haid 08

15th -16th September 2008 in Jyväskylä, Finland

Haptic and Audio Interaction Design,
3rd International Workshop

Proceedings of the

3rd International Haptic and Auditory Interaction Design Workshop

Volume II

Edited by: Andrew Crossan & Topi Kaaresoja

Conference Chairs: Antti Pirhonen & Stephen Brewster

Table of Contents

EXPLORING SOUND TO ENHANCE LEARNING OF ABSTRACT SCIENCE CONCEPTS

Taciana Pontual Falcão
London Knowledge Lab - Institute of Education
23-29 Emerald Street, WC1N 3QS, London UK
t.pontual@ioe.ac.uk

Sara Price
London Knowledge Lab - Institute of Education
23-29 Emerald Street, WC1N 3QS, London UK
s.price@ioe.ac.uk

ABSTRACT
Auditory feedback plays an important role for many educational technologies. This work is part of an on-going study that aims to explore whether the aggregation of sound to visuals on a tabletop may enhance users' interaction and be beneficial for the learning of abstract science concepts. In this paper, we present findings from our initial study in which a number of sounds were elicited, and their mappings to scientific concepts were explored and verified, both with and without visuals. Issues were raised regarding attributes (like pitch), levels of abstraction, duration and continuity of sounds, as well as connection between sounds and visuals.

Author Keywords
Auditory feedback, tangible learning environments, science concepts.

ACM Classification Keywords
H5.2. Information interfaces and presentation: User interfaces. K.3.m Computers and education: Miscellaneous.

CONTEXT
This study is part of an ongoing project investigating how tangible environments affect the way that learners interact with and understand scientific ideas [6]. Tangibles offer the opportunity to augment everyday interaction with the world, combining digital information, in the form of sound, narration, images, text or animation, with concrete objects. As part of this research a tabletop environment has been built. A first application allows for the exploration of physical processes like light absorption, reflection and transmission. Although the visual mode is often a predominant form of representation, the potential for audio and tactile modes in tangible computing requires a broader understanding of their role for learning. Auditory feedback is an important part of many educational technologies. In learning contexts sound has been used to: produce sound patterns through body movement [1]; assist the visually impaired in exploring graphical interfaces [8] or spatial information [5]; mediate understanding of large amounts of abstract data [3]; show progress in task solving activities [4]; and help conveying invisible concepts [7]. Physics of light is particularly challenging, as its behaviour cannot be naturally perceived (eg. specific frequencies of light being reflected or absorbed by a surface). Although visual augmentation can convey some aspects of invisible processes, the role of sound for supporting learning of abstract scientific phenomena is under-researched.

STUDY DESIGN
The main goal of this study was to inform the use of sound in the design of a tangible environment, allowing future investigation of its role in promoting learning. Although it is relatively easy to design sounds for referents that have direct mappings, presenting absolute data with sound is difficult [2]. In order to use sound in the context of the physics of light, types of sounds to represent light phenomena were first elicited (Phase 1). Next, the mappings of chosen sounds with the desired concepts were verified, with and without visuals (Phase 2).

Phase 1: Each participant from group 1 was asked for a verbal description about the "quality" [9] that a sound should have when representing reflection, absorption and transmission of light. Such qualities could include concepts of continuity, bouncing off, or prohibition. Each participant then selected sounds from a set of samples to associate with each of the phenomena being investigated, e.g. whistle (transmission), brake (absorption) and hitting a hard surface (reflection). The aim of Phase 1 was to elicit general descriptions for types of sounds people would a priori link to the concepts involving light, as well as mappings between specific sound samples and the phenomena.

Phase 2: Participants from group 2 listened to the sounds selected in Phase 1 and reported which phenomenon (reflection, absorption or transmission) they would associate to which sound. The aim of this task was to check if the sounds could evoke the expected phenomena. This is particularly important when considering the use of the system by the visually impaired. Participants from group 2 then worked with the tabletop, exploring the interface with and without the sounds chosen in Phase 1, to investigate the effect of combined sound and visuals on interaction.

RESULTS AND DISCUSSION
Preliminary studies were performed with 8 children aged between 7 and 16, to enable exploration of age-related issues to inform future studies. . The flexibility of the tangible environment allows for a wide age range, as different levels of complexity can be explored for similar activities, according to the forms of representation used. Verbal descriptions of qualities of sounds given in Phase 1

were quite uniform. Reflection was associated with the ideas of hitting and bouncing off; transmission was linked to passing or forcing through, in a continuous manner; absorption evoked ideas of evaporating, gathering, unclear mixture of sounds, or changes in pitch and volume. Sounds chosen in the next task generally matched verbal descriptions. For reflection, popular sounds were: cymbal, gunshot and spring bouncing. For transmission, continuous sounds of whistle, cars passing by, and laser swords were chosen. For absorption, chosen sounds, in general, were abstract and continuously changing pitch, except for a chimes sound chosen by the younger children. A number of interesting themes emerged from the studies:

Familiarity with concepts: sounds chosen for reflection had more similarities among participants. In Phase 2 children correctly identified which sound represented reflection. For absorption and transmission, choices were more varied and children did not easily associate the sounds to the expected phenomenon. It may be that children are more familiar with the concept of reflection (e.g. through mirrors) from school and everyday life, as even older children were unfamiliar with the concept of transmission.

Age differences: when chosen sounds were played with the visuals, younger children classified it as a good matching, while the older ones rejected some sounds they had chosen. They justified their reaction declaring that sounds needed to be more abstract, so that they would not be recognized as something concrete (like a doorbell), or evoke associations with real-world things not related to the concepts. Older children tended to go for varying pitches in sounds, which matched their preference for abstraction.

Duration, continuity and inference: short sound samples were used in the study, and participants sometimes took the duration of the sound into consideration, although the focus was on the type of sound. For example, sounds for transmission were supposed to be long, in their opinion. In addition, the visuals run continuously, whereas the sound samples were played for a very short time. When choosing a spring sound for reflection, participants were thinking of the idea of bouncing. However, the sound did not match the visual representation of reflection that consisted of the light beam reaching an object and arrows or ripples continuously coming off it. For absorption, children approved their choice, as they actually wanted a sound that would stop, as a way of showing light had been absorbed. This shows children had in mind a sequence in time to represent the phenomena, i.e., light travels from the source, then it reaches the object, then it is, say, absorbed, and finally the process "finishes". However, processes related to light rather occur in a continuous manner, and are represented in that way in our visual application. Exploring the use of sounds therefore called attention to a learning aspect that had previously gone unnoticed when only considering visuals.

FUTURE WORK

The preliminary findings presented here will be further investigated to define a set of sounds to be added to the visual application, considering characteristics like pitch, volume, duration, continuity, annoyance, and level of abstraction. Future studies could involve parallel (rather than disconnected) choice of visuals and sounds, aiming to find a proper match. An accurate mapping between sounds and visuals can also enable the use of the interface by the visually impaired, following the principles of universal design. As for these users auditory feedback may replace visuals (although the idea is not to design a solely audio-based system, which would have different implications), it is extremely important to find the most intuitive set of sounds to represent the concepts. Studies with blind children may be needed to verify the effectiveness of sounds in this context. Integration of sounds and visuals will then be more deeply analyzed concerning its effect on children's awareness, attention, interaction with information and understanding of science concepts.

ACKNOWLEDGMENTS

This research is supported by the EPSRC: EP/F018436.

REFERENCES

1. Antle, A. N., Droumeva, M., Corness, G. Playing with the Sound Maker: do embodied metaphors help children learn? In *Proc. IDC* (2008).
2. Brewster, S.A. Chapter 12: Non-speech auditory output. In Jacko, J. and Sears, A. (Eds.) *The Human Computer Interaction Handbook* pp 220-239 (2002).
3. Childs, E. The Sonification of Numerical Fluid Flow Simulations. *Proc. ICAD* (2001).
4. Droumeva, M., Antle, A. N., Wakkary, R. Exploring ambient sound techniques in the design of responsive environments for children. In *Proc. TEI* 2007.
5. Parente, P.; Bishop, G. BATS: The Blind Audio Tactile Mapping System. In *Proc. ACM SRC* (2004).
6. Price, S., Sheridan, J. et al. Towards a framework for investigating tangible environments for learning. IJART Special Issue on Tangible and Embedded Interaction (accepted 2008).
7. Randell, C., Price, S., et al. The Ambient Horn: Designing a novel audio-based learning experience. *Pers. and Ubiq. Computing Jour.* 8, 3, 144-161 (2004).
8. Sharlin, E., Watson, B., et al. The Tangible Pathfinder: Design of a Wayfinding Trainer for the Visually Impaired. In *Proc. Graphics Interface* (2004).
9. Somers, E. Abstract sound objects to expand the vocabulary of sound design for visual and theatrical media. In *Proc. ICAD* (2000)

A Pilot study on audio induced pseudo-haptics

Charlotte Magnusson
Department of Design Sciences, Lund University
P.O. Box 118, 221 00 Lund, Sweden
charlotte.magnusson@certec.lth.se

Kirsten Rassmus-Gröhn
Department of Design Sciences, Lund University
P.O. Box 118, 221 00 Lund, Sweden
kirre@certec.lth.se

ABSTRACT

In this paper we describe a small pilot study designed to explore the possibility of spatial pseudo haptic effects induced by artificial audio feedback. Four basic audio designs were investigated: pan, continuous since tone, virtual harp and noise. For reference a visual pseudo haptic effect together with real haptic effects were also included. The results indicate noise and spatially distributed discrete sounds like in the virtual harp to be promising, while both pan and continuous frequency changes did not appear to be distinct enough for this purpose.

Author Keywords

Pseudo haptics, audio, haptic, interaction design

ACM Classification Keywords

H5.2 User interfaces

INTRODUCTION

It is well established that manipulation of the visual feedback may generate pseudo-haptic illusions [1]. Since visual pseudo haptics is by design something that cannot be experienced by somebody with impaired vision (or not able to use vision for other reasons) it is interesting to see if similar effects can be obtained from manipulating auditory gestural feedback. That pseudo-haptic illusions may be generated by manipulating contact sounds has been investigated earlier also within the ENACTIVE community [2]. It has also been known for a long time that manipulation of friction sounds can cause haptic sensations [3, 4]. These types of sounds have also been used for pseudo haptic interaction design in [5].

This type of effects could potentially be used for making interfaces more accessible/easy to use both for persons with visual impairments and persons using mobile devices with small screens, and thus we find it interesting to investigate if one can also induce more spatial pseudo-haptic effects (similar to the visual effects reported in [1]).

RESEARCH QUESTION

As a first step in this direction we did a small pilot test with 12 persons. This pilot test was designed to provide us with preliminary insights on non visual pseudo haptic effects. The basic question to be answered was:

⇒ Which type of audio feedback can make users report that they can feel a haptic difference between different areas in space?

TEST SETUP

Since this was very much a first exploratory pilot, we decided on a very simple setup where the user basically only explored one dimension (along the x axis) and where different combinations of visual, auditory and haptic feedback was given. The visual and haptic feedback was included to get some comparative information with visual pseudo haptics and with real haptic effects. All in all, we explored seven different designs. For each design there was one case where the feedback did not change and one where it changed resulting in a total of 14 designs to be tested. The different feedback designs were presented in random order. For simplicity we only investigated one dimension and we used a PHANToM 1.0 premium for the tests. The use of a PHANToM enabled us to provide real haptic effects, although it was also used as simply a position sensor. The environment is shown in figure 1. The blue block was a virtual slab which provided a surface for the test persons to slide the PHANToM stylus along.

Figure 1. The haptic environment. The red ball followed the PHANToM cursor horizontally in the visual and haptic designs. The yellow transparent area indicate a possibly different area.

The persons were instructed to slide the PHANToM along the virtual surface in the left-right direction and to tell if they thought things felt different to the right of the middle line (the yellow area in figure 1).

Seven different designs were tested. Four of these involved audio feedback where changes were made in the pan, the frequency of a since tone, the string locations of a virtual harp and silence vs. noise. The visual feedback involved manipulating a ball which moved horizontally and the haptic effects were changing inertial forces or changing friction. The different designs are summarized in table 1.

Nr	Sense	Feedback	Change
1	Hearing	A sine tone propotional to the logarithm of the position and increasing left to right. No ball.	The position dependence to the right is 10% of the position dependence to the left.
2	Hearing	A virtual harp where the invisible strings were quite close. No ball.	Twice as long distance between the strings to the right.
3	Hearing	Panning of a musical sound source (a source we had used earlier for navigational tests). No ball.	The change in the pan rate to the right is 10% of the change in the pan rate to the left.
4	Hearing	Noise/silence. No ball.	Noise played to the right.
5	Vision	Ball moving with the stylus	Slower movement to the right (the ball now moves only 30% of the stylus distance).
6	Touch	Weak inertial force/no force. The ball follows the stylus.	Inertial force (0.5 times the velocity) to the right.
7	Touch	Friction. The ball follows the stylus.	Increased friction to the right. The static friction increases a factor 3 and the dynamic friction increases a factor 1.75.

Table 1. Different test designs.

RESULTS AND CONCLUSION

The results are summarized in figure 2. One can see that for the noise feedback (4), the visual (5) and the haptic feedback (6 and 7) more persons judge the ones that really are different to be different. The sine tone feedback and the pan feedback appeared not to be distinct enough for people to really note the differences, while there are a few more persons who note a difference for the virtual harp (2). The number of users in this small pilot was not enough to draw safe conclusions (the statistical analysis we did despite this showed the only significant difference was between the cases when there really was a force and when there was no force and no manipulation of the feedback).

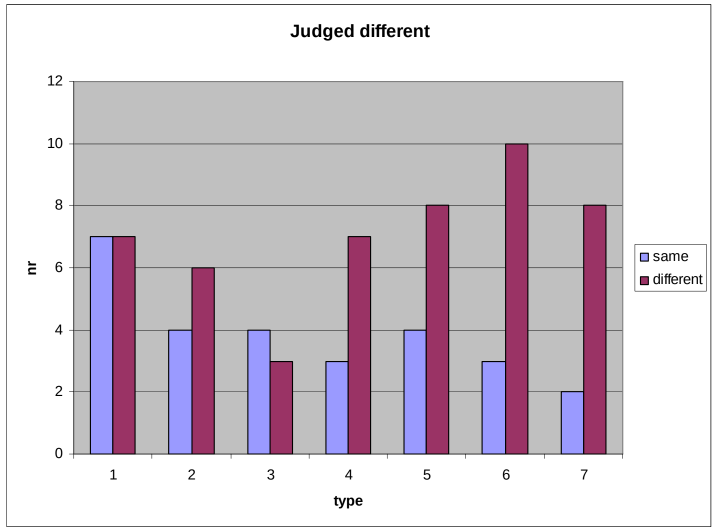


Figure 2. The number of persons judging things to feel different. For the blue bars the design did not actually contain any difference (same) while the red bars show the cases that really were different.

Still, the results combined with qualitative experience by the authors indicate that both playing noise (which has already been explored to some extent in [5]) and manipulating the distance between the strings of a virtual harp may be fruitful to investigate further. In this context it should be noted that eye-hand coordination is something all sighted persons train most of their lives. Experience also teaches us to associate sounds from impacts or touching to material properties, but to get some kind of spatial pseudo haptic effect for ear-hand coordination with artificial sounds, some training to establish the mapping is likely to be needed.

REFERENCES

1. Lécuyer, A., Burkhardt, J.M., Etienne, L. (2004) Feeling bumps and holes without a haptic interface: the perception of pseudo-haptic textures. In proc of CHI 2004, 239-246.
2. Rochesso, D., Avanzini, F., Rath, M., Bresin, R., Serafin, S., Contact sounds for continous feedback, Proceedings of the int. workshop on interactive sonification, Bienfield, 2004
3. Jousmaki V, Hari R (1998), Parchment-skin illusion: sound-biased touch. CURRENT BIOLOGY 8 (6): R190-R190
4. Guest S, Catmur C, Lloyd D, Spence C (2002) Audiotactile interactions in roughness perception. EXPERIMENTAL BRAIN RESEARCH 146 (2): 161-171
5. Fernstrom, M. Brazil, F. Bannon, L. (2005), HCI design and interactive sonification for fingers and ears, Multimedia, IEEE, 12 (2) : 36-44.

Haptices and Haptemes – Environmental Information through Touch

Riitta Lahtinen
Pastori Jussilaisentie 4 a 12, 00400, Helsinki
Finland
riitta.lahtinen@kolumbus.fi
mobile: + 358 40 522 4201

Russ Palmer
Pastori Jussilaisentie 4 a 12, 00400, Helsinki
Finland
rpalmer2@tiscali.co.uk

ABSTRACT

This research [1] is systematic, longitudinal process and development description of communication using touch and body with an acquired deafblind person. The research consists of observational and analysed written and video materials mainly from two informants´ experiences during period of 14 years using body and touch. The research describes the adaptation of Social-Haptic methods between a couple, and other informants´ experiences, which have been collated from biographies and through giving national and international courses.

A touching event can be an independent action or it may happen in interaction with another person. In an independent event a person touches an object or he/she is touched. In communication situation it is a question of an active, meaningful touch, when the sender touches the receiver. In this research haptic [2,3,4,5] is strongly connected to the sharing of environmental information and holistic orientation of the body. Haptics is a generic concept and tactile is a more limited subconcept. Social-Haptic touch includes interaction between skin, active touch and environmental orientation-touch.

Author Keywords

deafblindness, hapteme, haptic, haptice, movement, social-haptic communication, social-haptic confirmation system, tactile, touch

Analysing haptices and haptemes

When the hearing and sight deteriorates due to having an acquired deafblind condition, communication consists of multi-systematic and adaptive methods. A person's expressive language, spoken or Sign Language, usually remains unchanged, but the methods of receiving information could change many times during a person's lifetime.

The development of Social-Haptic communication was related to the emergence of new concepts. [6,7,8] Single messages shared by touch on the body are ***haptices***. Along with verbal information they are related to holistic movements, sharing of functionality and perceiving the environment as well as orientation. A haptice consists of variables of touch i.e. ***haptemes***. A hapteme is received through a body channel, in which the whole body is transmitting touch information. A hapteme is a grammatical variable related to touch, an element for building and identifying haptices and of separating them from each other. Figure 1 shows the relations between Social-Haptic communication, haptices and haptemes.

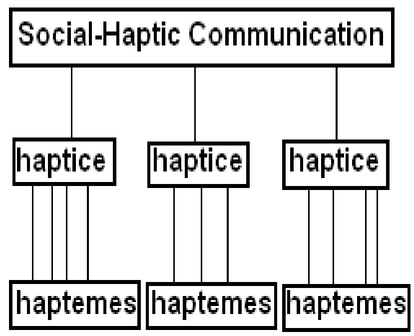


Figure 1. The relation between Social-Haptic communication, haptices and haptemes.

The use of haptices gained a new value when confirmation was connected. Three turns are included in confirmation and advancing of discussion; message, answer (response) and feedback [9]. This became evident in the use and development of haptices. The confirmation emerged for the first time in the use of feedback when producing double feedback, and later in the movements of different body parts [7]. In the use of haptices by other informants, the confirmation was evident between interactive parties. The message was received, interpreted and evaluated, after which feedback was given. The DeafBlind (DB)-informant expressed a message, the Hearing Sighted (HS)-informant answered and the DB-informant commented on answer. In terms of haptices*, the turns advanced as follows:

DB and HS enter the space using a guiding grip.

HS-informant: *"Auditorium."*

The DB-informant nods and looks around.

* Ħ symbol in describing haptices and haptemes

HS-informant: Ħ draws the ground plan of the auditorium onto the back of the DB- informant Ħ *"Door-lecturer-free seats."*

DB-informant: *"Let's stay here beside the door."*

HS-informant: Ħ feedback onto the shoulder with the hand Ħ

DB-informant: Ħ feedback onto the guiding hand with his hand Ħ.

The confirmation can be seen in various forms during the use of haptices; verbally, visual seeking of the target, acting in personal space, moving and feedback behaviour. These show that the confirmation will extend through the body. The bodily feedback functions as a comment and response to a given message. The place of touch can be pre-information on the action. Thus the contents of the message are related to the movement and action of the receiver's touch area, for instance touching the arm when "starting to leave" means, that the guiding arm is ready. Pauses are part of the structure of haptices. Pauses in touching communicate on waiting for feedback.

A single haptice can consist of many layers of messages at the same time (simultaneous multidimensionality); information on the emotions of the describer, environment and use of time (urgency). From a psycholinguistic perspective the feedback and other social quick messages given by touch are related in real time to the function of the sentence, situation or action. Expanding of the meaning indicates an expanded content of haptices. Along with the use of haptices, the messages started to get multi-dimensional meanings and sentence levels. The development of the use of haptices made body areas wider and movements smaller. Emotional expressions came along in addition to basic meanings of the messages. Five stages of emotional differentiation were identified as very light, light, medium, heavy and very heavy touch.

Haptices are made from haptemes that determine which regulations are analysed. When defining haptemes the definition, classification and varied meanings of touch were discovered. Haptemes include sharing a personal body space, meaning of touch-contact, context and using different communication channels. Communication distances are classified as exact distance, estimated distance and touch distance. Physical distance can be termed as very long, long, medium or very close. Social body space includes the body areas involved in sending and receiving haptices and applying different types of contacts.

One or two hands can produce messages by using different hand shapes and orientations. This research classifies how the body can be identified into different areas such as body orientation, varied body postures, body position levels, social actions and which side of the body is used. Spatial body space includes environmental and situational elements. Haptemes of movements are recognised as the direction of movements, change of directions on the body, directions between people, pressure, speed, frequency, size, length, duration, pause, change of rhythm, shape, macro and micro movements.

Research describes haptices in different situations enhancing sensory information and functioning also as an independent language. Haptices include social-haptic confirmation system, social quick messages, body drawing, contact to the people and the environment, guiding and sharing art experiences through movements. Haptices give the possibility to share art, hobby and game experiences.

A new communication system development based on the analysis of the research data is classified into different phases. These are experimental initiation, social deconstruction, developing the description of Social-Haptic communication and generalisation of the theory as well as finding and conceptualising the haptices and haptemes. The use and description of haptices is a social innovation, which illustrates the adaptive function of the body and perceptual senses that can be taught to a third party – and be inspiration for the design of haptic systems.

REFERENCES

1. Lahtinen, R. M. Haptices and haptemes. A case study of developmental process in social-haptic communication of acquired deafblind people. Doctoral Dissertation. Tampere, Cityoffset Oy, 2008.
2. Gibson, J.J. Observations on active touch. Psychological Review Vol 69(6), (1962) 477- 491.
3. Gibson, J. J. The senses considered as perceptual systems. Prospect Heights, I.L.Waveland Press, Houghton Mifflin Co, Boston, USA, 1966.
4. Gibson, J.J. The senses considered as perceptual systems. Greenwood Press, Connecticut, USA. Original 1966, Boston, Houghton Mifflin, USA, 1983.
5. Katz, D. The world of touch. Krueger, L.E. (Edit. and transl.). Lawrence Erlbaum Associates, USA, 1925/1989.
6. Lahtinen, R. Holistic and interactive communication methods. In Peckford, B. & Hawcroft, L. (Edit.). Proceedings of an International Symposium on Interpreting for Deafblind People. Durham, UK, Prontaprint, 64-65, 1999.
7. Lahtinen, R. Development of the holistic social-haptic confirmation system. A case study of the yes & no - feedback signals and how they become more commonly and frequently used in a family with an acquired deafblind person. Licenciate Theses, University of Helsinki, 2003
8. Lahtinen, R. Sosiaaliset pikaviestit. Suomen Kuurosokeat ry. Julkaisuja A4. Tampere, Cityoffset, 2005.
9. Schegloff, E., Jefferson, G. & Sacks, H. 1977. The preference for self-correction in the organization of repair in conversation. Language Vol 53(2), (1977), 361-382.

Pressure-Based Input for Mobile Devices

Stephen Brewster and Michael Hughes
Glasgow Interactive Systems Group, Department of Computing Science
University of Glasgow, Glasgow, G12 0AX, UK
stephen@dcs.gla.ac.uk www.dcs.gla.ac.uk/~stephen

ABSTRACT
This paper describes the design and evaluation of a pressure-based keyboard for a mobile device to investigate the possibilities of pressure as a new method of input. A soft press on the device's touchscreen generated a lowercase letter, a hard press an uppercase one. Results from an evaluation showed that different designs of pressure input could be fast or accurate and were not significantly affected by the movements cause by walking. This suggests that pressure could be a useful technique for future mobile interactions.

Author Keywords
Pressure, input, keyboard, text entry, multimodal.

ACM Classification Keywords
H5.2. User Interfaces: *Haptic I/O.*

INTRODUCTION
Interactions on touchscreen mobile devices are often constrained by the size of the devices; they must be small and light otherwise people will not carry them. This can seriously restrict user input and output, limiting the kinds of interfaces that can be offered to users. The provision of new mobile applications and services is forecast to be a large growth area, but will be restricted if users cannot use their devices in flexible ways. New techniques are needed to increase the bandwidth of communication between the user and the mobile device to support these new applications and services. In this paper we discuss the use of pressure as a new form of input and show how it can be used in a keyboard to improve user performance when typing.

Humans have very precise control over pressure, especially at the fingertips. It is key in tasks such as picking up objects, drawing or playing a musical instrument. Pressure for interaction has been studied in the context of graphics tablets [5] and mice [1] but not for touchscreen mobile devices where input is commonly limited to simple button presses and where there will be many benefits from a richer form of dynamic input.

BACKGROUND
Research into tactile (skin-based) interaction for mobile devices has grown over recent years due to the limitations of screen size and the fact that audio is not always appropriate for output. Touch provides a rich channel through which to communicate. There are several different sub-modalities within touch. Tactile feedback (commonly via vibrotactile stimulation of the skin) is the best understood and used in HCI. Pressure is part of the same sensory system and could be used for mobile input.

Srinivasan and Chen [6] studied force using the index finger. Participants had to control the force applied to a force sensor under a range of different conditions (including an anesthetized fingertip to examine the effect of removing tactile feedback). They suggest that pressure-based interfaces need to have a force resolution of at least 0.01N in order to make full use of human haptic capabilities. Mizobuchi *et al.* [4] suggest that ranges of 0-3N are comfortable and controllable and users can reliably apply around 5-6 levels of pressure [4, 5]. All of these tests have been in static situations using graphics tablets or specially adapted mice. The movements of walking or the bumping of a train may have a serious impact on the amount of force people can consistently apply and may significantly reduce the number of usable pressure levels for real-world mobile interactions.

Ramos *et al.* have done some of the key work in HCI on pressure input on graphics tablets. They looked at how pressure might be used in applications such as video editing and proposed a set of 'pressure widgets [5] for tasks such as zooming and selection based on pressure. Irani and colleagues [1] looked at adding pressure to a mouse for desktop interactions, and at how multiple pressure sensors could be used. Their results showed that users were slower when they had to press harder, they also showed that a click selection technique was faster than a dwell, although dwell was the most accurate. Li *et al.* [2] looked at pressure for mode switching in pen-based interfaces. One important aspect of pressure was that it could be used to mode switch whilst continuing to perform an action, such as writing. One factor they noticed was that one user used much lower overall pressure than everyone else. They suggest that different users may need different 'pressure spaces'.

HARDWARE
The device that we used for our work was the Nokia N800 Internet Tablet (see Figure 1). This is a small (75 x 144 x 13mm), light (206g) Linux-based mobile device with a touchscreen, normally operated with a finger or stylus. However, it is possible to read the pressure values generated when the user presses on the touchscreen, meaning that

we can use pressure for input without making any modifications to the standard device.

Figure 1: The Nokia N800 Internet Tablet.

EVALUATION

To test the usefulness of pressure as an input technique we designed a pressure-based keyboard where a soft press generated a lowercase letter and a harder press an uppercase one. This removed the need to move to the shift key, potentially reducing targeting errors when typing.

We implemented two of the pressure techniques shown to be the best by Ramos *et al.* [5]: Dwell and Quick Release. For Dwell the user had to apply the force for 0.5s before a selection was made. Audio feedback was given when pressure has been applied for the appropriate duration. For Quick Release the user pressed the keyboard with the appropriate pressure and released immediately. In this case, a different sound was played to confirm whether an upper or lowercase letter had been typed. We gave a dynamic graphical representation of the level of pressure being applied (and the case of the letter) with a pressure meter that popped-up beside the key being pressed (Figure 2).

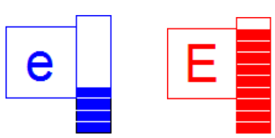


Figure 2: The graphical pressure meter.

In our evaluation we compared Quick Release, Dwell and a Standard keyboard which used a shift key to change letters in both static (sitting) and mobile (walking) settings, giving 6 conditions. The experiment used a repeated measures, within-subjects design with conditions presented in a counterbalanced order. The 12 participants had to type 6 randomly chosen phrases from MacKenzie's standard phrase set [3] in each condition. The phrase set does not contain uppercase letters, so we capitalized the first letter of each word in each of the phrases.

RESULTS AND CONCLUSIONS

Overall timing results showed that Dwell was significantly slower than the Quick and Standard keyboards (as was found by Ramos *et al.* in their study with graphics tablets). This is not surprising as the user had to stay on a key for 0.5s to make a selection. The Standard and Quick keyboards had similar performance, except when mobile where Quick was significantly faster than the Standard. Uppercase letters also took longer to enter than lowercase ones for the Quick and Standard keyboards, for Dwell they took the same time.

The Dwell keyboard produced significantly fewer errors than the Quick or Standard ones (again, as Ramos found). There was no effect for mobility; users were not less accurate when walking than sitting. For Dwell there was no difference in the error rates of upper or lowercase letters. For Quick and Standard more errors were made with lowercase than uppercase letters.

We measured subjective workload using NASA TLX workload scales and found no difference for workload between any of the keyboard types. This suggests that pressure does not significantly change the workload needed to operate a mobile device.

This work has shown that pressure can be an effective form of input for mobile devices. The Quick Release keyboard was fast and the Dwell keyboard caused fewer errors than a standard keyboard with a shift key. If speed and error rates can be optimized we may be able to create a pressure keyboard that would be better in both regards than a standard one. This performance also remained good when users were mobile. We had anticipated that the movements of walking might have made it harder to apply pressure consistently, with the device and user both moving. This turned out not to be the case, suggesting that pressure could be a useful method of interaction for mobile users.

ACKNOWLEDGMENTS

This work was supported by EPSRC Advanced Research Fellowship GR/S53244.

REFERENCES

1. Cechanowicz, J., Irani, P. and Subramanian, S., Augmenting the Mouse with Pressure Sensitive Input. *ACM CHI 2007*, ACM Press, Addison Wesley, 1385-1394.
2. Li, Y., Hinckley, K., Guan, Z. and Landay, J., Experimental Analysis of Mode Switching Techniques in Pen-based User Interfaces. *ACM CHI 2005*, ACM Press, 461-470.
3. MacKenzie, S. and Soukoreff, W., Phrase Sets for Evaluating Text Entry Techniques. *ACM CHI 2003*, ACM Press, 754-755.
4. Mizobuchi, S., Terasaki, S., Keski-Jaskari, T., Nousiainen, J., Ryynanen, M. and Silfverberg, M., Making an impression: force-controlled pen input for handheld devices. *ACM CHI 2005*, ACM Press, 1661-1664.
5. Ramos, G., Boulos, M. and Balakrishnan, R., Pressure Widgets. *ACM CHI 2004*, ACM Press, Addison-Wesley, 487-494.
6. Srinivasan, M. and Chen, J. Human performance in controlling normal forces of contact with rigid objects. *Winter Annual Meeting of the American Society of Mechanical Engineers, 49*. 119-125.

Expressive Interaction with Contact Sounds

L.-M. Reissell   Dinesh K. Pai
Computer Science
Rutgers University and University of British Columbia
lmre@cs.rutgers.edu   pai@cs.ubc.ca

ABSTRACT
This paper describes work in progress towards expressive control of contact sounds that arise during haptic interaction in multimodal environments. While previous work has focused on physically realistic sound synthesis, we show how similar methods can be used to modify the expressive content of sounds based on the haptic interaction. We present an algorithm for robust and efficient decomposition and resynthesis of sounds modeled as frames of modulated complex exponentials. Several examples will be demonstrated live at the conference, controlled by a Wii Remote device.

INTRODUCTION
Contact sounds are an important and perhaps essential part of multisensory virtual environments involving haptics [4]. Haptic interaction produces impact forces that result in perceptible and meaningful sounds that enhance that haptic interaction experience. Recently, methods for automatically measuring the sounds produced by haptic interaction have also been developed [6].

However, most previous work on designing appropriate sounds for haptic interaction has focused on synthesis of physically realistic sounds, and not on the emotional or expressive nature of the sounds. In many applications, such as in games, the interaction is part of a narrative environment in which the expressive nature of the sounds is as important as physical realism. For example, it would be useful if the sound of a rolling ball [5] could be made pleasant if representing a bowling ball in a game, but ominous if it represents an oncoming projectile. While it is difficult to automatically achieve these complex effects, providing users with a method to control the expressive nature of environmental sounds would be very helpful.

We describe our ongoing work towards constructing an interactive controller for the expressive content of sounds. We present an efficient algorithm for decomposing and resynthesizing environmental sounds. The resynthesis can be performed interactively using a 3D interaction device, for instance, a Wii Remote. Changes in sound parameters, such as tempo, pitch, roughness, can applied individually, or several changes can be grouped and applied together to alter the perceived character or "expressive" content following, for instance, the results in [2] obtained for musical sounds.

RELATED WORK
Some previous work on synthesis of environmental sounds in haptic interaction has been described above. A general reference on interactive sound sythesis is [3].

Expressive modification of sounds has mainly been explored for musical sounds. For instance, [2] has investigated the connection between sound parameters and musical expression, as well as the verbal characterization of performance expression. Expressive musical intentions are described by words such as "light", "heavy", and "energetic." Different expressive styles are related to fairly specific changes in parameters such as tempo, amount of legato, attack emphasis, note and phrase envelope characteristics. Additional parameters can include pitch, and the introduction of noise or variability into the parameters.

We are not aware of any previous work on expressive modification of environmental sounds.

MODULATED PIECEWISE EXPONENTIAL MODELS
The sounds used in this paper are synthetic models of realistic environmental sounds, but the method applies generally to any sounds that can be modeled as frames of complex exponentials. We allow continuous modulation of the complex amplitude by e.g. piecewise polynomial approximations and their exponentials. (Frequency modulation is then included in the framework.) The model is a more flexible extension of a sinusoidal model representation. It allows the representation of a wide range of sounds and also permits a natural implementation of sound modifications.

In particular, the model can provide accurate, reasonably sparse representations of the difficult to model inharmonic sounds which occur during haptic interaction.

Analogously to a sinusoidal model, we define a piecewise exponential model, for N time samples, consisting of $L-$many frames, to be the real or imaginary part of the complex signal:

$$x_w(t) = \sum_{f=1}^{L} \sum_{k=1}^{K(f)} C_f^k(t) \exp s_f^k t, t = 0, 1, ..., N-1,$$

where $\exp s_f^k = \exp(a_f^k + i\omega_f^k)$ are the complex poles in frame f, $K(f)$ is the number of exponentials ("model order") used in frame f. The real and imaginary parts of the complex amplitudes correspond to real amplitude and phase.

In the general model used in synthesis, the amplitude coefficient C_f^k is allowed to be a complex number whose real and imaginary parts, corresponding to real amplitude and phase, are time-varying functions, such as piecewise polynomials, or exponentials of piecewise polynomials.

For the model used in analysis, C_f^k are constant. The analysis will ideally locate frames at places where the model changes. We use subspace methods (similar to ESPRIT) for analysis. In these methods good frame location can be ensured by the use of a subspace tracking algorithm [1].

We will require some smoothness in the modulation and some continuity across frames, by requiring that the modulating functions have some continuous derivatives and that the exponential "partials" do not exhibit large discontinuities at frame boundaries. In practice this can be achieved for instance by blending nearby modes at frame boundaries.

METHODS: SOUND CHANGE ALGORITHMS

Sequences of environmental sounds have more limited expressive potential than musical sounds, but can be transformed in a similar way. In addition, in our model it is simple to consider other expressive effects, such as "busyness". The collection of variable parameters used for environmental sounds include: speed, pitch (e.g. rising or falling chirps, tremolo), attack emphasis, amplitude envelope shape, selection of modes used, mode decays, amount of legato, presence of spectral noise, number of "voices" (sound sequences playing simultaneously), and so on.

The changes in such parameters are immediately translated to simple changes in the underlying piecewise exponential model. Many changes are analogous to the ones for a sinusoidal model.

The haptic interaction is considered as the control input. There is some flexibility about whether the signal is thought of as a single sound, or whether it represents a description of the sound at a different hierarchical level, e.g., "phrases" of sounds. There is also some flexibility about the aspects of the interaction used for input. In our current implementation we use interaction velocity. The control input can affect one paramater only, which can be a basic parameter such as pitch, speed, and loudness, or an expressive parameter which is a composite of these basic parameters.

For instance to achieve pitch change, in the current frame, we change each individual mode by multiplying the per-frame/mode amplitude C by a chirping term $\exp i\Phi(t)$:

$$C_{new}(t) = C(t) \exp i\Phi(t) = C(t) \exp ict^2.$$

The constants c can be constrained to retain harmonicity.

Figure 1 shows an example. More examples, including the haptic-audio interaction, will be presented with the poster.

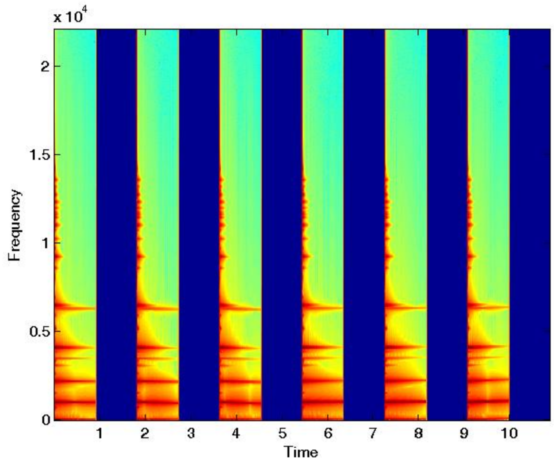


Figure 1: *An impact sound is transformed into different chirped versions (the first three with descending pitch and the last three ascending), based on the user interaction with Wii Remote. In the figure the effect is most visible at the lowest frequencies but is clearly audible. The ascending sounds appear more cheerful than the descending ones.*

ACKNOWLEDGEMENTS

This work was supported in part by the Canada Research Chairs Program, NIH grant 1R01NS050942 and NSERC.

REFERENCES

1. Roland Badeau. *Méthodes à haute résolution pour l'estimation et suivi de sinusoïdes modulées. Application aux signaux de musique.* Ph.D. Thesis, l'Ecole Nationale Supérieure des Télécommunications, 2005.
2. Di Federico R. Drioli C. Canazza S., De Poli G. and Rodà A. Symbolic and audio processing to change the expressive intention of a recorded music performance. In *Proc. of the 2nd COST G-6 Workshop on Digital Audio Effects (DAFx99), NTNU, Trondheim*, December 1999.
3. Perry Cook. *Real Sound Synthesis for Interactive Applications*. A.K. Peters, 2002.
4. Derek DiFilippo and Dinesh K. Pai. The ahi: An audio and haptic interface for contact interactions. In *UIST'00 (13th Annual ACM Symposium on User Interface Software and Technology)*, November 2000.
5. D. J. Hermes. Synthesis of the sounds produced by rolling balls. In *Internal IPO report no. 1226*, IPO, Center for User-System Interaction, Eindhoven, The Netherlands, 2000.
6. D. K. Pai et al. Scanning physical interaction behavior of 3D objects. In *SIGGRAPH*, pages 87–96, 2001.

"Silooets" - Audiotactile Vision-Substitution Software

David Dewhurst
HFVE.org
Darwins, Hillside, Odiham, Hampshire, UK
daviddewhurst@HFVE.org

ABSTRACT
This poster describes "work-in-progress" on "Silooets" (Sensory Image Layouts and Object Outlines Exhibited via Touch and Sound) - experimental vision-substitution software that uses audiotactile methods to present certain features of visual images to blind people.

Keywords
Blindness, deafblindness, Silooets, sensory-substitution, vision-substitution, audiotactile, haptic, Morse code, braille.

INTRODUCTION

At the 1st HAID Workshop in 2006, the "HiFiVE" (Heard & Felt Vision Effects) vision-substitution system was presented at a poster session [1]. The system has been further developed since then.

The HiFiVE system highlights features of visual images that are normally perceived categorically, and substitutes with coded sound effects and their tactile equivalents. By smoothly changing the pitch and binaural positioning of the sounds, they can be made to appear to "move", whether following a systematic path or describing a specific shape. Such moving effects are referred to as "tracers", and can be "shape-tracers", whose paths convey the shapes of items in an image; or "area-tracers", which systematically present the properties of parts of an image.

In the tactile modality, tracer location and movement are presented via force-feedback devices; and categorically-perceived features via braille, or via Morse code-like "tapping" effects.

Other work in the field includes tone-sound scanning methods that have been devised for presenting text [4], and for general images [5]; and software for presenting audiotactile descriptions of pixels in computer images [6]. Audio description is used to supplement television etc. (The merits of other approaches are not discussed in this poster.)

Potential applications

This project is not focussed on a specific application, but is trying various methods for presenting sequences of visual images via touch and sound. More straightforward material (such as simple shapes and diagrams) can also be presented.

Possible applications include:- presenting shapes and lines for instructional purposes; adding shape, colour and texture data to diagrams; providing ad-hoc information to users wishing to know the colour and shape of an item; and for specific tasks such as seeking distinctively-coloured items (for which corresponding sets of parameters are selected).

FEATURES OF THE SYSTEM

(This section recaps some of the previously-published features of the system [2,3], and describes enhancements.)

The Silooets software can present images as arrangements of pixels known as "layouts" Figure 1; and as the outlines of the objects in images Figure 2. In both cases the categorical properties of the areas are exhibited via groups of coded CV (Consonant-Vowel) speech sounds; or as braille dots; or via Morse code-like "tapping" effects.

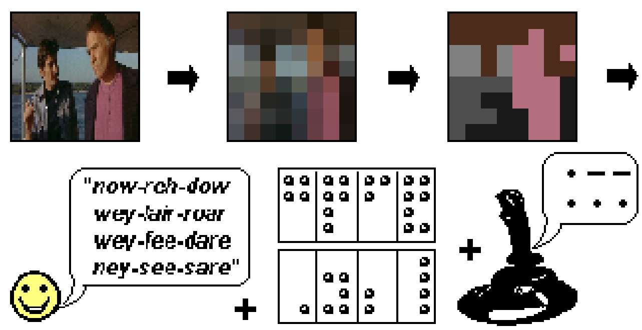


Figure 1. Diagram illustrating the conversion of the "layout" of an image into coded phonetics, braille, and "tap codes".

Figure 1 shows an example of an image being reduced to two colour shades in each image quarter / "panel", and some of the corresponding speech sounds, braille cells and tapping codes, which describe the 8 by 8 pixels shown. "Layouts" can be used to present any section of an image.

Alternatively, if the shapes of entities in an image can be determined (Figure 2), then audiotactile "shape-tracers" can exhibit those shapes and their corners, at the same time as the categorical properties are being presented.

Figure 2. Identified entity shapes, whose outlines and corners can be presented via audiotactile "shape tracers".

Colour mapping

The system normally uses two colour shades (e.g. “blue and red”) when presenting an area or entity, allowing the shades to be given via a modest number of corresponding effects. The two-colour-shade approach can simply present the two associated shades, or have the effect of painting a picture with one colour on a differently-coloured background.

A 15-colour-palette format allows two colours to be easily presented via a single “CV” syllable; via a single 8-dot braille cell; or via “tap codes” of 8 or less pulses.

Speech mappings are available for English speakers using the full range of English phonemes; or an “International” format can be used, which uses the sounds found in most languages to present combinations of 24 shades (Table 1).

Vowel sound 2nd ↓ — Consonant sound 1st ↓

↓	**S** (w)	**R** (L)	**K** (g)	**N** (m)	**D** (b,p,t)
I	Lt.Purple	Lt.Brown	White	Cream/24	(Special)
E	Pink	Yellow	Light Grey	Lt.Green	Light Blue
A	Red	Orange	Mid Grey	GrnYell/23	Turquoise
O	Purple	Brown	Dark Grey	Green	Blue
U	Dk.Purple	Dk.Brown	Black	Dk.Green	Dark Blue

Table 1. An example of colour-to-speech mapping.

Tactile effects and user interaction

Programmable 8-dot braille cells are available commercially, and are an effective way of presenting categorical data to blind people who are able to read braille.

Alternatively, coded pulses can be induced on a standard force-feedback device, presenting “tap codes” to the user. It is found that short, evenly-spaced pulses of two force levels are more straightforward to interpret than conventional Morse-code timings. “Tap codes” are relatively slow when compared to speech or braille, but may be useful for deafblind users who cannot read braille.

A force-feedback joystick makes an effective control and pointing device, as it can also be moved by the system, pushing and pulling the user's hand and arm, tracing out any shapes (and highlighting corners) that are to be presented.

Corners

Corners (such as the vertices of an octagon) produce a considerable effect in giving the impression of a shape to sighted people, and the system emphasises them by introducing effects within shape tracers at the appropriate point in time. For example, the system can momentarily stop the movement of a “shape-tracer” (by stopping a moving force-feedback joystick); and in the audio modality the system can alter the volume of the sound.

RECENT DEVELOPMENTS

At the 1st HAID Workshop, the system was shown presenting standard demonstration shapes with corners; and speech-like sound representing the “layouts” of live images. Since then, the following features have been progressed:-

Prepared media

Prepared media can be presented, with predetermined shapes etc. embedded in common image formats (e.g. bitmap, GIF or JPEG still images; or DVDs or AVI movie files), produced via a straightforward procedure. (Such media can be also be viewed on standard media players.)

Enhanced layouts and colour mappings

The Silooets software is designed so that “layout” configurations and colour shade mappings are flexible. New layout configurations include ones with greater resolution at the centre of the layout. Standard colour mappings have been simplified and enhanced.

Morse code-like tapping output

The tap-codes described above have been implemented, and can exhibit certain properties. Various timings and force levels have been tested, and are available as options.

Enhanced tactile effects

The system can now exhibit a notchy “grid” effect when a force-feedback joystick is used, so giving the user a tactile indication of where the joystick is currently located.

Two joysticks can be used. For example, the main joystick can be used as a pointer by the user, and by the system to indicate the location and size of the area being presented. The other device, for example a force-feedback mouse, can be used by the system to present any shapes and tap codes.

Activity-related processing

Users can rapidly switch between sets of parameters and options by selecting identifiers for particular activities.

SUMMARY

When completed, the Silooets software will allow visual features of images, ranging from basic properties, to object descriptions (if known), to be presented to blind people. Several features of the system can now be demonstrated.

REFERENCES

1. David Dewhurst, “An Audiotactile Vision-Substitution System” in *Proc. of First International Workshop on Haptic and Audio Interaction Design*, Vol.2 (2006) 17-20.
2. See website http://interactive-sonification.org/ISon2007/ISon2007Proceedings/papers/Dewhurst_ISon2007.pdf.
3. U.S. Patent Appl. No. US 2008/0058894 A1
4. E. E. Fournier d'Albe, “On a Type-Reading Optophone” in *Proc. of the Royal Society of London. Series A*, Vol. 90, No. 619 (Jul. 1, 1914), pp. 373-375.
5. See website http://www.seeingwithsound.com.
6. See website http://www.ifeelpixel.com.

Wiimote Music – Towards an intuitive gestural interface grounded in a theory of learning schemas

Alex Reed
Department of Music, University of Jyväskylä
PL35(M) 40014, University of Jyväskylä, Finland
reechka@gmail.com

Marc R. Thompson
Department of Music, University of Jyväskylä
PL35(M) University of Jyväskylä, Finland
mathomps@jyu.fi

Karolien Dons
Department of Music, University of Jyväskylä
PL35(M) University of Jyväskylä, Finland
karoliendons@hotmail.com

Tuukka Tervo
Department of Music, University of Jyväskylä
PL35(M) University of Jyväskylä, Finland
tuheterv@jyu.fi

ABSTRACT
This demonstration showcases a novel implementation of the Nintendo Wii gaming controller as an interface for musical activities. The Wii remote (Wiimote) interfaces with a computer through an open-source software named OSCulator, which in turn triggers musical events in the Ableton Live music sequencer. Using the Wiimote's built-in accelerometers and infrared camera, we have programmed the controller to react to bodily gestures based on children's learning schemas. The intention was to use gestures that are the most intuitive and innate to humans. Furthermore, by implementing learning schemas with widely available pre-existing software, the controller can be conveniently used in an educational setting.

Author Keywords
Wiimote, gesture, Bluetooth, learning schema, mapping.

ACM Classification Keywords
H5.2. User interfaces: haptic I/O, input devices and strategies, interaction styles.

INTRODUCTION

New interfaces for musical instruments are rarely introduced, and much less often widely accepted or commercialized as products. Most often software instruments are controlled with devices that are either derived from traditional mechanical and electro-mechanical instruments, such as the musical keyboard, or with generic (not musically oriented) input devices, such as the alphanumeric keyboard and the mouse.

The Nintendo Wii game controller has many ways of translating gestures into signals that can be sent to a computer. Apart from traditional game controls such as buttons and a joystick, it has accelerometers, which means that empty-handed gestures [6] can also trigger signals. Accelerometers have been used in previous expressive digital musical interface designs such as the Musical Playpen, Beatbugs, Fireflies and Squeezables [1] but these all had to be built from scratch, as not all parts were commercially available. The Wiimote, on the other hand, is widely available, wireless and although a game controller, and adaptable [7].

The other problem with Beatbugs was found to be wires that became unattached. Currently, there are many forms of wireless connectivity to computers. One such solution was the Optophone [4], which tracks user movements through a webcam. However, its resolution has proven to be relatively poor, and only works under specific lighting conditions. The Wiimote is commercially and cheaply available and uses Bluetooth and infrared connectivity, which provide superior resolution than the method of using webcams to track movements. In other words, the Wiimote allows for the tracking of a wider range of gestures with a larger pool of mapping strategies to choose from.

Rather than try mapping those gestures ad hoc to musical parameters, we propose to apply a theory of children's learning schema [2] to determine which gestures might be more intuitive for amateur or non-musicians such as children, or those enjoying music in a therapeutic context, to make if they were freed from the constraints of traditional musical interfaces.

We will be using three of the seven learning schemas proposed by Athey [2]: dynamic vertical; dynamic back and forth (side to side); and dynamic circular. These three learning schemas were chosen because they seem to be the most intuitively 'playful' in terms of making music. The schemas will be used to map the gestures the user is most comfortable with, to the three musical functions of *excitation, modification* and *selection* [3].

Excitatory functions can be instantaneous (e.g. hitting a drum) or continuous (e.g. bowing a cello). Modifying functions usually require neutral energy expenditure compared to excitation and can be continuous (e.g. bending pitch), or structural (e.g. depressing a valve or button). Selection of notes can be sequential: as in the case of monophonic sounds or parallel: as in the case of polyphonic sounds [3].

A gesture could be mapped to parameters of sound synthesis not previously possible in traditional musical interfaces, for instance increasing the virtual size of the body of an instrument while playing the same note. Also big gestures could control subtle changes in musical parameter, and equally small gestures could control larger changes if required.

METHOD

We are using existing software for a different purpose to that which is was originally designed for, and we are using learning schemas from an educational context to suggest musical gestures.

We are not proposing to reinvent the wheel, so to speak, but perhaps to put it on another machine. If the interface and interaction can relate to an existing schema that the person can then associate with a different musical function, than could be realised before, then our interface has successfully created a new metaphor in music: a gesture that was already embedded as part of a schema used in another context is now used in a musical context.

Our current set-up uses OSCulator [5] to map the signals from the Wiimote to the computer, and Ableton Live to then create and modify music from those signals, through a combination of manipulating audio files and MIDI instruments. And because this interaction is non-mechanical, thanks to Bluetooth or infrared connectivity, it is enough for it to be a gesture.

The challenge is to find intuitive and viable gestures that are not necessarily constrained by the more traditional musical interfaces, like the piano keyboard for instance. In this way people such as children or those with disabilities, who might otherwise have felt a lack of control or unmotivated by other musical interfaces are empowered, and can make meaningful and enjoyable music with the Wiimotes.

We have created a number of set-ups in OSCulator and Live, which we will first demonstrate individually, then as a group together. Each one will use a variety of gestures from any one or combination of the aforementioned three learning schemas. And we will also let people try out these patches for themselves. The interfaces will use a variety of combinations of haptic and/or visual feedback.

EQUIPMENT AND EXPECTED RESULTS

In terms of equipment we will only need sound amplification for four computers. Our demonstration hopes to show that the Wiimote can enable us to choose which gesture we map to which function of a musical instrument.

The choice of gesture will have been made before the demo, for the patch on each computer. For this reason it is important that on each computer we have installed patches that differ significantly in the mapping of gesture to a sound. This will show the versatility of the different set-ups. But also we will aim to make it so that all four computers can play together as a group.

The demonstration consists of a performance and explanation, which will be followed by a chance for the audience to try out the interface themselves. We expect there to be some amount of fun and music. The demo will be interactive and there will be someone with each computer to help and explain. Of course, only some gestures can be modelled effectively by Wii controllers in a musical interface, and these shortcomings will also be presented as part of the demonstration.

CONCLUSION

The demo aims to show how the Wiimote can map intuitive gestures and create an expressive new digital instrument that could be used in a variety of contexts to make music. The gestures expected to be used will be of a prototypical kind that can be associated with a learning schema and a set of simple functions. It is also postulated that this will make them easier and more intuitive to learn.

Gestures covered by the dynamic vertical schema can be expected to include jumping or hopping, and other forms of up and down movement. The dynamic back and forth schema includes gestures such as moving the Wiimote from side to side, and the dynamics circular schema could involved spinning the Wii, and the parameters of roll and yaw [5].

REFERENCES

1. Aimi, R.M. (2002). New expressive percussion instruments. *Masters thesis, Massachusetts Institute of Technology.* Retrieved 10 February 2008 from www.media.mit.edu/hyperins/papers/aimimasters.pdf.
2. Athey, C. (1991). Extending Thought In Young Children. London: Paul Chapman Publishing.
3. Cadoz, C. & Wanderley, M. (2000). Gesture- Music. From *Trends in Gestural Control of Music*. Paris: IRCAM (PDF version)
4. Optophone. http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/index#animations
5. OSCulator. http://www.osculator.net/wiki/Main/FAQ
6. Wanderley, M. & Miranda, E. (2006). New Digital Musical Instruments: Control and Interaction Beyond the Keyboard. Middleton, Wisconsin: A-R Editions, Inc.
7. Wii forum: http://www.wiimoteproject.com

Perception of rubbing sounds

Dik J. Hermes,
Kevin van der Pol,
Arthur Vankan, Frank Boeren,
and Onno Kuip

Eindhoven University of Technology
P.O. Box 513
NL 5600 MB Eindhoven
+31 40 2475214
d.j.hermes@tue.nl

ABSTRACT
When people interact manually with solid objects, the interactions are generally coupled with rubbing, scraping, or other impact sounds. In virtual reality, such sounds are mostly lacking, thus affecting the naturalness of the interactions with the virtual objects. This contribution deals with the perceived properties of rubbing sounds, generated by, e.g., gliding a sheet of paper on a table surface. This sound appears to consist of high-pass filtered noise. An experiment is described in which synthetic rubbing sounds were varied in respect of their high-pass cut-off frequency and their spectral balance. Subjects were asked to estimate the distance traveled during the rubbing movement, and to indicate the effort with which the paper was pressed on the table during the gliding movement. The results show that the perceived distance traveled is positively related with spectral balance, while the cut-off frequency did not affect perceived distance. The perceived pressure exerted is negatively related with both cut-off frequency and spectral balance. The results are applied on a tablet to synthesize natural sounding rubbing sounds when moving over the surface of the tablet.

Keywords
Contact sounds, rubbing, sound design, virtual reality.

1. INTRODUCTION
Sound is an essential component of the perceptual information people can use while interacting with their environment. Although the intensity may be low, the sound produced when touching an object gives information about its material and its surface. In virtual reality these sounds are absent if not synthesized.

2. EXPERIMENT
2.1 Stimuli
Gaussian white noise was high-pass filtered with cut-off frequencies of 2, 3 and 4.5 kHz. The sample frequency was 44.1 kHz. The spectral slope of each of these filtered noises was modified in the spectral domain by the overlap-and-add of 1025 sample Hanning-windowed noise frames, so that the spectral envelope decreased according to $f^{-0.8}$, remained constant, or increased according to $f^{0.8}$. The spectral slope will be indicated by this exponent of f. Thus, nine different realizations of noise were synthesized of which the cut-off frequency and the spectral balance were varied in an orthogonal design. The perceived loudnesses were equalized by asking ten participants to adjust the loudness of all eight remaining stimuli to the stimulus with a cut-off frequency of 3 kHz and a constant spectral slope.

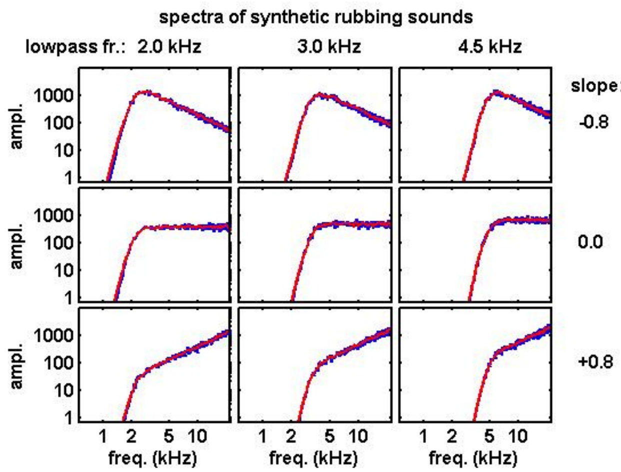


Figure 1: Spectra of the nine synthetic rubbing sounds varying independently in spectral slope and high-pass cut-off frequency.

The temporal envelope of these noise intervals increased, decreased, increased and decreased according to one period of a squared sinus wave, mimicking a rubbing movement back and forth between two points with a finger on a piece of paper over a table. An example is shown in figure 2. A pilot experiment showed that naive subjects could not distinguish between these synthetic and recorded rubbing sounds. Furthermore, when listening to these sounds more than half of the listeners, who were asked to indicate what they heard, indicated that they heard rubbing or sweeping sounds.

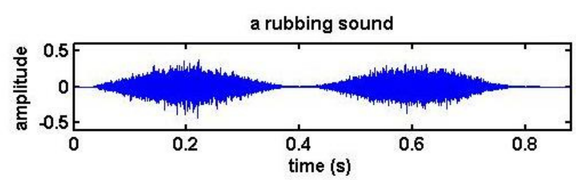


Figure 2: Oscillogram of a synthetic rubbing sound.

2.2 Experiment
The experiment consisted of presenting 18 listeners, who were paid for their participation, with a total of 30 of these sounds. The first nine presentations consisted of all nine sounds played in random order. These were practice stimuli. Then the experimental stimuli were presented twice, resulting in 18 presentations. The first series consisted of all nine stimuli

presented in random order; the second consisted of the same series but now presented in reverse order. The experiment was completed by three sounds randomly selected from the set of nine stimuli. The data presented here only contain the responses to the 18 stimuli from the second and the third presentations of these stimuli.

Subjects were told that the sounds represented the sounds produced when a rubbing movement was made back and forth between two points. In one experimental session, SPEED, they were asked to indicate how many centimeters there were between these two points. It is supposed that this estimation of distance gives a good representation of the perceived speed of the rubbing movement. In the other experimental session, EFFORT, subjects were asked to indicate on a scale from 1 to 5 how much effort they estimated was used in pressing on the table while making the rubbing movement. This is supposed to represent a measure for perceived effort.

3. RESULTS

3.1 Perceived effort

The results for the EFFORT experiment are presented in Figure 3. It can be seen that with increasing cut-off frequency and increasing spectral slope the effort is perceived as decreasing. These results are statistically significant as shown by a within-subjects analysis of variance in which Huynh-Feldt's correction for the lack of sphericity is applied. The main factor cut-off frequency is significant with an $F[2,34]=18.643$, $p<0.001$, while the main factor spectral slope is significant with an $F[2,34]=4.590$, $p=0.030$. The interaction was not significant ($F[2,34]=0.116$, $p=0.944$).

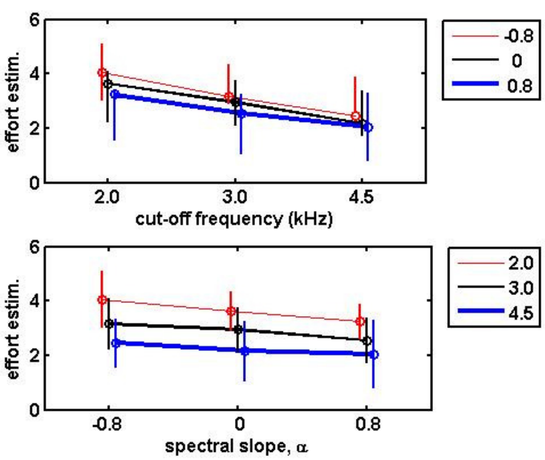


Figure 3: The estimated effort on a scale from 1 to 5 subjects reported to perceive. In the upper panel the data are presented as a function of cut-off frequency with the spectral slope as parameter shown in the legend. In the lower panel the same data are presented as a function of spectral slope with cut-off frequency in kHz as parameter. Data points are averages, the bars indicate one standard deviation up and one standard deviation down.

3.2 Perceived speed

At first instance, the estimations of the distances traveled during the rubbing movements varied largely from subject to subject. The within-subject analysis of variance yielded no statistically significant results. Apparently, the participants differed in estimating the absolute distance between the two points. The following analysis showed, however, that they agreed on the relative distances. Indeed, the results were normalized to z-scores for each subject by first subtracting from each estimation the average of all estimations and then dividing the result by the standard deviation. These z-scores are presented in Figure 4. The effect of cut-off frequency appeared not to be statistically significant, $F[2,34]=0.790$, $p=0.422$, but the effect of spectral slope was significant, $F[2,34]=6.746$, $p=0.006$. The interaction was not significant ($F[2,34]=0.604$, $p=0.437$).

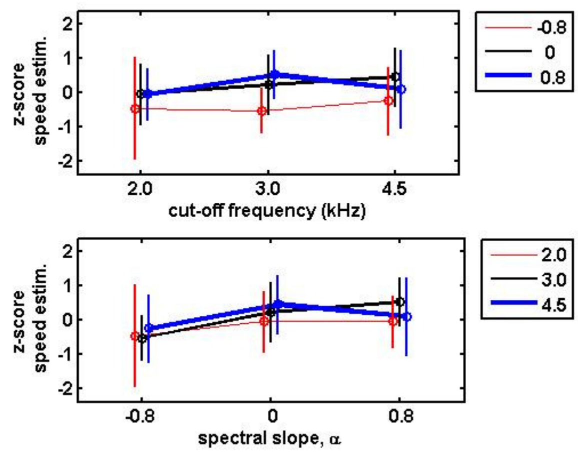


Figure 3: The z-scores of the estimated speed subjects reported to perceive. For details see Figure 3.

4. DISCUSSION

In the experiment two properties of the frequency spectrum of the high-pass noise were varied, the cut-off frequency and the slope of the spectrum envelope. These properties are supposed to relate to perceived brightness and perceived sharpness. Bismarck [1, 2] argues that sharpness increases with the upper and lower limiting frequency as well as the slope of the spectral envelope, and shows a relation between the weighted first moment of the loudness-critical band rate-pattern. These results show that these timbral properties play a role in the perception of rubbing sounds. Note, however, that these results do not conform to the results for speech sounds and musical sounds. For these sounds, in general, the greater the contribution of higher frequencies in the spectrum, the greater the perceived effort of the sounds produced. For speech, see, e.g., Bloothooft [3]. For rubbing sounds we find the opposite. The results of our experiments show that both cut-off frequency and spectral slope play a role in the perception of these sounds, but that the role is different for different kinds of sounds.

5. REFERENCES

[1] Bismarck, G. von 1974. Timbre of steady sounds: A factorial investigation of it verbal attributes. Acustica 30, 146-159.

[2] Bismarck, G. von 1974. Sharpness as an attribute of the timbre of steady sounds. Acustica 30, 159-172.

[3] Bloothooft, G. and Plomp, R. 1988. The timbre of sung vowels. JASA 84(3), 847-860.

New Tonalities with the *Thummer* and *The Viking*

Andrew Milne
University of Jyväskylä
Department of Music, P.O. Box 35
40014 Jyvaskyla, Finland
anjomiln@cc.jyu.fi

Anthony Prechtl
University of Jyväskylä
Department of Music, P.O. Box 35
40014 Jyvaskyla, Finland
aprechtl@gmail.com

ABSTRACT
In this paper we explain the theoretical background of Dynamic Tonality using the *Thummer*, a new musical interface, and *The Viking*, a software synthesizer written especially for it. *Dynamic Tonality* is a musical audio routine that allows for novel tunings and enables the user to relate – to an arbitrary degree – these tunings with the partials of their notes. *The Viking* features Dynamic Tonality and works with any MIDI instrument, but when paired with the *Thummer* (or another two-dimensional interface) it creates a system of fingering invariance across chords and tunings. Thus, the *Thummer* and *The Viking* render non-standard tunings more physically, pedagogically, and aesthetically accessible.

Author Keywords
Thummer, *The Viking*, Dynamic Tonality, fingering invariance, microtonality.

ACM Classification Keywords
H5.2. Information interfaces and presentation (e.g., HCI): User interfaces.

H5.5. Information interfaces and presentation (e.g., HCI): Sound and music computing.

INTRODUCTION

The *Thummer* is a new MIDI instrument currently being developed by Thumtronics, Inc. in Austin, Texas [10]. Its name reflects a set of thumb-specific buttons and joysticks designed to manipulate a variety of user-assignable parameters. Note selection is performed via two button lattices (see Figure 1), where the user finds – when in standard tuning – the flats on the left, the sharps on the right, and the diatonic notes in the middle. This layout can be regarded as two-dimensional in that its buttons are arrayed on both horizontal and vertical axes.

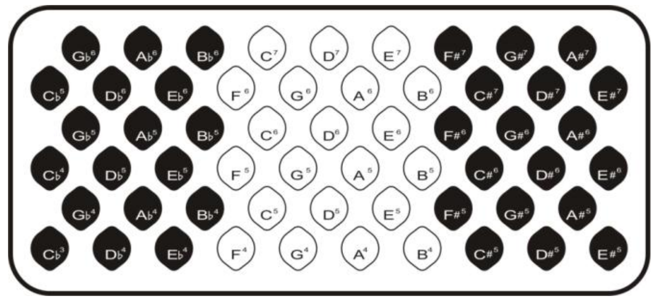


Figure 1. The *Thummer*'s button lattice.

In order to fully reap the benefits of a two-dimensional layout, we introduce *The Viking*, a software synthesizer designed to interpret both one- *and* two-dimensional MIDI interfaces. This paper will focus on that part of *The Viking*, namely Dynamic Tonality, which utilizes two-dimensionality to allow for – and provoke the exploration of – new tonalities.

DYNAMIC TONALITY

Dynamic Tonality uses a small number of parameters that enable the user to:

- Choose between a number of different tuning continua, each of which allow both conventional and novel scales.
- Freely move between a number of equal temperaments, non-equal temperaments, circulating temperaments, and closely-related just intonations.
- Use the same fingering pattern for all tonal intervals across all possible keys and tunings (when they are played on any musical controller that has a two-dimensional lattice of buttons or keys) [3].
- Freely move the timbre from being perfectly harmonic to perfectly optimized to the tuning.

Tuning Continua

A *tuning continuum* is a range of tunings over which all possible melodies built from harmonic consonances, and the voice-leading intervals between them, have the same contour [4]. Musically useful continua can be generated using only two intervals of variable size (such as a perfect fifth and an octave, which generate the familiar pentatonic, diatonic, and chromatic scales) [11,1].

Continua generated from quite different intervals result in radically different scale structures [2]; *The Viking* implements two non-standard continua: "Magic", which is generated by a variable major third and octave, and "Hanson", which is generated by a variable minor third and octave. Both of these continua produce scales that are relatively unfamiliar, but which still feature numerous well-tuned major and minor triads.

Dynamic Tuning

In a musical scale generated by a perfect fifth and an octave, varying the size of the fifth produces a wide range of useful tunings. For example, string and aerophone players often prefer Pythagorean tuning for expressive melodies, and ¼-comma meantone for sustained chords [8]. There are also numerous "non-standard" tunings used in non-Western music such as the 5-TET of Indonesian Slendro [9] and the 7-TET of traditional Thai music [5].

When playing *The Viking* with a *Thummer*, the size of the perfect fifth can be controlled simply by moving one of the thumb controllers. This enables *Thummer* players to easily mimic the intonation devices used by string and aerophone players, and to easily move between different tunings throughout a performance.

Fingering Invariance

Perhaps one of the most important aspects of Dynamic Tonality is that it allows for fingering invariance. On any musical controller with keys or buttons, intervals are played by a specific set of buttons that outline a geometric shape. However, on one-dimensional interfaces like the piano keyboard, a given interval often requires different shapes. For example, D–F# and F–A are each major thirds, but form different shapes on a piano keyboard. Contrarily, on a *Thummer*, all major thirds are exactly the same shape; in fact, all intervals on the *Thummer* require only one fingering. Similarly, piano keyboards require multiple fingerings across a tuning continuum, while *Thummers* only require one fingering. This property of invariance across tunings and keys is called *fingering invariance* [3].

With fingering invariance, musicians need only learn the fingering of a given interval or chord once, and thereafter apply that shape to all occurrences of that interval or chord, independent of its location within a key, across keys, or across tunings. This reduces rote memorization considerably and engages the student's visual and tactile senses in discerning the consistency of musical patterns. Though it has not yet been empirically tested, we expect that fingering invariance will be of benefit for both beginners and experts; hopefully of sufficient advantage to mitigate the initial effort required to learn a new interface.

Dynamic Timbre

As described above, scales and tunings can vary significantly throughout a tuning continuum, and even more significantly between different tuning continua. With fingering invariance, these scales and tunings can even be easy to play and learn. However, we have not yet addressed the aesthetic significance of Dynamic Tonality. There are many reasons why 12-TET is the standard, so why bother with non-standard tunings?

There is strong evidence that an instrument's timbre determines the tunings it plays in best, and that by taking the reverse approach each tuning also has *related* timbres that sound most consonant [7]. Most non-standard tunings have no access to a related timbre, due to the limitations of acoustic instruments. Dynamic Tonality addresses this issue by adjusting the timbre to match the current tuning specified by user. More precisely, it allows the user to specify to what degree the *partials* (also known as *overtones* or *harmonics*) should match the tuning. To date, this feature has been almost entirely unexplored in music and research, but we have found – and will demonstrate – that using a related timbre greatly mitigates the dissonance of non-standard tunings.

SYNTHESIS WITH *THE VIKING*

With regard to the many synthesis methods available today, Dynamic Tonality really only has one requirement: the frequency of each partial must be adjustable in real-time. *The Viking* uses additive synthesis, which generates each partial with its own sinusoidal oscillator. Frequencies are determined automatically by Dynamic Tonality parameters, while amplitudes are determined automatically by a traditional waveform selector and further modified by familiar subtractive synthesis filters and envelopes. *The Viking*, therefore, enables sounds similar to those produced by the "classic" subtractive synthesizers (e.g., Moog, ARP) to be adjusted spectrally.

The Viking was written in Outsim's *Synthmaker* [6]. It works with any MIDI interface, and is available for Windows as a VSTi at www.dynamictonality.com. A version for Mac OS X may be forthcoming in the near future.

CONCLUSION

In order to reap the many benefits of two-dimensional musical interfaces, there is a need for a computational routine that can understand both dimensions in a meaningful manner. Dynamic Tonality does this by interpreting two-dimensionality as a logical tuning mechanism that is both flexible and easy to learn. Moreover, in order to make alternate tunings more aesthetically accessible, it also enables the relation of timbre and tuning with a simple and intuitive interface. We hope that each of these offerings will eventually motivate a new exploration of tonality.

As mentioned previously, *The Viking* is an additive synthesizer; however, there are other synthesis methods capable of employing Dynamic Tonality. The forthcoming *TransFormSynth* uses analysis-resynthesis to allow both pre-recorded and live sound to be manipulated per the parameters described above. Information on all Dynamic

Tonality synthesizers as well as information on how to implement Dynamic Tonality in your own software can be found at www.dynamictonality.com.

REFERENCES

1. Carey, N., and Clampitt, D. Aspects of well-formed scales, *Music Theory Spectrum 11* (1989), 187-206.
2. Erlich, P. A middle path between just intonation and the equal temperaments, part 1, *Xenharmonikon 18* (2006), 159-199.
3. Milne, A., Sethares, W. A., and Plamondon, J. Isomorphic controllers and dynamic tuning: Invariant fingering over a tuning continuum. *Computer Music Journal 31*, 4 (2007), 15-32.
4. Milne, A., Sethares, W. A., and Plamondon, J. Tuning continua and keyboard layouts. *Journal of Mathematics and Music 2*, 1 (2008), 1-19.
5. Morton, D. The Music of Thailand. In May, E., ed. *Music of Many Cultures.* University of California Press, Berkeley, California (1980) 63-82.
6. Outsim Ltd. http://synthmaker.co.uk/.
7. Sethares, W. A. *Tuning, Timbre, Spectrum, Scale, 2nd ed.* Springer-Verlag, London, 2004.
8. Sundberg, J., Friberg, A., and Frydén, L. Rules for automated performance of ensemble music. *Contemporary Music Review 3*, 1 (1989), 89-109.
9. Surjodiningrat, W., Sudarjana, P. J., and Susanto, A. *Tone Measurements of Outstanding Javanese Gamelan in Yogyakarta and Surakarta*. Gadjah Mada University Press, Yogyakarta, Indonesia, 1993.
10. Thumtronics Ltd., *Thumtronics: The new shape of music*. http://www.thummer.com.
11. Wilson, E. Letter to John Chalmers pertaining to moments of symmetry/Tanabe cycle (1975). http://www.anaphoria.com/mos.PDF

Demonstration of a universally accessible audio-haptic transit map built on a digital pen-based platform

Steven Landau
Touch Graphics, Inc.
330 West 38 Street Suite 1204
New York, NY 10018 USA
sl@touchgraphics.com
212-375-6341

Gene Bourquin
Touch Graphics, Inc.
330 West 38 Street Suite 1204
New York, NY 10018 USA
gb@touchgraphics.com
212-375-6341

Andy van Schaack
Vanderbilt University
268 Wyatt Center
Nashville, TN USA
andy.vanschaack@vanderbilt.edu
615-343-6780

Joshua Miele
The Smith-Kettlewell Eye Research Center
2318 Fillmore Street
San Francisco, CA USA
jam@ski.org
415-345-2113

ABSTRACT
In this demo, conference attendees will try out a new system for providing multi-sensory transit and way-finding information about New York City subways to riders who are blind, visually impaired, or otherwise print disabled. The system includes a booklet of raised-line and textured maps of train routes; users explore the maps through various combinations of vision and tactile sense, and then touch the tip of a special pen to locations on the map to hear station names and other information spoken aloud. Train lines are rendered as narrow channels that help to guide a user's hand as he or she moves the pen along a route, and small depressions within these channels mark individual station stops. This novel combination of tactile graphics and a haptic audio probe may provide an intuitive and information-rich interface that could help disabled individuals travel around the city with greater self-confidence, safety, and independence.

Author Keywords
Tactile, cartography, pen, computer, blind, subway, map, orientation

ACM Classification Keywords
H.5.2 User Interfaces (Haptic I/O, Audio (non-speech) feedback)

INTRODUCTION
Visually impaired travelers have difficulty accessing information that their sighted counterparts take for granted. For example, subway maps provide sighted riders with an overview of an entire network of underground trains, and also offer details about individual routes and stations. Without maps and other tools for acquiring up-to-date information prior to setting out on a journey, and for consultation *en route*, blind and low vision individuals experience navigation difficulties. This project embodies ideas that might help to reduce levels of inconvenience, frustration and hazard that often interfere with capable individuals' efforts to reach their employment and education objectives. The authors have developed a practical system for making and distributing accessible transit maps that are portable and inexpensive to manufacture. This work combines static tactile images with dynamic audio content, and adds a proprioceptive component that may make the system easy and intuitive to use for transit riders with a wide range of capabilities and preferences.

Background
The New York City subway is one of the most extensive public transportation systems in the world, with 468 stations and 842 miles (1355 km) of track [1]. Understanding the layout of the system and planning routes between stations is a daunting exercise for many riders, with and without disabilities. *The Map*, originally developed in 1979, and continually updated since then, is appreciated by many transit riders, because it presents the system's complexity in a surprisingly comprehendible way [2]. The map thoughtfully balances diagrammatic clarity and spatial accuracy [3], and it is the outcome of an evolutionary design process over more than 100 years. Without easy access to *The Map*, which is ubiquitously displayed in train cars and stations, and widely distributed to the public at no charge, navigating complex routes would be unmanageable for many riders. Yet, this is the exact situation faced on a daily basis by those who cannot see well enough to make sense of *The Map* in its current print-only format.

Project Description
The pen-enabled talking tactile subway map that is the subject of this demonstration appears to be the first portable and comprehensive cartographic system that is accessible to

riders who cannot read print documents. In the current implementation, only the 1-2-3 lines are shown (see figure 1); in future versions, the entire system will be presented in a bound booklet of maps. Each page will show a single route, along with a simplified outline of the geographical context of the four boroughs, and major parks and airports. While the map can be used by itself (two-character Braille abbreviations identify the most important transfer points and a legend for each subway line explains their meanings), its full potential as a navigational tool is only revealed when the map is used in conjunction with the computer-pen (Livescribe's Pulse Pen [3]). This powerful, compact device, developed as a consumer product and used mainly as part of a smart note-taking system, includes a tiny video camera in its tip. The camera "sees" a very fine (almost invisible) *Anoto* dot pattern on the map's surface, and through analysis of the position of the dots, the pen's on-board microprocessor is able to determine the precise location of the nib on the map's surface [4]. After consulting a program that associates each x,y coordinate position on the map with place names, the pen's audio system plays recorded messages that include names of the stations touched. Additional useful information is embedded as a series of layers that are accessed when the user taps multiple times on a station. These layers include descriptions of platform and stations physical layout; information on transferring to connecting services, such as bus and ferry; data on routes and schedules; and instructions about what to do in case of an emergency. In Phase 1, this material will be collected for all 84 stations along the 1-2-3 line, output through a speech synthesizer, and saved to the pen's 1 GB flash memory for playback on demand during map exploration.

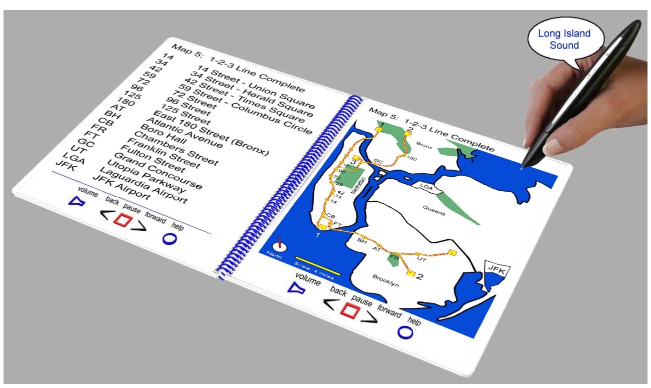


Fig.1: A map showing the 1-2-3 lines of the New York City subway system.

The subway map system demonstrates a new approach to delivering information through a static haptic display. This approach relies on some of the same sensory mechanisms that adept blind pedestrians rely upon as they move about in the world. Most blind travelers use long canes to interrogate the environment and to draw inferences about terrain conditions prior to taking each step, and also to detect and negotiate local obstacles [5]. They do this by interpreting haptic data transmitted from the cane's tip, through the shaft and grip of the cane, and into the palm of the hand. In the case of the subway maps, in similar fashion, the map's three-dimensional form encodes information that travels up the pen's body and into the user's hand. The raised parallel ridges that form channels guides the users hand along the train's route in an approximation of the actual track configurations, possibly creating mental linkages between the abstraction of the map and real-world conditions. Small dimples mark the location of each station, creating haptic events as the pen drops into each one. Users may find it easy to count the number of stations along a route. Furthermore, the dimples may stabilize the pen's tip at a single point, which is helpful if the user has difficulty with fine motor control, or if he or she experiences hand tremor.

CONCLUSION

This demonstration will introduce conference attendees to an innovative application of interactive audio-haptics. The New York City subway map system discussed here is an example of universal design, because it sets out to be usable by a very broad audience, it presents information in multiple and redundant formats, and it adapts to, or can be reconfigured for, a large variety of user profiles. If upcoming trials reveal that the subway maps are effective for blind and low vision transit riders, we envision other applications that leverage this technology, including DAISY-compliant illustrated digital talking books in categories such as children's literature; technical manuals; tourist data; emergency preparedness information; standardized assessments; maps of all kinds; and text books and curricula for mathematics, history and the sciences.

ACKNOWLEDGMENTS

The product demonstrated here is being developed through a Phase 1 SBIR grant administered through US Department of Education, National Institute on Disability and Rehabilitation Research (H133S080030).

REFERENCES

1. New York City Transit Information website. http://www.mta.info/nyct/facts/ffsubway.htm
2. Ovenden, M. (2007). **Metro Maps of the World**. Viking-Penguin.
3. Wood, M., Marggraff, J., Brown, M., & Fishbach, M. Interactive learning applicance. US Patent No. 6801751 Filed Aug 4, 2000, issue Oct 5, 2004.
4. Luff, P., Heath, C., Norrie, M., Signer, B. & Herdman, P. (2004). Only touching the surface: creating affinities between digital content and paper, *Proc. Conference on Computer Supported Cooperative Work (CSCW 2004)*
5. Blasch, B., Wiener, W. & Welsh, R., eds. (1997) **Foundations of Orientation and Mobility** (second edition). AFB Press, New York.

Audio-Tactile Interaction at the Nodes of a Block-Based Physical Sound Synthesis Model

Cumhur Erkut, Antti Jylhä, Matti Karjalainen
Dept. Signal Processing and Acoustics
Helsinki University of Technology, Finland
{Firstname.Lastname}@tkk.fi

M. Ercan Altınsoy
Chair of Communication Acoustics
TU Dresden, Germany
Ercan.Altinsoy@ias.et.tu-dresden.de

ABSTRACT

We investigate how to link a physically-based digital sound synthesis structure to a haptic interface for constructing flexible and dynamic virtual audio-tactile environments and musical instruments. Our block-based physical model (BBPM) imposes physicality on the audio signals by transforming them to a variable pair of effort and flow. With this pair, we can define the notions of energy and impedance/admittance for object blocks, and manage their physically-based interaction by using special elements we call nodes. We have recently ported a subset of the BBPM environment to PureData (PD), where we can construct dynamic audio environments akin to a scene graph. Moreover, we can connect off-the-shelf haptic controllers to the entry node of a BBPM, excite the model with one physical variable, provide haptic feedback with the dual variable, and concurrently synthesize audio. We will demonstrate this system with two prototypes during the HAID'08 Workshop.

Author Keywords

Block-based physical modeling, digital sound synthesis, sonic interaction design, prototype, vibrotactile commodities.

ACM Classification Keywords

H.5 HCI, H.5.5 Sound and Music Computing, Signal analysis, synthesis, and processing

INTRODUCTION

In this contribution we investigate how to link a block-based physical model (BBPM) [1] for digital sound synthesis with a haptic interface [2]. Auditory cues frequently occur when we interact with everyday objects. Physically-based methods preserve this interaction fidelity in synthetic sound, and we can naturally control the physical parameters and variables of the models by our gestures and actions. These properties have made physical models favorable in integrated haptic and audio interfaces [3, 4]. Typical implementations, however, support interaction between two objects in a static virtual world, based on a single physical modeling paradigm (usually *modal synthesis*). A BBPM supports dynamic audio environments and multiparadigm physical modeling [1]. It imposes physicality on the audio signals by transforming them to a variable pair of effort and flow. With this dual variable pair, we can define the notions of energy, impedance, and admittance for object blocks, and manage their physically-based interaction by using special elements we call *nodes*.

In [5], construction and playing of virtual musical instruments based on BBPM techniques were investigated. For instance, a virtual xylophone that consists of data gloves, a magnetic tracker for head and hands positions, and audio-visual rendering in a cave-like virtual reality (VR) environment has been reported. The difficulty of implementing tactile feedback at the VR interface and the negative effects of latency have been pointed out. Our aim, therefore, is to experiment with simplified systems and off-the-shelf interfaces that can enhance audio-tactile interaction with low latency and low demands on the rendering architecture.

Here, we present our preliminary prototypes on audio-tactile interaction, i.e., synthesizing audio and haptic feedback at the same time by using a BBPM. We specifically deploy two prototypes that couple a BBPM environment implemented in PureData (PD)[1] with off-the-shelf controller devices. Our first prototype is built upon DIMPLE [6]; an implementation[2] that can render haptics (3D haptic rendering, physical dynamics and collision detection, plus soft-realtime graphics) and audio separately. DIMPLE allocates demanding resources and timing accuracy to the haptics part and enable an asynchronous, event-based communication protocol based on the Open Sound Control[3] (OSC) between these processes.

IMPLEMENTATION AND PROTOTYPES

Both prototypes are built upon *nodes*[4], which enable us to construct dynamic audio environments akin to a scene graph. The nodes can also be interfaced to the external physical world via *terminals* or *ports* [1]. As a haptic experience, a bidirectional port makes us feel a grounded resistance and feedback, whereas a terminal can be used to supply non-grounded feedback (such as impulses and vibration).

A three-port K-node and an abstract dynamic audio environment are depicted in Fig. 1. This structure can be, for example, a part of a taut string terminated by an impedance Z_1, and excited (plucked or struck) by a haptic device connected to the K-port 3 in the figure. and the common junction velocity V_{J} is calculated based on the impedances of the haptic device and other branches (string and termination). This provides feedback immediately, as well as when the reflections from the terminations arrive. Alternatively, a force F_{ext} can be applied to a terminal for nongrounded interac-

[1] `http://puredata.info`
[2] `http://www.idmil.org/software/dimple`
[3] `http://opensoundcontrol.org`
[4] Technically, there are two types of nodes: the *wave and K-nodes*, but here we focus on the K-nodes only [1].

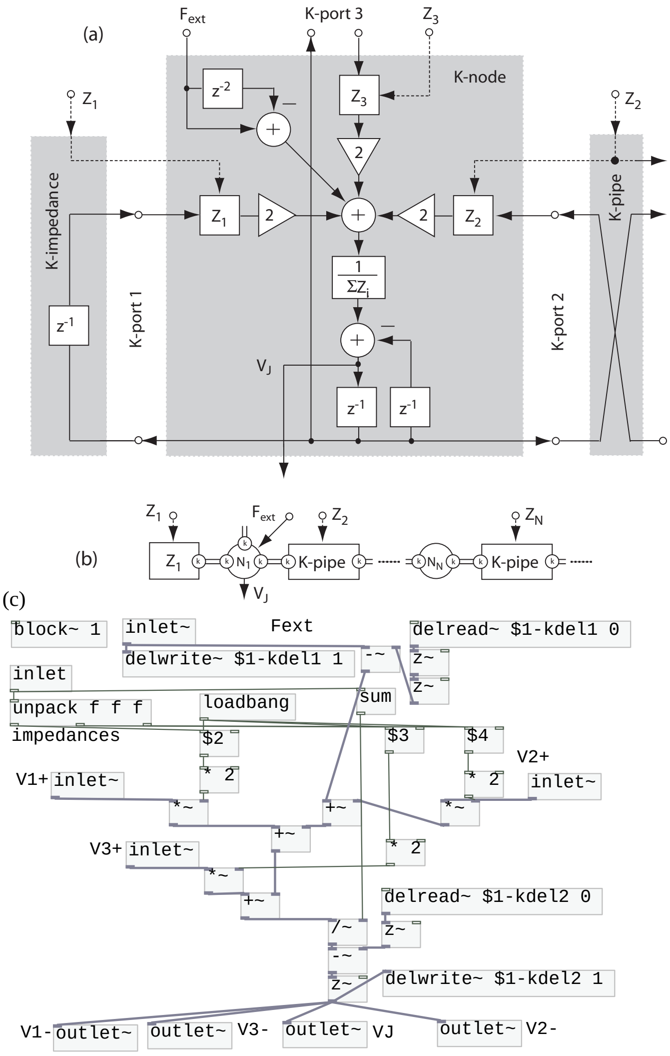


Figure 1. (a) A 3-port K-node with port impedances Z_i, after [5]. A haptic device that can be described by an impedance can be connected to a K-port. Otherwise, a force F_ext can be applied to a terminal to calculate the total velocity V_J. (b) Abstract representation of the K-node in (a). (c) PD implementation of the 3-port K-node in (a). Notice the block-size of one.

tion. The unit delays around V_J implement finite-difference approximation of the ideal wave equation. See [1] for further details.

Recently we have ported the nodes and some other BBPM elements to PD (see Fig. 1(c)). We run the models with minimum audio block-size for bidirectional interaction between the blocks. Albeit inefficient, this heavy duty computation is within the capabilities of current computer systems. We alter the signal propagation by supplying different impedance values, or by dynamically adding new ports or nodes with the internal messaging mechanism of PD.

In our first prototype we extend DIMPLE's OSC namespace to initialize and communicate with our BBPM via a port within PD, as illustrated in Fig. 2. The interface between DIMPLE and the haptic device is well-described in [6], we create a default BBPM with the following command: `/object/bbpm/create myModel`.

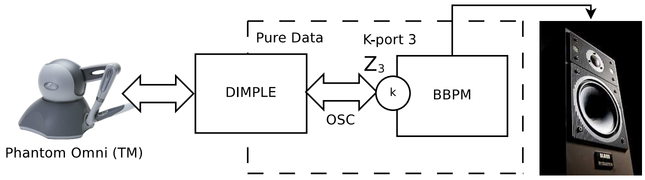


Figure 2. Our BBPM system linked with an impedance-based haptic device via a port (K-port 3 of Fig. 1).

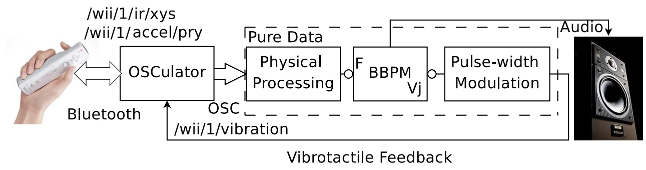


Figure 3. Our BBPM system linked with a Wiimote for simplistic audio-tactile interaction via a terminal force (F_ext in Fig. 1).

The default BBPM thus created is a 3-port K-node previously depicted in Fig. 1. Note that we need to convert the OSC control stream to an audio stream before applying it to a port. Each port except the one that interfaces DIMPLE via OSC can be nested. Port impedances may be changed with messages within PD in runtime.

Our second prototype, illustrated in Fig. 3, couples the Wiimote - the Bluetooth-based controller of the Nintendo's Wii console - to a terminal of our BBPM. We first convert the orientation, position, and the acceleration of the Wiimote (tracked by the OSCulator[5]) to the terminal force F_ext and apply it to the BBPM. We then modulate the only parameter available, namely the pulse-width of the signal feeding the vibrotactile motor by the calculated junction velocity for a nongrounded feedback substituting the first prototype.

REFERENCES

1. R. Rabenstein, S. Petrausch, A. Sarti, G. De Sanctis, C. Erkut, and M. Karjalainen, "Block-based physical modeling for digital sound synthesis," *IEEE Sig. Proc. Mag.*, vol. 24, no. 2, pp. 42–54, 2007.
2. V. Hayward and K. MacLean, "Do it yourself haptics: Part I," *IEEE Robot. Automat. Mag.*, vol. 14, no. 4, pp. 88–104, 2007.
3. D. DiFilippo and D. Pai, "The AHI: An audio and haptic interface for contact interactions.," in *Proc. ACM Symp. UIST*, (San Diego, CA USA), pp. 149–158, Nov. 2000.
4. F. Avanzini and P. Crosato, "Integrating physically based sound models in a multimodal rendering architecture," *Computer Animation and Virtual Worlds*, vol. 17, pp. 411–419, July 2006.
5. M. Karjalainen and T. Mäki-Patola, "Physics-based modeling of musical instruments for interactive virtual reality," in *Proc. Int. Workshop on Multimedia Signal Processing*, (Siena, Italy), pp. 223–226, Sep 2004.
6. S. Sinclair and M. M. Wanderley, "Using PureData to control a haptically-enabled virtual environment," in *Proc. PureData Conv.*, (Montreal, Canada), Aug 2007.

[5] http://www.osculator.net

An Audio-Haptic Sound Browser Using Distance Cues

Delphine Devallez
University of Verona
Strada Le Grazie 15
37134 Verona, Italy
delphine.devallez@univr.it

Federico Fontana
University of Verona
Strada Le Grazie 15
37134 Verona, Italy
federico.fontana@univr.it

Davide Rocchesso
IUAV of Venice
Dorsoduro 2206
30123 Venezia, Italy
roc@iuav.it

ABSTRACT

Spatialization of sound sources in depth allows a hierarchical display of multiple audio streams and therefore may be an efficient tool for developing novel interfaces for menu navigation and audio browsing. An interface concept will be demonstrated that allows audio browsing based on distance cues. The haptic interface includes a linear position tactile sensor made of conductive material. For consumer applications it may consists of a touchpad that is nowadays commonly available on portable media players and personal digital assistants. The touch position on the ribbon is mapped onto the listening position on a rectangular virtual membrane, modeled by a bidimensional Digital Waveguide Mesh and providing distance cues of four equally spaced sound sources. Therefore, by moving the finger on the ribbon the user may navigate through four audio files. For a practical use with a longer playlist, the virtual auditory environment acts as an audio window which can be moved along the whole playlist with the knob of a MIDI controller. Subjects involved in an auditory search task found the interface intuitive and entertaining.

Keywords

Audio-haptic interface, auditory navigation, distance perception, spatialization, digital waveguide mesh

Categories and Subject Descriptors

H.5.2 [**Information Interfaces and Presentation**]: User Interfaces

1. INTRODUCTION

While most research dedicated to the design of new auditory interfaces focuses on directional spatialization of multiple sound sources, we propose an interface based on distance information. The motivation is driven by the ability of depth information to provide a hierarchical relationship between objects and therefore bring the attention of the user on the closest sound source while still hearing the other ones in the background. In 1990, Ludwig [2] already suggested that techniques used in the music industry, such as reverberation and echo, could be valuable to the ordering of multiple sound sources in auditory interfaces.

2. MODELING OF THE AUDIO SPACE

In recent years, research on the Digital Waveguide Mesh (DWM) has enabled the use of discrete-time simulations

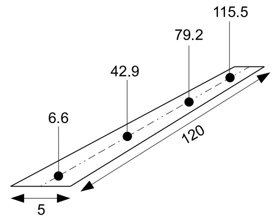


Figure 1: *Positions of the sound sources on the virtual membrane. All sizes are in centimeters.*

of acoustic propagation as a model for real-world acoustic spaces. Fontana and Rocchesso [1] have recently demonstrated the effectiveness of a DWM modeling a rectangular parallelepiped to provide auditory distance cues. Our sound spatialization model consists of a two-dimensional rectangular DWM. Each internal junction is connected to four other junctions via waveguides, thus providing acoustic wave transmission. Besides, reflections at the mesh boundaries are modeled by Digital Waveguide Filters, whose coefficients have been tuned to model specific reflective properties of surfaces. Finally, the number of nodes can be converted into the corresponding membrane dimensions once the speed of sound and the sampling frequency of the simulation have been determined. Due to software restrictions, the mesh dimensions are chosen to be 110×5 nodes, which correspond to a $120 \times 5\,\text{cm}^2$ membrane. Then, it was found by an informal user study on audio browsing that at most four concurrent audio streams could be discriminated in the resulting virtual environment. The four positions of the sound sources are equally spaced along the main axis of the membrane, placed at 6.6, 42.9, 79.2 and 115.5 cm respectively (See Fig. 1). The present model allows to render dynamic distance information of a sound source, thanks to the decrease of the intensity and of the direct-to-reverberant energy ratio with distance. Figure 2 shows the variation of the intensity with distance. By comparing with the energy decrease in open space, characterized by a reduction of 6 dB per distance doubling, it can be seen that the overall intensity on the membrane decreases significantly less. This behavior allows users to hear even the farthest sound sources at any location on the membrane. In addition, the volume can be manipulated by users, which might make the intensity cue unreliable for judging distance. Another reason for limitating the intensity cue for distance judgment is that the level of direct sound varies both with distance and with the energy emitted from the sound source, so that the listener needs some

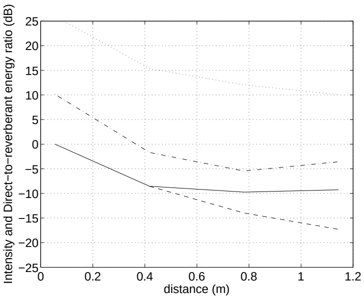


Figure 2: *Intensity in the mesh (Solid line) and in open space (Dashed line). Direct-to-reverberant energy ratio in the mesh (Dashdot line) and in a natural environment (Dotted line).*

a priori knowledge about the sound source level in order to evaluate its egocentric distance. Figure 2 also shows the direct-to-reverberant energy ratio in the mesh as a function of distance. For comparison the direct-to-reverberant energy ratio v was computed for a natural environment, modeled by Zahorik with the function $v = -3.64 \log_2(r) + 10.76$ [3]. The 2D mesh has a direct-to-reverberant energy ratio much lower than in the natural auditory space, which means that the amount of reverberation is exaggerated in the virtual environment.

3. THE HAPTIC USER INTERFACE

The user haptic interface consists of a ribbon controller, the Infusion Systems *SlideLong*, inspired by music controllers. It has an active area of $210 \times 20\,\mathrm{mm}^2$, and therefore exhibits a main dimension along which distance information may be explicitly understood. The position tactile sensor gives a value corresponding to the touch position on the ribbon. A gamepad is used as a fast, simple and low-cost interface sensor, and is connected to the USB-port of an Apple MacBook Pro computer. After rescaling, the incoming value from the ribbon controller sets the listening position input to the computation of the auditory signals in the DWM. Since both the ribbon and the DWM have a rectangular geometry, a coherent mapping is performed between the touch position on the ribbon and the position of a virtual microphone in the DWM, and by moving the finger on the ribbon the user may explore the virtual environment where different audio streams are being attributed different positions. Like music controllers, the touch sensor intends to provide an interface that is intuitive to use with immediate and coherent response to user's gesture. Since only four sound sources can be rendered in the DWM (See section 2), a second haptic interface is added in order to allow audio browsing among more sound files. For this purpose, we use one of the rotary encoders of a keyless MIDI controller (the Novation *ReMOTE ZeRO SL*). The encoder has discrete steps and therefore may be easily manipulated to switch between discrete levels. Incrementing the value of the encoder enables to move three sound files forward or backward in the playlist. The overlap of one sound file between consecutive audio windows is intended to model advancing in a linear space and consequently to make the user explore the whole set of sound files as a monodimensional space, which may be accessed through a window of four sound sources. Figure 3 shows the complete setup.

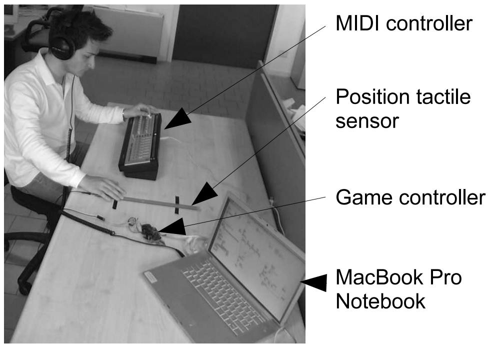


Figure 3: *Picture of the setup.*

4. DISCUSSION

A user study showed the ability of the tool to find audio files, in an entertaining and intuitive manner. The acoustic environment provided by the DWM needs however further investigations to evaluate its “naturalness”, and could be compared to other techniques such as crossfading or stereo panning. Though, spatialization in depth outputs a monaural signal and therefore does not rely on the reproduction hardware configuration.

The user interface presented in this paper offers various research directions, both in terms of interaction design and auditory perception. Among the numerous issues, the nature and the number of concurrently playing sound sources in the audio window most probably affect the performance results in a search task. A compromise should be found between auditory overload and performance. Finally, other uses of the interface may be explored. In particular, an auditory menu may originate from carefully sonified menu items, spatialized in depth using the DWM, and accessed through a linear tactile interface such as a touchpad. Originally used as a substitute for a computer mouse, touchpads are now invading the world of portable media players and personal digital assistants. They are used as a control interface for menu navigation on all of the currently produced iPod portable music players. Another illustrative example is the iPhone that may be used as a touchpad to wirelessly control a computer. The proposed spatialization in depth of sound sources for navigation could therefore benefit from the already existing hardware.

5. ACKNOWLEDGMENTS

This research work has been supported by the European project FP6-NEST-29085 CLOSED - Closing the Loop of Sound Evaluation and Design.

6. REFERENCES

[1] F. Fontana and D. Rocchesso. A physics-based approach to the presentation of acoustic depth. In *Proc. of the 2003 International Conference on Auditory Display*, pages 79–82, July 2003.

[2] L. Ludwig. Extending the notion of a window system to audio. *Computer*, pages 66–72, 1990.

[3] P. Zahorik. Direct-to-reverberant energy ratio sensitivity. *J. Acoust. Soc. Am.*, 112(5):2110–2117, November 2002.

Pressure Player: Combined Pressure and Audio Interaction

David McGookin, Stephen Brewster
Department of Computing Science
University of Glasgow, Glasgow G12 8QQ
{mcgookdk, stephen}@dcs.gla.ac.uk

ABSTRACT
Interacting with touch screen devices demands that users look at the screen for even trivial operations. We propose a music player that allows users to preview following tracks by applying pressure on the player screen. We couple this with dynamic modification of the music being played. We present two alternative audio designs, and outline our future evaluation plans.

INTRODUCTION
There has been a recent trend with many mobile device manufacturers to move away from physical controls and buttons, and towards touch screen interaction. Devices such as Apple Inc's iPhone and iPod touch, have shown many of the advances that touch screen interaction affords, such as a dynamically reconfigurable interface allowing different functions of the device to use an optimised interface. For example, the iPod touch uses "disappearing" controls for the movie player function, allowing the entire display to be used to show the movie being played, with the controls overlaid only when the user wishes to interact. There are however drawbacks when using such interaction. Notably, there is no tactile relief on controls. The device must be removed from the user's pocket or bag and looked at for all operations, including those that could previously be done "in pocket" using tactile feedback of the controls. Other researchers have proposed augmenting screens with vibrotactile feedback to provide a tactile sensation when operating virtual buttons [2]. However in this paper, we propose that pressure input could be used to provide a limited subset of operations on touch screen music players. These could be coupled with other haptic technologies, such as vibrotactile feedback, to create a dynamic and useful "eyes-free" interaction.

RELATED WORK
Previous work on pressure based interaction has provided only visual feedback to the user about the level of pressure being applied, and the meaning of that pressure in the interface. Ramos *et al.* [3] carried out various studies involving a Wacom pressure sensitive graphics tablet, where participants selected a target with a cursor controlled by varying amounts of pressure applied via the tablet pen. Ramos *et al.* identified several important findings from their work, including that feedback on the pressure applied, and its meaning in the interface, should be continuous rather than discrete, and that a dwell-based selection mechanism (where the user maintains the pressure applied in a particular range for a duration of time) was the most accurate, but not the fastest technique. Chechanowicz *et al.* [1] found that the number of effective pressure levels depended on the force sensing technology used.

Figure 1. A Screenshot of the Pressure Player Interface.

PRESSURE PLAYER
Whilst work on pressure looks promising, there is a lack of work on both mobile devices, and providing non-visual feedback. We intend to investigate how audio can be used as feedback for pressure based interaction, with a particular view on the situations where current touch screen music players fail (see the introduction). Our test bed application, Pressure Player, is a simple touch screen mp3 player (see Figure 1) that runs on a mobile device. The player allows users to listen to a playlist of songs. Information about the songs is displayed, and users can skip forward or back through tracks, as well as play and pause the music. The player is implemented on a Nokia N800 internet tablet. The N800 is equipped with a touch screen capable of detecting around 500 useful pressure levels. Pressure Player uses these pressure levels to allow the user to easily preview following tracks by applying a force to the touch screen. In our initial implementation we allow the user to preview and/or select the following two tracks in the playlist using pressure.

Mapping Pressure to Sound
Our aim, in providing a sonification of forthcoming tracks, is to provide some feedback to the user of how much more pressure needs to be applied to "pop through" to the next track. As noted by Ramos *et al.* [3], dynamic feedback is more useful than discrete feedback. Providing feedback in such a way should help the user to vary the pressure applied to stay on a particular track. To map feedback to pressure levels we use a probability density function (pdf). This is a function which can be used to calculate the distribution of values for a random event when the mean is know. For each of the three audio sources that the user can browse (the currently playing (or selected) track as well as the following two), we center that source on a particular pressure value. Using that pressure value as the mean, we calculate a proba-

bility density value for each track at each possible pressure value. This acts as a crude measure of probability that the user is trying to preview that track (see Figure 2). The closer the input pressure is to this value, the better the preview of that track will be, and the ability of the user to select it as the playing track. E.g. if the current touch screen pressure applied is around 101, there is greater probability that the user is trying to preview track 2 rather than tracks 1 or 3 (see Figure 2). We use this to drive the audio feedback. We discuss more on the alternative audio designs in the following sections. In order to select a particular track as the currently playing one (e.g. track 2), the user must perform the selection operation within a smaller pressure "bin", again centred on the target pressure for that track. Although Ramos *et al.* [3] found that selection by dwelling in the pressure bin was the most accurate, it is unlikely to be an appropriate mechanism here, as the user may dwell in the pressure bin simply to preview an audio track. Instead we use the quick release metric of selection based on the rate at which pressure falls. I.e. the rate at which the user removes his or her finger from the touch screen.

Dynamic Audio Feedback Designs

Our initial implementation of Pressure Player uses a simple pressure to volume mapping. However we are considering how another technique, granular synthesis, might be applied.

Volume Mapping

Volume mapping is conceptually the simplest design. Here all audio tracks are played simultaneously, but the volume of each is mapped to that track's probability density function at the current pressure level. E.g. if the user applies a pressure of 101 (see Figure 2), track 1 will be played at about 5% of its volume, track 2 at 35% and track 3 at 0%. Therefore the track with the highest probability will be played with the highest volume. Volume is limited to the user selected global volume. I.e. if the user is playing music at 50% of the level the device is capable of, a probability of 1 from the probability density function will be mapped at 50% of the overall device volume level.

Granular Synthesis

An alternate option for audio feedback, granular synthesis [4], is more complex. Here small samples of each audio stream (grains) are selected from each audio track and randomly combined creating a discordant noise. As a user approaches the selection bin of a track (and the probability of that track increases (see Figure 2)) more grains are chosen from that track causing it to "emerge" from the audio mix. Metaphorically this is like tuning an old fashioned radio to a different station; as the pressure level changes, the audio breaks to static before "tuning" into the next track [1].

Each of these options offers advantages and disadvantages for interaction. The volume control system is the most obvious to understand. However even if the tracks used are normalised for volume, there will still be quiet and loud parts in different songs that may confuse the user as to the pressure

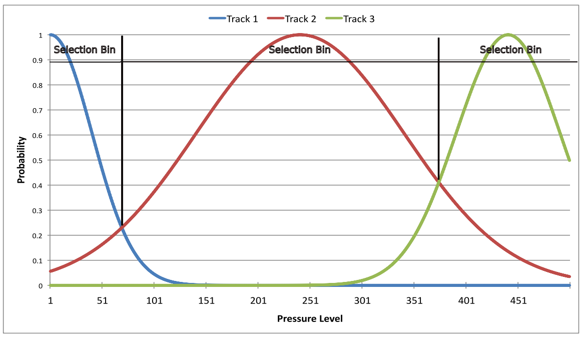


Figure 2. Probability density functions are calculated at each possible pressure value to determine the probability that the user is trying to preview a particular track.

level being applied. This may be less of an issue if the user is familiar with the track being played. The granular synthesis technique, whilst overcoming these problems has other potential issues, as the playing of white noise (as will occur when the probability of all of the audio streams is low) may be annoying to the user. To measure the extent of these problems we intend to carry out an experiment to compare and contrast each technique.

CONCLUSIONS

Whilst touchscreen interaction provides many compelling benefits for mobile devices, it requires visual attention from the user for even simple operations, which may not be convenient or possible. In this paper we have discussed the possibility of using pressure based interaction to effect non-visual useful interaction in a music player. We are currently considering alternate audio mappings, and also designing an experiment to evaluate our different audio designs as well as the number of pressure levels that are possible. We believe the results of this will show that pressure based interaction can be a useful and compelling mobile haptic and audio interaction design.

ACKNOWLEDGMENTS

This work is supported by Nokia and EU framework 6 OpenInterface (IST-35182) and framework 7 HaptiMap (ICT-224675) projects.

REFERENCES

1. Cechanowicz, J., Irani, P., and Subramanian, S. Augmenting the mouse with pressure sensitive input. In *Proceedings of CHI 2007* (San Jose, California, USA, 2007), ACM, pp. 1385–1394.

2. Hoggan, E., Brewster, S., and Johnston, J. Investigating the effectiveness of tactile feedback for mobile touchscreens. In *Proceeding of CHI2008* (Florence, Italy, 2008), ACM, pp. 1573–1582.

3. Ramos, G., Boulos, M., and Balakrishnan, R. Pressure widgets. In *Proceedings of CHI 2004* (Vienna, Austria, 2004), ACM, pp. 487–494.

4. Williamson, J., Murray-Smith, R., and Hughes, S. Shoogle: excitatory multimodal interaction on mobile devices. In *Proceedings of CHI 2007* (San Jose, California, USA, 2007), ACM, pp. 121–124.

[1] For this to be successful the probability across all audio sources at each pressure level must add up to 1. This does not always happen so the remainder will come from another "track" of white (static) noise, which is not shown in Figure 2

Rotational Dynamics for Multimodal Mobile Feedback

Steven Strachan
Hamilton Institute
NUI, Maynooth
Maynooth, Ireland
steven.strachan@nuim.ie

Roderick Murray-Smith
Dept. of Computing Science
University of Glasgow
Glasgow, Scotland
rod@dcs.gla.ac.uk

ABSTRACT

Rotational dynamic system models can be used to enrich tightly-coupled embodied control of movement-sensitive mobile devices, and support a more bidirectional, negotiated style of interaction. A simulated rotational spring system is used for natural eyes-free feedback in both the audio and haptic channels using twisting and tilting motions to interact with content. Users perceive the effect of varying parameters in the simulated dynamic system.

Author Keywords

Rotational dynamics, tilt-input, vibrotactile, accelerometer, multimodal, mobile, mobile spatial interaction.

INTRODUCTION

We present an interaction metaphor that uses the simulation of a rotational dynamic system driven by inertial sensor data, to facilitate and enrich interaction with a mobile device. Such tilt-based interaction is well-established [1], but typically depends on visual feedback, which in tilting interactions can cause difficulty due to poor screen visibility. Our intention in using a tangible physical metaphor for interaction is that users instantly possess a natural intuition for the effects that their movements have on a system. Feedback is provided via audio and haptic rendering of the internal states of the simulated system and which aspects of the mechanical system state you choose to feedback be it torque, velocity, friction or a combination, depends on the metaphor presented.

The physical modelling approach enables the provision of continuous interaction and rich feedback during the gesture-like interaction, which has advantages when learning a new interaction, or using the system in adverse conditions. This approach can be applied to any tilt-based system as a way to enrich the interaction experience but as an example we introduce here the simulation of a rotational spring system for twisting-based interaction with a mobile device and we show how basic aspects of this kind of system can be associated with information, and can be manipulated by a user in an intuitive fashion.

BACKGROUND

Eyes-free interfaces rely heavily on the provision of effective audio and vibrotactile sensations. Yao and Hayward [2] investigated the simulation of physical systems with audio and vibrotactile feedback, recreating the sensation of a ball rolling down a hollow tube via the haptic and audio modalities. Using apparatus that simulated the physics and provided audio and haptic cues, they found that when subjects were asked to estimate the position of the ball rolling inside a tubular cavity, they used their natural intuition of objects falling under the influence of gravity to accurately estimate the position. Similarly, Rath and Rochesso [3] created a convincing sonification of the physical motion of a ball along a beam, finding that subjects were able to perceive the ball motion from the sonification alone. Shoogle [4] enables the sensing of the state of a mobile device via the simulation of a physical system which responds to gestural input. By modelling the relatively simple dynamics of some balls inside a box and the quite intuitive effects of a users shaking of this box, information can be conveyed, such as the battery life of the device or number of new text messages via the use of auditory impact sounds and haptic rendering. Each new text message is represented by one simulated ball sensed only by the shaking of the device

ROTATIONAL SPRING SYSTEMS

There are a wide range of physical systems that could be used as metaphors for interaction and rotational spring systems are just one example. This kind of system has a number of features that make it appropriate for interaction design. Sensing the orientation state of the device via accelerometers allows us to use movement of the device to control the interaction. Twisting the mobile device, sensed via changes in roll angle, or tilting, sensed via changes in the pitch angle can be easily sensed and used to provide eyes-free feedback about the state of the device to the user. Some rotational metaphors that we can simulate using the dynamics of this kind of system include winding a clock, opening a door knob, turning a key or opening a box, all completely natural everyday metaphors for which people have a natural intuition and which can enhance and enrich the process of interaction. The four important characteristics of this kind of mechanical system from an interaction design perspective are *torque*, *friction*, *stiffness* and *mass*, which can be used to feedback device states to the user. We focus on torque in this example. Torque is important because it provides us with a measure for the amount of force present in the system. If this is fed back to the user in some way, via the audio or haptic channels, it can provide the user with a sense of how the system is reacting to certain events or movements.

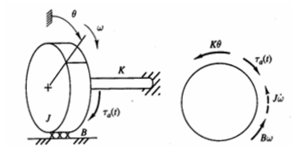


Figure 1. one-disk system and free body diagram showing all forces acting on the disk and illustrating the basic notation.

For a linear torsional spring or flexible shaft $\tau = K\Delta\theta$ where θ is the angular displacement and K is the spring constant. Altering the value of K can then have an effect on the overall feel of the system. For example, higher K values result in a more stiff system. Figure 1 shows a basic roational spring system with 1 disk connected by a flexible shaft. A free body diagram shows the forces acting on this disk, which may be summed to produce an equation for this system:

$$j\dot{\omega} + B\omega + k\theta = \tau(t)$$

Where J is the moment of inertia in the disk, B is the friction between the disk and the surface, ω is the angular velocity and θ is the angular displacement of the disk. We may now view the mobile device as being a minimal inertia element, coupled with a rotational system via a rotational stiffness element. Angle changes in the orientation of the device, sensed from accelerometers, act as reference values which drive the rotational system of interest, with the states of that system again fed back to the user via vibration or audio.

TWO DISK SYSTEM

The 1 disk system provides us with a very simple mechanical example but by adding another disk and reforming the system we have much more flexibility in our interaction design

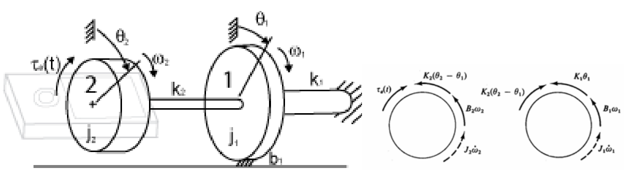
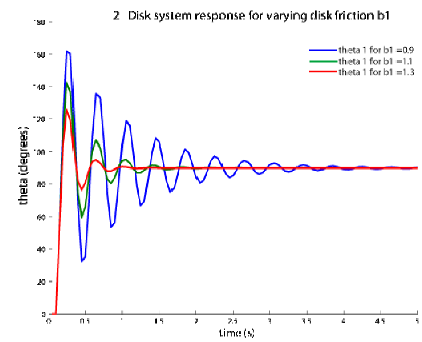


Figure 2. two-disk system and free-body diagram showing various forces acting on the two disks.

The system we have chosen to simulate here, the `Two-Disk' system is illustrated in figure 2. In this system we treat the angular displacement θ on disk 2 as an input to the system in order to observe the effects on θ_1 and ω_1 on disk 1. By again summing the forces present in the free-body diagram for disk 1 and converting to a state space model as described in [6] we may represent this system as follows:

$$\begin{bmatrix} \dot{\theta}_1 \\ \dot{\omega}_1 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -(K_2 + K_1)/J_1 & -B_1/J_1 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \omega_1 \end{bmatrix} + \begin{bmatrix} 0 \\ K_2/J_1 \end{bmatrix} \theta_2$$

where K_1 and K_2 are the stiffness constants in shaft 1 and shaft 2 respectively, B_1 is the friction element for disk 1 and J_1 is the moment of inertia for disk 1. If we imagine our mobile device to be represented by disk 2 and we exert some kind of roll-axis rotation on the device, this will induce a reaction in disk 1, the exact nature of which depends on the values chosen for K_1, K_2 and B_1. Figure 3 shows a typical response in the displacement angle θ for disk 1 from the two-disk system for varying values of B_1 after disk 2 has been twisted through 90°. As the friction parameter is increased, the simulated response of disk 1 to the input from disk 2 becomes increasingly damped. This more damped response, when fed back via the haptic and audio channels, can be clearly perceived by the user of the mobile device. Similar responses are observed for the varying of the K_1 and K_2 parameters of this model.



CONCLUSION

We have introduced and demonstrated the use of a mechanical dynamic system for the provision of rich feedback to the user of a mobile device. By taking advantage of user's natural familiarity with the dynamics of a rotational spring mechanical system, as would be found in a door-handle, we have shown that it is possible to produce an eyes-free multimodal display using a solid theoretical foundation as the basis for the interaction. Coupling such rich continuous-feedback models with real-time inference about the content and relevance of information sources on the mobile device allows designers to make the user aware of subtle variations in the nature of the content they are engaging with and manipulating

REFERENCES

1. K. Hinckley. Input technologies and techniques. In *Handbook of Human-Computer Interaction*, pages 151{168. Lawerence Erlbaum Associates, 2002.
2. H.-Y. Yao and V. Hayward. An experiment on length perception with a virtual rolling stone. In *Eurohaptics 06*, 2006.
3. M. Rath and D. Rocchesso. Continuous sonic feedback from a rolling ball. *IEEE MultiMedia*, 12(2):60-69, 2005.
4. J. Williamson, R. Murray-Smith, and S. Hughes. Shoogle: excitatory multimodal interaction on mobile devices. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 121{124, New York, NY, USA, 2007. ACM.
5. P. Eslambolchilar and R. Murray-Smith. Model-based, multimodal interaction in document browsing. In *Multimodal Interaction and Related Machine Learning Algorithms*, 2004

Crossmodal Combinations: Using Piezo-Electric, Vibrotactile and Audio Feedback

Eve Hoggan[1], Stephen Brewster[1] and Topi Kaaresoja[2]

Glasgow Interactive Systems Group
University of Glasgow, Glasgow, G12 8QQ, UK
{eve,stephen}@dcs.gla.ac.uk

Nokia Research Center
Helsinki, Finland
topi.kaaresoja@nokia.com

ABSTRACT

In this paper we introduce intramodal tactile combinations where two different types of tactile feedback are presented simultaneously using different actuators and the use of interchangeable audio/tactile feedback using CrossTrainer, a fully crossmodal mobile touchscreen application.

Author Keywords

Tactile feedback, audio feedback, crossmodal interaction.

ACM Classification Keywords

H5.m. Information interfaces and presentation: Haptic I/O, Auditory (non-speech) feedback.

INTRODUCTION

Audio and tactile feedback are becoming prevalent features in mobile touchscreen devices and recent studies [1-4] have indicated that such feedback can be beneficial to users, increasing typing speeds and reducing errors.

So far, however, the majority of audio and tactile research has been unimodal; there has been little consideration of any crossmodal possibilities. In some situations it can be beneficial for users to be able to switch from audio to tactile feedback and vice versa. For example, on a building site with high noise levels, tactile feedback may be more appropriate whereas on a bumpy train ride, audio feedback may be more suitable.

Furthermore, this research tends to focus on design parameters and the type of information encoded in each modality. There are few complete multimodal or crossmodal applications in existence as yet. This paper presents CrossTrainer: a mobile touchscreen game based on traditional IQ training, which makes full use of crossmodal audio and tactile feedback allowing modalities to become interchangeable, i.e. to provide the same interaction feedback, enabling users to select the most appropriate modality given their usage context or personal preference.

CrossTrainer makes use of a mobile touchscreen device that can provide tactile feedback in two ways: through the vibrotactile actuator or piezo-electric actuators [5]. We exploit this advancement in technology by using an intramodal combination [6] i.e. combining feedback from both types of actuator, creating new types of tactile texture not possible before.

CROSSTRAINER

We chose to include crossmodal feedback in a game because CrossTrainer requires a great deal of interaction with the user through many different types of widget and UI events. This allows us to incorporate and study a wide range of crossmodal audio and tactile feedback whilst remaining an enjoyable experience for the user.

CrossTrainer has been implemented on the Nokia 770 Internet Tablet, a commercial device which has been augmented with piezo-electric actuators [5] and a standard vibration motor. Tactile stimuli were created with a proprietary script language implemented on the device while the audio stimuli use standard wave files played through the device's stereo speakers.

There are 200 tasks in CrossTrainer (Figure 1) all of which are designed to test and train the user's IQ. Users are presented with random sets of 20 tasks each with a time limit of 40 seconds. There are five types of task involving different audio/tactile feedback: mathematics, true or false, reaction speeds, logical reasoning and general knowledge. Users are required to enter answers via the crossmodal touchscreen widgets. Upon completion, users are informed of their CrossTrainer IQ score in terms of brain age.

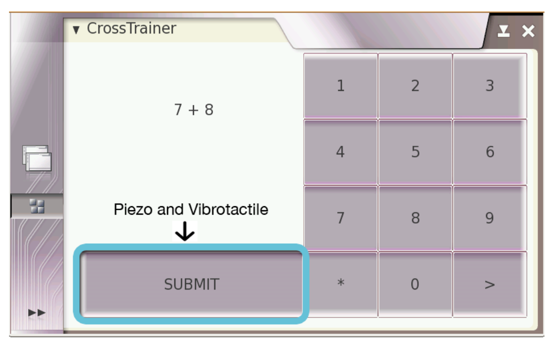


Figure 1: CrossTrainer screenshot with example task.

CrossTrainer uses audio and tactile feedback design based on crossmodal icons [7], which are abstract icons that can be instantiated in one of two equivalent forms (auditory or tactile). Three parameters have been chosen for the feedback design: rhythm, texture, and location. The type of CrossTrainer widget is encoded in the rhythm (QWERTY button, number button, radio button, scroll bar and notification dialogue), the widget's location on the display is encoded in spatial location (if the buttons are on the left of the screen, the audio feedback will be panned to the left and the tactile feedback will be provided by the piezo actuator under the left-hand section of the screen) and urgency is encoded in texture (i.e. as every 10 seconds pass and the time limit for the task runs out, the feedback provided by the widgets increases in roughness and intensity). Therefore, 5 different rhythms and 4 different levels of texture produce a set of 20 crossmodal icons: 20 Earcons and 20 Tactons each capable of providing the same feedback at different spatial locations.

The crossmodal rhythms and spatial location are based on parameters previously used in research on multi-dimensional icons [7]. The most novel aspect of design in CrossTrainer is the different audio and tactile textures used in the crossmodal feedback. The creation of these new sensations is the focus of this paper.

Texture

Two tactile textures were created using different waveforms established in previous work [8] and investigations into the use of frequency and intramodal tactile textures led to the creation of two completely new textures.

Task urgency is encoded in the texture of each widget. For example, when pressing the number keypad buttons in tactile mode, a 2 beat rhythm is used and it becomes increasingly rough as the task time limit approaches. This allows users to keep track of how much time is left before an answer must be submitted without having to switch their visual focus away from the task to look at a small clock or other type of alert displayed visually on the screen.

Time (secs)	40	30	20	10
Texture	Smooth	Semi Rough	Rough	Very rough, high intensity
Tactile	Sine wave	Square wave	Random increasing frequencies	Intramodal combination (piezo and vibrotactile)
Audio	Flute	Tremolo Horn	Guiro	Piano and cello

Table 1: Urgency and Texture Mapping in CrossTrainer

As shown in Table 1, with 40 seconds remaining, a tactile rhythm is presented using a smooth piezo-electric pulse much like a sine wave while the audio rhythm is played by a flute. With 30 seconds remaining, the same tactile rhythm occurs when a widget is touched but this time with a rougher texture shaped like a square wave from the piezo-electric actuators and the audio rhythm is played by a tremolo (soft vibrating) horn. Then, when there are 20 seconds to go, a much more rough version of the rhythm is presented. This is created using a piezo-electric pulse made up of random increasing frequencies ranging from 1 to 400Hz. The audio is a 10ms burst from a guiro (a percussion instrument played using a scraping motion).

Using an Intramodal Tactile Design

Lastly, to create a very urgent sensation during the last 10 seconds of each task, a very rough and intense (almost bouncy) sensation has been created using a novel technique involving the use of intramodal combinations. Now that devices can incorporate multiple types of tactile actuator, the use of intramodal combinations to produce new textures and sensations is possible. Piezo-electric actuators can create short display-localised tactile bursts, by moving the touch screen display module [5]. The piezo is also able to generate single-transients resembling the tactile feedback in physical buttons while the conventional vibrotactile motor is optimized for longer vibrations, where the whole device mass shakes without any localisation. In this case, both the vibrotactile actuator and piezo-electric actuator are activated simultaneously. This leads to a sharp piezo bump combined with long rough vibrations (Figure 2). The piezo-electric actuator maintains the spatial location parameter while extra strength is added through the vibrotactile actuator.

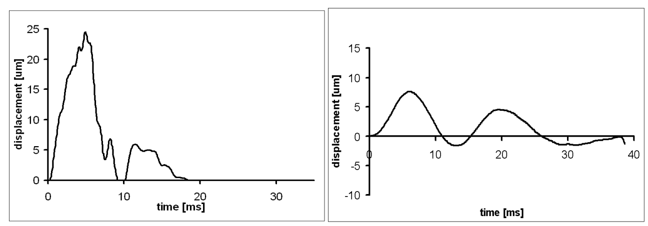


Figure 2: Example piezo-electric and vibrotactile output

Combining two different types of tactile feedback is similar to the use of chords in the audio modality played by two different instruments. In this case the audio feedback consists of a chord played by a piano (a sharp mid-range note) and cello (a long tremolo low-range note).

CONCLUSIONS

This paper has described the use of crossmodal audio and tactile feedback in a mobile touchscreen application called CrossTrainer. By applying all of the previous work on crossmodal icons we have shown that completely crossmodal applications can be created where both modalities can provide the same interaction feedback, and are therefore interchangeable. Furthermore, the use of intramodal combinations to create different tactile textures has been introduced. It has been found that combining the feedback from piezo-electric actuators and vibrotactile actuators can result in a completely new type of tactile sensation. In this case, a powerful, 'bouncy' but rough texture was created.

So far, initial pilot studies have been conducted and have indicated that it is possible to switch between modalities during the game without increasing difficulty for users. It is hoped that future studies of CrossTrainer will show that although crossmodal audio and tactile feedback may not actually improve game scores, it will aid users in entering answers quickly and accurately using a variety of different widgets.

REFERENCES

[1] Brewster, S. A., Wright, P. C. and Edwards, A. D. N. An Evaluation of Earcons for Use in Auditory Human-Computer Interfaces. In *Proc InterCHI 1993*, ACM Press (1993), 222 - 227.

[2] Hoggan, E., Brewster, S. A. and Johnston, J. Investigating the Effectiveness of Tactile Feedback for Mobile Touchscreens. In *Proc CHI' 08*, ACM Press (2008), 1573 - 1582.

[3] Kaaresoja, T., Brown, L. M. and Linjama, J. Snap-Crackle-Pop: Tactile Feedback for Mobile Touch Screens. In *Proc Eurohaptics '06*, (2006), 565 - 566.

[4] Poupyrev, I. and Maruyama, S. Tactile Interfaces for Small Touch Screens. In *Proc 16th Annual ACM Symposium on User Interface Software and Technology*, ACM Press (2003), 217 - 220.

[5] Laitinen, P. and Mäenpää, J. Enabling Mobile Haptic Design: Piezoelectric Actuator Technology Properties in Handheld Devices. In *Proc HAVE'06*, (2006),

[6] Lederman, S. J. and Klatzky, R. L. New Directions in Touch. *Canadian Journal of Experimental Psychology 61, 3* (2007), 169 - 170.

[7] Hoggan, E. and Brewster, S. A. Designing Audio and Tactile Crossmodal Icons for Mobile Devices. In *Proc Proceedings of the 9th international conference on Multimodal interfaces*, ACM (2007), 162 - 169.

[8] Hoggan, E. and Brewster, S. A. New Parameters for Tacton Design. In *Proc Vol II CHI 2007*, ACM Press (2007), 2417 - 2422.