Towards evaluation of performance, control, and preference in physical and virtual sensorimotor integration

Georg Essl * Charlotte Magnusson† Joakim Eriksson† Sile O'Modhrain‡
(*) Deutsche Telekom Laboratories & TU-Berlin, Germany
(†) Certec, Lund University of Technology, Sweden
(‡) SARC, Queens University, Belfast, UK
E-mail: georg.essl@telekom.de, charlotte.magnusson@certec.lth.se, joakim.eriksson@design.lth.se, s.omodhrain@qub.ac.uk

Abstract

This paper presents a pilot study on interfaces utilizing the integration of sensory modalities in the presence of motor action. This sensorimotor integration *is studied in various guises using both physical and virtual display of force and auditory stimuli in three complex object manipulation tasks. Participants are asked to judge number of solid objects under variation of haptic and auditory presentation in the absence of vision, describe believability and likeability of different stimuli settings and estimate how much control they have in a noisy interaction. We see evidence that participants are good at estimating how much control they have as well as at blindfolded counting. Believability and likeability responses give some indication that a relationship between motor action, haptic sensation and auditory response plays a role.*

1. Introduction

Sensorimotor integration is a key concept for designing enactive interfaces. The particular way in which the senses play together in the context of user actions may well influence the success of the actions, whether it is subjectively, in terms of preference and enjoyment, or objectively, in terms of performance in externally defined tasks.

In order to find some preliminary indications of the working of sensorimotor integration we have conducted a pilot study, which test for both objective and subjective measures while additionally estimate the users' judgement of being in control in the settings.

To this end three different experimental settings — two basic technologies and one natural — were used. One technology was a new interface for interactions with granular objects called PebbleBox [11]. PebbleBox allows for modification of the played sound while maintaining the visual and tactile behavior of a physical setting and hence varies only one parameter of this sensorimotor integrated experience. The other parameters remain physical. The other technology was a phantom interface optionally combined with a visual and audio display. In this case the different modal contributions to a sensorimotor experience can all be varied and are all virtual.

2. Related work

The contribution of our various senses to our perception of the world is a subject with a large body of literature in its own right.. (for a recent comprehensive review see [3]).

For a specific recent example of multimodal interaction design in the context of HCI, and references to literature on auditory display in interaction tasks see [8]. Object manipulation considering haptics also has been explored in experimental investigations. For example there is work giving evidence for the importance of haptics in object identification tasks [7]. However, the amount of studies involving the combination of senses, in particular the integration of audio and haptics is rather limited. Of particular relevance is the work of Barrass and Adcock [1] who consider the use of haptic display device in combination with granular synthesis. Two of the present authors also discuss the joining of haptics and audio with respect to new interface designs [11]. Related

Proceedings of ENACTIVE05
2nd International Conference on Enactive Interfaces
Genoa, Italy, November 17th-18th, 2005

technology for contact interactions were also presented DiFilippo and Pai [4]. In all three cases user studies of these interactions are yet missing. This is not to say that the importance of merged senses in interactions is not well recognized and experimentally investigated (see Ernst and Bülthoff [5] for a review). Bresciani et al [2] show evidence for sensory integration between auditory and haptic modalities in tapping interactions. McGee's thesis [9] gives indications for the modifying function of different modalities to the perception of roughness in interactions.

3. Pilot test, collisions between multiple objects

We are interested in a setting where users manipulate multiple objects at once. The scale of the objects under consideration is between one 5th and one half of the width of an average adult hand. The objects are assumed to be non-brittle and to create audible impact sounds upon collision.

Situations of this type offer a somewhat complex result for all involved senses. The interactions are assumed to involve direct manipulation of the objects themselves or may possibly be mediated by a pen. In general the interactions seem to emphasize the tactile and auditory sensation in such interactions. The study of interactions involves multiple dynamic objects, focusing on just the auditory and haptic aspects of the interaction, is an area which has received little prior attention.

Primarily non-visual environments containing multiple dynamic objects, is little investigated. Thus this pilot was an opportunity to get a first insight into future research directions within this area.

Each participant conducted a 3-part test in varying order. The parts were: " Number of objects. The subject was blindfolded, and asked to assess the number of objects inside a box. " Preferences. The subjects were presented with different feedback sounds, and asked about their opinion about a particular sound's believability and likeability. " Sense of control. The subjects were exposed to auditory feedback, and asked to assess how much in control he or she felt.

The pilot tests were conducted on 7 volunteers (5 males, and 2 females) for the control and believability experiments and 6 volunteers (5 males, and 1 female) for the counting experiment. The age of the participants ranged from 25 to 65.

Figure 1: Interaction with the use of a PHANToM.

4. Experiment 1: Estimating Number of Objects

4.1 Motivation

In this experiment we were interested in getting a first rough sense of the impact of varying sensory contributions in a complex sensorimotor task. Hence a specific goal was given to the subjects with varied forms of sensory output being available.

4.2 Procedure

During this test, the subjects were wearing glasses with blinds, thereby removing the visual input. The subjects were asked to assess the number of objects (from 0 to 20) inside a box using a pen or pen-like haptic device as probe[1]. Each subject performed four different cases of interaction: 1. Virtual objects, feedback from audio only. 2. Virtual objects, feedback from haptics only. 3. Virtual objects, feedback from audio+haptics. 4. Natural objects (both haptics and audio).

Each of case 1 to 4 consisted of 8 trials. During a trial the subject had 15 seconds to freely stir around the objects in the box, after which the subject had to report a number on how many objects were thought to be inside the box and state, on a 3-point scale how confident they felt about their estimate.

To reduce the bias of learning effects, the subjects performed case 1-4 in different order.

[1] For experimental investigation of the the relation of object-mediated versus whole-hand interactions see for example [10].

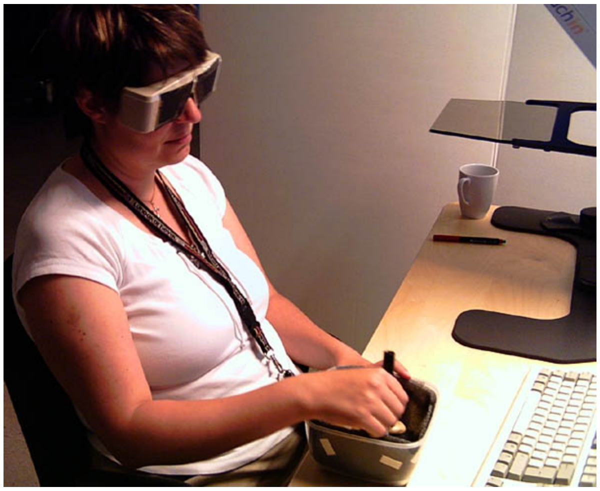


Figure 2: Interaction with natural objects.

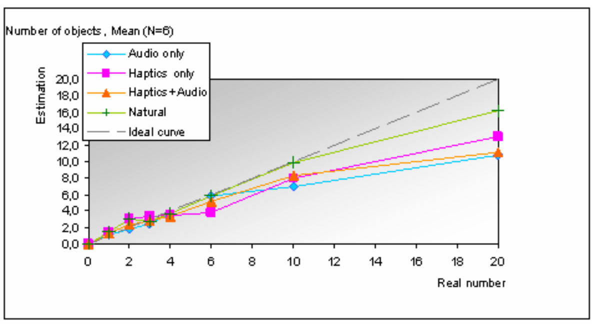


Figure 3: The averaged estimates for the six subjects, for the different setups (audio only, haptics only, audio + haptics and natural).

4.3 Equipment

For the virtual haptic interaction, a PHANToM, model Premium (Sensable Technologies Inc.), was used (see Figure 1). The software was developed around Sensable's OpenHaptics API, combined with dynamics simulation based on the open-source code Open Dynamics Engine (www.ode.org). For the sound feedback, Direct Sound (Microsoft corp.) was used. The natural interaction was carried out with a foam padded plastic box, a thick pen, and some round stones (see Figure 2).

4.4. Results, number of objects

The task of identifying the number of objects without visual feedback is hard. Still, as can be seen in Figure 3, the six subjects performing this task did fairly well on the average.

This is somewhat in contrast with the stated level of confidence (Figure 4), which on the average level shows that the subjects were better at estimating the number of

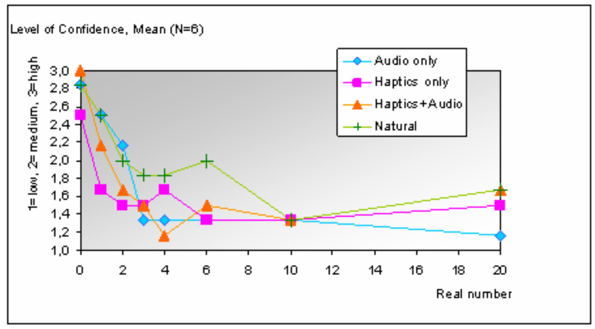


Figure 4: Level of confidence.

objects than they thought.

As expected low numbers are easier to estimate than high, but it is interesting that the curve falls off very rapidly. Already at 3-4 objects subjects on the average rate the level of confidence as below medium. The comparatively high confidence for N=20 is with great probability due to the fact that subjects were told the maximum possible number of objects. It should also be noted that for the 'natural' case some additional cues were present - the sound the box produced when placed on the table changed noticeably as the box was full of stones. Also, the stones did not slip away the way they did in the virtual environment. In the 'natural' setup the stones lay still, and some subjects actually attempted counting them with the pen.

If we look at the subjects' individual performances (Figure 5) the picture becomes less clear. We can still say that people can judge the qualitative number of objects reasonably well - i.e most curves have an upward slope in Figure 5, but it is seen that the accuracy in the estimates break down early - already at 1-3 objects we start to see errors. But the errors are not random, the users can discriminate between "many" and "few" (although they may be wrong in their estimate of the exact number). It is also interesting to note that it is not obvious that haptics + audio produces the best estimates. And there is an indication that different people rely very differently on audio vs. haptics.

4.5 Discussion, number of objects

We see that people on the average are fairly good also at the very difficult task of estimating the number of objects when the objects are invisible and moving. There was no clear pattern as to which type of feedback was most efficient - instead this varied from subject to subject. During the experiment we noted a few problems. One was that there was no audio feedback from the collisions with the walls. The environment had been de-

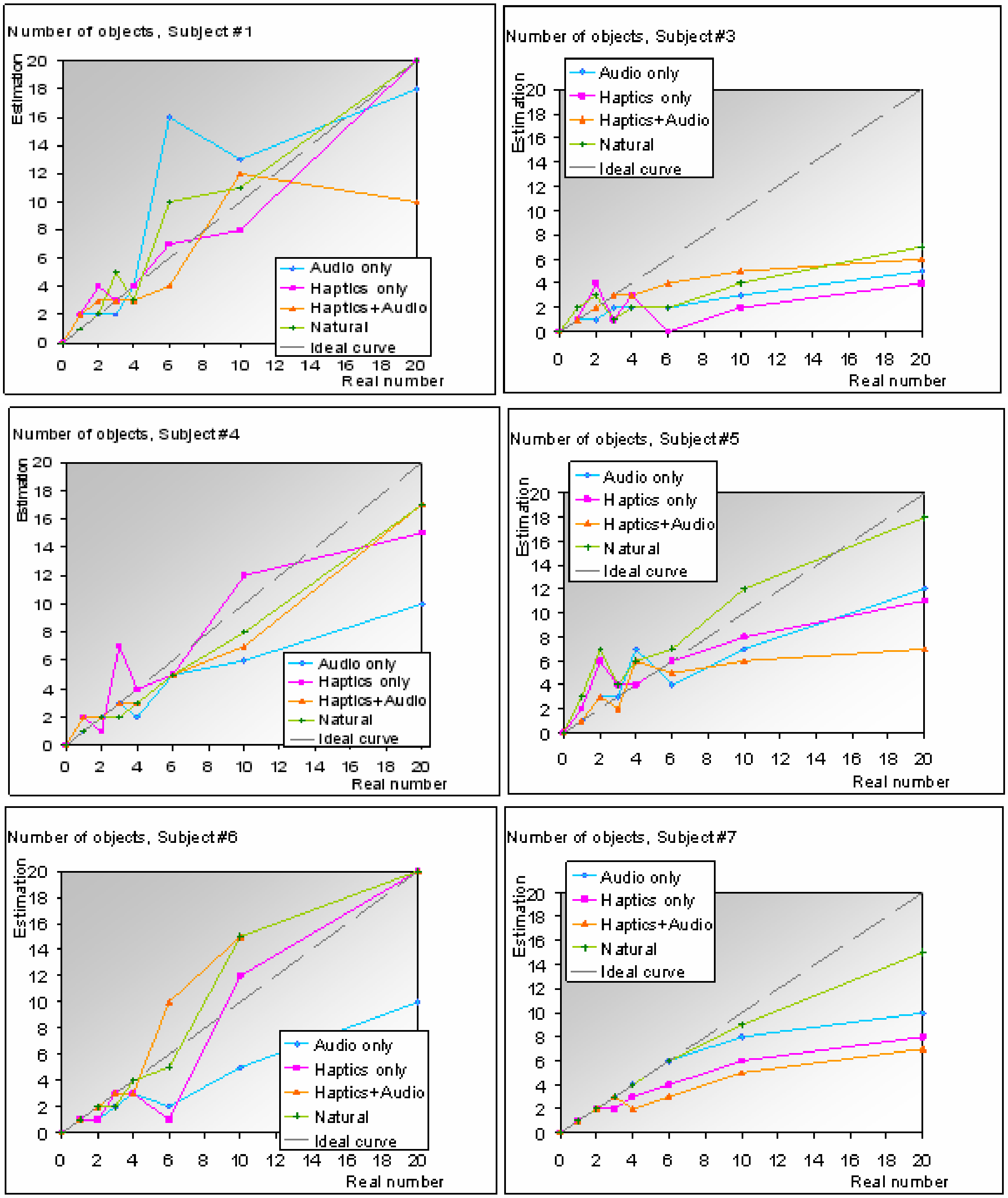


Figure 5: Individual results of all the six subjects.

signed this way for better comparison with the PebbleBox, but several subjects complained about the loss of wall collision sounds. The other was the design of the ”natural” setup. The objects in this setup did not behave very similarly to the real objects, and it is clear that this setup needs to be redesigned to reduce the potential impact of this dissimilarity.

This pilot test still shows that it would be interesting to investigate this type of environment further. The type of environment is complex and dynamic, and points of interest in future experiments are

- which are the key features of the feedback (collision sounds, spatial location of sounds, possible echos, friction, forces at impact etc), and how do they affect given number estimates?
- how should tools that help the user understand this type of environment be designed?
- how does the virtual environment compare to a corresponding real environment and what are the minimum requirements for believability?

5. Experiment 2: Preferences

5.1 Motivation

This experiment was designed to find evidence whether or not subjects prefer real or virtual sensorimotor settings on the one hand and to find evidence when sensorimotor integration works favorably in an emotive sense. In particular, both interfaces, the PHANToM and PebbleBox allow for the sound connected with collisions to be varied arbitrarily. But what sounds are best suited for this situation? PebbleBox was designed for a broad class of sounds that relate to hand gestures and hand gesture related impact sounds. Is this a sensible design goal of the interface?

5.2 Procedure

In this part, the subjects were to interact with two different kinds of devices: the PebbleBox, see Equipment below, and the PHANToM, see previous section. For each device, the subject carried out 4 rounds, each with a different sound for the feedback. The sound-sources used were: (1) hands in water, (2) chewing apple, (3) birds' song, and (4) fencing, i.e. metal-to-metal contact sounds. The first and last were chosen as sounds related to typical hand-action mediated sounds. The second and third were chosen because of their lack of or very remote relation to hand-action mediated sounds.

Figure 6: The PebbleBox.

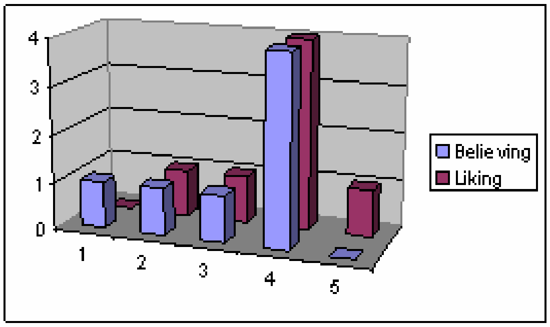
Each round lasted 15 seconds, where the subject could freely play with the objects in order to produce collisions (which, in turn, generated the sound-feedback). Afterwards, the subject was asked how they would rate the experienced interaction:

- Did you find the interaction experience *believable*?
- Did you find the interaction experience *likeable*?

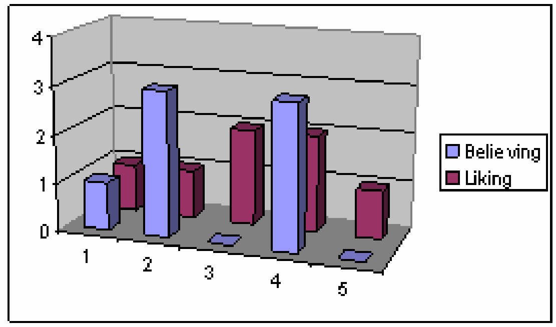
5.3 Equipment

The PebbleBox (see Figure 6) is physically a wooden-box with foam padded interior. Embedded in the foam is a small microphone connected to a PC. When playing with glass-beads or small stones, the small collisions are picked up by the microphone. Running on the PC, a program detects collisions in the sound signal (see [11]) returns onset time, amplitude and spectral information then uses these parameters to synthesize an arbitrary sound based on multiple samples. The software is randomly selecting and playing a sound-sample from a set of up to 10 variants of a particular sound (e.g. hands in water). The sound-output is then presented to the user through headphones in order to shut-out direct sounds from the glass-beads or stones.

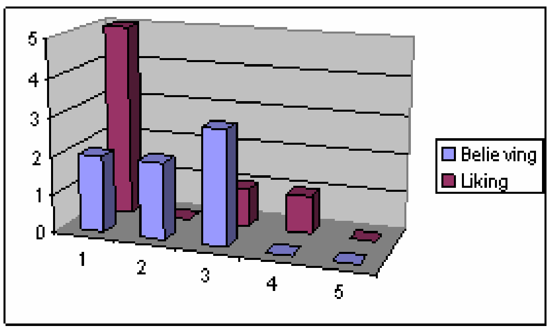
For the corresponding test-rounds with the PHANToM, the same hardware and software was used as in the ”Number of objects”-experiment, although the code was modified to utilise the same randomly chosen, multiple sound-samples as the PebbleBox, although with spatially distributed (3D) sounds.



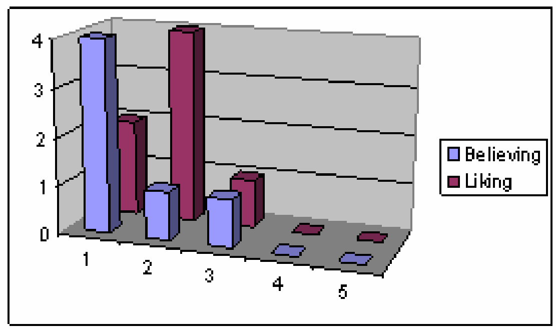
PebbleBox sound, water



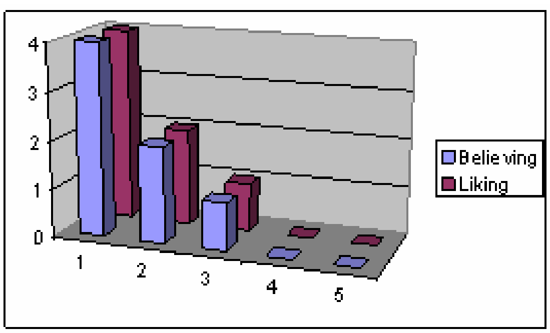
PHANToM sound, water



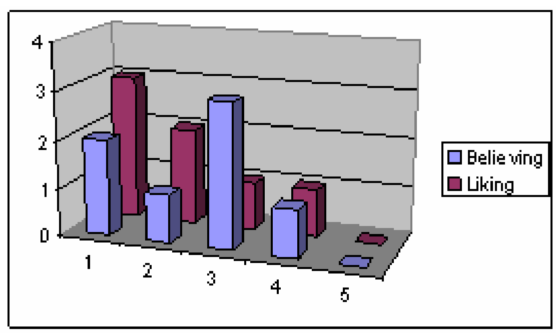
PebbleBox sound, apple



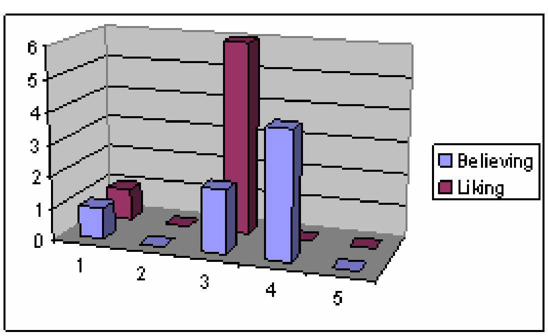
PHANToM sound, apple



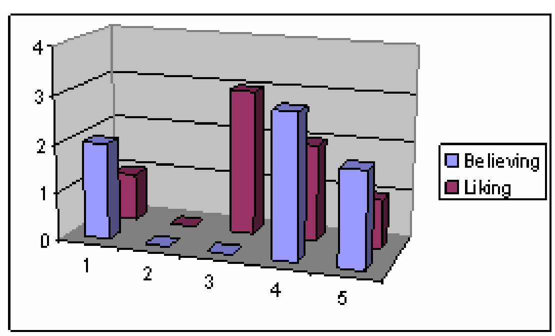
PebbleBox sound, bird



PHANToM sound, bird



PebbleBox sound, metal



PHANToM sound, metal

Figure 7: The number of persons (vertical) assigning different believability and likeability ratings (horizontal) for different sounds and setups. The scale for the ratings was one to five, where five implies ”very much”.

5.4 Results, preferences

In the preferences experiment people were asked to rate believability and likeability on a subjective scale. The outcome for the different types of sounds can be seen in Figure 7. Although there is a difference between the PebbleBox and the virtual PHANToM environment, on the whole the two setups seem to produce somewhat similar ratings. The water was to some extent believed, as was the metal (sound from fencing). The apple eating and the bird sounds were less believed, although the bird seemed to be easier to believe in the virtual setting.

This was a fairly expected result - one may produce splashing sounds by playing with stones in water, and clanking metal sounds may be produced by metal balls hitting each other. Bird song or apple eating sounds are usually not produced by this type of interaction.

On the whole the ratings for believability and likeability follow each other, although several sounds appeared a little better liked in the virtual setting.

5.5 Discussion, preferences

It was interesting to note some differences in the ratings between the PebbleBox and the virtual setup. Particularly the more "unnatural" sounds (i.e. unnatural for this type of interaction) seemed a bit easier to accept in the virtual setting.

The ratings were done totally subjectively, and with each person's own notion of believability. It is an interesting task to explore more in the future how the believability notion for this type of environments could be developed.

And, of course, which classes of sounds that are acceptable/believable for which classes of interaction techniques. As well as how the reality/virtuality of the interaction affects the results.

Overall, due to the highly subjective nature of the experiments, a much larger subject pool would be required.

6 Experiment 3: Sense of control

6.1 Motivation

How in control do subjects feel when in a complex interaction? This question is at the core of the last experiment.

6.2 Procedure

Also in this part, the subjects were to interact with two different kinds of devices: the PebbleBox, and the PHANToM. The subject was instructed to assess the

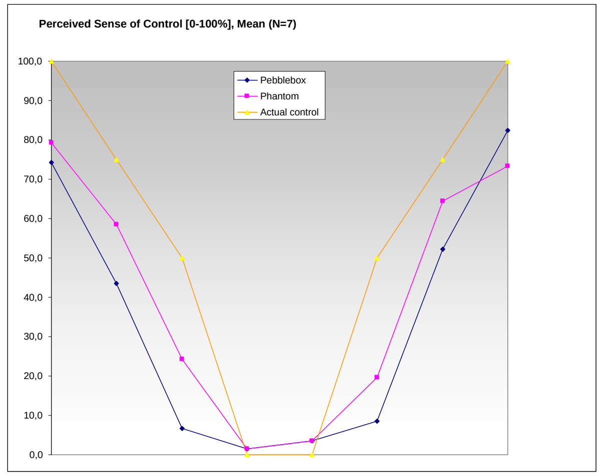


Figure 8: Sense of control for the PebbleBox and the PHANToM virtual setup. The sound used as feedback was the water sound. The numbers given for the 'actual control' is the probability that collision sounds actually generate sound feedback.

perceived sense of control he/she had on the sound-feedback (the water sound was used for this test). The rating was in the form of percentage (0-100%).

For each device, the subject worked with 8 rounds of 15 seconds. These rounds had different levels of added false random events, as well as true events randomly suppressed. The probability that a user initiated event should generate sound feedback in the 8-round sequence the subjects were working with, was as follows:

100% 75% 50% 0% 0% 50% 75% 100%

6.3 Equipment

The same hardware and software was used as above, but with modification in the code that added a certain level of random false events, as well as suppressing a certain amount of true events.

6.4 Results, sense of control

The results for the sense of control experiment is summarized in Figure 8 below.

The numbers given for actual control is not quite as straightforward as it appears in the Figure 8. The amount of control depends on two factors:

1. how many of the events the user generates that produces the intended feedback

2. how often feedback is produced even though the user does nothing

Probability for a user initiated event to generate sound feedback	Average nr of user independent events per second
100%	0
75%	2
50%	11
0%	22

Table 1: Parameters for the generation of sound events.

In the Figure 8 the numbers are actually only the number for the 1) type of control. The increasing noise level will of course also influence the perceived level of control. The figures for the 2) type of lack of control is shown in table 1.

To give an indication of how large fraction of the total number of sound events that will be false events, we let the program calculate the ratio for a trial run containing three samples. The fraction of false sound events (noise) was on the average 4% for the 75% level and 28% at the 50% level.

On the whole though, the shape of the curves in Figure 8 show that the subjects are fairly good at estimating the amount of control they have - they sense when they have a lot of control, and they know when they have less. Looking at the data in more detail, it is seen that subjects sometimes state that they have no control even though their input actually does change the output. And the opposite also occurs - some users believe they are a little in control although the actual level of control is zero. It should also be noted that the users on the average do not feel that they are 100% in control even when they are.

6.5 Discussion, sense of control

The outcome of this test very likely depends a lot on both sound design and the choices for the sound generation parameters as described above. The water sound is not so distinct, which may explain why the perceived amount of control was less than 100% at the maximum level.

The sense of control is also tied to the concept of believability. It is interesting to investigate in more detail how different sounds and different parameter settings influence the way users think they are able to control an environment.

7 Conclusion

These pilot experiments have shown that this type of dynamic environments with colliding multiple objects is an interesting subject of study. First indications of comparative measures in terms of performance tasks, believability and control with respect to the absence, presence and fusion of senses can be seen. The data indicates that users in general have a rather good sense of the amount of control they have in an interaction and will trace increases and decreases in control rather reliably. This is independent of the presentation of tactile sensations and specifics of the motor action. A first indication of categories of believability may be seen, though further studies are necessary to arrive at reliable evidence. Participants performed rather well in counting tasks even though their self-evaluation in terms of confidence does not reflect their performance for high number of objects. When it comes to the importance of haptic versus auditory qualities of the output, it can be seen that individual users rely on these quite differently. Some appear to do better with only haptic, some with only audio and some with haptic and audio. This indicates the importance to allow for different interaction styles also for the haptic audio combination.

Given these results there are however a number of questions that remain to be explored:

- which are the key features of the feedback (collision sounds, spatial location of sounds, possible echoes, friction, forces at impact etc), and how do they affect given number estimates?
- how should tools that help the user understand this type of environment be designed?
- how does the virtual environment compare to a corresponding real environment and what are the minimum requirements for believability?
- how should believability be measured for this type of environments?
- which classes of sounds that are acceptable/believable for which classes of interaction techniques?
- how does the reality/virtuality of the interaction affects the results?
- how do different sounds and different parameter settings influence the way users think they are able to control an environment?

Acknowledgements

We are grateful for helpful discussion with Nicolas Castagne. The reported work was carried out with the financial assistance of the European Commission which co-funded the NoE ENACTIVE - “ENACTIVE Interfaces”, FP6-2002-IST-1. The authors are grateful to the EU for the support given in carrying out these activities.

References

[1] Barrass S and Adcock M (2002) Interactive Granular Synthesis of Haptic Contact Sounds, AES 22nd International Conference on Virtual, Synthetic and Entertainment Audio, Helsinki University of Technology, Espoo, Finland, 15th-17th June 2002.

[2] Bresciani, J.P., M.O. Ernst, K. Drewing, G. Bouyer, V. Maury and A. Kheddar (2005): Feeling what you hear: auditory signals can modulate tactile taps perception. Experimental Brain Research 162, 172-180.

[3] Calvert, G. (2004). The Handbook of Multisensory Processes. Cambridge, MA, MIT Press.

[4] DiFilippo, D. K. & Pai, D. (2000). Contact Interaction with Integrated Audio and Haptics. In Proceedings of the International Conference on Auditory Display (ICAD), 2000.

[5] Ernst, M.O. and H.H. Bülthoff (2004): Merging the Senses into a Robust Percept. Trends in Cognitive Sciences 8(4), 162-169.

[6] Jansson, G. (1993). Perception of the amount of fluid in a vessel shaken by hand. In S.S. Valenti & J.B. Pittinger (Eds.), Studies in perception and action II. Posters presented at the VIIth International Conference on Event Perception and Action. August 8-13, 1993. University of British Columbia, Vancouver, BC Canada (pp. 263-267) Hillsdale, NJ: Erlbaum.

[7] Jansson, G. & Monaci, L. (2005). Improving haptic displays: Providing differentiated information at the contact areas is more important than increasing the number of areas. Poster at Worldhaptics05, Pisa, March 18-20 2005.Available in Proceedings on CD-ROM.

[8] Marentakis, G. and Brewster, S.A. (2005) A Comparison of Feedback Cues for Enhancing Pointing Efficiency in Interaction with Spatial Audio Displays. To appear in Proceedings of MobileHCI 2005 (Saltzburg, Austria). ACM Press.

[9] McGee, M.R. (2002). Investigating a Multimodal Solution for Improving Force Feedback Generated Textures. Dissertation, Department of Computing Science, University of Glasgow, September 2002.

[10] O'Modhrain, S. (1999), Restricted Access: Exploritory Procedures and Object Properties, Proceedings of the ASME International Mechanical Engineering Congress: Dynamic Systems and Control Division, Vol. 2 (Haptic Interfaces for Virtual Environments and Teleoperator Systems) DSC-Vol. 61, 1999.

[11] O'Modhrain, S. and Essl, G., (2004), PebbleBox and CrumbleBag: Tactile Interfaces for Granular Synthesis, Proceedings of the International Conference for New Interfaces for Musical Expression (NIME), 2004.