axes).

Advantage of ANOVA. The ANOVA expresses the level of difference (i.e., significant difference or non significant difference) between two emotions according to a given dimension. If there is a significant difference between measured values for two different emotions, we consider that the mean value of this measure for each emotion is a discriminative feature for this pair of emotions.

Drawbacks of ANOVA. The ANOVA approach presents two limitations. First, ANOVA can not be applied to non-homogeneous populations, and can not deal with subpopulations). For instance, some emotions can be expressed, by some

Table 1. Number of measures presenting significant differences when applied to two emotions. The number of similarities can be obtained by calculating the difference between the total number of measures (i.e., 27 measures) and the number of measures presenting significant differences. Bold and italic values correspond to pairs of emotions presenting more and less differences respectively than the overall mean number of differences (mean of 9.4 measures).

	Elation	**Disgust**	**Contempt**	**Anxiety**	**Fear**	**Irritation**	**Rage**
Joy	*8*	*7*	**19**	**16**	**14**	**17**	*5*
Elation	↳	**12**	**14**	**15**	**15**	**21**	*2*
Disgust	-	↳	*1*	*5*	*3*	*9*	*9*
Contempt	-	-	↳	*1*	*4*	*5*	**13**
Anxiety	-	-	-	↳	*1*	*1*	**11**
Fear	-	-	-	-	↳	*5*	**11**
Irritation	-	-	-	-	-	↳	**18**

participants, with a first category of movements (e.g., slow movements), while the rest of the participants express the same emotion with a different category of movement (e.g., fast movements). In this case, the ANOVA can not find statistical differences between emotions, even if the different populations of movement include specific features for the same emotions (e.g. elation is expressed with two categories of movements: vertical and horizontal movements).

The second limitation of ANOVA concerns the number of emotions compared, which is limited to two emotions at a time. If a multivariate ANOVA can correct this (if the measures are not too correlated), it remains not possible to simultaneously compare a given emotion to several other emotions. For instance, it is irrelevant to compare the speed of expressions of "Irritation" and "Rage" simultaneously to the speed of "Joy" by mixing the expressions of the first two emotions for the ANOVA. Indeed, there is a huge difference between the speed of expressions between "Irritation" and "Rage" (means of $= 0.16\text{m.s}^{-1}$ and $= 0.33\text{m.s}^{-1}$ respectively).

3.2 Principal Components Analysis

Compared to ANOVA, a PCA simultaneously deals with all quantitative measures and emotions. Besides, it also highlights linear correlations between linear measures. For instance, it highlighted in our data set an inverse correlation between the weight of major and second axes of movements. High weight for the major axis implies low weight for the second major axis. Using these correlations, a PCA can reduce the dimensionality of the data set by mixing elementary correlated measures in new axes called factorial axes. This operation simplifies the set to keep only useful information to describe the data set.

In a second step, we applied a PCA to keep only the two main factorial axes. These two factorial axes cover 55% of the whole information contained in all quantitative measures. This revealed that barycenters of "Joy", "Elation" and

Table 2. Distance between the barycenters of each pair of emotion category in the space computed by the PCA. Distances higher than the mean are displayed in bold.

	Elation	**Disgust**	**Contempt**	**Anxiety**	**Fear**	**Irritation**	**Rage**
Joy	**1.629**	1.414	1.431	**1.624**	**1.734**	**1.633**	**1.822**
Elation	↳	**1.740**	**1.961**	**1.528**	**1.905**	**1.698**	**1.483**
Disgust	-	↳	1.139	1.045	1.138	1.304	1.422
Contempt	-	-	↳	1.091	1.199	1.220	**1.740**
Anxiety	-	-	-	↳	1.241	0.915	**1.457**
Fear	-	-	-	-	↳	1.055	**1.463**
Irritation	-	-	-	-	-	↳	**1.528**

”Rage” are far from the barycenters of other emotions. This corroborates the results obtained by the ANOVA, which displayed many significant differences for these emotions compared to other emotions.

Advantages of PCA. The PCA presents two advantages compared to ANOVA. First, it enables the computation of explicit Euclidean distances between the barycenters of emotions in the space provided by the PCA (i.e., with uncorrelated dimensions, see Table 2). This distance, including all factorial axes, explicitly determines how much the haptic expressions of different emotions are close.

The second contribution of the PCA is the highlighting of subpopulations of haptic expressions for a given emotion. The haptic expressions for a given emotion might not concentrate on a single point but spreads across several ones.

Drawbacks of PCA. The PCA approach presents also two limitations. First, a PCA creates factorial axes which are the combination of different measures. Thus, this method must deal with at least two different measures. For example, contrary to ANOVA, PCA can not deal with only the measure of duration.

The second limitation of PCA is that it can not process the highlighted subpopulations (i.e. multiple expressions for a given emotion). For example, the two main factorial axes revealed two main ways used to express for ”Rage” : one is near from expressions of other negative emotions, while one is isolated from all other emotions. But we need an external algorithm to determine to which set each expression belongs to.

3.3 EM-Cluster Analysis

Considering the limitations of PCA, we investigate here a clustering approach that enables the identification of subsets in the same group of data. This method groups in clusters haptic expressions presenting similar features (i.e. similar on most values of our measures), regardless of the emotion labels.

We decided to use the EM algorithm implemented in the Weka platform for clustering since this algorithm enables the estimation of the optimal number of clusters from the data. In our case, this number is not known a priori.

The clustering results of the haptic expressions in the HAPTEMO data set are displayed in tables 3 and 4. Table 3 shows the percentage of an emotion in a cluster against the total number of emotions in this cluster. Table 4 shows the distribution of each emotion across the different clusters.

Table 3. Percentage of emotions per cluster. The most representative emotions in each cluster are highlighted.

	clust. #0	**clust. #1**	**clust. #2**	**clust. #3**	**clust. #4**	**clust. #5**
Joy	5%	11%	11%	14%	**29%**	0%
Elation	0%	11%	**32%**	14%	**18%**	9%
Disgust	0%	**18%**	18%	9%	6%	0%
Contempt	16%	**20%**	4%	0%	12%	11%
Anxiety	**21%**	5%	18%	5%	6%	14%
Fear	16%	**20%**	11%	9%	9%	14%
Irritation	**32%**	5%	4%	5%	6%	**34%**
Rage	11%	9%	4%	**45%**	15%	17%
TOTAL	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

This technique produced six different clusters for the eight emotions. In Table 3, the most representative emotion of each cluster is highlighted (white font, dark background). The predominance of an emotion in a cluster means that the haptic expressions of that emotion differ from those of other emotions. For instance, Table 3 shows that cluster #4 includes mainly expressions of positive emotions (”Joy” and ”Elation”). However, cluster #2 mainly includes expressions of ”Elation”. This means that cluster #2 represents haptic expressions that are specific to ”Elation”, while cluster #4 includes haptic expressions of ”Elation” that are closer to expressions of the ”Joy” emotion.

Table 4. Percentage of clusters by emotion. The most representative clusters of each emotion are highlighted.

	clust. #0	**clust. #1**	**clust. #2**	**clust. #3**	**clust. #4**	**clust. #5**	TOT.
Joy	4%	26%	13%	13%	**43%**	0%	**100%**
Elation	0%	22%	**33%**	11%	22%	11%	**100%**
Disgust	0%	**53%**	26%	11%	11%	0%	**100%**
Contempt	13%	**48%**	4%	0%	17%	17%	**100%**
Anxiety	20 %	15%	**25%**	5%	10%	**25%**	**100%**
Fear	11%	**41%**	11%	7%	11%	19%	**100%**
Irritation	24 %	12%	4%	4%	8%	**48%**	**100%**
Rage	7%	17%	3%	**34%**	17%	21%	**100%**

In contrast, some clusters do not present a dominant emotion (i.e., they contain haptic expressions of several emotions). This means that the mixed emotions in that cluster display similarities in their haptic expressions. This is the case of cluster #1 that contains expressions of ”Disgust”, ”Contempt” and ”Fear”.

Table 5. Advantages and limitations of each method we studied in this paper

Method	Advantages	Limitations
ANOVA	• Gives specific features for two emotions if there is a significant difference between these features.	• Can not highlight non-homogeneous populations. • Can not compare groups of emotions
PCA	• Determines distances of several emotions explicitely. • Highlights subpopulations of haptic expressions for a given emotion.	• Deals with at least two measures. • Can not process subpopulations.
EM	• Analyze simultaneously several haptic expressions and measures. • Determine specific haptic expressions for a given emotion.	• Can not identify features of emotions that are not dominant in any cluster.

Advantages of Clustering. Clustering presents two advantages. First, this approach simultaneously analyzes a large number of haptic expressions and measures. However, an extraction of the principal features from data (as correlations between measures) would improve the EM procedure.

Second, clustering enables the determination of specific haptic expressions characterizing a given emotion. EM is based on the similarities between haptic expressions and not on the emotion labels to identify the clusters. This explains that a cluster could include a mix of haptic expressions corresponding to different emotions without the prevalence of one emotion. In contrast, the dominance of expressions of a single emotion can correctly describe this emotion.

Drawbacks of Clustering. The main limitation of the clustering approach is the difficulty to identify emotions that are not dominant in any cluster. In that case, it is impossible to determine the main features of corresponding haptic expressions. There are two emotions which are not represented by a single cluster: "Disgust" and "Anxiety". In order to detect which cluster is the most representative of these emotions, it is helpful to calculate the percentage of these emotions in each cluster. Table 4 displays those percentages. On one hand, we observe that 53% of haptic expressions of "Disgust" are in cluster #1 which is the cluster for "Contempt" and "Fear". On the other hand, we found that "Anxiety" is spread in cluster #2 and cluster #5 (25% in each one of these two clusters). This means that the algorithm could not find a clear pattern for the expressions of this emotion. Besides, these two emotions display the two lowest quantities of haptic expressions: only nineteen and twenty expressions respectively. Few participants were confident in the way they expressed these two emotions. This could also explain why a cluster for these emotions could not be found.

4 Conclusion

This paper has compared three statistical methods to determine similarities and differences between haptic expressions corresponding to different emotions.

These methods have produced different but complementary results. Table 5 summarizes the advantages and drawbacks of each method.

It could be interesting to combine those methods to overcome their individual limitations. For instance, the PCA is an efficient pre-processing before applying the EM-cluster analysis as it is designed to extract the principal features of a data set. Then, by considering the resulting clusters, the ANOVA analysis (or a derivative method) should be more efficient than a single ANOVA to highlight differences between clusters for the examined measures.

Future works consist in collecting more spontaneous haptic expressions of emotions and comparing them with the acted data described in this article. Features extracted could be used to recognize emotions from users's haptic expressions, or to synthesize haptic expressions for the eight investigated emotions.

References

1. Bailenson, J.N., Yee, N., Brave, S., Merget, D., Koslow, D.: Virtual Interpersonal Touch: Expressing and Recognizing Emotions Through Haptic Devices. Human-Computer Interaction 22(3), 325–353 (2007)
2. Bickmore, T.W., Fernando, R., Ring, L., Schulman, D.: Empathic Touch by Relational Agents. IEEE Transactions of Affective Computing 1(1), 60–71 (2010)
3. Bonnet, D., Ammi, M., Martin, J.C.: Improvement of the recognition of facial expressions with haptic feedback. In: IEEE Haptic Audio Visual Environments and Games International Workshop, Hebei, China, pp. 81–87 (October 2011)
4. Castellano, G.: Movement expressivity analysis in affective computers: from recognition to expression of emotion. Ph.D. thesis, University of Genoa, Italy (2008)
5. Dael, N., Mortillaro, M., Scherer, K.R.: Emotion Expression in Body Action and Posture. Emotion 12(5), 1085–1101 (2011)
6. Dempster, A.P., Laird, N.M., Rubin, D.B.: Maximum Likelihood from Incomplete Data via the EM algorithm. Journal of the Royal Statistical Society. Series B (Methodological) 39(1), 1–38 (1977)
7. Gaffary, Y., Eyharabide, V., Martin, J.C., Ammi, M.: Comparison of Statistical Methods for Analysis of Affective Haptic Expressions. In: ACM Symposium on Applied Perception, Los Angeles, California, U.S.A., p. 128 (2012)
8. Golan, O., Baron-Cohen, S., Hill, J.: The Cambridge Mindreading Face-Voice Battery: Testing Complex Emotion Recognition in Adults with and without Asperger Syndrome. Journal of Autism and Developmental Disorders 36(2), 169–183 (2006)
9. Klema, V., Laub, A.: The singular value decomposition: Its computation and some applications. IEEE Transactions on Automatic Control 25(2), 164–176 (1980)
10. Pearson, K.: LIII. On lines and planes of closest fit to systems of points in space. Philosophical Magazine Series 6 2(11), 559–572 (1901)
11. Russell, J.A., Mehrabian, A.: Evidence for a three-factor theory of emotions. Journal of Research in Personality 11(3), 273–294 (1977)
12. Scherer, K.R.: What are emotions? And how can they be measured? Social Science Information 44(4), 695–729 (2005)
13. Surakka, V., Hietanen, J.K.: Facial and emotional reactions to Duchenne and non-Duchenne smiles.. International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology 29(1), 23–33 (1998)
14. Tsalamlal, M.Y., Ouarti, N., Ammi, M.: Psychophysical study of air jet based tactile stimulation. In: to appear in IEEE World Haptics, Daejon, Korea (April 2013)

Plucked String Stiffness Affects Loudness Perception

Mounia Ziat[1], Ilja Frissen[2], Gianni Campion[3], Vincent Hayward[4],
and Catherine Guastavino[2]

[1] Department of Psychology, Northern Michigan University, Marquette, MI, USA
[2] McGill University & Centre for Interdisciplinary Research on Music,
Media and Technology, Montreal, QC, Canada
[3] McGill University, Montreal, QC, Canada
[4] UPMC Univ Paris 06, Institut des Systèmes Intelligents et de Robotique, Paris, France

Abstract. A great variety of interactions between senses, and between motor production and senses, have been reported in previous research. In the present study, we tested whether the mechanics of a plucked string affected how the sound it produced was perceived. To test this hypothesis, we simulated the feel of a plucked string using a high fidelity haptic force-feedback device and simultaneously simulated its acoustic emission. This way, we could independently manipulate the two sensory inputs — how it felt and how it sounded — together with physical correct haptic interaction and with accurate synchronization. This arrangement makes it very plausible that the two sensory inputs came from a common source. We used a two-interval forced-choice discrimination procedure to determine the point of subjective equality of the loudness between a stiff and a soft plucked string. When the stiffness of the string was low, the sound was perceived to be softer. Interestingly, this effect was found only when the first string was less stiff than the second string plucked during a comparison. The results are consistent with the inverse effectiveness principle of multisensory integration.

Keywords: loudness perception, haptic stiffness, auditory-tactile integration.

1 Introduction

Playing a musical instrument is a sensorimotor task. During this activity, it is expected that numerous interactions take place between the sensory channels, touch, vision, audition, and motor control. Several studies explored the impact of the auditory feedback on haptic perception during musical performance and how performance can be affected by haptic feedback from the instrument [1-3]. Most of these studies focused on timing and motor performance [4-5], rather than on how the produced sound is perceived by the player. Other studies found that touch was affected by audition [6-8], although not necessarily in a musical context. Very few studies investigated the effect of touch on auditory perception, despite the fact that the haptic feedback provided by the instrument is always present when playing an instrument. There is always some physical contact, even if the sensory-motor loop can be very loose as in the case of electronic or computer-based instruments. A well known and

I. Oakley and S. Brewster (Eds.): HAID 2013, LNCS 7989, pp. 79–88, 2013.

frequently discussed example is the case of the digital piano. The sound may be accurately synthesized, but the feel of the keyboard remains an issue for many performers. It is permitted to think that the light touch of the electronic keyboard, compared to the complicated dynamics of the escapement mechanism of a traditional piano, may affect how the sound is perceived by the player. Here we are interested in characterizing such as effect in the case of a plucked string, the feel of which is strongly dependent on its tension at rest.

Another hint of the importance of the link that exists between the feel given by an instrument and its acoustic qualities can be found in the degree of inharmonicity of the sound of a plucked string. Inharmonicity, i.e., the degree to which overtone frequencies deviate from perfect integral multiples of the fundamental is known to add "warmth" to the sound [9]. The stiffer the strings are, the more inharmonicity they exhibit [10], which, in turn, affects loudness perception. This and other strong correlations of this type between the feel and the sound may be internalized by the auditory systems of performers and listeners alike as perceptual invariants.

A possible outcome of our research could be suggestions aimed at improving the playability of digital instruments, which are often commented to provide a impoverished user interface for the player, compared to traditional instruments. Researchers suggested several devices to compensate for the quasi-elimination of haptic feedback in digital instruments, for a review see [11].

Here, we focus on the effects of touch on loudness judgment. Previous studies have found that a more intense haptic interaction with the instrument increased the perceived loudness of the produced sounds. For instance, with pianists, the perceived loudness increased exponentially with the force applied on the piano keys [12]. More recently, Okazaki et al. found that complex, broadband auditory stimuli tend to be perceived louder when the same signal was simultaneously heard and felt rather than when it was heard only [13]. Other results corroborated the effect of vibrotactile inputs on loudness perception, such as whole-body vibrations delivered through the chair while listening to a sound [14], or with irrelevant tactile, stimuli [15]. Schürmann et al., found that when tactile vibrations were delivered simultaneously with a sound, the latter was perceived 12% softer than when the sound was played alone [16], suggesting an increase in loudness perception due to the presence of the tactile stimulus.

With the exception of [12], these previous studies typically used musically irrelevant signals delivered in static conditions, that is, eschewing the likely effects of sensori-motor couplings. Besides, participants were not actively creating the vibration that was the source of the sound and of the feel, as in a real instrument. Several recent studies have shown that actively generated movement affect haptic perception [17-18] and in some cases auditory-haptic perception [19].

2 Pilot Study

In the present pilot study, we employed a virtual string paradigm to measure perceived loudness under controlled conditions. A plucked string was simulated with reasonable realism in order to assess participants' loudness perception when the tactile feedback changed as a result of controlled modifications of the string stiffness. We used a range of sound loudness between 35.4 dB and 51.7 dB. In each trail, participants were presented with a pair of sounds, each associated with particular stiffness, and asked to indicate which of two sounded louder. We hypothesized that sounds presented at the same sound level would nevertheless be perceived to be louder when associated with greater stiffness, and vice versa, namely that sounds associated with lower stiffness would be perceived as quieter.

2.1 Participants

Ten right-handed participants from McGill University completed the experiment and received $10 for their participation. None of them had any extensive experience with psychophysical procedures. Participants gave their informed consent before participating. Procedures were in accordance with the guidelines set out in the Declaration of Helsinki and approved by the McGill University research ethics board.

2.2 Apparatus and Stimuli

The main apparatus was a Pantograph, a high-performance haptic device (Fig. 1) developed for rendering virtual surfaces [20]. The Pantograph can produce forces of up to 2 N in a two-dimensional workspace of 100 × 60 mm and has a flat response from DC to 400 Hz. The torque commands were processed by a low pass reconstruction filter, so that the commands to the motors matched the mechanical bandwidth of the system.

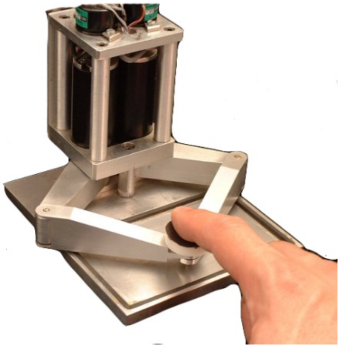


Fig. 1. The Pantograph: A planar parallel mechanism with a nonslip plate on which the finger pad rests

The auditory stimulus was generated using two delay lines following a variation of the well-known Karplus–Strong string synthesis algorithm [21]. The audio output was the sum of the two samples coming out of the delay lines at each sample of time. These two delay lines were terminated by low-pass Butterworth filters. The first delay line had a second order filter set at 500 Hz cut-off frequency, a depth of 50 samples, and a decay of 0.995. The second delay line had a 10th order filter with a cut-off at

400 Hz, a depth of 70 samples and a decay of 0.995. The digital simulation ran at a sampling frequency of 20 kHz. The input to the algorithm was an impulse triggered at the instant of release of the virtual string tension. The mono audio stimulus was presented to the two ears over sound-isolated headphones (Direct Sound EX-29).

The subject was instructed to move the plate of the haptic device from left to right. Two strings were simulated side by side. The direction of the force returned by the device was always oriented to the left. The force intensity, $|f|$, increased from a rest position, x_0, up to a maximum value, x_{max}, at which point it was set to zero discontinuously to simulate the plucking effect. The simulation was repeated for the second string. For a stiffness, K,

$$|f| = \begin{cases} 0 & \text{if } x < x_0 \\ K(x - x_0) & \text{if } x_0 < x < (x_0 + x_{max}) \\ 0 & \text{if } (x_0 + x_{max})\, x < x_1 \\ K(x - x_1) & \text{if } x_1 < x < (x_1 + x_{max}) \\ 0 & \text{otherwise} \end{cases}$$

where x_0 and x_1 are the positions of the strings in the workspace and $x_1 > (x_0 + x_{max})$.

2.3 Procedure

After reading the instructions and having given their informed consent, the participants placed their right index finger on the pantograph's plate interface. Throughout the experiment, their arm was comfortably supported by a soft gel pack placed near the right elbow. The participants entered their answers on a standard keyboard using their left hand.

The participants engaged in an unspeeded, two-interval force-choice (2-IFC) auditory loudness discrimination task. On each trial the participant plucked the virtual string two times in succession and heard the corresponding auditory stimulus. In one of the intervals a standard (i.e., reference) stimulus was presented and in the other a comparison stimulus, and the task was to judge which one of the two strings sounded louder.

We manipulated the loudness of the auditory stimulus as well as the stiffness of the string. Two standards were created by combining a string stiffness of 4 N/cm with one of two acoustic levels, 40.6 or 46.2 dB. For each of these standards six comparison acoustic levels were used, three below and three above the standard and equidistant (i.e., steps of 2.45 dB and 2.73 dB for the 40.6 and 46.2 dB standard, respectively). The comparisons were tested at one of three values of string stiffness, 1, 4, or 7 N/cm. Each combination of 2 (loudness intensity) x 6 (acoustic levels) x 3 (stiffness) was tested 8 times in completely random order. To avoid fatigue the repetitions were divided over two equal sessions.

Performance was measured using the method of constant stimuli, from which we could determine the "point of subjective equal loudness" (PSE), which gives the acoustic level of the comparison that is perceived equally loud as the standard. Since we hypothesized that when the string was softer the sound would be perceived as quieter, we expected that the level of the comparison would need to be higher than the standard to compensate for the perceptual difference. This would be reflected in a PSE that was larger than the standard loudness. The opposite effect would be observed when the virtual string was stiffer.

2.4 Data Analysis

Figure 2 illustrates how the dependent measures were obtained. We first pooled the raw data for the two runs. We calculated, for each comparison and for each string stiffness, the proportion of trials in which the comparison had been perceived as the louder of the two intervals. Individual psychometric functions were obtained by fitting cumulative Gaussians using the software package 'psignifit' [22]. From the fits we calculated the PSE. Statistical analyses were conducted using RStudio running R version 2.15.1. For repeated measures ANOVAs we employed the 'car' package [23]. We used a significance level of 0.05.

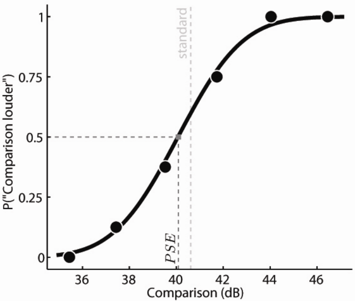


Fig. 2. An illustrative example of the analysis of the psychometric function of one participant for the condition with standard 40.6 db and string stiffness 4 N/cm for both standard and comparison stimulus. The figure demonstrates how the PSE are extracted from cumulative Gaussian fits to the response data. On the abscissa are the comparison acoustic levels and on the ordinate, the proportion of times that the auditory stimulus was perceived louder.

2.5 Results

Figure 3a shows the PSEs as a function of the stiffness of the comparison string for the S-C order (a) and the C-S order (c), from which a number of observations can be made. First of all, there were clear effects of the ordering of the stimuli on the overall pattern of results. Whereas C-S showed apparent effects of stiffness on PSE this was not the case for S-C. A 2 (Order) x 2 (Acoustic Reference) x 3 (Stiffness) ANOVA with Order as a between subjects factor and Acoustic Reference and Stiffness as within subject factors showed that all main and effects and interactions were significant (all p-values < 0.05). We therefore discuss the results separately for the two order groups.

The effects most pertinent to the current question were observed in the C-S group. Changing the stiffness of the string appeared to change its perceived loudness, although this effect was apparently restricted to when the stiffness was lowered to 1 N/cm. When the comparison stiffness was lower than the standard, the comparison sound had to be presented at a higher level in order to be perceived at equal loudness. In other words, consistent with the hypothesis, sounds associated with a softer string were perceived to be quieter. The reverse, however, was not observed. Making the comparison string stiffer did not increase the perceived loudness of the sounds. For the S-C group on the other hand there was no apparent effect of stiffness at all. This appreciation of the results was followed up with a separate 2 (Acoustic Reference) x 3 (Stiffness) repeated measures ANOVA for each order group. Indeed, for the C-S group there was a significant effect of Stiffness ($F(2,8) = 21.85$, $p < 0.01$), Acoustic Reference ($F(1,4) = 241.6$, $p < 0.001$), and their interaction ($F(2,8) = 4.69$, $p < 0.05$). For the S-C group, on the other hand, there was only a significant effect of Acoustic Reference ($F(1,4) = 134.1$, $p < 0.001$), but none for Stiffness ($F < 1$) or the interaction term ($p = 0.14$).

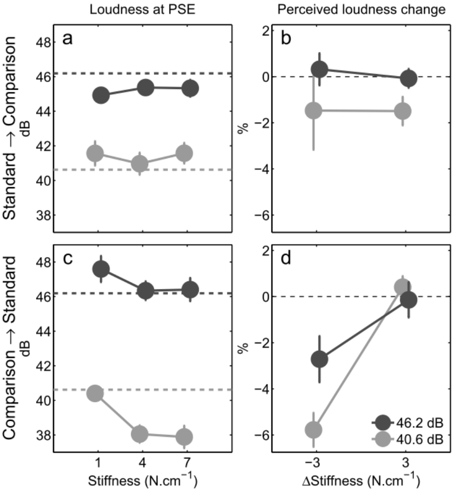


Fig. 3. Results shown separately for the group that had the standard always before the comparison (panels a and b) or the comparison always before the standard (panels c and d). The left column shows the PSEs as functions of the stiffness of the comparison string for the two auditory reference levels (40.6 and 46.2 dB). The right column shows the relative change in perceived loudness as a function of the change in string stiffness (±3 N.cm^{-1} with respect to the standard stiffness). Negative values indicate that the sound was perceived to be less loud, and vice versa.

To further illustrate the effects of changing string stiffness we calculated the difference in PSE between the standard stiffness and the two different levels of stiffness such that a negative value corresponds to a reduction in the perceived loudness and a positive value an increase in perceived loudness. These results are plotted as a function of the change in stiffness between the standard and comparisons in Figure 3b and d.

Thus for the C-S case, when the stiffness decreased by 3 N/cm, the perceived loudness decreased, whereas it did not change for an increase in stiffness. One-sample t-test showed that the shift in perceived loudness was significantly different from zero for the 40.6 dB stimulus [$t(4) = 7.89$, $p < 0.01$], and close to significant for the 46.2 dB stimulus ($p = 0.053$). No significant differences were observed for any of the other cases (all p-values > 0.07).

3 General Discussion

In this pilot experiment, we investigated the effect of the stiffness of a virtual string on the perceived loudness of an actively generated musical tone using standard psychophysical methods. The results showed a significant effect of changes in string stiffness on the perceived loudness of the sound created by plucking the string. They also show that this effect is strongly moderated by whether the string stiffness increases or decreases. Thus, for the stiffness value of 1 N/cm, the participants perceived the comparison sound as being softer than the standard sound where string stiffness was 4 N/cm. In other words, the second sound had to be louder to be perceived as equally loud as the first sound. This effect was larger for the 40.6 dB standard sound than the 46.2 dB. Increasing string stiffness had no perceptual effect whatsoever. Moreover, there was a very strong stimulus order effect. The effect of string stiffness was only observed if the comparison was always presented before the standard. We will address these findings in turn.

3.1 Effect of String Stiffness on Perceived Loudness

At this point we will restrict our discussion of the effects of changes in string stiffness on perceived loudness to the results obtained with the group of participants for which the comparison stimulus was always presented before the standard. We defer the discussion of stimulus order effects to a later section.

The fact that a softer string leads to it being perceived as less loud is in accordance with our hypothesis. Interestingly, whereas earlier studies, such as [16], found an average increase in perceived loudness due to a tactile stimulus of 12% using arbitrary stimuli, the effect observed here was less than 6% which is close to the typical loudness discrimination threshold of around 5% [25]. Before concluding that this then trivializes the effect, we should keep in mind the procedural context of the experiment. Whereas loudness thresholds are typically determined for entirely passively received acoustic stimuli, the stimuli in our experiment were created through an active involvement on the part of the participant. We already mentioned in the introduction that voluntary active movements can affect haptic perception [17] and auditory-haptic interactions [19], the nature of which are still far from understood [17]. We therefore have to consider the possibility that the previously observed facilitatory, or "boosting", effect of tactile stimuli on loudness perception can be modulated by the voluntary actions on the part of our participants.

The fact that the effect of string stiffness was larger for the softer of the two auditory standards can be understood from the principle of inverse effectiveness which states that multisensory integration is stronger for weaker signals [24]. It can be argued that the responsiveness to softer tones is smaller than for the louder ones and is therefore more susceptible to crossmodal influences, in this case from the tactile sense. If this holds then we would expect larger effects of stiffness on increasingly softer tones. This notion should be tested in future experiments. Inverse effectiveness, however, cannot easily explain why we did not observe an effect of an increase in stiffness since the auditory input per se was physically identical to the one when there was a reduction in stiffness. We can speculate as to the cause of this asymmetry. It can be that a reduction in stiffness is more salient than an increase. In other words, -3 N/cm effectively presents a larger change than +3 N/cm, and therefore can exercise a larger crossmodal effect. Alternatively, Odgaard et al. found that for a visual-auditory interaction there was no interaction between loudness and brightness when the masking stimulus was kept constant in the two intervals [26]. In our experiment, we explicitly instructed the participants to focus on stimulus intensity which might explain the absence of effect of the highest stiffness on loudness. Since the order of standard and comparison was fixed, it is possible that some participants might have been susceptible to this bias for the strongest stiffness. Since very little research has been done on this it remains unclear how the response bias affects loudness perception specifically.

3.2 Effects of Stimulus Presentation Order

A remarkable result of our pilot experiment was that we obtained significant effects of changes in string stiffness when the comparison was presented before the standard (C-S) but none when the presentation order was reversed (S-C). We there-fore need to consider the ramifications of stimulus presentation order. Although the order was randomized across participants for each individual participant the order was fixed. This could mean that we did not observe a change in participants' perception of the auditory stimulus but a modification in their response criterion, which could lead to biased results.

The interval-bias hypothesis, for instance, states that the influence of presentation order affects performance accuracy in a discrimination task, a phenomenon known as the time-order error (TOE) and is supposedly an attentional effect [27] in which the first stimulus cues attention to the occurrence of the second stimulus. A consequence is that participants start favoring one interval over another. These kinds of biases are apparently complicated further by other factors such as sensory modality. For instance, Okazaki et al [13] tested a number of standard and comparison conditions. On half the conditions the standard was haptic and in the other the standard was auditory. Participants would always perceive the second stimulus as more intense than the first, irrespective of the actual order of the stimuli. However, this response pattern was not observed when the standard was auditory. Although, in our case both standard and comparisons are always auditory-haptic, a condition not considered in [13], it remains unclear to what extent these kinds of response order effects played a role.

We need to keep in mind that although these kinds of extraneous influences can introduce "false results" or cancel out genuine perceptual effects they do so by adding constant offsets to the results. Close inspection of the present results, however, shows that they cannot be entirely explained by simple offsets. This becomes most apparent when considering the differences between the two acoustic references. Consider Figure 3a and assume for the sake of argument that participants tend to favor the second interval as they response, which in this case would be the comparison. The influence of such a bias would lead to a systematic underestimation of (i.e., decrease in) the PSE, which is what we observe for the 46.2 dB reference. However, it is immediately clear that this bias cannot explain the tendency for the PSE to be higher than the reference for the 40.6 dB reference. Moreover, a similar appreciation can be made when we consider Figure 3c. In short, no single bias can explain the present order effects. We cannot at this point exclude the influence of a complex of biases. Further studies should consider testing procedures that minimize stimulus order effects, such as completely randomizing the presentation order of the standard and the reference or direct comparison and adjustment procedures such as used by Schürmann et al [16].

4 Conclusion

In conclusion, we have shown that changing the stiffness of a virtual string can change the perceived sound pressure level of sounds associated with that string. Although subject to methodological complications, this finding is not only of theoretical interest in our further understanding of multisensory perceptual processes but also is encouraging for auditory-haptic applications related to music performance. Indeed, the cognitive processes of music are different from listening to a single sound or speech. Playing a musical instrument not only requires the auditory feedback but also requires basic motor functions such as timing, sequencing and spatial organization of the movement [28] as well as possibly strong attentional efforts.

References

1. Repp, B.H.: Effects of auditory feedback deprivation on expressive piano performance. Music Perception 16, 409–438 (1999)
2. Pfordresher, P.Q., Palmer, C.: Effects of hearing the past, present, or future during music performance. Perception and Psychophysics 68, 362–376 (2006)
3. Drake, C., Penel, A., Bigand, E.: Tapping in time with mechanically and expressively performed music. Music Perception 18, 1–23 (2000)
4. Goebl, W., Palmer, C.: Tactile feedback and timing accuracy in piano performance. Experimental Brain Research 186, 471–479 (2008)
5. Palmer, C., Koopmans, E., Loehr, J.D., Carter, C.: Movement-related feedback and temporal accuracy in clarinet performance. Music Perception 26, 439–449 (2009)
6. Avanzini, F., Crosato, P.: Haptic-auditory rendering and perception of contact stiffness. In: McGookin, D., Brewster, S. (eds.) HAID 2006. LNCS, vol. 4129, pp. 24–35. Springer, Heidelberg (2006)
7. Soto-Faraco, S., Deco, G.: Multisensory contributions to the perception of vibrotactile events. Behavioural Brain Research 196, 145–154 (2009)
8. Guest, S., Catmur, C., Lloyd, D., Spence, C.: Audiotactile interactions in roughness perception. Experimental Brain Research 146(2), 161–171 (2002)
9. Fletcher, H., Blackham, E.D., Stratton, R.: Quality of piano tones. J. Acoust. Soc. Am. 34(6), 749–761 (1962)

10. Järvelälinen, H., Välimäki, V., Karjalainen, M.: Audibility of inharmonicity in string instrument sounds, and implications to digital sound synthesis. In: Proc. Int. Computer Music Conf., Beijing, China, pp. 359–362 (1999)
11. Marshall, M., Wanderley, M.: Vibrotactile feedback in digital musical instruments, Internat. In: Conference on New Interfaces for Musical Expression (NIME 2006), pp. 226–229 (2006)
12. Kinoshita, H., Furuya, S., Aoki, T., Altenmüller, E.: Loudness control in pianists as exemplified in keystroke force measurements on different touches. Journal of Acoustical Society of America 121, 2959–2969 (2007)
13. Okazaki, R., Kajimoto, H., Hayward, V.: Vibrotactile Stimulation Can Affect Auditory Loudness: A Pilot Study. In: Isokoski, P., Springare, J. (eds.) EuroHaptics 2012, Part II. LNCS, vol. 7283, pp. 103–108. Springer, Heidelberg (2012)
14. Merchel, S., Leppin, A., Altinsoy, E.: Hearing with your body: the influence of whole-body vibrations on loudness perception. In: Proceedings of the Sixteenth International Congress on Sound and Vibration (ICSV16), Kraków, Poland, July 5-9 (2009)
15. Gillmeister, H., Eimer, M.: Tactile enhancement of auditory detection and perceived loudness. Brain Research 1160, 58–68 (2007)
16. Schürmann, M., Caetano, G., Jousmäki, V., Hari, R.: Hands help hearing: facilitatory audiotactile interaction at low sound-intensity levels. J. Acoust. Soc. Am. 115(2), 830–832 (2004)
17. Frissen, I., Ziat, M., Campion, G., Hayward, V., Guastavino, C.: The effects of voluntary movements on auditory-haptic and haptic-haptic temporal order judgments. Acta Psychologica 141, 140–148 (2012)
18. Vitello, M.P., Ernst, M.O., Fritschi, M.: An instance of tactile suppression: Active exploration impairs tactile sensitivity for the direction of lateral movement. In: Proceeding of Eurohaptics, Paris, July 3-6, pp. 351–355 (2006)
19. Kitagawa, N., Kato, M., Kashino, M.: Assessing the effect of voluntary action on sensitivity to temporal asynchrony between auditory and somatosensory events. In: 10th International Multisensory Research Forum (IMRF), New York, June 29 - July 2 (2009)
20. Campion, G., Wang, Q., Hayward, V.: The Pantograph mk-II: A haptic instrument. In: Proceedings of the Internat. Conf. on Intelligent Robots and Systems 2005, pp. 723–728 (2005)
21. Karplus, K., Strong, A.: Digital Synthesis of Plucked String and Drum Timbres. Computer Music Journal 7(2), 43–55 (1983)
22. Wichmann, F.A., Hill, N.J.: The psychometric function: I. Fitting, sampling, and goodness of fit. Perception & Psychophysics 63, 1293–1313 (2001)
23. Fox, J., Weisberg, S., An, R.: An R companion to applied regression, 2nd edn. Sage Publications, Thousand Oaks (2011)
24. Stein, B.E., Meredith, M.A.: The merging of the senses. MIT Press, Cambridge (1993)
25. Coren, S., Ward, L.M., Enns, J.T.: Sensation and perception. Harcourt Brace, Fort Worth (1999)
26. Odgaard, E.C., Arieh, Y., Marks, L.E.: Cross-modal enhancement of perceived brightness: Sensory interaction versus response bias. Perception & Psychophysics 65(1), 123–132 (2003)
27. Hairston, I.S., Nagarajan, S.S.: Neural mechanisms of the time-order error: An MEG study. Journal of Cognitive Neuroscience 19, 1163–1174 (2007)
28. Zatorre, R.J., Chen, J.L., Penhune, V.B.: When the brain plays music: auditory-motor interactions in music perception and production. Nat. Rev. Neuroscience, 547–558 (2007)

Perceptual and Technological Issues in the Design of Vibrotactile-Augmented Interfaces for Music Technology and Media

Marcello Giordano and Marcelo M. Wanderley

Input Devices and Music Interaction Laboratory,
Centre for Interdisciplinary Research in Music Media and Technology,
Schulich School of Music, McGill University,
Montréal, Québec, Canada
`marcello.giordano@mail.mcgill.ca`, `marcelo.wanderley@mcgill.ca`

Abstract. In this paper we present tactile feedback and stimulation design principles for applications in music technology and media. We discuss features and limitations of the human sense of touch, in the context of conveying musical content solely via the tactile sense. These factors should be firmly taken into account when designing a tactile-augmented interface. Applications of tactile displays in the field of music and media are then presented using a three-fold taxonomy of tactile feedback.

Keywords: haptics, vibrotactile feedback, vibrotactile stimulation, sensory substitution, music technology.

1 Introduction

In the last century, the number of "new" systems for generating sound and music has witnessed a constant growth which reached its climax with the availability of computer-based synthesizers [1]. In this context, the idea of defining a new family of instruments, called digital musical instruments (DMIs), rose naturally to indicate a class of instruments in which control interfaces and sound generators are physically separated, the sound generator usually being a computer-base synthesizer. The control interface is a "gestural controller" [2], capable of sensing performer's gestures which are then mapped to the sound synthesis engine; in this context, the mapping process becomes the core of the instrument design.

The decoupling of control and sound generation in two independent units also has a secondary effect, it breaks the tactile and kinesthetic feedback loop coming from the resonating parts of the instrument. This feedback has been proven to be an essential component in the player-instrument interaction [3] especially in expert performance [4] and has been suggested to be the only form of feedback fast enough to control articulation [5].

In DMIs, haptic feedback becomes a design factor [6] in a way that was unknown to traditional instruments. The type of actuators used to deliver information to the user, their placement and the signals used to drive them become a

I. Oakley and S. Brewster (Eds.): HAID 2013, LNCS 7989, pp. 89–98, 2013.

fundamental component in the architecture of the instrument. The use of haptic and particularly tactile technology in the musical domain is of course not only limited to DMIs; tactile-capable devices have been used for creating learning interfaces for music education [7], notification tools for live performances [8], displays for hearing-impaired people [9] and so on.

The purpose of this paper is to provide both technological and perceptual guidelines that, in our experience, need be taken into account when designing interfaces for music and media applications.

2 The Sense of Touch: Physiology and General Properties

Tactile perception operates through a network of cutaneous receptors present in human skin (17,000 receptors have been estimated to be present in each hand [10]) and is responsible of sensations like pressure, temperature, texture, orientation, vibration and many others. Its role is crucial in motor control and in the execution of many simple and complex tasks.

In glabrous skin we can identify four *mechanoreceptors*: Pacinian and Meissner corpuscles, Ruffini end-organs and Merkel disks [11]. These detect skin deformations and are connected to two families of afferent nerve fibers: the Fast Afferent (FA) family, ending in Meissner (FA I) and Pacinian (FA II) corpuscles, and the Slow Afferent (SA) one, connected to Merkel (SA I) and Ruffini (SA II) end-organs. FA fibers cease responding very rapidly after the end of the stimulation, while the SA ones keep responding for a longer period of time. This behavior is called the *adaptation property* [11]. Each of these four separate sensory channels is sensitive to very specific features in a tactile stimulus [12], and the FA I and FA II are considered to be the ones responsible for vibration sensation [13]. The firing behavior of Pacinian cells in particular is very similar to the way in which the auditory system reacts to a stimulus [14]. This seems to be evidence to the fact that the FA II channel is the one to be mainly exploited for the mediation of musical vibrotactile events [15].

The frequency response of glabrous skin has a characteristic U-shaped form which spans from 40 Hz to 1000 Hz with a peak sensitivity at around 250 Hz for every kind of tactor used in the tests and at any amplitude [11]. A set of equal sensation magnitude curves, relating perceived intensity of a tactile stimulus to frequency, has also been identified [11]. These curves are very similar to Fletcher-Munson curves for auditory perception [16].

For what concerns hairy skin, experimental studies have shown [17] that hairy skin presents higher threshold and lower peak sensitivity values compared to glabrous skin.

3 Conveying Information through the Sense of Touch

3.1 Temporal Domain

Pitch. When designing interfaces that rely on *vibrotactile pitch* discrimination, a few important points should be considered. Several experiments have tried to define the ability of tactile sense to elaborate pitch information. Rovan and Hayward, for instance, [18] have informally suggested that the glabrous skin is capable of discriminating between 3 to 5 different values for a continuous change in frequency from 2 Hz to 300 Hz and from 8 to 10 values when going from 70 Hz to 1000 Hz. Branje et al. [19] measured the ability of participants to discriminate vibrotactile frequencies, using large voice-coils placed in contact with the back; results showed that participants in the study could effectively discriminate frequency separated by 1/3 of an octave, ranging from 67 Hz to 1047 Hz.

These and other studies [11] suggest that, compared to auditory perception, frequency discrimination performance using solely tactile information is poor, but the tactile channel is nonetheless able to determine gross frequency changes. Tactile frequency discimination is also dependent on several factors, such as: the amplitude of the vibration [20]; the presence of an adapting stimulus [21]; and training [22].

Rhythm. A few studies suggest that the sense of touch performs surprisingly well in *rhythm recognition.*

Kosonen and Raisamo [23] and Jokiniemi et al. [24] have compared performance of participants in same/different judgments of rhythmical patterns presented unimodally in the auditory, tactile and visual modality. Results showed that the tactile modality performed better than the visual one, and close to the auditory. Rhythmical patterns have also been used by Brown et al. [25] as a parameter for defining tactile icons (Tactons), a set of messages for delivering information through the tactile channel: a series of three different rhythms were used to deliver messages about the interaction with a mobile phone. Results showed a rate of recognition of the pattern up to 90%.

Roughness. Different authors have reported that it is possible to convey information about *roughness* of a stimulus through the tactile sense. There is however some ambiguity about what tactile roughness means: based on Weisenberger [26], Brown et al. [25] characterized “roughness” using a sinusoidal stimulus modulated in amplitude by another one. They showed that the perception of roughness can be adjusted acting on the frequency of the modulation signal (the higher the frequency, the rougher the final signal is perceived), and that this parameter can be used for encoding information about the priority of an event in the interaction with a mobile phone with a recognition rate that goes up to 80%. Other authors provide a different definition for tactile roughness: for Rovan and Hayward for example [18], a rough signal is one that has a richer spectrum, consisting of many partials, such as a square wave, as opposed to smooth sine wave which only consists of one partial. Verrillo [11] reported that participants in his experiments defined as “smooth” a stimulus of higher frequency, while low frequency ones were identified as more “buzzy”. Recent experiments by Park and Choi [27] indicate that amplitude modulation techniques seem to be more suitable for modeling tactile roughness.

Timbre. In a recent experiment, Russo and colleagues [28] investigated the capability of participants to distinguish between cello, trombone and piano samples, matched for fundamental frequency and perceived magnitude, only by the sense of touch. The results indicate a discrimination level that is above chance, indicating that timbre information could be conveyed by means of tactile-only information.

3.2 Spatial Domain

Acuity, Pattern Recognition and Numerosity. Evaluating spatial resolution of the tactile sense is crucial in the design of devices that depend on the user's capability of distinguishing the part of the body to which a tactile stimulus has been applied. Cholewiak and colleagues [29,30] performed a series of experiments to test participants' ability to discriminate the location where a vibrotactile stimulus was presented. The possibilities consisted of seven points of the forearm, three points on the upper arm, two points on the shoulder and seven points around the lower torso. Results showed poor performance for what concerns the fore-arm, with results superior to 70% only in two points, the elbow and the wrist. Better results were achieved for the torso, for which all the points were identified more than 70% of the time, with peaks up to almost 100%.

Other studies seem to confirm these results; van Erp [31] showed that the torso has a spatial acuity of about 3 cm, remarking however that the discrimination is highly dependent on two temporal factors as well: the duration of the simuli and the temporal offset between two consecutive stimuli. Piateski and Jones [32] presented vibrotactile patterns using a tactile display on the forearm and on the lower back. While the recognition rate attained almost 100% for patterns displayed on the back, performances were unsatisfactory for the fore-arm.

Another important aspect to consider is the capacity of the tactile sense to make numerosity judgements. Gallace et al. [33] conducted an experiment in which participants had to identify how many tactile stimuli were presented simultaneously on their body, while wearing a tactile display composed of seven different actuators. Results show that the error rate increases to more than 50% when more than two stimuli are presented at the same time.

Tactile Illusions. A particularly interesting phenomenon is represented by tactile illusions, i.e. phantom tactile sensations created by tuning the onset of two or more real tactile stimulations applied at different sites [34]. The most well-known is the so-called "cutaneous rabbit" illusion [35]. A specific temporal and spatial pattern of stimulation between two points on the skin can give rise to the emergence of the sensation of a moving point that flows between the two stimulated points.

Attention. Vibrotactile stimulation can be used to attract attention to a specific part of the body. Spence and Gallace [36] present evidence showing that after the presentation of a tactile stimulus to a part of the body, subsequent stimuli in the same location, in the tactile or also other modalities, can more easily be detected. Tactile stimulation can so be used to direct users' spatial attention to events taking place in other modalities.

4 Technical Remarks

When designing interfaces that rely on the communication of tactile events, the aforementioned perceptual properties and limitations that are inherent to tactile perception must be considered. These perceptual issues reflect precise choices that should be taken into account from a technological and a design point of view when defining the system's architecture.

4.1 Choice of Actuators

Many different kinds of vibrating actuators exist on the market. A review of the properties of commercially available actuator technologies is beyond the scope of this paper. A general, thorough review on available tactile-displays and actuators can be found in [3] and [13].

An analysis of the most popular actuators used in DMIs and tactile displays for music and media can be found in [3]; this survey shows that the type of actuators mostly used in the field are loudspeaker-like ones.

Generally speaking, the choice of the actuators to be used is mainly determined by the role of vibrotactile feedback in the interface design, by other factors such as size or power consumption and by the information to be conveyed.

Simpler actuators, such as unbalanced masses or solenoids for instance, have the advantage of requiring low power and they can be driven by very simple signals such as pulse-width modulation [3]. Loudspeaker-like actuators usually require proper amplification (i.e. suitable for an audio signal), which requires in turn proper power source, but they can provide independent control of amplitude and frequency of vibrations [3].

4.2 Frequency Response

When using loudspeaker-like actuators, the signals used to drive the chosen actuators should be equalized [15] to compensate for the previously mentioned equal-sensation magnitude curves [11] that relate frequency of vibration to perceived intensity. Moreover, the actuator itself, together with the circuitry needed to drive it (i.e. power amplifiers) and the whole interface in which it is eventually embedded, possess a specific frequency response. This has to be taken into account if a "perceptually-flat" frequency response of the system needs to be achieved; in the design phase, the choice of materials with a linear mechanical response should be preferred to simplify the compensation process.

4.3 Placement and Activation Patterns

When designing distributed, whole-body displays, relying on the stimulation of multiple body sites at the same time, the placement of the actuators is a crucial factor to ensure that the desired information is successfully conveyed to the final user.

Results presented in the previous section indicate that the sense of touch can proficiently be used to give directional information or to give the user specific directives, associated with pre-determined patterns. As we have seen, however, this ability varies considerably with the chosen body sites. For instance, better results were achieved on the waist and torso than on the forearm for example[31,32].

On the other hand, the amount of information that can be processed at the same time, when only relying on the tactile sense, is modest, at least if the tactile stimuli are not presented in a specific pattern [33].

5 Applications

We present exemplifying tactile displays chosen from both our previous work and from the most relevant literature on the topic. These displays show the relevance of the perceptual and technological issues previously described. Using a functional organization, we will distinguish between tactile notification, translation and synthesis. We believe that these categories describe the three main roles that tactile feedback can play in the context of designing interfaces for musical expression and media. They determine the choice of the actuators, their placement and the synthesis algorithms needed to produce the signals used to drive them.

Our taxonomy aims to clearly state what this roles are, what receptive capabilities of the sense of touch they address, and which technology approach is better suited for each category.

5.1 Tactile Notification

The most straightforward application of tactile stimulation is for notifying the user about events taking place in the surrounding environment or about the results of their interaction with a system. Usually, the stimuli needed for this kind of application do not need to have any specific characteristics other than being supra-threshold; they only need to direct user attention without necessarily having to convey any other extra information. For this reason, the actuators used in this context are generally simple rotating eccentric masses, which require very low power (usually 3V or less) and can be driven, for instance, directly from the PWM output of an Arduino[1] microcontroller.

Michailidis et al. [8] used these devices investigate ways of conveying haptic feedback in live-electronics performance. Using small vibrating motors, the authors managed to give musicians valuable information about the successful triggering of effects in a live-electronics performance, using an augmented trumpet.

---

[1] `http://www.arduino.cc`

Following a similar approach, we are developing a tactile notification tool to be used in conjunction with *CLEF*[2] (CIRMMT Live-Electronics Framework). This tool, consisting of two vibrating motors placed on the ankles, can give the performer feedback about the successful activation of events in CLEF by means of simple clicks or "buzzes". The internal state of a particular effect can also be represented; for instance, leveraging the cutaneous rabbit illusion, a tactile "flow" can be induced (from left to right or vice versa) to give the performer information about the panning of the sound as perceived by the audience.

5.2 Tactile Translation

In sensory substitution, a stimulus usually addressed to one sensory channel is transformed in a way that it can be received and processed by another sensory channel. Tactile translation is a form of sensory substitution. Leveraging the neuro-physiological similarities between auditory and tactile sensory channels [14], one can *translate* an auditory stimulus to the tactile channel by means of frequency rescaling and cross-modal mapping between features in the original sound and in the target tactile stimulus.

Loudspeaker-like actuators are evidently better suited to render stimuli as (spectrally) complex as those produced by a tactile translation process. As we mentioned in Sect. 4, proper amplification and frequency compensation techniques must be considered when using this kind of devices. For the tactile stimulation to be meaningful, the actuating signals need to be tuned to the receptive range specific to the part of the body where the actuators are to be applied.

Birnbaum's work [15] is one of the first examples of tactile translation in a purely musical-related domain. A flute-like controller for breakbeat loops, the BreakFlute [15], was augmented with small voice-coils placed in the keyholes. A tactile translation environment called FA/SA performed the translation of the sound output into tactile stimuli played by the voice-coils, with the aim of recreating a feedback loop between sound output and haptics in a DMI. Performers informally reported a greater degree of "control" while interacting with the instrument.

In our previous work [7] we have studied the possibilities given by vibrotactile stimulation, to enhance the learning process of novice guitar players. Our prototype consisted of a vibrotactile whole-body display composed of 10 loudspeaker-like vibrating actuators. Following Birnbaum's approach a multi-channel version of the FA/SA environment was developed to drive the suit. A tactile translation of a base-track (base-line and drums) was spatially mapped onto the body of the performers. Novice players were asked to play along with the base track; they remarked that the presence of the display, particularly because of the actuators on the back (mapped to the drum-kick), gave valuable information about the tempo, allowing them to keep a better focus on the instrument.

Karam et al. [9], developed a prototype of augmented chair (the "Emotichair") embedded on the back with an array of speakers. The aim was to create a

---

[2] `http://clef.sf.net`

display for deaf people to enjoy music through vibrations. Following their "Model Human-Cochlea", a model of physical translation of the cochlear critical band filter on the back, the authors mapped different frequency bands of a musical track, rescaled to fit into the frequency range of sensitivity of the skin, to each of the speakers on the chair.

5.3 Tactile Synthesis

Tactile synthesis is the attempt to create compositional languages solely addressed to the sense of touch, so as to be able to convey more complex information than in the tactile notification paradigm. This information is not issued from a direct transposition of a signal normally addressed to another sensory modality, as is the case for tactile translation.

An example in a musical context is the work of Gunther [37] who developed a musical compositional language for the sense of touch: the "Skinscape" system is composed of a tactile display and a tactile composition environment. The building blocks of this language (also at the base of the "Cutaneous Grooves" project [38]) are frequency, intensity, envelope, spectral content and spatial position on the body of the user; the aim is to create a language that could be used to accompany music. The author does not provide formal results about the effectiveness of his tactile display, but, based on perceptual evidence [33,39], it seems that amount of information he plans to deliver to the user could become quickly overwhelming, not allowing to attend to all the stimuli presented simultaneously.

A DMI featuring tactile synthesis capabilities is the Viblotar by Marshall et al. [40]. The sound output from the instrument can be redirected to the embedded speakers or to an external sound system. In this case, the internal speakers can be used to generate additional vibrotactile feedback and the authors hypothesize the use of the internal speakers to simulate the frequency response of another instrument.

6 Conclusions

Vibrotactile feedback and stimulation can be implemented in interfaces for musical expression and media with a great variety of uses: restoring the intimacy in instrumental interaction with a DMI, providing notification cues and alerts, allowing hearing-impaired people to experience music.

In this paper we provide a review of the most important features and limitations of human tactile perception, with special attention to the factors most relevant to the communication of musical-tactile information. We believe that the knowledge of these perceptual aspects is fundamental for achieving coherent tactile experiences that reflect the designer's intentions. Relevant technological issues are also presented as a natural counterpart to the perceptual properties of the sense of touch. Finally, we present a taxonomy of tactile feedback and stimulation, to provide exemplary applications of tactile-augmented interfaces in the domain.

References

1. Miranda, E.R., Wanderley, M.M.: New Digital Musical Instruments: Control and Interaction beyond the Keyboard. The Computer Music and Digital Audio Series. A-R Editions (2006)
2. Wanderley, M.M.: Gestural control of music. Proceedings of the IEEE (2001)
3. Marshall, M.T.: Physical Interface Design for Digital Musical Instruments. Doctoral thesis, McGill University (2008)
4. Keele, S.W.: Attention and Human Performance. Goodyear Pub. Co. (1973)
5. Puckette, M., Settel, Z.: Nonobvious roles for electronics in performance enhancement. In: Proceedings of the International Computer Music Conference, p. 134 (1993)
6. MacLean, K.E.: Designing with haptic feedback. In: Proceedings of the International Conference on Robotics and Automation (ICRA), vol. 1, pp. 783–788. IEEE (2000)
7. Giordano, M., Wanderley, M.M.: A Learning Interface For Novice Guitar Players Using Vibrotactile Stimulation. In: Proceedings of the Sound and Music Computing Conference, SMC (2011)
8. Michailidis, T., Berweck, S.: Tactile Feedback Tool: Approaching the Foot Pedal Problem in Live Electronic Music. In: Proceedings of the Internation Computer Music Conference, ICMC (2011)
9. Karam, M., Russo, F.A., Fels, D.I.: Designing the model human cochlea: An ambient crossmodal audio-tactile display. IEEE Transactions on Haptics 2(3), 160–169 (2009)
10. Vallbo, Å.B., Johansson, R.S.: Properties of cutaneous mechanoreceptors in the human hand related to touch sensation. Human Neurobiology 3(1), 3–14 (1984)
11. Verrillo, R.T.: Vibration sensation in humans. Music Perception 9(3), 281–302 (1992)
12. Johnson, K.O.: The roles and functions of cutaneous mechanoreceptors. Current Opinion in Neurobiology 11(4), 455–461 (2001)
13. Choi, S., Kuchenbecker, K.J.: Vibrotactile Display: Perception, Technology, and Applications. Proceedings of the IEEE, 1–12 (2012)
14. Makous, J.C., Friedman, R.M., Vierck, C.J.: A critical band filter in touch. The Journal of Neuroscience: the Official Journal of the Society for Neuroscience 15(4), 2808–2818 (1995)
15. Birnbaum, D.M., Wanderley, M.M.: A systematic approach to musical vibrotactile feedback. In: Proceedings of the Internation Computer Music Conference, ICMC (2007)
16. Fletcher, H., Munson, W.A.: Loudness, Its Definition, Measurement and Calculation. The Journal of the Acoustical Society of America 5(2), 82–108 (1933)
17. Verrillo, R.T.: Vibrotactile Thresholds for Hairy Skin. Journal of Experimental Psychology 72(1), 47–50 (1966)
18. Rovan, J., Hayward, V.: Typology of tactile sounds and their synthesis in gesture-driven computer music performance. In: Wanderley, M.M., Battier, M. (eds.) Trends in Gestural Control of Music, pp. 297–320 (2000)
19. Branje, C., Maksimouski, M., Karam, M., Fels, D.I., Russo, F.: Vibrotactile Display of Music on the Human Back. In: Third International Conference on Advances in Computer-Human Interactions, pp. 154–159. IEEE (2010)
20. Morley, J.W., Rowe, M.J.: Perceived Pitch of Vibrotactile Stimuli: Effects of Vibration Amplitude, and Implications for Vibration Frequency Coding. Journal of Physiology 431, 403–416 (1990)

21. Goble, A.K., Hollins, M.: Vibrotactile adaptation enhances frequency discrimination. The Journal of the Acoustical Society of America 96(2 pt. 1), 771–780 (1994)
22. Harris, J.A., Harris, I.M., Diamond, M.E.: The topography of tactile learning in humans. The Journal of Neuroscience 21(3), 1056–1061 (2001)
23. Kosonen, K., Raisamo, R.: Rhythm perception through different modalities. In: Proceedings of Eurohaptics, pp. 365–370 (2006)
24. Jokiniemi, M., Raisamo, R., Lylykangas, J., Surakka, V.: Crossmodal Rhythm Perception. In: Pirhonen, A., Brewster, S. (eds.) HAID 2008. LNCS, vol. 5270, pp. 111–119. Springer, Heidelberg (2008)
25. Brown, L.M., Brewster, S.A., Purchase, H.C.: A First Investigation into the Effectiveness of Tactons. In: First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, pp. 167–176 (2005)
26. Weisenberger, J.M.: Sensitivity to amplitude-modulated vibrotactile signals. The Journal of the Acoustical Society of America 80(6), 1707–1715 (1986)
27. Park, G., Choi, S.: Perceptual space of amplitude-modulated vibrotactile stimuli. In: 2011 IEEE World Haptics Conference, pp. 59–64 (June 2011)
28. Russo, F.A., Ammirante, P., Fels, D.I.: Vibrotactile discrimination of musical timbre. Journal of Experimental Psychology. Human Perception and Performance 38(4), 822–826 (2012)
29. Cholewiak, R.W., Brill, J.C., Schwab, A.: Vibrotactile localization on the abdomen: effects of place and space. Perception & Psychophysics 66(6), 970–987 (2004)
30. Cholewiak, R.W., Collins, A.A.: Vibrotactile localization on the arm: effects of place, space, and age. Perception & Psychophysics 65(7), 1058–1077 (2003)
31. van Erp, J.B.F.: Vibrotactile Spatial Acuity on the Torso: Effects of Location and Timing Parameters. In: Proceedings of Eurohaptics, pp. 80–85. IEEE (2005)
32. Piateski, E., Jones, L.: Vibrotactile Pattern Recognition on the Arm and Torso. In: First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, pp. 90–95. IEEE (2005)
33. Gallace, A., Tan, H.Z., Spence, C.: Numerosity judgments for tactile stimuli distributed over the body surface. Perception 35(2), 247–266 (2006)
34. Hayward, V.: A brief taxonomy of tactile illusions and demonstrations that can be done in a hardware store. Brain Research Bulletin 75(6), 742–752 (2008)
35. Geldard, F.A., Sherrick, C.E.: The Cutaneous "Rabbit": A Perceptual Illusion. Science 178(4057), 178–179 (1972)
36. Spence, C., Gallace, A.: Recent developments in the study of tactile attention. Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expèrimentale 61(3), 196–207 (2007)
37. Gunther, E.: Skinscape: A tool for composition in the tactile modality. Master's thesis, Massachusetts Institute of Technology (2001)
38. Gunther, E., O'Modhrain, S.: Cutaneous Grooves: Composing for the Sense of Touch. Journal of New Music Research 32(4), 369–381 (2003)
39. Gallace, A., Tan, H.Z., Spence, C.: The Body Surface as a Communication System: The State of the Art after 50 Years. Presence: Teleoperators and Virtual Environments 16(6), 655–676 (2007)
40. Marshall, M.T., Wanderley, M.M.: Examining the Effects of Embedded Vibrotactile Feedback on the Feel of a Digital Musical Instrument. In: Proceeding of the New Interfaces for Musical Expression Conference (NIME), pp. 399–404 (2011)

jamTable: Can Physical Interfaces Support the Collaboration between Novice and Experienced Musicians?

Augusto Esteves[1], Filipe Quintal[1,2], and Ian Oakley[1,2]

[1] CCCEE, University of Madeira, Funchal, Portugal
[2] Madeira Interactive Technologies Institute, Funchal, Portugal
`{augustoeae,filp3q,ian.r.oakley}@gmail.com`

Abstract. This paper introduces jamTable, a system that enables the collaboration between users playing a standard musical instrument and users interacting with a tangible musical sequencer. In an introductory study both qualitative and quantitative data were collected from eight participants in two setup conditions: Musician-Musician and Novice-Musician pairs. By comparing the performance of participants in these two groups, this paper gathers relevant insights regarding the ability of a tangible musical application such as the jamTable to support musical collaborations between novice and experienced musicians – in both learning or performance activities.

1 Introduction

Music has been moving from a social to a personal experience, from the concerts halls to the iPods, while technology has been doing exactly the opposite – moving towards the Web and towards the group [15]. One of the latest instantiations of this trend in technology has been the tangible interaction paradigm, where the digital is mapped to the physical, allowing for several concurrent users to engage directly with information through a rich sensorial experience (see Fig. 1). As such, tangible interaction has been associated with several benefits that seem appropriate for a musical context, such as: support for seamless collaboration, visible interaction, increased levels of engagement and enjoyment, and of being highly effective in learning scenarios [11, 18]. Several tangible musical applications have been developed that strive for these qualities, including the popular reacTable [10] or the Audiopad [17]. Research on such applications normally reports on their ability to support a wide range of users in learning and performance tasks [e.g. 3, 5, 10, 12, 15], and while they do provide an unique musical experience, most of these claims stem from subjective and qualitative metrics. In order to look into how quantitative data can be obtained to back up such claims, this paper presents both the jamTable, a tangible system designed to support the collaboration between novice and experienced musicians, and an introductory study that captures both quantitative and qualitative data to gather insights in the true benefits of tangible musical applications.

I. Oakley and S. Brewster (Eds.): HAID 2013, LNCS 7989, pp. 99–108, 2013.

Fig. 1. On the left: A set of physical artifacts to control pitch in MoSo Tangibles, a system for children's musical education [2]; on the right: the Audiopad [17], a tangible music controller that employs knob-based tokens for fine control over the interaction

2 Related Work

Links between music and HCI can be traced back to as early as the 1970s and 1980s, mostly in the work of William Buxton [e.g. 4]. Similarly, musical applications are amongst the oldest and most popular forms of tangible interaction [18]. This can be explained by the fact that interactive tabletops have been found to seamless support collaboration between users, allow several pieces of information to be controlled simultaneous and in real-time, and ultimately provide an environment for skilled, creative and explorative interaction [9]. As with the system described in this paper, these benefits have normally been explored in the context of two specific areas within tangible interaction: systems for musical education, and for musical performances. The remainder of the section will briefly review these two fields.

Researchers have long linked the physical manipulation of objects with the development of children's cognitive capabilities [13]. As such, many studies in tangible interaction [e.g. 6, 20] have discussed the potential of this interaction paradigm to promote learning and development, either by contributing to self-reflection [24], providing more appealing experiences [1], or by facilitating autonomous and personal activities [16]. As a result, numerous tangible applications have focused exclusively on musical education for children. One prominent example is Tangible Notes [21], a system that allows children to explore musical notes and scales through tokens on an interactive tabletop. Another example is MoSo Tangibles [2], a tangible application that presents abstract audio concepts through different body-based mappings. With this system, children can physically and haptically manipulate sound properties like pitch, volume, or tempo in order to learn about them (see Fig. 1). Similarly, SOUNDGARTEN [22] is an interactive toy that allows children to apply sound effects to either pre-recorded sound samples, or to sounds they record in their environment. Children can create and arrange different sound scenarios, and control the volume and pitch of the final result. Finally, Marble Tracks [7] combines electronic artifacts that can be added to physical marble tracks to create a musical sequencer that children can manipulate through play.

These examples illustrate systems that were designed and studied in regards to their capabilities to actually motivate musical learning in children. While to some extent beyond their scope, these systems nevertheless fail to explore how they could be used by adults in pursuit of musical education. Furthermore, the remainder of this section includes several examples of musical applications that argue for applicability in learning scenarios for adults. Despite this, they lack the focus and grounding generally found in systems that promote music learning in children [e.g. 3, 10].

As such, this section closes with musical performance applications, a popular application domain for tangible interaction. Such applications are normally divided into three categories: musical sequencers, which allow audio samples to be mixed and played; musical controllers, which serve as interface for random sound synthesizers; and musical instruments, capable of synthesizing and generating sound [18].

Examples of musical sequencers include Audio d-touch [5], a tangible framework that enables collaborative sequencing through customizable tokens; and Block Jam [15], a tangible polyrhythmic sequencer that relies on 26 tokens to promote collaboration and social interaction. In terms of musical controllers, a popular example is Xenakis [3] (which can also be seen as a musical sequencer). Xenakis uses mathematical models to create music, relying on tangible interaction to offer users a simple and accessible interface to its complex automatic composition techniques. Another popular example of a musical controller is Audiopad [17], a sophisticated, real-time musical application that combines the precision of knob-based tokens with the unconstrained interaction enabled by interactive tabletops (see Fig. 1). The last category of musical instruments is well characterized by the work of Levin [12], who developed a spectrographic instrument capable of supporting both pondered composition and live improvisation. Another prominent example is the reacTable [10], arguably the most successful tangible musical application to date. The reacTable offers users a robust interactive tabletop environment for both electronic and acoustic musical composition. This is complemented with very strong collaborative capabilities, allowing for several users to play simultaneously, either locally or remotely (through the use of additional reacTables).

A common thread with these examples regards their proposed target population. Most conclude that their tangible music application supports both novices and experienced musicians, and that it provides users with a musical platform that affords both learning and performing activities [e.g. 3, 5, 10, 12, 15]. This paper argues that such conclusions are normally drawn from qualitative studies or questionnaires, and suggests they comment on the natural qualities of tangible interaction – not on the particular characteristics of the tangible application in question. Examples of such conclusions that might indeed be rooted in the natural benefits of physical interfaces include: reports of intuitive and easily accessible interfaces [18]; of a fun and improved user experience [23]; and of interaction that is visible and expressive [18]. While these systems rely on the particular qualities of tangible interaction to deliver unique musical experiences, this paper argues that more empirical evidence is still required to validate their claims of universal applicability. But if indeed physical interfaces can mediate the collaboration between novice and experienced musicians, the design of these systems must rely on a broader, and still to be attained body of work. Only then can both types of users experience a seamless and sensory-rich interaction between themselves, and between them and the system.

3 jamTable

The jamTable is a tangible system that was developed with the goal of lowering the entry bar for novice users who want to perform and learn with more experienced musicians. This is enabled by allowing two users to collaborate in real-time over an interactive tabletop: while one plays a musical instrument, the other controls a simple tangible musical sequencer (in the likes of the Audio d-touch [5]). The interface for this sequencer relies on a grid of 24 music tiles, seven control tiles (one for each of the music tile columns), and a record tile (see Fig. 2). The interaction relies solely on tangible actions, with users having two sets of tokens to manage: music and control tokens. Users can start or stop recording an instrument's output by placing or removing a music token on the recording tile. Likewise, users can play recorded sounds by simply adding the corresponding music tokens to any vacant music tiles. Furthermore, by using a control token, users can change the volume and pitch, or apply the popular drive effect to any set of music tokens playing in parallel above the control tile just occupied (see Fig. 2).

The jamTable was implemented using conventional technologies, such as the Processing programming language, the reacTIVision tracking software, and the TUIO messaging protocol. It runs on an interactive tabletop based on the Diffused Surface Illumination (DSI) method, relying on a near throw projector and an IR camera positioned underneath the tabletop's screen, and a set of 850nm IR diodes around it. The surface has an area of 120 by 70 centimeters, and each token is a cuboid of 8 by 8 by 1 centimeters. These are built out of standard PVC, with iconic labels affixed to their uppermost surface and reacTIVision fiducial markers on their base. Twenty tokens were deployed: 15 music tokens that enable users to record and play sound clips; and five control tokens that can alter the playback through higher and lower pitch and volume, or through a drive effect (see Fig. 2). Finally, recording is enabled through a small directional microphone capable of recording most musical instruments. Its sound is played through four small speakers and a subwoofer.

Fig. 2. From left to right: the jamTable's interface with 24 music tiles (on top), seven control tiles (on the bottom), a record tile (on the bottom right corner), two music tokens playing in sequence (highlighted in red), and a control token (highlighted in blue) applying a higher pitch to the sound being played by the first music token of the sequence; the popular drive effect being applied to two music tokens playing in parallel; and, a closer look at a control token to lower the pitch of a recording, and three music tokens

4 Method

The goal of this paper is to present an introductory study that can provide researchers with insights regarding the real applicability of tangible applications in musical performances between novice users and experienced musicians. This was achieved by examining the performance of two distinct participant setups: Musician-Musician (MM), and Novice-Musician (NM) collaborations.

4.1 Experimental Design and Participants

The study followed a between subjects design based on two setup conditions: MM and NM paired collaborations. There were eight participants in total: six experienced musicians, and two novice users. As such, there were two MM and two NM pairs. From the eight participants all but one were male, with their ages ranging from 17 to 58 (*M* = 28.38, *SD* = 12.69). The musicians experience with guitars ranged from 2 to 43 years (*M* = 14.83, *SD* = 14.30). Finally, four of the participants were university students, two were programmers, one a guitar teacher, and one a researcher.

4.2 Procedure

Each study session consisted of two participants: one playing an instrument while the other operated an interactive tabletop. A session would start with the assignment of these roles by the researcher, respecting the experiment design and previous information regarding the participants' musical expertise – information that was not shared amongst participants. Both participants would then be given a quick introduction to the jamTable, both in terms of the tabletop's interface and how it can interact with a musical instrument. Both an acoustic and an electric guitar were made available to participants, while they were also allowed to bring their own guitars. Participants would then be allowed to casually test the system for five minutes, before the study commenced. Once started, participants had 15 minutes to accomplish a single qualitative goal: to produce a music sequence to their liking. At the end of the session participants completed a range of subjective measures.

4.3 Measures

To better gain insights in the differences in performance between pairs of MM and NM participants, both quantitative and qualitative measures were obtained in this study, as described bellow:

Quantitative. Several events were automatically recorded by the system, such as: the highest number of music tokens used simultaneously by a participant (only music tokens containing a sound clip and located on a music tile were accounted for); the number of control tokens applied; how many unique control tokens were used; the amount of time these control tokens were active (time spent in a control tile); the number of recordings performed; and the duration of such recordings (time spend by a music token in a record tile). Participants also completed the NASA TLX [8], Hart and Staveland's six-item workload questionnaire.

Qualitative. Participants were asked to report on their experience by writing up to three positive and negative remarks. They were also video-recorded while performing, with the focus being on how pairs of participants communicated between themselves, and what kind of collaborative profiles they adopted [19].

5 Results

This section includes the quantitative results from the study presented in this paper. Being an introductory study with a limited amount of participants, this section limits itself to reporting on individual values (with the exception of the mean results from the NASA TLX, shown in Fig. 3).

Table 1. Individual values for all participants interacting on the interactive tabletop (in both MM and NM pairs). Results include interaction with music and control tokens, and how sound clips recorded. All mean times are in seconds, with standard deviation in brackets.

	Musician-Musician		Novice-Musician	
	Pair 1	**Pair 2**	**Pair 3**	**Pair 4**
Most music tokens used simultaneously	2	3	6	4
Control tokens used	6	0	5	4
Unique control tokens	1	-	3	4
Duration of effects used	98.17 (104.63)	-	49.2 (55.58)	11 (4.08)
Recordings performed	8	26	11	23
Duration of recordings	24.86 (20.78)	6.57 (2.56)	15.08 (6.23)	15.25 (12.97)

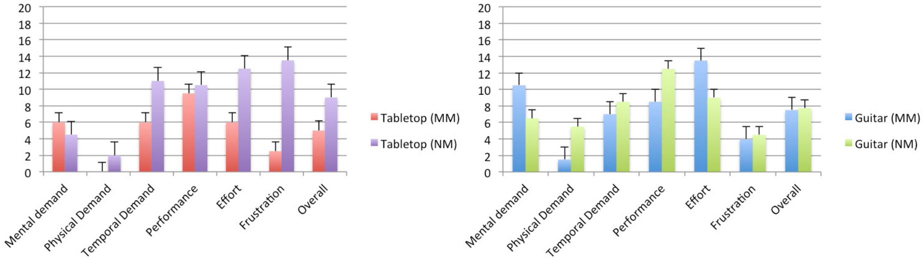


Fig. 3. The mean results of the NASA TLX. On the left the data from participants interacting on the tabletop, on the right the data from participants that played guitar.

6 Discussion

As with other work in the field of tangible musical applications [e.g. 3, 10, 12], the reaction to jamTable was generally very positive. Users found the interaction easy and fun, and appreciated how it enables and promotes the communication and collaboration between participants. But as this paper argued before, it is hard to judge the nature of such reports. While users are describing their experience with the jamTable, they are also be reporting on the natural benefits of tangible interaction (e.g. improved user experience, seamless collaboration) [18], or are simply being influenced by the novelty effects of physical interfaces [23]. As such, the following three sections provide relevant insights through the combination of qualitative and quantitative data gathered from the study with the jamTable. These are aimed at helping other researchers in the design of future studies and the musical prototypes to support them.

6.1 Learning with the jamTable

Shaer et al. [19] developed four collaboration profiles to help understand the process of collaboration on a tabletop. While these profiles are directed at participants collaborating in the same interaction space, they seem appropriate in the context of the jamTable. Through video analysis it was clear that all four pairs of participants adopted what is called as the Driver-Navigator. In this profile both users are engaged, with the driver (the participant at the tabletop) listening and discussing with the navigator (the participant playing an instrument) before committing most actions. The navigator adopts a series of actions to ground the collaboration, which in the case of the jamTable included: direct voice commands (e.g. "can you remove the last token?"); simple gestures (e.g. a nod or pointing to instruct a start or end of a recording); or a combination of both (e.g. defining a count down to start recording). This consistent behavior from participants seems to indicate that music applications like the jamTable can support a collaborative profile akin to a Student-Teacher relationship. Furthermore, by supporting a symbiotic interaction between students and teachers, the jamTable can strengthen both the motivation and commitment to learning or teaching [14].

The next natural question is of what musical concepts can actually be taught through the use of tangible musical applications such as the jamTable. While the data presented is not conclusive, it still provides valuable insights. The first refers to the different use of the control tokens by MM and NM pairs. Individual data from both participants operating the tabletop in MM pairs shows that only one used control tokens at all, and always the same one: the control token for "Higher volume". This can be partially explained by the fact that this participant was collaborating with the only musician in the study that played an acoustic guitar. On the other hand, both participants operating the tabletop in NM pairs used at least three different control tokens (from a maximum of five). Furthermore, through video analysis it is possible to observe that the decision to use or experiment with such effects came, on most occasions, entirely from these novice participants. A last observation relates to how some pairs opted to record multiple music tracks (of sometimes over 60 seconds each), while others confined themselves to recording small guitar riffs or solos (of as a little as four seconds each). This suggests that, apart from music effects, concepts such as tracks, arrangements or riffs can also be taught through the jamTable.

6.2 Performing with the jamTable

This paper argues that the key to understand how successful the jamTable was in supporting collaborative performances between novice and experienced musicians is by looking at how different were the interactions between the participants in the MM and NM pairs. Through video analysis it is possible to conclude that, despite the task's open-ended goal, all four pairs of participants managed to create a sensible musical arrangement within the time limit. Furthermore, there was no observable differences found in the Performance means of the NASA TLX: for both the novice and experienced participants using the tabletop, and the musicians playing guitar. There was an observable difference in the way paired participants would communicate though. MM pairs would rely mostly on musical terms (e.g. "Take the electric guitar out so that we have only the main riff", "Can you play that track again?"), while NM pairs would rely mostly on terms pertaining to the interface (e.g. "Remove the third token", "Leave the one in your hand out"). This seems to indicate that the physicality of the interaction (e.g. by simply being visible [11]) was indeed of great importance in supporting collaboration during the task, as it provided novice and experienced musicians with a common ground for their communication.

6.3 Future Work

Several improvements can be suggested to the jamTable. Through video analysis it was possible to observe the guitar players experience sporadic difficulty in understanding the current state of the tabletop (e.g. by peeking over). Thus, additional feedback could be created for the system, which in conjunction with an additional visual output such as a projection could improve the harmony between the users at the tabletop, the musicians, and ultimately the audience. On the other hand, by looking at the participants' qualitative feedback two more improvements can be drawn. The first relates to the control the tabletop has over the musical outcome, with participants commenting on the lack of fine tune options available (e.g. being able to trim a recording, apply more than one effect to music clips playing in parallel, or apply an effect to only one of the music clips playing in parallel). While the jamTable should remain approachable and intuitive for novice users, both multi-touch controls and a dynamic interface (instead of tile-based) could be deployed to provide features that address these comments. The second improvement is based on qualitative feedback from the experienced musicians at the tabletop who indicated that they felt unchallenged and even unnecessary at times. While the jamTable used in this study relied on only one microphone, the interface could be easily adapted to support multiple units and thus multiple musicians playing together. This adaption would create a richer and more demanding interaction between the musician at the tabletop and those on traditional instruments; the increased challenge and complexity this would result in may ameliorate such concerns.

While the possibilities mentioned above are relevant for the improvement of the jamTable, pressing future work should extend the preliminary study presented in this paper so that formal statistical analysis can be conducted. Furthermore, the contrasting NASA TLX means between novice and experienced musicians that operated the tabletop needs to be addressed (see Fig. 3). These include higher perceived effort (12.5 to 6), frustration (13.5 to 2.5), and overall workload (9 to 5, respectively). Future studies need to isolate why such contrast exists, either by exploring other learning scenarios and performance activities, or how appropriate really are physical interfaces in enabling the collaboration between novice and experienced musicians.

7 Conclusion

This paper presented the jamTable, a tangible musical application that aims to create an appealing experience that connects novice and experienced musicians. The study presented focused on how these two types of users collaborated in an unconstrained composition task. While the initial reaction to the system was broadly positive, preliminary quantitative data provided additional insights on how appropriate tangible interaction is in the support such activities. While the study presented still lacks the necessary depth (and participant numbers) to fully characterize this broad and complex topic, it provides what it argues is a necessary step towards empirical validation of many of the benefits still taken for granted in the field of tangible musical applications.

References

1. Antle, A.N.: The CTI framework: informing the design of tangible systems for children. In: Proceedings of the 1st International Conference on Tangible and Embedded Interaction, pp. 195–202. ACM, Baton Rouge (2007)
2. Bakker, S., Hoven, E.V.D., Antle, A.N.: MoSo tangibles: evaluating embodied learning. In: Proceedings of the Fifth International Conference on Tangible, Embedded, and Embodied Interaction, pp. 85–92. ACM, Funchal (2011)
3. Bischof, M., Conradi, B., Lachenmaier, P., Linde, K., Meier, M., Pötzl, P., André, E.: Xenakis: combining tangible interaction with probability-based musical composition. In: Proceedings of the 2nd International Conference on Tangible and Embedded Interaction, pp. 121–124. ACM, Bonn (2008)
4. Buxton, W., Patel, S., Reeves, W., Baecker, R.: Objed and the Design of Timbral Resources. Computer Music Journal 2, 32–44 (1982)
5. Costanza, E., Shelley, S.B., Robinson, J.: Introducing Audio d-touch: A Tangible User Interface for Music Composition and Performance. In: Proceedings of the 2003 International Conference on Digital Audio Effects, London, United Kingdom (2003)
6. Dourish, P.: Where The Action Is: The Foundations of Embodied Interaction. The MIT Press (2004)

7. Fischer, T., Lau, W.: Marble track music sequencers for children. In: Proceedings of the 2006 Conference on Interaction Design and Children, pp. 141–144. ACM, Tampere (2006)
8. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research
9. Jordà, S.: On stage: the reactable and other musical tangibles go real. International Journal of Arts and Technology 4, 268–287 (2008)
10. Jordà, S., Geiger, G., Alonso, M., Kaltenbrunner, M.: The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces. In: Proceedings of the 1st International Conference on Tangible and Embedded Interaction, pp. 139–146. ACM, Baton Rouge (2007)
11. Klemmer, S.R., Hartmann, B., Takayama, L.: How bodies matter: five themes for interaction design. In: Proceedings of the 6th Conference on Designing Interactive Systems, pp. 140–149. ACM, University Park (2006)
12. Levin, G.: The Table is The Score: An Augmented-Reality Interface for Real-Time, Tangible, Spectrographic Performance. In: Proceedings of the 2006 International Conference on Computer Music, New Orleans, USA, pp. 151–154 (2006)
13. Montessori, M., Gutek, G.L.: The Montessori Method: The Origins of an Educational Innovation. Rowman & Littlefield (2004)
14. Muller, C.: The Role of Caring in the Teacher Student Relationship for At Risk Students. Sociological Inquiry 2, 241–255 (2001)
15. Newton-Dunn, H., Nakano, H., Gibson, J.: Block Jam: A Tangible Interface for Interactive Music. Journal of New Music Research 4, 383–393 (2003)
16. O'Malley, C., Stanton Fraser, D.: Literature Review in Learning with Tangible Technologies
17. Patten, J., Recht, B., Ishii, H.: Audiopad: a tag-based interface for musical performance. In: Proceedings of the 2002 Conference on New Interfaces for Musical Expression, National University of Singapore, Dublin, Ireland, pp. 1–6 (2002)
18. Shaer, O., Hornecker, E.: Tangible User Interfaces: Past, Present, and Future Directions. In: Foundations and Trends in Human-Computer Interaction, pp. 1–137 (2010)
19. Shaer, O., Strait, M., Valdes, C., Feng, T., Lintz, M., Wang, H.: Enhancing genomic learning through tabletop interaction. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2817–2826. ACM, Vancouver (2011)
20. Ullmer, B., Ishii, H.: Emerging frameworks for tangible user interfaces. IBM Systems Journal 3(4), 915–931 (2000)
21. Wang, T.-H.: Tangible Notes: Exploring Virtual Information via Physical Objects
22. Wolf, M.: SOUNDGARTEN: A tangible interface that enables children to record, modify and arrange sound samples in a playful way. University of Applied Sciences Cologne (2002)
23. Zaman, B., Abeele, V.V., Markopoulos, P., Marshall, P.: Editorial: the evolving field of tangible interaction for children: the challenge of empirical validation. Personal and Ubiquitous Computing 4, 367–378 (2012)
24. Zuckerman, O., Arida, S., Resnick, M.: Extending tangible interfaces for education: digital montessori-inspired manipulatives. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 859–868. ACM, Portland (2005)

A Film-Type Vibrotactile Actuator for Hand-Held Devices*

Sang-Youn Kim[1] , Ki-baek Kim[2], Jaehwan Kim[2,**], and Ki-uk Kyung[3,***]

[1] Interaction Lab., Advanced Technology Research Center, Koreatech
(Korea University of Technology and Education), 1600 Chungjeolno,
Byeongchunmyun, Cheonan, Chungnam Province, Korea
`sykim@koreatech.ac.kr`
[2] Department of Mechanical Enginnering, Inha University,
100 inha-ro, Nam-gu, Incheon, Korea
`kimroll@naver.com, jaehwan@inha.ac.kr`
[3] Electronics and Telecommunications Research Institute, Daejeon, Korea
`kyungku@etri.re.kr`

Abstract. Vibrotactile actuators for small consumer electronic products, such as mobile devices, have been widely used for conveying haptic sensation to users. One of the most important things in vibrotactile actuators is to be developed in the form of thin actuator which can be easily embedded into mobile devices and to provide vibrotactile signals with wide frequency band to users. Thus, this paper proposes a thin film type haptic actuator with an aim to convey vibrotactile information with high frequency bandwidth to users in mobile devices. To this end, a vibrotactile actuator which creates haptic sensation is designed and constructed based on cellulose acetate material. A cellulose acetate material charged with an electric potential can generate vibration under the AC voltage input. It is found that the motion of the actuator can have concave or convex shape by controlling a polarity of both charged membranes and the actuator performance can be modulated by increasing level of biased electric potential. The experiment clearly shows that the proposed actuator creates enough output force to stimulate human skin with a large frequency bandwidth and to simulate various vibrotactile sensations to users.

Keywords: Vibrotactile, haptic actuator, transparent actuator, mobile device.

1 Introduction

Due to the advancement of information technology, various mobile devices have been developed and commercialized. These mobile devices usually have mechanical keypads which are primarily used for providing commands. Hence, the physical sensation for interaction with mobile devices is always the same regardless of its

---

[*] This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (grant number: 2012-0004550).

[**] Corresponding author.

[***] Co-Corresponding author.

I. Oakley and S. Brewster (Eds.): HAID 2013, LNCS 7989, pp. 109–116, 2013.

application. However, in the case where users interact with graphical contents in a mobile device, they want to feel the realistic touch sensations. As such, the haptic technology strives to enable users to feel a variety of touch sensations depending on applications. In general, haptic feeling consists of tactile sensation (sensory information acquired by pressure receptors in the skin) and kinesthetic sensation (sensory data obtained by receptors in joints, muscles, and ligaments).

Over the past decade, mechanical keypads mobile devices have been replaced by touch screens in order to increase the size of a visual display unit. Many researchers focused on tactile actuators to realize haptic sensations in a touch screen. The reason is that the tactile actuators can easily be embedded into mobile devices [1][2][3][4][5]. The first successfully commercialized actuator for mobile devices is an eccentric rotary motor (ERM). The ERM has been widely used as a vibration motor for creating alert signal in mobile devices. However, there are several critical problems in the ERM to create a variety of vibrotactile patterns and to convey the patterns to users. The most critical problem is that the response time of the ERM is too slow (over 100msec) to be perceived by users. Another problem is that the frequency and amplitude of vibration are coupled with each other.

In order to overcome the performance of ERM, researchers have been actively developing a new vibrotactile actuator which has fast response time enough to create haptic sensation in real-time. Recently, a linear resonant actuator (LRA) is a representative haptic actuator for mobile devices [6]. The LRA improves the response rate of ERM to 25ms based on the linear movement. The vibration force of LRA is strong enough to stimulate human's mechanoreceptors at its resonant frequency. However, the LRA also has a limitation in stimulating human's mechanoreceptors because a user can be conveyed haptic sensation at only around the resonant frequency. Some researchers have been studied a piezoelectric motor which is a type of electric motor based upon the physical change in shape of a piezoelectric material when an electric field is applied [7][8][9]. The piezoelectric motor can be deformed according to the amount of electric field. Since the piezoelectric motor can actuate over a broad range of frequency, these actuators can generate various tactile sensations. However, it is difficult to adopt the piezoelectric actuators in mobile devices because of its brittleness and weak transmitted force. Another typical actuator is an electroactive polymer (EAP) actuator which has low weight, high power density, and fast response [10][11]. However the actuators require considerable amount of efforts to be transparent. Opaque actuators are placed beneath a visual display unit so that a user senses vibrotactile feedback from backside of a mobile device.

Several polymer materials such as polyvinylidene fluoride (PVDF) and crystal copolymers have been intensively studied for actuators [12][13]. Kolesar et al. successfully demonstrated a polymer tactile transducer using 40 μm thick PVDF film on monolithic silicon integrated circuits [14]. Compared to an industrial PVDF polymer, as a natural polymer, cellulose has been renewed by the discovery of a piezoelectric electroactive paper (EAPap) regenerated from natural cellulose and classified as a novel electroactive polymers. A cellulose-based EAPap film has polarization behavior under applied external fields due to its dipolar structures. In this paper, we suggest a transparent film-type vibrotactile actuator, which is enough to be embedded into mobile devices, based on electroactive paper.

2 Film-Type Haptic Actuator

The Electro-Active Paper (EAPap) based on natural polymer cellulose has actuation mechanisms consisting of piezoelectricity, ion migration effect, electrets, electro static force and so on. Cellulose, which is one of the piezoelectric materials, has a molecular dipole linear chain array structure, and it can accumulate electric charge through interfacial polarization because it is porous material. The property of cellulose makes it possible to generate mechanical deformation because there is polarizing action in the cellulose. In Fig. 1, the film is a just plate with inner space empty when input voltage is not applied. However, when it is charged by input voltage, repulsion or attraction force is generated by polarization in cellulose.

EAPap is classified as semi-crystalline electrets. In the cellulose crystalline producing the piezoelectric effect, since there can be remnant polarization which disturbs charging /discharging of electrostatic phenomena, we used cellulose acetate consisting of fully amorphous to fabricate a film-type haptic actuator.

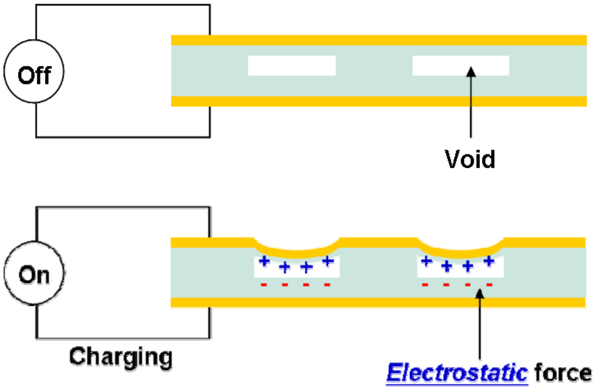


Fig. 1. Electret property of the cellulose

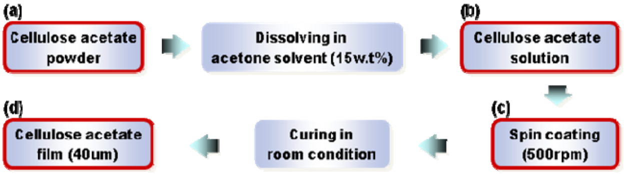


Fig. 2. Manufacturing process for fabricating cellulose acetate film

Fig. 2 shows the manufacturing process for fabricating cellulose acetate film. At first, cellulose acetate solution was made from cellulose acetate powder dissolved in acetone solvent with 15 w.t%. Well dissolved cellulose acetate solution was spin casted with 500rpm spinning speed. After curing in room condition, 40um thickness cellulose acetate film which was transparent and had good morphology could be obtained.

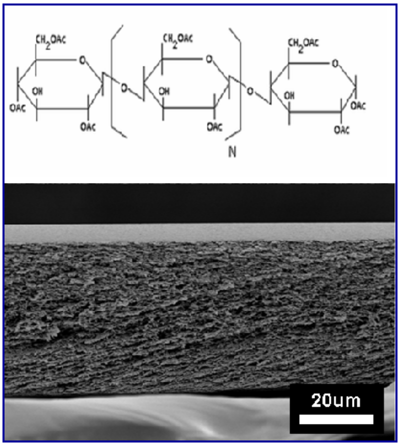


Fig. 3. Schematic and Film cross section Scanning Electron Microscope image of cellulose acetate

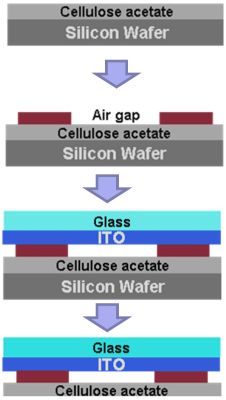


Fig. 4. Fabrication mechanism for the proposed actuator

Fig. 3, which is the scanning electron microscope image, shows the cross section of fabricated cellulose acetate films. A dense and small air gap layer, where there is easily charged with electricity, is observed in cellulose acetate. In order to maximize the air gap which can contain electric charge, we designed a new structure. Fig. 4 shows fabrication mechanism for constructing the developed film-type actuator. For fabricating film-type actuator, we placed the fabricated cellulose acetate films on a silicon wafer and applied a PDMS (Polydimethylsiloxane which is most widely used silicon-based organic polymer) mold on the films. By putting an Induim tin oxide (ITO) glass on the PDMS mold, the proposed actuator was finally constructed.

Fig. 5 shows the structure of the proposed haptic actuator based on cellulose acetate. The developed actuator has three important layers: 1) a cellulose film; 2) an ITO coated glass for actuation; and (3) a pillar (a PDMS mold) for composing air gap.

To provide electricity power to the cellulose, an electrode was coated by gold sputtering on one side of the fabricated films. In this structure, a pillar used for the formation of the air gap was produced by photo resistors on the film with the MEMS process, and an ITO coated glass was attached using polyvinyl alcohol (PVA) which is one of strong adhesive polymers. As voltage between the electrodes becomes zero, the electrostatic force between the cellulose acetate film and the ITO coated glass does zero. As the voltage between the two electrodes increases, the resulting electrostatic force between the parallel cellulose film and ITO coated glass pulls one another. Fig. 6 (a) shows the fabricated film type haptic actuator whose actual size is 40x10x0.27 mm and Fig. 6(b) shows the gold coated actuator for providing voltage input. If we use transparent electrode for providing control input to the proposed actuator, we can maintain the degree of the actuator's transparency.

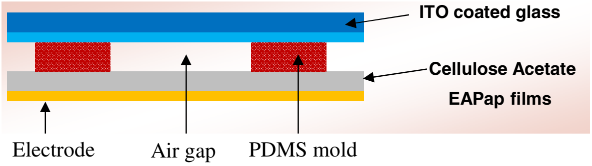
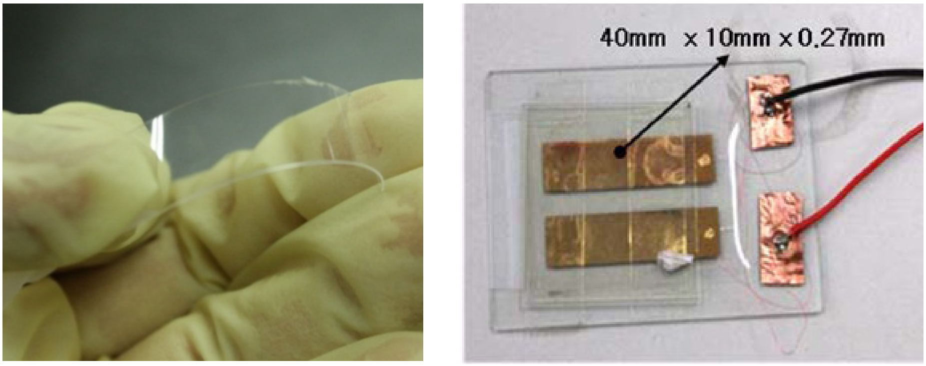


Fig. 5. Structure of film-type haptic actuator



(a) (b)

Fig. 6. (a) Fabricated film-type haptic actuator whose actual size is 40x10x0.27 mm, (b): the gold coated actuator for connecting power

3 Experiments and Results

In the haptic actuator for mobile devices, transparency can be one of the most important factors. Therefore, we measured the transparency of the proposed actuator using a UV-VIS spectrophotometer. Fig. 7 shows the transparency of the proposed actuator according to the number of spin coating. Its transparency varies from 86% to 72% as its thickness increases. If the number of spin coating becomes 20 times, the transparency of the proposed actuator rapidly goes down. Therefore, we selected the number of spin coating as 15 times.

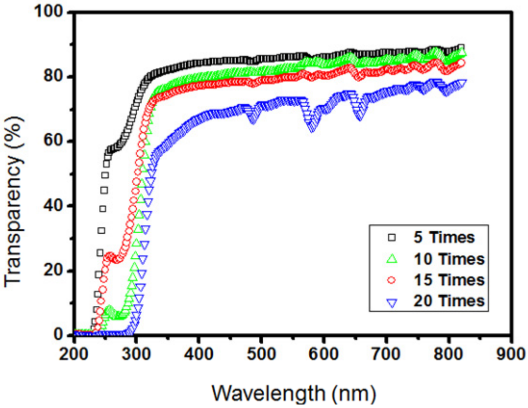
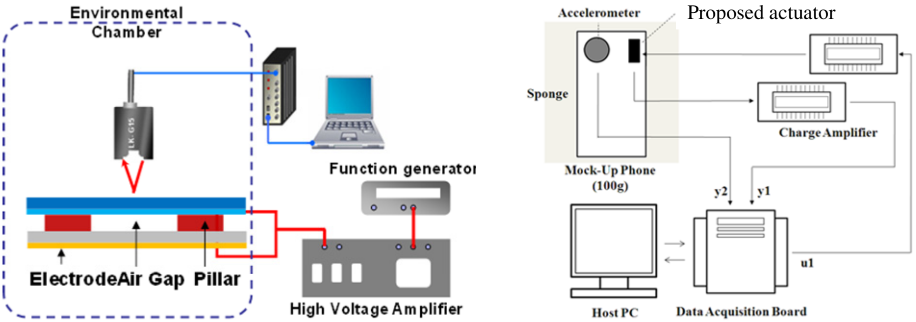


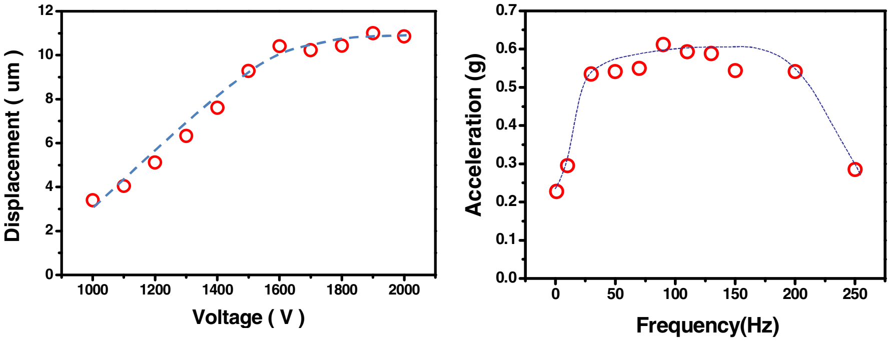
Fig. 7. The transparency of the proposed actuator according to the number of spin coating



(a) (b)

Fig. 8. (a): Displacement evaluation system consisting of laser displacement, sensor, electric field applying systems and data acquisition systems (b) : Experimental setup to evaluate the vibration force of the proposed actuator.

We constructed a displacement evaluation system as shown in Fig. 8(a). A film-type haptic actuator based on cellulose acetate was analyzed by using a laser displacement sensor (Keyence LK-G15) and data was gathered by using a signal conditioner (B&K Pulse analyzer) with a PC. To measure the displacement of the proposed haptic actuator, an AC electric potential of square wave was applied to the both sides of the sample by using a function generator (Agilent 33220A) which is connected with a high voltage power amplifier (trek 20/20). Due to environment condition sensitivity of unpackaged cellulose material, the conditions were kept in environment chamber which is about 25% relative humidity and 23°C temperature. Fig. 8(b) is the block diagram of experimental set-up to evaluate the acceleration of the proposed actuator. We measured the acceleration of the proposed actuator. A measurement system consists of a data acquisition board (NI USB-6259) connected by a host PC, which is equipped with analog-to-digital and digital-to-analog conversion channels, a charge amplifier (STT-200S-01), a mock-up of 100g which is equipped with the proposed actuator and an accelerometer (SCA1020), and a power supplier (Face International Co. TD-2). The charge signal generated at the PVDF in the proposed actuator is converted into voltage signal using the set of charge amplifier. This signal is fed into the data acquisition board via the analog-to-digital conversion channel. The control signal is calculated and converted to an analog signal. At this time, the sampling frequency is 10kHz, which is high enough in comparison with the frequency of vibrations.



(a) (b)

Fig. 9. (a) Displacement of the proposed actuator (b) Acceleration of the proposed actuator

Since the proposed actuator is operated by voltage input (not current), we changed the magnitude of the voltage in order to investigate the displacement according to the voltage input. In this experiment, we fix the input frequency at 1Hz. Figure 8 (a) shows the displacement result according to the voltage input. The displacement of the proposed actuator was increased from 3.3um (1000V) to 10.4um (1600V). Even though the voltage input is larger than 1600V, there is little change in the displacement output. Since the accelerated behavior of a haptic actuator is one of the most important factors for mobile devices, we investigated the acceleration according to the input frequency. The input frequency was varied from 1Hz to 250Hz. Figure 8 shows the accelerated behavior of the haptic actuator. The acceleration value was about 0.6g (g = $9.8m/s^2$) when the input frequency is from 30Hz to 200Hz.

4 Conclusion

In this research, a new film-type haptic actuator based on a cellulose acetate EAPap has been investigated. The electrostatic effect of the cellulose is associated with electrets properties which is charging/discharging phenomena electrically. The result of the proposed haptic actuator shows that the response time is about 15um, maximum displacement is about 11.4um, and vibration force is about 0.6 g (g = $9.8m/s^2$). Taken together, as enduring a little decreasing output displacement, the developed film-type haptic actuator is enough to create vibrotactile sensation. From this research, we can apply this result to mobile devices with improving its performance and responsibility. Currently, we are investigating the material characteristics of the proposed actuator in more detail to increase the transparency and vibration force. Even though, the vibrotactile force of 0.6 g is insufficient to provide alert signal in mobile devices, the force of 0.6 g enough to convey a variety of haptic effect to users. Therefore, it may not be so critical to create a variety of haptic sensation.

References

[1] Siewiorek, D., Smailagic, A., Furukawa, J., Krause, A., Moraveji, N., Reiger, K., Shaffer, J., Wong, F.L.: SenSay: a Context-Aware Mobile Phone. In: Proceeding of 7th IEEE International Symposium on Wearable Computers (ISWC 2003), New York, USA, pp. 248–249 (2003)

[2] Brewster, S., Chohan, F., Brown, L.: Tactile Feedback for Mobile Interactions. In: Proceeding of International Conference for Human-Computer Interaction (CHI 2007), San Jose, CA, pp. 159–162 (2007)

[3] Brown, L., Brewster, S., Purchase, H.,, H.: Multidimensional Tactons for Non-Visual Information presentation in Mobile Devices. In: Proceeding of 8th Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI 2006), Helsinki, Finland, pp. 231–238 (2006)

[4] Luk, J., Pasquero, J., Little, S., MacLean, K., Levesque, V., Hayward, V.: A role for Haptics in Mobile Interaction: Initial Design using a Handheld Tactile Display Prototype. In: Proceeding of Premier International Conference for Human-Computer Interaction (CHI), Montreal, Canada, pp. 171–180 (2006)

[5] Kwon, D.S., Yang, T.H., Cho, Y.J.: Mechatronics Technology in Mobile Devices. IEEE Industrial Electronics Magazine 4(2), 36–41 (2010)

[6] Kweon, S.D., Park, I.O., Son, Y.H., Choi, J., Oh, H.Y.: Linear vibration motor using resonance frequency. In US PATENT no. 7,358,633 B2, Assignee Samsung Elecctro-Mechanics Co., Ltd. (2008)

[7] Ernst, M.O., Banks, M.S.: Humans Integrate Visual and Haptic Information in a Statistically Optimal Fashion. Nature 415, 429–433 (2005)

[8] Wood, R.J., Steltz, E., Fearing, R.S.: Optimal Energy Density Piezoelectric Bending Actuators. Sensors and Actuators A 119, 476–488 (2005)

[9] Kim, D.H., Kim, B., Kang, H.: Development of a Piezoelectric Polymer Based-sensorized Microgripper for Microassembly and Micromanipulation. In: Microsystem Technology, vol. 10, pp. 275–280 (2004)

[10] Sarban, R., Oubaek, J., Jones, R.: Closed-Loop Control of a Core Free Rolled EAP Actuator. In: Proceeding of SPIE 7287, 72870G (2009)

[11] Ozsecen, M.Y., Sivak, M., Mavroidis, C.: Haptic Interfaces Using Dielectric Electroactive Polymer. In: Proceeding of SPIE 7647 (2010)

[12] Bune, A.V., Zhu, C., Bucharme, S., Blinov, L.M., Fridkin, V.M., Palto, S.P., Petukhova, N.G., Yudin, S.G.: Piezoelectric and Pyroelectric Properties of Ferroeletric Langmuir-Blodgett Polymer Films. Journal of Applied Physics Letter 85, 7869–7873 (1999)

[13] Ambrosy, A., Holdik, K.: Piezoelectric PVDF Films as Ultrasonic Transducer. Journal of Physics. E: Sci. Intrum. 17, 856–859 (1984)

[14] Kolesar, E.S., Dyson, C.S.: Object Imaging with a Piezoelectric Robotic Tactile Sensor. Journal of Microelectromech. System 4, 87–96 (1995)

Strike a Pose: Directional Cueing on the Wrist and the Effect of Orientation

Sabrina Panëels, Lucie Brunet, and Steven Strachan

CEA, LIST, Sensorial and Ambient Interfaces Laboratory
91191 - Gif-sur-Yvette CEDEX, France
`{firstname.surname}@cea.fr`

Abstract. Many wearable haptic devices have been developed for providing passive directional cues, in the form of belts or back displays but these systems have so far failed to make an impact in the public domain. One other potential solution is a light, discrete and aesthetically acceptable vibrotactile bracelet. However, contrary to these other systems, the wrist is subject to rotations, therefore a controversial locus for vibrotactile feedback in a navigational context. This paper presents a set of experiments aimed at both determining the basic feasibility of using this kind of bracelet and to examine to what extent the orientation of the user's wrist affects their perception of directional cues both in static and mobile conditions. It was found that changes in orientation have little negative effect overall, distraction being more of a concern.

Keywords: Vibrotactile, mobile, wearable, bracelet, haptic, wrist device.

1 Introduction

Mobile navigation has become increasingly common with the rise of GPS enabled mobile devices, which has led to the development of an abundance of pedestrian navigation applications. These applications usually deliver turn-by-turn navigational instructions via audio and visual feedback. For a mobile device user on the move, this can be a considerable distraction from the world around them. For example, visual feedback can prevent a tourist from being immersed in and discovering their new surroundings while for a pedestrian at a crossing, it can be dangerous.

Haptics offers an interesting alternative or a complementing modality [1] and using it as a means of providing eyes-free directional information is still an emerging area in terms of commercially viable applications. The transmission of tactile information has been shown to not only improve efficiency and safety but also to reduce distraction from the surrounding environment [2]. Strachan et al. [3] combined spatial audio and tactile feedback, finding that all users were able to navigate from one end of an unknown trajectory to the other using only audio and tactile feedback. Tactile feedback has also proved successful when used alone. For instance, Robinson et al. [4] developed a system aimed at supporting the discovery of geo-located information using sweeping and tilting gestures with an inertial device, with users receiving haptic

I. Oakley and S. Brewster (Eds.): HAID 2013, LNCS 7989, pp. 117–126, 2013.

feedback depending on the presence of information in the locality. They found that although the design of the interaction had some shortcomings, users were able to find targets using vibrotactile feedback alone. Pielot et al. [5] described a study with a tactile compass to convey geospatial locations with a single vibration motor and without requiring any explicit interaction. They provide direction and distance information using different patterns of vibration finding that cueing spatial locations in vibration patterns can form an effective and efficient navigation aid.

All of these systems involve a very direct pointing style of interaction with a mobile device and vibrotactile feedback delivered to the palm of the hand. This style of interaction may not be suitable in cases where a user's hands are full, whilst carrying bags or performing another task such as cycling, for example [6]; instead tactile wearable devices may be more suitable. They have often been used to provide feedback in a non-obtrusive manner and come in many flavours including gloves, shoes and belts. For example, Tsukada et al. [7] presented a tactile belt, equipped with multiple vibration motors and a GPS device, to deliver directional information around the waist. And similarly Heuten et al. [8] used a tactile belt for eyes-free navigation, finding that users were able to navigate effectively using the tactile feedback alone. Interestingly, Vélazquez et al. [9] exploited the sensitivity of the feet to convey navigational information through actuators embedded in shoes and obtained promising initial results. However, these systems have been mostly confined to the research realm; the wrist could be a solution more widely accepted as it provides a light and discrete solution, but its utility has been debated for conveying tactile feedback.

After surveying the issues surrounding the use of the wrist to provide tactile feedback and existing systems, we present a set of studies aimed first at confirming the feasibility of using the wrist as a means of providing simple tactile directional cues; and second at quantifying to what extent, if any, the change of wrist orientation hinders the ability of a wrist-mounted tactile device to aid user navigation in several realistic scenarios.

2 Wrist as a Locus

2.1 Tactile Perception and Existing Systems

Interaction involving feedback to the wrist is less common than other parts of the body and in fact there has been some debate as to its utility for receiving tactile feedback. Oakley et al. [10] described experiments examining the limitations of a vibrotactile display placed on the forearm and concluded that different arrangements of tactors can result in different levels of performance and that increasing the size of the stimulated area results in an increased perception of intensity. Their overall conclusion about the use of the forearm for tactile input is rather positive. Karuei et al. [11] explored the potential and limitations of vibrotactile displays in wearable applications, finding that wrists were one of the preferable body locations, particularly for navigation applications. Whereas Lee et al. [12] showed that the reaction time to perceive alerts on the wrist was not deteriorated by visual distraction, thus potentially making wrist-mounted tactile displays appropriate for enabling mobile multitasking.

A few wrist systems have been developed. Tsetserukou and Tachi [13] introduced BraTact, a tactile bracelet with six symmetrically arranged vibration motors, designed to alert the operator of a teleoperation system to object collisions. Schätzle et al. [14] presented VibroTac, a similar system but ergonomically improved and designed principally for forearm use. Panëels et al. [15] developed and evaluated a tactile device in the shape of a wristwatch. They found that participants encountered some difficulty discriminating directional patterns due to the small inter-distance between actuators but could effectively identify patterns based on rhythm. However, these systems were either not evaluated in a mobility context [15,14] or rely on dynamic patterns using solely rhythm and intensity [13], therefore not tackling the potential orientation issue.

2.2 The Orientation Problem

One problem with using the wrist to deliver directional instructions is that the orientation is constantly changing thus increasing the potential for confusion in the user's perception of the direction being presented. On the contrary, belt systems do not suffer from this problem since it is placed around the waist and has a consistent orientation for all users. For example with the wrist, whilst talking on the phone with the right hand, is a vibration on the right side of the wrist considered as a prompt to go right? Or is it actually to move forward? Gleeson and Provancher [16] investigated the effect of different orientations of the hand when directional shear forces are delivered to the tip of the index finger. They found that users could successfully identify directional stimuli quickly and accurately, even when the stimuli were rendered in a rotated reference frame, suggesting that the use of such stimulations on continually reorienting mobile devices is feasible. To our knowledge, the issue of wrist orientation on the perception of simple directional cues has not yet been researched.

3 Evaluation

In the context of an application designed to deliver directional cueing on the wrist, a set of studies aimed at validating the use of the wrist were conducted. First, a preliminary study tested our tactile device and the discriminability of simple cues in a static condition. The following two experiments tested the effect of changing wrist orientation on the user's perception of the directional cue in static and mobile conditions.

For all the user studies, vibrotactile cueing was provided via a band wrapped around the wrist with four actuators attached as displayed in Fig. 1. Each actuator is composed of a commercially available coin motor (Precision Microdrives 310-113), a microcontroller and a power circuit to control vibration amplitude. All the actuators are linked in series to a supervisor microcontroller which regulates the actuation level and timing; it also ensures the battery management and the Bluetooth communication with the computer. Each actuator, numbered 1 to 4, was used to provide tactile stimulation on the top, right, bottom and left of the wrist, respectively. Directional cues were presented in the form of discrete vibrations of 0.5s in length at 210Hz with an amplitude of 1.6g. They were sent to the bracelet from a tablet running Windows 7.

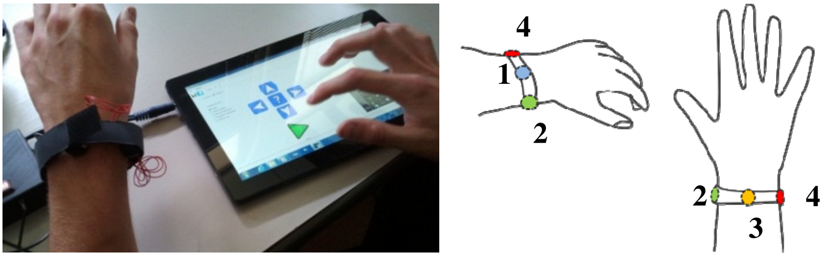


Fig. 1. Left: a user wearing the bracelet during the pilot experiment. Right: the placement and number of actuators. (1=front, 2=right, 3=back, 4=left)

3.1 Preliminary Study

To test the basic feasibility of providing directional cues with our tactile bracelet, a simple preliminary study was performed with 10 participants (8m/2f) aged between 25 and 36, all of whom were right handed and had some experience with tactile devices. The bracelet was mounted on the non-dominant wrist, to leave the dominant hand free for tapping answers on the tablet, as illustrated in Fig. 1.

Participants were first given a brief period of training with the interface, with a number of example vibrations. They subsequently felt 40 random vibrations they had to identify, i.e. 10 vibrations for each actuator corresponding to one of the directions, see Fig. 1. All participants were asked to wear headphones in order to block the sound of the vibrations and any potential effect this audio feedback may have otherwise had.

The recognition rates were 93% for the top actuator and 95% for the others: participant recognition of the four locations was excellent in these controlled conditions. This confirms that this bracelet can be used in fixed orientations to stimulate the wrist in an informative way.

3.2 Change in Orientation in Static Condition

This experiment was designed to test the effect of changing wrist orientation/user pose on the user's perception of the directional cue. The four poses, as illustrated in Fig. 2, were chosen to mimic typical mobile situations: hands in the pockets (Pose 1), talking on the phone (Pose 2), looking at the phone (Pose 3) and holding a bag (Pose 4). As in the first experiment the bracelet was strapped to the participant's non-dominant wrist. The participants were asked to hold the phone in the non-dominant hand for poses 2 and 3, justified by the fact that people regularly switch hands to write notes while using the phone, for example.

9 new male participants, aged between 22 and 45 were recruited. They received the same patterns as in the preliminary study and were asked to identify them, except that this time they would also assume the poses described above. They were introduced briefly to the vibrations and trained with the poses. They were divided in two groups: group 1 (4 participants) was asked the part of the wrist that the vibration was on, while group 2 (5 participants) was asked the direction that they felt the vibration was indicating to them, i.e. their perceived direction. Mappings of actuator to pointing direction for each pose were selected and used as a baseline from which to compare the participant performance.

Fig. 2. From left to right: The default ‘rest’ position, the ‘hands in pockets’ pose, the ‘talking on phone’ pose, the ‘looking at phone’ pose and the ‘holding bag’ pose

Both groups were prompted to change pose between each trial rather than perform a batch of trials with one pose. This was both to stop the user adjusting to a particular pose and to more accurately mimic the instantaneous response of a user to a directional cue, as is more likely to happen in a real life situation.

Results. Group 1, who were asked simply to locate the source of the vibration on their wrist, took less time in general than group 2, who were asked to indicate their perceived direction (see Fig. 3). This is particularly pronounced for pose 1 (hands in pockets) and much less pronounced for pose 4 (holding bag). This is likely due to the fact that for pose 1 the wrist is slightly rotated in the pocket and so the direction was not clearly defined in the participant’s mind. For pose 4, holding the bag naturally aligned the actuators to the directional axis making the participant’s decision easier.

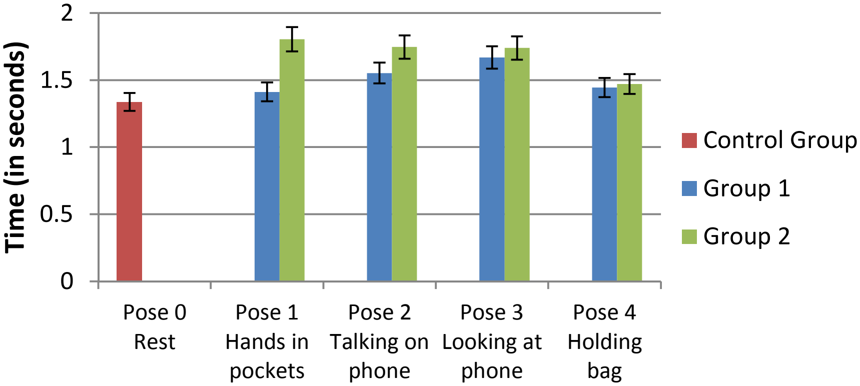


Fig. 3. The average time for the preliminary study (control group), group 1 (asked to locate the actuator position on the wrist) and group 2 (asked to indicate the perceived direction)

Fig. 4 shows the average user performance for both groups. The accuracy of the answers from group 2 was calculated from the estimated baseline (e.g. for talking on phone, a left vibration indicated front). The scores for the participants in group 1 were very similar to the "static" control group and were much more accurate than group 2. This difference is again particularly pronounced for pose 1 (hands in pockets) where the directional mapping was not clear to the participant and much less pronounced for pose 4 (holding bag) where the alignment of the wrist was well defined.

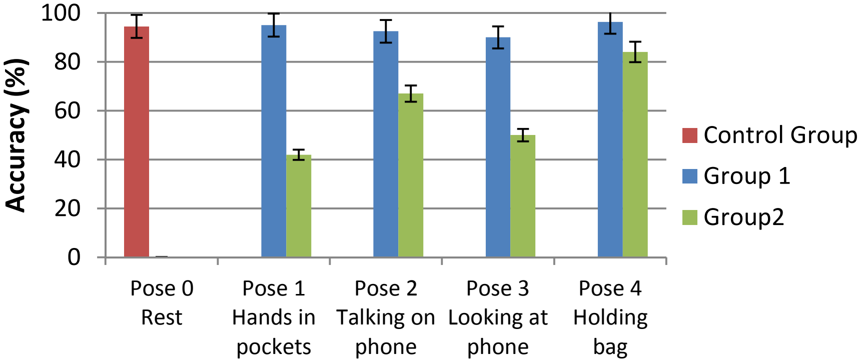


Fig. 4. The average accuracy of answers for the preliminary study (control group) and the two groups (group 1 - cues, group 2 - direction compared to the baseline) for the experiment

These initial results show that while users have little trouble locating the source of a vibration on their wrist in both static and changing poses (with scores very similar in these two conditions), some confusion is experienced when the user needs to associate the vibration to a perceived direction. Overall, these results tend to indicate that the extra cognitive work involved in associating a perceived direction does have an effect on performance. We hypothesize that locating vibrations in a frame of reference relative to the user's wrist (physical location on their own body) rather than relative to the world (location in space) is less cognitively demanding and less subject to ambiguity.

The findings of this study lean towards the absence of degradation of cue perception with wrist orientation changes. To further demonstrate this result, an evaluation in more realistic mobile settings was conducted next.

3.3 Mobility Study

To pursue the investigation of the influence of wrist orientation in the delivery of directional cues, the same study was performed while walking. 13 participants (5f/8m) aged between 24 and 56 were recruited to take part; most had some experience with tactile devices. Three of them were left-handed.

Each participant was given an introduction to the bracelet and the corresponding directions indicated by the four actuators. The users were required to wear the bracelet on their dominant wrist and walk along a set trajectory outdoors accompanied by an experimental coordinator. During the walk each user received vibrational cues on their wrist, after each of which they were required to orally indicate the direction to the coordinator or if they simply did not know, while continuing to walk. As the user did not need to answer directly on the tablet, they were permitted to wear the bracelet on the dominant wrist as would be the case in a real setting. The walk was divided into five phases. The first phase involved walking with a natural style, with the bracelet by the side. The second and third involved both carrying and dragging a suitcase, the fourth involved holding a phone at the ear and the fifth involved writing a text message (SMS). All were performed with the bracelet wearing arm and all while receiving the vibrotactile cues. The last condition was added in order to compare the effects of simple orientation changes to an orientation change with additional distractor tasks. The two distractor tasks involved writing the numbers 1 to 10 in letters and the second involved writing the name of five colleagues on the participants' own phone. 20 vibrations were sent per phase (5*20=100 stimuli). At the end, the participants were asked to fill a questionnaire.

Results. Fig. 5, showing the users' correct responses, indicates a reasonably good performance overall (poses 93.3%, SMS 79.6%). The normality of the distribution was tested using the Kolmogorov-Smirnov test which showed that for three of the positions (normal walking, holding suitcase and pulling suitcase) there is a significant deviation from normality, therefore violating the assumption of parametric data. As a consequence, the non-parametric Friedman test was used to compare the different conditions. The accuracy scores did significantly differ for the five conditions: $\chi^2(4) = 22.33$ $p<.05$. Given the small score difference between the four positional conditions, we suspected the significant difference was due to the distractor condition. Therefore the Friedman test was run again on the first four positional conditions and showed no significant difference between them $\chi^2(3) = 7.564$, $p>.05$. Subsequent post hoc Wilcoxon tests were conducted comparing the SMS condition with the other four. The two-tailed significance for all four comparisons is < 0.0125 (using the Bonferroni correction) meaning the SMS condition significantly differs from all of the poses.

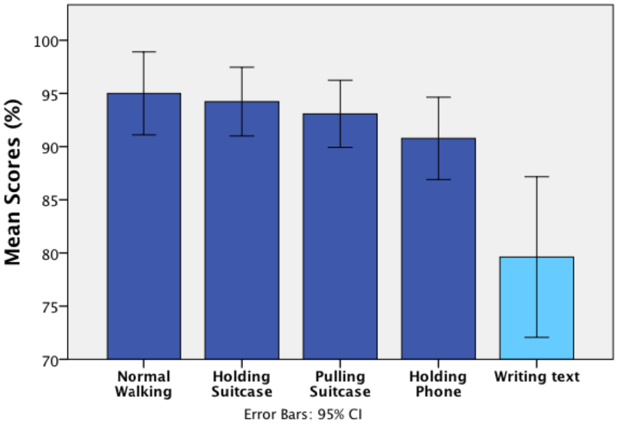


Fig. 5. User performance for each position

The overall performance is very similar in both static and mobile conditions: group 1 of the static study reached an overall performance of 93.4%, while it reached 93.3% in the mobile study with the distractor condition excluded.

The confusion matrices in Fig. 6 show that there is little confusion between patterns (<= 10%) with most recognition rates above 84%. The most common confusions for the positional conditions (≥4.6%) are between top and left for normal walking, pulling the suitcase and answering the phone, bottom and right for normal walking and answering the phone (and right/bottom), and left and bottom for answering the phone. These confusions could be due to the spatial distribution with the right/left actuators being close to the top/bottom actuators and covering less skin. Apart from the left/bottom confusion when holding the phone, none of these conditions seem to be due to the orientation issue (in the phone condition, the "left" actuator is orientated downwards). For the distractor condition (SMS), there was a lot more confusion. Top was largely confused with left (=20%), right with top (>9%), bottom with right and left (>6%). Amongst these, only the last one could be due to an orientation issue but the confusion rate is very low. It is likely then that the distraction of the SMS task played a significant role.

The questionnaire included questions about the difficulty of the main task, the localization of vibrations and their association with particular directions. Participants were asked to rate these on a scale of 1 to 6, the latter being the most difficult. Results, with averages between 2 and 3, indicated that in general participants had no problem recognising where the vibrations originated from and associating each of them with a corresponding direction (average score of 2 out of 6). The two conditions considered the most difficult were writing the SMS, due to the concentration required, and pulling the suitcase as the vibrations from the suitcase rolling interfered with the perception of the vibrations from the device. Nearly all (N=12/13) the participants replied positively to whether they thought these kind of cues would enable them to navigate more easily. They reported on the advantages of not looking at the screen constantly and focusing on potential surroundings or dangers. Nearly all (N=12/13) participants suggested the provision of repetition, either on demand or as part of the pattern design. Many (N=9/13) suggested the introduction of a signal announcing the pattern to grab the user's attention before any kind of recognition is required, in particular when the user is distracted. Finally, many (N=8/13) participants thought that the patterns could be better designed for discrimination, for instance a more complex combination of tactile parameters could be used (amplitude, location, rhythm).

These qualitative findings support the quantitative results in the sense that the orientation alone does not seem to pose a problem for the recognition of direction, and provide some leads for improvement of the feedback design.

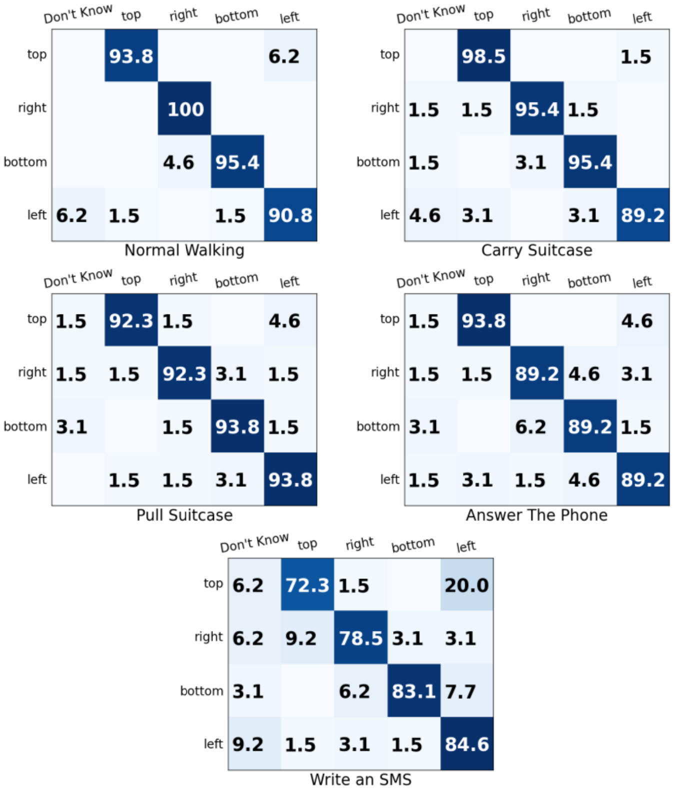


Fig. 6. Confusion Matrices for the 5 poses described. User responses are from left to right

4 Conclusion

We have investigated the feasibility of using a vibrotactile bracelet to convey simple directional cues. In particular, we have conducted a set of studies examining the potential issue of orientation change both in static and mobile conditions. The results show that the orientation of the user's wrist does not have a strong effect on the presentation of tactile directional cues. There was however, a clear effect of distraction as discussed in [10]. In future studies, more emphasis will be placed on the role of distraction in a real navigation scenario as it is likely to be an important factor in such tasks. Overall, we can conclude after this experimental validation, combined with positive feedback from our participants, that the wrist is an excellent candidate for the provision of passive vibrotactile feedback.

References

1. Heikkinen, J., Olsson, T., Väänänen-Vainio-Mattila, K.: Expectations for User Experience in Haptic Communication with Mobile Devices. In: Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI), pp. 28:1–28:10. ACM, Bonn (2009)

2. Pielot, M., Poppinga, B., Heuten, W., Boll, S.: Pocketnavigator: studying tactile navigation systems in-situ. In: SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 3131–3140. ACM, Austin (2012)
3. Strachan, S., Williamson, J., Murray-Smith, R.: Show me the way to Monte Carlo: density-based trajectory navigation. In: SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 1245–1248. ACM, San Jose (2007)
4. Robinson, S., Eslambolchilar, P., Jones, M.: Sweep-Shake: Finding Digital Resources in Physical Environments. In: Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI), pp. 12:1–12:10. ACM, Bonn (2009)
5. Pielot, M., Poppinga, B., Heuten, W., Boll, S.: A tactile compass for eyes-free pedestrian navigation. In: Campos, P., Graham, N., Jorge, J., Nunes, N., Palanque, P., Winckler, M. (eds.) INTERACT 2011, Part II. LNCS, vol. 6947, pp. 640–656. Springer, Heidelberg (2011)
6. Crossan, A., Williamson, J., Brewster, S., Murray-Smith, R.: Wrist rotation for interaction in mobile contexts. In: Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI), pp. 435–438. ACM, Amsterdam (2008)
7. Tsukada, K., Yasumura, M.: ActiveBelt: Belt-type Wearable Tactile Display for Directional Navigation. In: Mynatt, E.D., Siio, I. (eds.) UbiComp 2004. LNCS, vol. 3205, pp. 384–399. Springer, Heidelberg (2004)
8. Heuten, W., Henze, N., Boll, S., Pielot, M.: Tactile Wayfinder: A Non-Visual Support System for Wayfinding. In: Nordic conference on Human-computer interaction: building bridges (NordiCHI), pp. 172–181. ACM, Lund (2008)
9. Vélazquez, R., Bazan, O., Magaña, M.: A Shoe-Integrated Tactile Display for Directional Navigation. In: IEEE/RSJ International Conference on Intelligent RObots and Systems (IROS), pp. 1235–1240. IEEE Press, St. Louis (2009)
10. Oakley, I., Kim, Y., Lee, J., Ryu, J.: Determining the Feasibility of Forearm Mounted Vibrotactile Displays. In: Haptic Symposium, pp. 27–34. IEEE Press (2006)
11. Karuei, I., MacLean, K.E., Foley-Fisher, Z., MacKenzie, R., Koch, S., El-Zohairy, M.: Detecting vibrations across the body in mobile contexts. In: SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 3267–3276. ACM, Vancouver (2011)
12. Lee, S.C., Starner, T.: BuzzWear: Alert Perception in Wearable Tactile Displays on the Wrist. In: SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 433–442. ACM, Atlanta (2010)
13. Tsetserukou, D., Tachi, S.: Efficient Object Exploration and Object Presentation in TeleTA, Teleoperation System with Tactile Feedback. In: WorldHaptics, pp. 97–102. IEEE Press, Salt Lake City (2009)
14. Schätzle, S., Ende, T., Wüsthoff, T., Preusche, C.: Vibrotac: An ergonomic and versatile usable vibrotactile feedback device. In: International Symposium on Robot and Human Interactive Communication (RO-MAN), pp. 670–675. IEEE Press, Viareggio (2010)
15. Panëels, S., Anastassova, M., Strachan, S., Van, S.P., Sivacoumarane, S., Bolzmacher, C.: What's Around Me? Multi-Actuator Haptic Feedback on the Wrist. In: WorldHaptics. IEEE Press, Daejeon (2013)
16. Gleeson, B., Provancher, W.: Mental rotation of directional tactile stimuli. In: Haptics Symposium, pp. 171–176. IEEE Press, Vancouver (2012)

The Time Machine – An Inclusive Tourist Guide Application Supporting Exploration

Kirsten Rassmus-Gröhn, Delphine Szymczak, and Charlotte Magnusson

Department of Design Sciences, Lund University, PO Box 118, 221 00 Lund, Sweden
`{kirre,delphine.szymczak,charlotte}@certec.lth.se`

Abstract. In the present paper we describe an inclusive tourist application, The Time Machine, that allows users to explore an environment or be guided along a trail while experiencing virtual sounds and get information (visual and auditory) at key locations in the environment. We report the application design and preliminary results from evaluations in a city environment. Results indicate that the Time Machine is fun, stimulating and usable for a wide range of users.

Keywords: mobile, interaction, haptic, audio, augmented reality, navigation, tourism, user experience.

1 Introduction

The small screen on mobile devices poses problems both in perceiving the information from a small visual source, but also in keeping focus on the environment instead of on the mobile screen. By making better use of additional (non-visual) modalities, it is possible to create applications where the user can keep attention on the environment even when being “on the go”. In the present paper we describe an inclusive tourist application, The Time Machine, that allows users to explore an environment or be guided along a trail while experiencing virtual sounds and get information (visual and auditory) at key locations in the environment. We report the application design and preliminary results from evaluations in a city environment. This article extends on [18] by looking at explorative usage. The application has also been tested in rural/hiking environments but these tests lie outside the scope of the present paper.

2 Related Work

When an augmented reality application is intended to be used “on the go” it is clear that the looking at and attending to the screen is a problem (since you have to attend to/look at your environment in order to avoid obstacles, people etc). Despite this, the bulk of mobile augmented reality applications developed rely primarily on screen-based presentations. That this is not unproblematic is exemplified by [2] where the authors report “Paradoxically the game encourages looking at the screen more than the surroundings”. Even “Backseat playground” [1] which makes use of audio for

I. Oakley and S. Brewster (Eds.): HAID 2013, LNCS 7989, pp. 127–136, 2013.

most of its interaction use on-screen visual elements for feedback, and the authors report that this drew the eyes of its users towards the screen rather than to the surrounding environment. Another problem often occurring in visual augmented reality is the difficulty of having a smooth juxtaposition of virtual elements on the real world image. GPS, compass and other sensor inaccuracies cause the virtual parts of the image to move around in an unconvincing way [2, 20].

Current navigation systems in mobile phones are based on screen interaction. The user is usually expected to look at the map to find out where to go (e.g. Google Maps). The interest in non-visual modalities to guide the navigation is increasing in the research community, explained in part by the need to reduce the load on visual attention in mobile situations [11]. Several systems have been devised using sound as guidance. An early attempt was the Audio GPS[3]. The Swan project [21] gives auditory feedback about routes and context aimed for visually impaired persons. The ONTRACK [4] system uses 3D audio and music to guide the user, while the Soundcrumbs [6] uses chosen audio tracks of varying volume according to the user's phone bearing. Audio Bubbles [8] gives auditory feedback about near-by landmarks. Others have explored vibrations to convey the information. Sweep-Shake [14] uses vibration feedback also to let users get information on close-by points of interest. It was then evolved to support users' navigation as described in "I did it my way" [15]. The Tactile Wayfinder [12] explores the use of a vibrating belt to give directional information. PointNav [5] gives both orientation and navigation support through vibrations and speech feedback. More recent work include the NiviNavi inclusive audio game [7] and the TimeMachine [18] which reports on an inclusive tourist application making use of haptic and auditory information. In DigiGraff [9] it is explored how location based information can be used in a more social setting. Specific studies have also been made on how tactile information can provide both distance and directional information [17]. A more playful approach where users explore without guidance was used in the Virtual Excavator [10]. For more exploratory navigation, different kinds of soundscapes have been created, by communities or artists. The Urban Sound Garden [19] and the Tactical Sound Garden [16] are two examples.

3 The Time Machine Application

The Time Machine is a virtual tourist guide application. It allows for the following of pre-designed trails, which also can be shared, using a GPX format with a HaptiMap XML schema to store trails. Trail following is based on pointing and scanning with a mobile device and receiving feedback with sound and/or vibrations. Sound windows playing localized sounds are used to help global navigation and get a feeling for the history of the environment. Additionally, the Time Machine app allows for free exploration of the surrounding points of interest, also loaded through a sharable GPX file. It is then possible to choose one point of interest and be guided to it, with the same modalities that the trail following provides. A haptic map is also implemented, which makes it possible to explore the map by touch, feeling vibrations on streets and your own position and hearing the street names read out loud. This implementation is based on the Touch Over Maps design [13]. Both users with visual impairments and full vision can follow trails and hear pre-recorded information about historical sites as well as explore the surrounding points of interest and can be guided to them, by using the multimodal components explicitly developed for the Time Machine application. The application furthermore contains an accessible multiple choice quiz game.

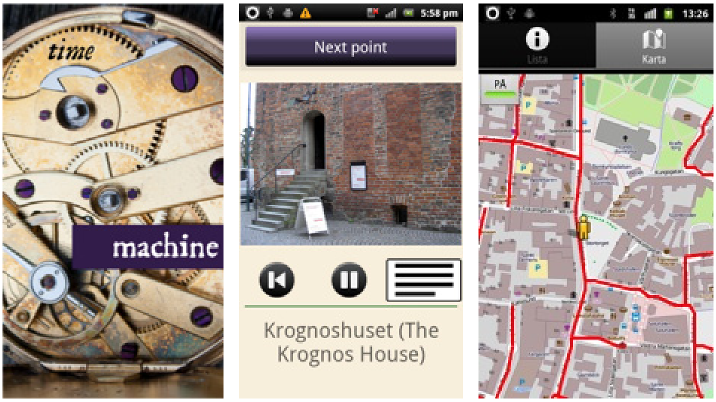


Fig. 1. Time Machine User Interface Screens: Splash screen, trail point screen and map screen

To be guided to a location you use the phone to point in front of you, rotating to find your target direction (figure 2a). The direction of your goal will be indicated by vibration bursts that repeat more frequently as you get closer. The default guiding is by vibration. It is possible to select sound or speech guiding in the settings menu. On arrival your phone will notify you with vibrate and the arrival sound. Your virtual guide will start telling you about the location. You can pause and repeat the guide. You can choose to read the text on screen as well by switching to text presentation. You click ”Next point” when you want to continue.

For the exploration you your finger on the screen and point the device in different directions (Figure 2b) while keeping the finger on the screen.

Fig. 2. a) Guiding (left) b) exploration (left)

While the finger is in touch with the screen the device acts as a virtual torch or scanner, telling you which point of interest can be found in the direction the phone is pointing. To explore points close to you, you keep your finger close to the torch (bottom of screen). To explore further away, you slide the finger away from the torch.

To select a point, you remove your finger from the screen. The last point of interest found is selected, and you can choose to be guided to the point or receive information about it. The guiding works in the same way as for a trail. The arrival behaviour is also the same (except there is no “next point”).

4 Evaluation

The Time Machine has been tested by in total 9 visually impaired adults and 2 sighted participants in the city centre of Lund, Sweden, making use of the exploratory version of the applications: scanning for points-of-interest and walking towards them. Since the Time Machine app aims at users with a variety of visual capabilities we have done several tests, with sighted, visually impaired and elderly users. In the test reported in the present paper users were asked to use the free explore function, where the user scans (by pointing with the mobile phone) for historical points of interest and chooses one to be guided to. The evaluation was carried out in the medieval city centre of Lund. Visual observation during test was carried out. The test was also documented by SenseCam pictures, hung on the participant, to see what he/she was turned towards, an audio recording to record the surrounding sounds and also the conversation between the participant and an accompanying person. The application itself also logged variables. The points of interest were collected with the help of the archaeology department, the culture historical museum and the building preservation program in the city of Lund. The test task was to use the Time Machine app to scan around for points of interest (see figure 3), choose a point, and be guided to it. There were 34 different points to choose from, and there were 46 sound windows scattered among the POIs. The phones used were SonyEricsson (now Sony) Xperia Neo and Arc. These were chosen because of the physical buttons for back, home and menu.

Participants were carrying out the evaluation by walking together with an accompanying person. This was done for two reasons: first, for the safety of the user with visual impairment, and second to aim for capturing conversation between the participant and the accompanying person about the usage of the application, rather than asking a single person to think aloud. The accompanying person was in most cases a test leader who could also help with technical problems, but 2 participants had accompanying persons with them.

An introduction was given to the participants about the Time Machine and also the use of the “back” button in Android. This introduction was adapted to the user’s previous knowledge of smart phones. These instructions were recorded in order to collect information about the improvement of the help section of the application, and also to document the learning process. Then, the user carried out two learning tasks, one to scan and choose a specified point and let themselves be guided to it, and another to scan and choose a historical sight to examine from afar.

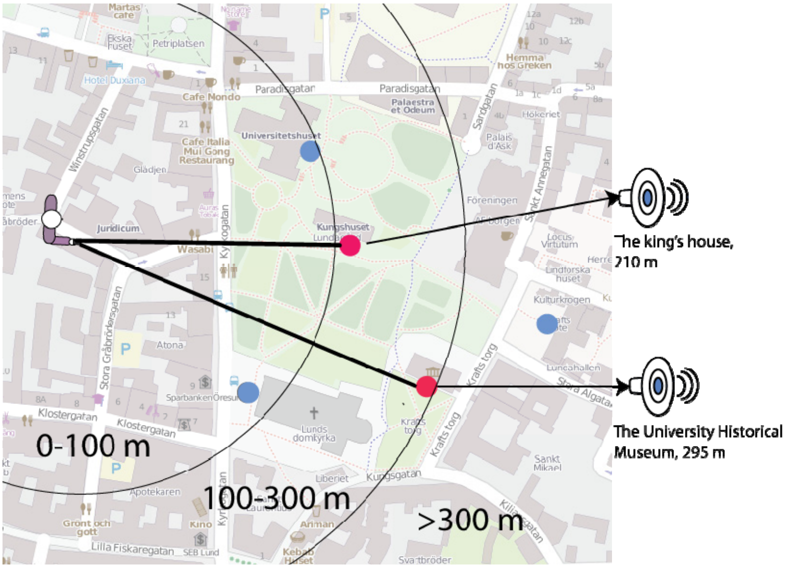


Fig. 3. Scanning concept in the explore mode

Half of the participants were given a phone configured with turn-by-turn routing. The routing employed was the adapted kind of routing that routes over open areas as well as in the road network. The other half of the participants were given a version with directions given "as the crow flies".

After the learning/training session, the participants explored freely, and were asked to chose and walk to at least 2 different points of interest in the city. The final point after the 2 freely chosen ones was the place for re-union for the debriefing and interview. The test procedure in the city took approximately 1 hour for the introduction and actual test and between 1 and 1.5 hours for the post-test interviews and questionnaires.

After the test, the NASA RTLX workload was employed, a demographic questionnaire filled in, and a semi-structured interview carried out. Additionally, a word list was used, aiming to capture the more subjective feelings toward the app.
The results presented in this paper are based on the visual observations done during the test combined with the post test materials – NASA TLX, questionnaire, interview and word list.

4.1 Participants

The participants with visual impairment who carried out the evaluation were recruited through the local visual impairment organization (SRF Skåne). Two sighted users participated as well, internally recruited, one of which uses a wheelchair. This user used a SHAKE SK7 attached to a hat for the pointing to allow for hands free use of application while moving. In total 11 persons tested the application. Of these 9 had varying degrees of more severe visual impairments (low vision including one blind person), 3 in addition had hearing problems and as was stated above one person used a wheelchair. The ages of the test participants ranged from 42 to 73 years old. 6 men and 5 women did the test.

5 Preliminary Results

Of the 11 primary participants (the accompanying persons are not counted), all users except one carried out all test tasks. The one person who did not succeed, had technical problems with the hardware in the phone. Three participants with visual impairment commented particularly on the need for a longer training period, while several others also indicated that they needed more time to really get used to the app. Also, the question "On a scale from 1-7 how would you rate your wish to use the app again?" (1 indicating low wish, and 7 a high wish) was answered with *5-7* by all of the participants with visual impairment.

All participants filled in the NASA-RTLX task load rating for the entire activity. The perceived mental demand on doing the tasks and using the app varied between 2 and 6. The perceived physical demand is rated low (1-3), except for the two sighted users. The temporal demand is rated quite low (1-3) by all except one participant, who had technical difficulties and wasn't able to fulfil all tasks. The participants' subjective rating on performance is generally high (4-7), with the same outlier as for the temporal demand. In general, participants felt they needed to put some, but not an excessive amount of effort into the use of the app (3-5). Two participants rate the frustration higher than middle, the participant with technical problems and the participant who was using a wheelchair and needed both hands to use it while simultaneously using the touch screen to choose scanning distance. The guiding with vibration was considered the least demanding task. This is in line with the qualitative answers that users gave.

The users were asked to check all words in a list of 106 words (balanced to contain both positive and negative properties and feelings), that they felt applied to the app. From these they were asked to mark the 5 most important words. The most prominent describing words, using a weighted figure with the important words counted twice, were: Fun, Stimulating and Usable.

6 Discussion

Although several users with visual impairment would have wished for more training, in the interview they commented on their learning while using the app. While rating the concentration level using the guiding, one user said: "*In the beginning, I would need to concentrate on a 7 level to follow the guiding, in the end only a 2 level*". Another user, while commenting the use of the haptic/touch user interface, said: "*It was hard in the beginning, otherwise OK*".

Holding a device and pointing with it while walking while also using a cane or guide dog, proved to be possible, but it had drawbacks. When participants were using a cane, it was held in the dominant hand. This forced the visually impaired participants to hold and interact with the phone in the non-dominant hand, being somewhat awkward, but still possible. The thumb was mostly used for interaction with the screen, but while scanning (standing still) two hands could be used. One user was afraid that the mobile phone would be stolen, since she was pointing with it and holding it quite lightly.

To scan the user had to keep one finger in touch with the screen (scanning stopped when the finger was lifted). This design was selected based on results from the PointNav application [5] where buttons activating the scanning had caused problems, but the new design was also problematic. One user complained that he got too warm and wanted to lift the finger once in a while, resulting in a point being selected. Another user wanted to actively select the scanning range by lifting the finger when the speech synthesis read the distance, also resulting in a selection event. Participants quite quickly learned the use of the Android "Back" button, and recovered from those errors easily. The participant using a wheel chair had another problem. Needing two hands to rotate the wheel chair to scan around, it wasn't possible to put a finger on the screen. Thus, several different ways of actively selecting points and also using the scanning without constant screen contact are needed.

The pointing around oneself with the phone to explore what is there (introduced already in the Sweep-Shake [14]) is relatively intuitive, although unusual for the users. It was observed at times, that participants didn't turn all 360 degrees, and also that they tried to move the finger on the screen sideways to scan. This was also seen when points were close to each other. The distance filtering by moving the finger up and down on the screen was well received, but still the points of interest sometimes occluded each other. This led to problems finding a specified point, or finding it again when you lost it. Additional filtering is needed, and two users suggested list search and one user regular keyboard search. Since it seems that users intuitively move the finger on the screen for scanning, a combination of point-scanning and screen scanning seems like an interesting approach to pursue in the future.

Rotating imposes a problem for users with visual impairment. The turning forces them to give up their reference direction. There are some solutions to the problem, one being to add a compass with cardinal directions, which was suggested by at least one user during the interview. Another solution could be to add personalized reference points in the database of points-of-interest to allow users to better orient themselves (that POIs can be used in this way by persons with visual impairments was seen already in the PointNav evaluation [5]).

The vibration guiding used by all visually impaired participants and all except one sighted participant (who used sounds instead) was well received. Comments like: "*Strong and good*" (1 user), "*Pleasant*" (2 users) and "*The vibrations were good*" (1 user) were uttered during the interview, and 8 users spontaneously answered "the guiding" when asked about the easiest part in the application. One user occasionally had problems feeling the vibrations. This was on a particularly windy day, and the problems seemed to occur at windy places.

Regarding routing, it seemed that users with severe visual impairment in general preferred to have routing on. In the case where users didn't have the routing condition, they would comment on the weirdness of being led through houses. However, even when routing is on, this can of course occur due to the inaccuracies in

GPS positioning, and also occurred at times. This preference may also be a matter of previous experience – existing GPS applications route turn by turn.

Both routing and showing directions "as the crow flies" allowed users to cross over open areas, such as squares. For a visually impaired user, crossing a square is unusual. To orient themselves adequately, they usually follow the outline of the square, and if a regular GPS device is used, it will route along the perimeter of the square. One user particularly commented on it: "*I was able to walk across the square. It was great.*"

To be less vulnerable to GPS inaccuracies, the POI arrival has a 10 m offset (i.e. in a radius of 10 m from the position the user is considered to have "arrived"). This is a weakness if you want to guide users very close to an object, but has nothing to do with the interaction as such - it is a consequence of current GPS inaccuracies. With better position accuracy users could be guided closer to the objects.

The sound windows were active on all test occasions. Three of the users reported hearing problems, but only one of them asked for the sound windows to be removed when asked about improvements. No other participants reported that they in general were annoyed with the sounds.

Some improvements were asked for, since sounds illustrated different activities that had been going on at a particular place in historical times. Due to the exploratory usage the sounds were no longer part of a trail, and since they were separate from the POIs, there was no natural connection between the sounds you heard while exploring and the goal that you had chosen. Participants in the city were curious about the sounds, and wanted to be able to explore the sounds and understand what they were and why they were there (in the earlier tests [18] where the sounds were part of a trail this had not been a problem since the sounds connected to the content of the trail).

The users had several suggestions for improvement. They requested for example:

- Information about the sound windows. They captured the user's interest, yet there was no additional information about them unless you found the POI where they were mentioned.
- Information about what was on the points of interest in present time, not just the historical information.
- Regular detailed information about distance to target (in metres). For sighted users the increasing intensity in the vibration patterns has been seen to work well, but several visually impaired users wanted more specific distance information.
- A better "where am I?"-function (the touch over map had this function but it was not used)
- Navigation in speech also (to notify of turns ahead for example). The application has simple speech based guiding (keep right/left, straight ahead and turn around) as an alternative, but this does notify the user of turns ahead).
- A function that guides the user back to the starting point, to be invoked at any time

7 Conclusion

In this paper we report on a novel way of combining scanning as in the Sweep-Shake [14] with on-screen gestures to filter points depending on distance. The application in general is seen to be appreciated by the test users, but there is also room for improvement. The guiding was seen to be working well also for users with visual impairments (previous tests [18] had shown this to be the case for elderly sighted users). The scanning/exploration was well liked, but improvements are needed to be able to better deal with occlusion and close lying objects. The results indicate that directions both "as the crow flies" and through turn by turn routing (at least with the more advanced pedestrian routing used) are useful. To sum up the Time Machine is a good example of how it is possible to make an inclusive tourist guide application that is fun, stimulating and usable for a wide range of users.

Acknowledgements. We thank the EC which co-funded the IP HaptiMap (FP7-ICT-224675). We also thank VINNOVA for additional support.

References

1. Bichard, J., Brunnberg, L., Combetto, M., Gustafsson, A., Juhlin, O.: Backseat Playgrounds: Pervasive Storytelling in Vast Location Based Games. In: Harper, R., Rauterberg, M., Combetto, M. (eds.) ICEC 2006. LNCS, vol. 4161, pp. 117–122. Springer, Heidelberg (2006)
2. Carrigy, T., Naliuka, K., Paterson, N., Haahr, M.: Design and evaluation of player experience of a location-based mobile game. In: Proceedings of the Nordic Conference on Human-Computer Interaction, NordiCHI. ACM (2010)
3. Holland, S., Morse, D.R.: Audio GPS: spatial audio in a minimal attention interface. In: Proceedings of the 3rd International Workshop on HCI with Mobile Devices (2001)
4. Jones, M., Jones, S., Bradley, G., Warren, N., Bainbridge, D., Holmes, G.: ONTRACK: Dynamically adapting music playback to support navigation. Personal Ubiquitous Comput. 12(7), 513–525 (2008)
5. Magnusson, C., Molina, M., Rassmus-Gröhn, K., Szymczak, D.: Pointing for non-visual orientation and navigation. In: Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries. ACM (2010)
6. Magnusson, C., Rassmus-Gröhn, K., Breidegard, B.: Soundcrumbs - Hansel and Gretel in the 21st century. In: Altinsoy, M.E., Jekosch, U., Brewster, S. (eds.) HAID 2009. LNCS, vol. 5763. Springer, Heidelberg (2009)
7. Magnusson, C., Waern, A., Rassmus-Gröhn, K., Bjernryd, Å., Bernhardsson, H., Jakobsson, A., Salo, J., Wallon, M., Hedvall, P.-O.: Navigating the world and learning to like it - mobility training through a pervasive game. In: Proceedings of the MobileHCI. ACM (2011)
8. McGookin, D., Brewster, S., Priego, P.: Audio Bubbles: Employing Non-speech Audio to Support Tourist Wayfinding. In: Altinsoy, M.E., Jekosch, U., Brewster, S. (eds.) HAID 2009. LNCS, vol. 5763, pp. 41–50. Springer, Heidelberg (2009)
9. McGookin, D.K., Brewster, S.A., Christov, G.: DigiGraff: Considering Graffiti as a Location Based Social Network. In: CHI 2012. ACM, Austin (2012)

10. McGookin, D., Vazquez-Alvarez, Y., Brewster, S., Bergstrom-Lehtovirta, J.: Digitally Glueing the Past: Multimodal Interaction for Un-Stewarded Historical and Archaeological Sites. In: Digital Engagement 2011, Newcastle, UK, November 15 - 17 (2011)
11. Oulasvirta, A., Tamminen, S., Roto, V., Kuorelahti, J.: Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM (2005)
12. Pielot, M., Boll, S.: Tactile Wayfinder: Comparison of Tactile Waypoint Navigation with Commercial Pedestrian Navigation Systems. In: Proceedings of the Pervasive Computing. Springer (2010)
13. Poppinga, B., Magnusson, C., Pielot, M., Rassmus-Gröhn, K.: TouchOver Map: Audio-Tactile Exploration of Interactive Maps. In: Accepted for Poster Presentation at MobileHCI 2011, Stockholm, Sweden (2011)
14. Robinson, S., Eslambolchilar, P., Jones, M.: Sweep-Shake: finding digital resources in physical environments. In: Proceedings of the MobileHCI 2009: 11th International Conference on Human-Computer Interaction with Mobile Devices and Services. ACM (2009)
15. Robinson, S., Jones, M., Eslambolchilar, P., Murray-Smith, R., Lindborg, M.: “I did it my way”: moving away from the tyranny of turn-by-turn pedestrian navigation. In: Proceedings of the MobileHCI 2010: 12th International Conference on Human Computer Interaction with Mobile Devices and Services. ACM (2010)
16. Shepard, M.: Tactical sound garden toolkit. In: Proceedings of the ACM SIGGRAPH 2007 art gallery. ACM (2007)
17. Szymczak, D., Magnusson, C., Rassmus-Gröhn, K.: Guiding Tourists through Haptic Interaction: Vibration Feedback in the Lund Time Machine. In: Isokoski, P., Springare, J. (eds.) EuroHaptics 2012, Part II. LNCS, vol. 7283, pp. 157–162. Springer, Heidelberg (2012)
18. Szymczak, D., Rassmus-Gröhn, K., Magnusson, C., Hedvall, P.-O.: A real-world study of an audio-tactile tourist guide. In: Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI 2012). ACM, New York (2012)
19. Vazquez-Alvarez, Y., Oakley, I., Brewster, S.: Urban Sound Gardens: Supoprting Overlapping Audio Landmarks in Exploratory Environments. In: Proceedings of the Multimodal Location-based Techniques for Extreme Navigation Workshop, Pervasive (2010)
20. Wetzel, R., Blum, L., Broll, W., Oppermann, L.: Designing Mobile Augmented Reality Games. Springer (2011)
21. Wilson, J., Walker, B.N., Lindsay, J., Cambias, C., Dellaert, F.: SWAN: System for Wearable Audio Navigation. In: Proceedings of the 11th IEEE International Symposium on Wearable Computers (2007)

Author Index