

Tactual Articulatory Feedback and Gestural Input

Bert Bongers

Human-Computer Interaction, Multimedia & Culture
Department of Computer Science
Vrije Universiteit
De Boelelaan 1081a, Amsterdam, The Netherlands
e-mail: bertbon@xs4all.nl

Gerrit van der Veer

Human-Computer Interaction
School of Computer Science
Open University
Valkenburgerweg 177, Heerlen, the Netherlands
e-mail: gerrit@acm.org

Abstract The role of the sense of touch in Human-Computer Interaction as a channel for feedback in manipulative processes is investigated. The paper discusses information as generated by the computer as presentation and feedback, and focuses on the feedback that supports the articulation of human gesture. Several experiments are described that investigate the use of (redundant) tactual articulatory feedback. The results presented show that a significant improvement of effectiveness only occurs when the task is sufficiently difficult, while in some other cases the added feedback can actually lower the effectiveness. The qualitative data show a positive effect for added tactual feedback. The discussion includes suggestions for further research, particularly investigating the effect in free moving gestures.

1 Background

Many everyday interactions with physical objects that we manipulate are guided by a variety of (redundant) feedback modalities, such as touch, sound, smell or vision. This is true not only in the real world, it can also be expected to be true in the 'virtual world' as generated by the computer. However in the current computer interface paradigm based on windows, icons, menus and pointer (WIMP), the human sense of touch is not actively addressed by the system. All one can feel are things like the texture of the mouse pad or the resistance of keys, which is only limited as a source of information.



Fig. 1. Mouse with the vibrotactile element.

The computer has a wide variety of information to present, such as feedback about processes, information about objects, representation of higher level meaning, indications of relevant aspects of the system state. Currently, the desktop computer is mainly presenting this information visually, and in some cases with sounds. This paper describes a research project investigating if there is an advantage in using tactual feedback as secondary notation or added feedback when manipulating the virtual objects in the computer

(such as widgets, icons and menus). We are looking particularly at feedback that supports the *articulation* of a gesture. A standard mouse input device is used with an added vibrotactile feedback element based on a small loudspeaker placed under the tip of index finger of the preferred hand.

Section 1.1 and 1.2 of this paper present an information based approach partially informed knowledge from the field of communication studies. Levels of interaction are discussed in Section 1.3 and 1.4, placing the level of physical interaction and articulation in a context. A short summary of the various sub-modalities of tactual perception is given in Section 1.5. Section 1.6 gives an overview of haptic interaction devices and related research in this field, and section 1.7 describes previous work on vibrotactile feedback.

Section 2 describes the experimental setup, method and results. In section 3 the results are discussed and there are concluding remarks and future research directions presented in section 4.

1.1 An Information Based Approach

In the context of interactive systems and communication, it is useful to view the world around us (including all the objects and processes) as information. Objects and processes can be physical or mental, and they can be real or virtual - that is, computer generated, synthetic.

Objects inform the attentive perceiver of all sorts of things, from just being there and not to be bumped into, to revealing things about themselves - what they are made of, what one potentially can do with them etc. The latter are called *affordances*, informing the perceiver about what can be done with the object and how it can be used (Gibson 1979, Ch. 8).

Several levels or classes of *signs*, through which information is communicated, can be described (Rosengren, 2000; Littlejohn, 2002). There is information from objects themselves, and what is called their indices or *signals*, information that they provide. Processes are often revealed through their *symptoms*. When manipulating things in the world around us, these kinds of information or signs give us feedback on our actions. Living things add further signs to our world and so do human artefacts, particularly those that are made to convey information. Signs can be made that mimic the object that is signified, for instance the pictograms found in train stations, or a gesture that imitates a certain action. These kinds of signs are called *icons*, which in the case of the computer interface can have their own behaviour or processes. Abstract signs that have no resemblance to the object or concept they signify are called *symbols*, such as in written or spoken language or a musical score. Symbols can be organised in language, such as our speech, and then written down. Additionally, signals and symbols can have emotional value, supporting or contradicting the information and meaning.

Summary of information signs

- all the things in the world, their affordances, signals and symptoms
- icons that mimic and resemble the thing that is signified
- symbols — organised in languages (or codes) — and written down

A language can be verbal as well as non-verbal, such as gestures, often accompanying each other. Non-verbal signs can be uttered unconsciously. The information can be conveyed through a variety of modalities, such as speech, tactile symbols such as Braille, visual modalities such as images, written language, musical scores, and olfactory signs (Eco, 1976).

1.2 Presentation and Feedback

All these signs (of objects and processes) and symbolic languages are applied in the most complex artefact developed by humans so far, the computer. Computer generated phenomena are called virtual as opposed to real - that is, these synthetic phenomena are known not to be real. Often virtual phenomena can be discriminated from real phenomena because they are lacking tangible properties, but this is rapidly changing. There is a philosophical side to this debate, from Plato's cave to Daniel Dennett's 'Brain in the Vat' (Dennett, 1991), which is beyond the scope and relevance of this article.

There are two main forms of information generated by computers: *presentation* and *feedback*. Through its displays (visual, auditory, haptic) it presents what objects and processes are offered and what affordances they have. When being used, being manipulated, being communicated with by the user, the system conveys

information about the process(es) at hand, by feeding back information to the user, in different temporal scales. This can have the form of messages (verbal) or signals. If the feedback is presented in order to guide the users actions, to support them in articulating their intentions, it can be called *articulatory feedback* (Hix & Hartson, 1993, p 40) or synchronous feedback (Erickson, 1995). Articulatory feedback is particularly relevant in gestural interfaces, often using tactual modalities as shown in musical instruments (Chafe, 1993; Chafe & O'Modrain, 1996, Bongers, 1998).

When the system is generating information that actively draws the user to something, it is generally referred to as *feed-forward*. The strongest examples of this can be found in haptic systems, that actively 'pull' the user towards some location.

Feedback can come from elements specially designed for that purpose, such as message boxes and widgets that allow manipulation, or from the content itself (this is sometimes called functional feedback). Information that is presentation in one moment, can be feedback in another situation or even at the same time.

Most feedback is actively generated by the system, but some feedback can come from passive elements of the system. An example of such passive feedback, sometimes also called inherent feedback (Schomaker et al, 1995) (Wensveen et al, 2004), is the mouse click felt when the button is pressed - it is useful information but not generated by the system. The system therefore cannot be held responsible for misinterpretations - the click will still be felt even if there is nothing to click on the screen, the machine has crashed, or is not even on. In fact this is a bit of real world information, blending in an often useful way with the virtual world information. Mixing real and virtual phenomena is a good thing as long as one is aware of it, and it is designed as a whole. In the example of the mouse click it can be said that it is estimated information - usually true, but not always.

Presentation and feedback information can be displayed in various ways, addressing various human senses. In each case, particularly as part of parallel and multitasking situations, the most appropriate modality has to be chosen to convey the information.

1.3 Levels of Interaction

In order to understand the interaction between human and technology, it is useful to discern various *levels of interaction*. An action is usually initiated in order to achieve some higher order goal or intention, which has to be prepared and verbalised, and finally presented and articulated through physical actions and utterances. The presentation and feedback by the computer passes through several stages as well before it can be displayed, possibly in various modalities including the haptic, in order to be perceived by the user. The actual interaction takes place at the physical level. A dialogue is often described in three levels: semantic, syntactic, and lexical (Dix et al, 1993, Chapter 8), but for more specific cases more levels can be described. Nielsen's virtual protocol model (1986) is example of this, specifying a task and a goal level above the semantic level, and an alphabetical and physical level below the lexical level. This can also be applied on direct manipulation interface paradigms. The Layered Protocol is an example of a more complex model (Taylor, 1988), particularly to describe the dialogue using the speech modality, but also applied on general user interface issues (Eggen, Haakma and Westerink, 1996). When more sensory modalities are included in the interaction, models often have to be refined. Applying the Layered Protocol on the interaction which includes active haptic feedback, introduces the idea of (higher level) E-Feedback which has to do with expectations the user has of the system, and the I-Feedback which communicates the lower level interpretations of the user's actions by the system (Engel, Goossens & Haakma, 1994).

A framework for analysing and categorising interaction by levels of interaction, modalities and modes has been developed and presented elsewhere (Bongers and van der Veer, 2006).

The articulatory feedback (or interpretation feedback) which is studied in the research described in this paper, takes place at the physical level and extends into the semantic levels.

1.4 Articulation

In the final phase of the process of making an utterance or a gesture we rely on feedback in order to shape our actions. Any action is a continuous process of acting, processing feedback, and adjusting. When speaking, we rely on the auditory feedback of our utterances and continuously adjust our articulation. When

manipulating objects or tools, we rely on the information conveyed to our senses of vision, touch, sound, smells, including our self-perception, in order to articulate. The mainstream computer interface paradigm relies almost entirely on the visio-motor loop, sometimes with an added layer of auditory feedback such as sound-sets built into the operating system. However, in a real world 'direct manipulation' action the closest sense involved in that act is usually the sense of touch. Musicians know how important the touch feedback on articulation is, and so does the craftsman. Computers seem to have been conceived as a tool for intellectual processes such as mathematics rather than as a 'tool' for manual workers, and have inherited a strong tendency of anti-crafts, and therefore anti-touch (McCullough, 1996).

Anticipating a further development and emphasis on gestural control of computers, including the fine movements and issues of dexterity, there will be a need for the feedback to be generated properly in order to be adequate. That is, we need feedback that supports the articulation of the gesture, enabling the user to refine the action on the fly.

1.5 Tactual Perception

The human sense of touch gathers its information through various interrelated channels, together called *tactual perception* (Loomis & Leederman, 1986). These channels and their sub-channels can be functionally distinguished, in practice however they often interrelate.

The tactile perception receives its information through the *cutaneous* sense, from the several different mechanoreceptors in the skin (Sherrick and Cholewiak, 1986). *Proprioceptors* (mechanoreceptors in the muscles, tendons and joints) are the main input for our kinaesthetic sense, which is the awareness of movement, position and orientation of limbs and parts of the human body.

A third source of information is 'efference copy', reflecting the state of the nerves that control the human motor system. In other words, we know that we are moving and this is a source of input as well. This is called active touch (Gibson, 1962) as opposed to passive touch.

Haptic perception uses information from both the tactile and kinaesthetic senses when actively gathering information about objects outside of the body. Most of the interactions in the real world involve haptic perception, and it is the main modality for applying in HCI.

The feedback discussed in this paper mainly involves the tactile sense. There are four sense systems involved relating to four types of mechanoreceptors in the skin, making up all four combinations of the parameters adaptivity (slow and fast, having to do with frequency) and spatial sensitivity (diffuse and punctuate) (Bolanowski et al, 1988). The experiments described in this paper are particularly addressing the fast adapting and diffuse system. This is often called the Pacinian system (named after the Pacinian corpuscles that are the relevant mechanoreceptors), and is important for perceiving textures but also vibrations – its sensitivity overlaps with the audible range.

1.6 Tactual Feedback Devices

Several devices have been developed to actively address the human sense of touch, and many studies have shown the improvement in speed or accuracy in the interaction.

Existing devices like the mouse have been retrofitted with solenoids for tactile feedback on the index finger and electromagnets acting as brakes on an iron mousepad for force-feedback (Akamatsu and Sato, 1992, 1994) (Münch and Dillmann, 1997), or solenoids on the side of the mouse for vibrotactile feedback (Göbel et al, 1995). The principle of electromagnetic braking mouse was also applied in an art interface, an interesting application where the emphasis was not on improving efficiency (Kruglanski, 2000). The now discontinued Logitech force-feedback mouse, and its predecessor the Immersion FEELit mouse, have two motors influencing the mouse movement through a pantograph mechanism. These have been used in several studies (Dennerlein et al, 2000), including studies which are a continuation of earlier work with the vibrotactile device as described in this paper but then using force-feedback (Hwang et al, 2003). All show performance improvements by applying various forms of feedback as described above.

Several motor-based force-feedback joysticks have been used for generating virtual textures (Hardwick et al, 1996, 1998)), and other experiences (Minsky et al, 1990). These joysticks have become cheaply available for gaming applications, such as the devices from Microsoft and Logitech using the Immersion TouchSense protocol.

To investigate the notion of E- and I-Feedback as described above, a force-feedback trackball was developed using computer controlled motors to influence the movement of the ball, enabling the system to generate feedback and feed-forward (Engel, Goossens & Haakma, 1994). This device was further used for studies in tactual perception (Keyson & Houtsma, 1995; Keyson, 1996), and the development of multimodal interfaces (Keyson & van Stuivenberg, 1997; Bongers et al, 1998).

The Phantom is a well known device, it is a multiple degree-of-freedom mechanical linkage that uses motors to generate force-feedback and feed-forward on the movements. It has been used in many studies that investigate the advantages of tactual feedback (Oakley et al, 2000, 2002).

1.7 Vibrotactile feedback

Somewhat similar to the solenoid is the (miniature) relay coil that can generate tactile cues. Such a relay coil has been used in the Tactile Ring developed for gestural articulation (Bongers, 1998). The advantage of using a miniature loudspeaker however is that many more pressure levels and frequencies can be generated. This flexibility and potential for rich display of vibrotactile information is the rationale behind the developments described in this paper and earlier research (Bongers, 2002b). This led to the formulation of our research question of whether vibrotactile feedback could be used to improve articulation.

2 Investigating Tactual Articulatory Feedback

As stated before, in the current computer interaction paradigm articulatory feedback is given visually, while in the real world many movements and manipulative actions are guided by touch, vision and audition. The rationale behind the work as described in this paper is that it is expected that adding auditory and vibrotactile feedback to the visual articulatory feedback improves the articulation, either in speed or accuracy, and that of these two the vibrotactile will give the greatest benefit as it is the most natural form of feedback in this case.

The work reported here focuses on the physical level of the layered interaction as described before, where the articulation of the gesture takes place. It is at this stage limited to a one degree-of-freedom movement, of a mouse on a flat surface. An essential part of our approach is to leave some freedom for the participants in our experiments to explore.

The research method is described in section 2.1. Section 2.2 describes the experiment, which is set up in five phases. Phase 1 (2.2.1) is an exploratory phase, where the participants' task is to examine an area on the screen with varying levels of vibrotactile feedback, as if there were invisible palpable pixels. This phase is structured to gain information about the participants' sensitivity to the tactual cues, which is then used in the following phases of the experiment. In the second phase (2.2.2) a menu consisting of a number of items has been enhanced with tactual feedback, and performance of the participants is measured in various conditions. In the third and fourth phase (2.2.3 and 2.2.4) an on-screen widget, a horizontal slider, has to be manipulated. Feedback conditions are varied in a controlled way, applying different modalities, and performance is measured. The last phase (section 2.2.5) of the experiment consisted of a questionnaire. Section 2.3 gives the results of the measurements and statistical analyses.

2.1 Method

In the experiments the participants are given simple tasks, and by measuring the response times and error rates in a controlled situation differences can be detected. It must be noted that, in the current desktop metaphor paradigm, a translation takes place between hand movement with the mouse (input) and the visual feedback of the pointer (output) on the screen. This translation is convincing due to the way our perceptual systems work. The tactual feedback in the experiments is presented on the hand, where the input takes place, rather than there were the visual feedback is presented. To display the tactual feedback a custom built vibrotactile element is used. In some of the experiments auditory articulatory feedback was generated as well, firstly because this often happens in the real world, and secondly to investigate if sound can substitute for tactual feedback. Mobile phones often have such an audible key click.

A gesture can be defined as a multiple degree-of-freedom meaningful movement. In order to investigate the effect of the feedback a restricted gesture was chosen. In its simplest form, the gesture has one degree-of-freedom and a certain development over time. As mentioned earlier, there are in fact several forms of tactual feedback that can occur through the tactual modes as described above, discriminating between cutaneous and proprioceptive, active and passive. When interacting with a computer, cues can be generated by the system (active feedback) while other cues can be the result of real world elements (passive feedback). When moving the mouse with the vibrotactile element, all tactual modes are addressed, actively or passively; it is a haptic experience drawing information from the *cutaneous*, *proprioceptive*, and *effeference copy* (because the subject moves actively), while our system only actively addresses the cutaneous sense. All other feedback is passive, ie. not generated by the system, but it can play some role as articulatory information.

It was a conscious decision not to incorporate additional feedback on the reaching of the goal. This kind of 'confirmation feedback' has often been researched in haptic feedback systems, and has proven to be an effective demonstration of positive effects as discussed in section 1.6. However, we are interested in situations where the system does not know where the user is aiming for, and are primarily interested in the feedback that supports the articulation of the gesture - not the goal to be reached, as it is not known. For the same reason we chose not to guide (pull) the users towards the goal, which could be beneficiary although following Fitts' Law what happens in such cases is that the target size gets actually increased (Akamatsu & MacKenzie, 1996).

2.2 Experimental Set-Up

The graphical, object based programming language Max/MSP (2005) was used. Max/MSP is a program originally developed for musical purposes and has therefore a suitable real time precision (the internal scheduler is set to 1 ms.) and many built-in features for the generation of sound, images, video clips, and interface widgets. It has been used and proved useful and valid as a tool in several psychometric and HCI related experiments (Vertegaal, 1998, p. 41; Bongers 2002a). Experiments can be set up and modified very quickly.

The software for this experiment run on an Apple PowerBook G3 with OS 8.6, used in dual-screen mode; the experimenter uses the internal screen for monitoring the experiment and the subject uses an external screen. The subject carries out the tasks using a standard Logitech mouse, with the vibrotactile element positioned under the index finger of the subject where the feedback would be expected. For the right handed users this was on the left mouse button, for the left handed users the element could be attached to the right mouse button. The vibrotactile element is a small loudspeaker (\varnothing 20 mm, 0.1W, 8Ω) partially covered and protected by a nylon ring, as shown in Fig. 2. The ring avoids the user to press on the loudspeaker cone which would make the stimulus disappear. The vibrotactile element is covered by the user's finger so that frequencies in the auditory domain are further dampened. The signals sent to the vibrotactile element are low frequency sounds, making use of the sound generating capabilities of Max/MSP.



Fig. 2. The vibrotactile element in position

Generally the tactual stimuli, addressing the Pacinian system which is sensitive between 40 and 1000 Hz, are chosen with a low frequency to avoid interference with the audio range. The wavelengths of the (triangular) tactual pulses were 40 and 12 milliseconds resulting in a base frequency of 25 and 83.3 Hz respectively. The repeat frequency of the pulses would be depending on the speed of the movement made with the mouse, and in practice is limited by the tracking speed. To investigate whether the tactual display really would be inaudible, measurements were performed with a Brüel & Kjaer 2203 Sound Level Meter. The base frequency of the tactual display was set to 83.3 Hz and the amplitude levels at 100 % and at 60% as used in the experiment. At a distance of 50 cm and an angle of 45 degrees, the level was 36 dB(A) (decibel adjusted for the human ear) when the device was not covered, and 35 dB(A) when the device was covered by the human finger as in the case of the experiments. At an amplitude of 60% these values were 35 dB(A) and 33 dB(A) respectively. So in the most realistic situation, ie. not the highest amplitude and covered by the finger, the level is 33 dB(A), which is close to the general background noise level in the test environment (30 dB(A)) and furthermore covered by the sound of the mouse moving on the table top surface (38 dB(A)).

Auditory feedback can be presented through a small loudspeaker next to the subject's screen.

In total 35 subjects participated in the trials and pilot studies, mainly first and second year students of Computer Science. They carried out a combination of experiments, as described below, using their preferred hand to move the mouse and explore or carry out tasks, and were given a form with open questions for qualitative feedback. In all phases of the experiment the subjects could work in their own pace, self-timing when to click the "next" button to generate the next cue.

Not all subjects participated in all experiments. Four out of 35 were pilot studies, the 31 main ones all participated in the most important phases of the experiment (2, 3 and 5) and 11 of them participated in the 4th phase which was an extension.

Special attention was given to the user's posture and movements, as the experiment involved lots of repetitive movements which is a recipe for RSI while the overall goal of the research is to come up with paradigms and interaction styles that are more varied precisely to avoid such complaints. At the beginning of each session therefore participants were instructed and advised on strategies to avoid such problems, and monitored throughout the experiment.

2.2.1 Phase 1: Threshold Levels

The first phase of the experiment was designed to investigate the threshold levels of tactile sensitivity of the participants, in relationship to the vibrotactile element under the circumstances of the test set-up. The program generated in random order a virtual texture, varying the parameters base frequency of the stimulus (25 or 83.3 Hertz), amplitude (30%, 60% and 100%) and spatial distribution (1, 2 or 4 pixels between stimuli) resulting in 18 different combinations. A case was included with an amplitude of 0, eg. no stimulus, for control purposes. In total therefore there were 24 combinations. The software was developed by the first author at Cambridge University in 2000 and is described in more detail elsewhere (Bongers, 2002b).

The participant had to actively explore an area on the screen, a white square of 400 x 400 pixels, and report if anything could be felt or not. This could be done by selecting the appropriate button on the screen, upon which the next texture was set. Their responses, together with a code corresponding with the specific combination of values of the parameters of the stimuli, were logged by the programme into a file for later analysis. In this phase of the experiment 26 subjects participated.

This phase also helped making the subjects more familiar with the use of tactual feedback, in case they weren't.

2.2.2 Phase 2: Menu Selection

In this part of the experiment, participants had to select an item from a list of numbers (a familiar pop-up menu), matching the random number displayed next to it. There were two conditions, one visual only and

the other supported by tactual feedback where every transition between menu items generated a tangible 'bump' in addition to the normal visual feedback. The 'bump' was a pulse of 12 ms. (83Hz, triangular waveform), generated at each menu item.

The menu contained 20 items, a list of numbers from 100 to 119 as shown in Fig. 3. All 20 values (the cues) would occur twice, from a randomly generated table in a fixed order so that the data across conditions could be easily compared - all distances travelled with the mouse were the same in each condition.

Response times and error rates were logged into a file for further analysis. An error would occur if the participant selects the wrong number, the experiment carries on. In total 30 subjects participated in this phase of the experiment, balanced in order (15 -15) for both conditions.

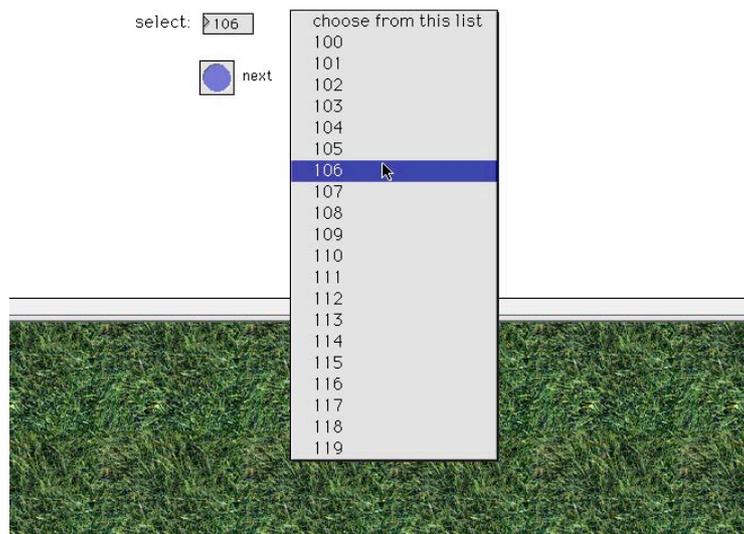


Fig. 3. The list of menu items for selection

2.2.3 Phase 3: Slider Manipulation

In this experiment a visual cue was generated by the system, moving the horizontal system slider on screen to a certain position. The participants were instructed to place their slider to the same position, as fast and as accurate as they could. Articulatory feedback was given in four combinations of three modalities: visual only (V), visual + tactual (VT), visual + auditory (VA), and visual + auditory + tactual (VAT). The tactual feedback consisted of a pulse for every step of the slider, a triangular wave form with a base frequency of 83.3 Hz. The auditory feedback was generated with the same frequency, making a clicking sound due to the overtones of the triangular wave shape.

These combinations could be presented in 24 different orders, but the most important effect in the context of user interface applications, is from Visual Only to Visual and Tactual and vice versa. It was decided to choose the two that would reveal enough about potential order effects: V-VT-VA-VAT and VAT-VA-VT-V.

Fig. 4 shows the experimenter's screen, with several buttons to control the set up of the experiment, sliders and numbers that monitor system and participant actions, and numeric data indicating the number of trials. All this data was automatically logged into files. The participant's screen has only a system slider (giving the cues), their response slider, and some buttons for advancing the experiment.

All 40 cues were presented from a randomly generated, fixed order table of 20 values, every value occurring twice. Values near or at the extreme ends of the slider were avoided, as it was noted during pre-pilot studies that participants developed a strategy of moving to the particular end very fast knowing the system would ignore overshoot. The sliders were 600 pixels wide, mouse ratio was set to slow in order to make the participants really move. Through the use of the fixed values in the table across conditions, it was ensured that in all conditions the same distance would be travelled in the same order, thereby avoiding having to bother with Fitts' law or the mouse-pointer ratio (which is proportional to the movement speed, a feature which cannot be turned off in the Macintosh operating system).

The cues were presented in blocks for each condition, if the conditions would have been mixed subjects would have developed a visual only strategy in order to be independent on the secondary (and tertiary) feedback. Their slider was also colour coded, each condition had its own colour. 31 subjects participated in this phase, and the orders were balanced as well as possible: in 15 times the V – VAT order and 16 times the VAT – V order.

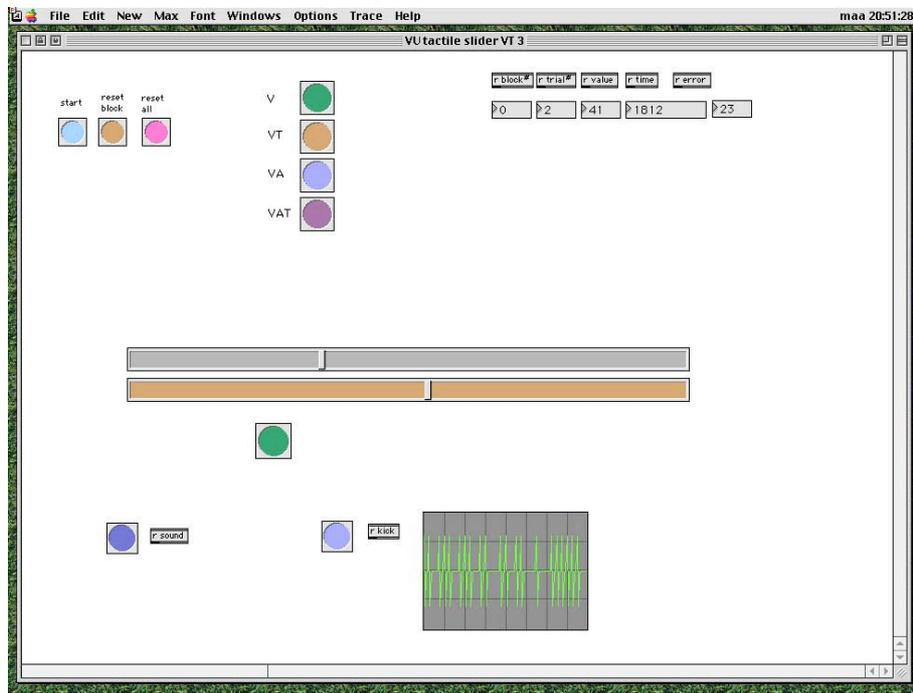


Fig. 4 The experimenter's screen of phase 3, the Slider experiment, showing the controls and monitoring of the experiment and the user's actions.

2.2.4 Phase 4: Step Counting

In this experiment the participants were prompted with a certain number of steps to be taken with a horizontal slider, similar to the one in the previous experiment. The range was set to 20 rather than 120 in this case, to keep it manageable. Conditions were visual only (V) and visual combined with touch (VT). No confirmative feedback was given, the subject would press a button when he or she thought that the goal was reached and then the next cue was generated.

In this phase of the experiment 11 subjects participated, balanced between the orders in 5 V – VT and 6 VT- V order.

2.2.5 Phase 5: Questionnaire

The last part of the session consisted of a form with questions to obtain qualitative feedback on the chosen modalities, and some personal data such as gender, handedness and experience with computers and the mouse. This data was acquired for 31 of the subjects.

2.3 Results and Analysis

The data from the files compiled by the Max patch was assembled in an spreadsheet. The data was then analysed in SPSS.

All phases of the experiment showed a learning effect. This was expected, so the trials were balanced in order (as well as possible) to compensate for this effect.

The errors logged were distinguished in two types: misalignments and mistakes. A mistake is a miss hit, for instance when the subject 'drops' the slider too early. A mistake can also appear in the measured response times, some values appeared that were below the physically possible minimum response time, as a result of the subject moving the slider by just clicking briefly in the desired position so that the slider jumps to that position (they were instructed not to use this 'cheating' but occasionally it occurred). Some values were unrealistically high, probably due to some distraction or other error. All these cases contain no real information, and were therefore omitted from the data as mistakes. The misalignment errors were to be analysed as they might reveal information about the level of performance. In the sections below the results of each phase are summarised, and the significant results of the analysis are reported.

2.3.1 Phase 1: Threshold Levels

Of the 19 presented textures (averaged over two runs) twelve were recognised correctly (including four of the six non-textures) by all the subjects, and a further ten with more than 90% accuracy (including two non-textures). There were two lower scores (88% and 75%) which corresponded with the two most difficult to perceive conditions, the combination of the lowest amplitude (30%), low frequency (25 Hz) and the lowest spatial distributions of two respectively four pixels.

2.3.2 Phase 2: Menu Selection

Table 1 shows the total mean response times. All trials were balanced to compensate for learning effects, and the number of trials was such that individual differences were spread out, so that all trials and subjects (N in the table) could be grouped. The response times Visual+Tactual condition were slightly higher than for the Visual Only condition, statistical analysis showed that this was not significant. The error rates were not statistically analysed, in both conditions they were too low to draw any conclusions from.

Table 1. Mean response times and standard deviations for the Menu Selection

| Condition | N | Mean response time (ms) | Standard deviation (ms) |
|-----------|------|-------------------------|-------------------------|
| V | 1120 | 1806 | 624 |
| VT | 1120 | 1842 | 531 |

2.3.3 Phase 3: Slider Manipulation

The table below gives an overview of the means and the standard deviations of all response times (for all distances) per condition.

Table 2. Mean response times and standard deviations for the Slider.

| Condition | N | Mean response time (ms) | Standard deviation (ms) |
|-----------|------|-------------------------|-------------------------|
| V | 1239 | 1436 | 530 |
| VT | 1240 | 1512 | 572 |
| VA | 1240 | 1468 | 595 |
| VAT | 1240 | 1547 | 593 |

The response times were normally distributed and not symmetrical in the extremes, so two-tailed t-tests were carried out in order to investigate whether these differences are significant. Of the three possible degrees-of-freedom the interaction between V and VT as well as VA and VAT was analysed, in order to test the hypothesis of the influence of T (tactual feedback added) on the performance.

Table 3. The results of the t-tests

| Condition | t | Significance |
|-----------|--------|--------------|
| V-VT | -3.435 | 0.001 |
| VA-VAT | -3.307 | 0.001 |

This shows that response times for the Visual only condition were significantly faster than the Visual + Tactual condition, and that the Visual + Auditory condition was significantly faster than the similar condition with tactual feedback added.

The differences in response times between conditions were further analysed for each of the 20 distances, but no effects were found. The differences in error rates were not significant.

2.3.4 Phase 4: Step Counting

It was observed that this phase of the experiment was the most challenging for the subjects. It relies a bit more on the participants' cognitive skill than on the sensori-motor loop alone. The total processing time and the standard deviation are reported the table 4.

Table 4. Mean response times and standard deviations for the Counter

| Condition | N | Mean processing time (ms) | Standard deviation (ms) |
|-----------|-----|---------------------------|-------------------------|
| V | 200 | 3569 | 2281 |
| VT | 200 | 3078 | 1868 |

In order to find out if the difference is significant an analysis of variance (ANOVA) was performed, to discriminate between the learning effect and a possible order effect in the data. The ANOVA systematically controlled for the following variables:

- between subject variable: sequence of conditions (V - VT vs. VT - V);
- within subject variables: 20 different distances.

The results of this analysis are summarised in the table below.

Table 5. The results of the comparison for the two conditions of the Counter phase

| Condition | df | error df | F | Significance |
|-----------|----|----------|--------|--------------|
| V / VT | 1 | 8 | 12.432 | 0.008 |

From this analysis follows that the mean response times for the Visual + Tactile condition are significantly faster than for the Visual Only condition.

The error rates were lower in the VT condition, but not significantly.

2.3.5 Phase 5: Questionnaire

The average age of the 31 subjects who filled in the questionnaire was 20 years, the majority was male (26 out of 31) and right handed (27 out of 31). The left handed people used the right hand to operate the mouse, which was their own preference (the set-up was ready to be modified for left hand use). They answered a question about their computer and mouse experience with a Likert scale from 0 (no experience) to 5 (lot of experience). The average from this question was 4.4 on this scale, meaning their computer and mousing skills were high.

The qualitative information obtained by the open questions was of course quite varied, the common statements are however categorised and presented in the table 6.

Table 6. Categorised answers on open questions, out of 31 subjects

| | tactual | auditory |
|------------------------------------|---------|----------|
| effective | 22 | 17 |
| effective but potentially annoying | 7 | 3 |
| annoying / irritating | 1 | 5 |
| not effective | 1 | 6 |

There was a question about whether it was thought that the added feedback was useful. For added sound 16 (out of 31) people answered "yes", 4 answered "no" and 11 thought it would depend on the context (issues were mentioned such as privacy, blind people, precision). For the added touch feedback 27 people thought it was useful, and 4 thought it would depend on the context. In table 7 the results are summarised, in percentages.

Table 7. Categorised answers on question about usefulness

| Added feedback | auditory | tactual |
|----------------------|----------|---------|
| Useful | 52% | 87% |
| Not useful | 13% | 0% |
| Depending on context | 35% | 13% |

When asked which combination they preferred in general, 3 (out of 31) answered Visual Only, 13 answered Visual+Tactual, 3 answered Visual+Auditory, 7 Visual+Auditory+Tactual, and 4 thought it would depend on the context. One subject did not answer. The results are shown in percentages in table 8.

Table 8. Preferred combination of articulatory feedback

| Preferred combination | V | VT | VA | VAT | depending on context |
|-----------------------|-----|-----|-----|-----|----------------------|
| | 10% | 42% | 10% | 23% | 13% |

3 Discussion

The results from the Threshold test (Phase 1) show that our subjects had no difficulty recognising the stimuli generated, apart from the cases of the really low levels. For the experiments (Phases 2 – 4) combinations of the stimuli were chosen that were far above these thresholds.

The results from the menu selection (Phase 2) show no significant difference. This task was easy and familiar for the participants.

The results from the slider experiment (Phase 3) show that people tend to slow down when the tactual feedback is added. This can be due to the novelty effect of it, as people are not used having their sense of touch actively addressed by their computer. This slowing down effect may be explained by the factor of perceived physical resistance, as has been found in other research (Oakley et al, 2001).

When the task reaches a sufficient level of difficulty, the advantage of the added tactual feedback can be shown. This has been asserted by the step counter (Phase 4 of the experiment), where task completion times were significantly shorter with the added feedback.

The questionnaires show that people generally appreciate the added feedback, and favour the tactual over the auditory. Sound feedback is sometimes perceived as irritating, also depending on the context. Note however, that computer generated tactual feedback can be unpleasant as well. Not much research has been carried out investigating this rather qualitative aspect and its relationship to the choice of actuator (eg. a motor or a speaker).

The participants who carried out the experiments were involved in a lecture series on Information Representation at the department, and they were therefore very involved in the experiments. It can be argued that this would bias their responses, particularly in the qualitative parts of the session, but it must be stressed that the work here is carried out in the context of human-computer interaction research and not as pure psychometric experiments. Phase 5 of the experiment was therefore more an expert-appraisal.

Moreover, the core idea behind the research into tactual feedback as described in this paper was not part of the lectures.

The response times between subjects vary largely, which is quite fascinating and a potential subject for further investigation. Clearly, people all have their own individual way of moving (what we call the 'movement fingerprint') and these results show that even in the simplest gesture this idiosyncrasy can manifest itself.

The observation that in some cases some participants actually seem to slow down when given the added tactual feedback has a lot to do with the tasks, which were primarily visual. The tactual feedback can be perceived by the user as resistance. This is interesting, as we often experience in real world interactions (particularly with musical instruments) that indeed effort is an important factor in information gathering of our environment, and for articulation. Dancers can control their precise movements while relying more on vision and internal kinaesthetic feedback, but only after many years of training. It is still expected that the greatest benefit of adding tactual feedback to a gesture will be found in a free moving gesture, without any passive feedback as is the case when moving the mouse. This has to be further investigated. Some preliminary experiments we carried out both with lateral movements as well as with rotational movement in free space show promise. The potential benefit of this has already been shown elsewhere, for instance at Sony Labs (Poupyrev, 2002) where a tilting movement of a handheld device did benefit of the added or secondary feedback (here somewhat confusingly labelled as 'ambient touch').

4 Conclusion and further research

In the research described in this paper we have investigated the influence of added tactile feedback in the interaction with the computer. It has been found that in simple tasks the added feedback may actually decrease performance, while at more difficult tasks an increase in effectiveness can occur.

The vibrotactile element, based on a miniature loudspeaker and low-frequency 'sounds', proved to be a cost-effective and flexible solution allowing a wide variety of tactile experiences to be generated using sound synthesis techniques.

In the experiments the added feedback was redundant. In all these cases the various feedback modalities were used to convey the same information, following the recommendations of the Proximity Compatibility Principle (PCP) (Wickens, 1992 pp. 98-101), (Wickens and Carswell, 1995). In a multitasking situation the feedback would perhaps be divided differently according to PCP, but in our experiment they are tightly coupled. We are not trying to replace visual feedback by tactual feedback, but add this feedback to make the interaction more natural and possibly more effective. It is expected that in some cases this may lead to a performance benefit. The widget manipulations we chose to investigate are standard user interactions, the subjects were very familiar with it, and they could devote all their attention to fulfilling the tasks. Only when the task is made less mundane, such as in the last phase where the subjects had to count steps of the slider widget, adding tactile feedback helps to improve the interaction. This may seem irrelevant in the current computer interaction paradigm, which is entirely based on the visio-motor loop, but our research has a longer term goal of developing new interaction paradigms based on natural interaction. Multitasking is an element of such paradigms, and a future experiment can be developed where the participants have to divide their attention between the task (ie. placing the slider under various feedback conditions) and a distracting task.

The set-up as described in this paper mainly addresses the Pacinian system, the one out of the four tactual systems of cutaneous sensitivity that has the rapidly adapting and diffuse sensitivity (see section 1.5), and that can pick up frequencies from 40 Hz to 1000 Hz (Verrillo, 1992). Other feedback can be applied as well, conveying different kinds of information.

In the current situation only pure articulatory feedback was considered, other extensions can be added later. This would also include the results from the ongoing research on virtual textures, resulting in a palette of textures, feels and other tangible information to design with. A logical extension of the set-up is to generate feedback upon reaching the goal, a 'confirmative feedback' which has been proven to produce a strong effect. This could greatly improve performance as it is known that the gesture can be described by dichotomous models of human movement in a pointing task, which imply an initial ballistic movement towards a target followed by finer movements controlled by finer feedback (Mithal, 1995) (Philips and Triggs, 2001). This was observed in some cases in our user trials.

Other improvements that are thought of are: to have multiple points of feedback addressing more fingers or parts of fingers (Max/MSP works with external hardware to address multiple output channels) and to make the vibrotactile element smaller, so that more can be put on one finger. This can be used to simulate direction, such as in the effect of stroking velvet. The palette eventually should also incorporate force-feedback, addressing the kinaesthetic sense.

Multiple degree-of-freedom input is another expected element of new interaction paradigms.

We are interested to see the effects of added tactual feedback on a free-moving gesture, when the passive feedback (kinaesthetic guidance from the table top surface in the 2D case) is absent. It is expected that adding active feedback can compensate for this. Such an experiment has been developed and is currently tested.

Acknowledgements

The authors wish to thank the student interns Ioana Codoban and Iulia Istrate for their assistance in setting up and carrying out the experiments and their work on the statistical analysis of the data, and Jos Mulder for his help with the sound level measurements. Also thanks to the anonymous reviewers for their critical comments, which helped the authors to make a number of improvements to the article.

References

- Akamatsu, M. and Sato, S. (1992). Mouse type interface device with tactile and force display. *Proceedings of the International Conference on Artificial Reality and Telexistence*, pp. 177 – 182.
- Akamatsu, M. and Sato, S. (1994). A multi-modal mouse with tactile and force feedback. *International Journal of Human-Computer Studies*, **40**, 443-453.
- Akamatsu, M. and MacKenzie, I. S. (1996). Movement characteristics using a mouse with tactile and force feedback. *International Journal of Human-Computer Studies*, **45**, 483-493.
- Bolanowski, S. J. Jr., G., Gescheider, G. A., Verrillo, R. T. and Checkowsky, C. M. (1988). Four Channels Mediate the Mechanical Aspects of Touch. *Journal of the Acoustic Society of America*, **84/5**, 1680-1694.
- Bongers, A. J. (1998). Tactual Display of Sound Properties in Electronic Musical Instruments. *Displays Journal*, **18**, 129 – 133
- Bongers, A. J., Eggen, J. H., Keyson, D. V. and Pauws, S. C. (1998). Multimodal Interaction Styles. *HCI Letters Journal*, **1/1**, 3-5.
- Bongers, A. J. (2002a). Using electronic musical instrument design techniques for measuring behavior. In: *Proceedings of the 4th International Conference on Methods and Techniques in Behavioral Research*. pp. 22 – 25.
- Bongers, A. J. (2002b). Palpable Pixels, A Method for the Development of Virtual Textures. In: Ballesteros, S. and Heller, M. A., Eds. *Touch, Blindness and Neuroscience*, Madrid: UNED Press.
- Bongers, A. J. and Veer, G. C. van der, Towards a Multimodal Interaction Space, categorisation and applications. *Journal of Personal and Ubiquitous Computing*, special issue on Movement-Based Interaction [accepted for publication], 2006
- Chafe, C. (1993). Tactile Audio Feedback. *Proceedings of the ICMC International Computer Music Conference*, pp. 76 – 79.
- Chafe, C. and O'Modhrain, S. (1996). *Proceedings of the ICMC International Computer Music Conference*. pp. 428 – 431.
- Dennerlein, J. T., Martin, B. M. and Hasser, C. (2000). Force-feedback improves performance for steering and combined steering-targeting tasks. *Proceedings of the CHI 2000 Conference on Human Factors in Computing Systems*, pp. 423-429.
- Dennett, D. C. (1991). *Consciousness Explained*. US: Little / Brown.
- Dix, A., Finlay, J., Abowd, G and Beale, R. (1993). *Human-Computer Interaction*. Prentice Hall.
- Eco, U. (1976). *A Theory of Semiotics*. Bloomington: Indiana University Press.
- Eggen, J. H., Haakma, R. and Westerink, J. H. D. M. (1996). Layered Protocols: hands-on experience. *International Journal of Human-Computer Studies*, **44**, 45-72.
- Engel, F. L., Goossens, P. and Haakma, R. (1994). Improved efficiency through I- and E-feedback: a trackball with contextual force feedback. *International Journal of Human-Computer Studies*, **41**, pp. 949-974.
- Erickson, T. (1995). Coherence and Portrayal in Human-Computer Interface Design. In: R.J. Beun, M. Baker and M. Reiner, Eds. *Dialogue and Instruction*. pp. 302-320. Heidelberg: Springer Verlag.
- Gibson, J. J. (1962). Observations on active touch. *Psychological Review*, **69/6**, 477-491.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston, MA: Houghton Mifflin.
- Göbel, M., Luczack, H., Springer, J., Hedicke, V. and Rötting, M. (1995). Tactile Feedback Applied to Computer Mice. *International Journal of Human-Computer Interaction*, **7/1**, 1-24

- Hardwick, A., Rush, J., Furner, S. and Seton, J. (1996). Feeling it as well as seeing it – haptic display within gestural HCI for multimedia telematic services. *Progress in Gestural Interaction*, Springer-Verlag, pp. 105-116.
- Hardwick, A., Furner, S. and Rush, J. (1998). Tactile access for blind people to virtual reality on the world wide web. *Displays Journal*, **18**, 153-161
- Hix, D. and Hartson, H. R. (1993) *Developing User Interfaces*. New York: John Wiley & Sons, Inc.
- Hwang, F., Keates, S., Langdon, P. and Clarkson, P. J. (2003). Multiple haptic targets for motion-impaired users. *Proceedings of the CHI 2003 Conference on Human Factors in Computing Systems*, pp. 41-48.
- Keyson, D. V., and Houtsma, A. J. M., (1995). Directional sensitivity to a tactile point stimulus moving across the fingerpad. *Perception & Psychophysics*, **57/5**, 738-744.
- Keyson, D. V., (1996). *Touch in User Interface Navigation*. PhD Thesis, Eindhoven University of Technology, The Netherlands.
- Keyson, D. V. and Stuivenberg, L. van, (1997). TacTool v2.0: an object-based multimodal interface design platform. *Proceedings of the HCI International Conference*, San Francisco, pp. 311-314.
- Kruglanski, O. (2000). As much as you love me. In: Leopoldseeder, H and Schöpf, C. eds., *Cyberarts International Prix Arts Electronica*, pp. 96-97, Springer –Verlag.
- Littlejohn, S. W. (2002). *Theories of Human Communication*. Belmont CA: Wadsworth Thomson.
- Loomis, J. M. and Leederman, S. J. (1986). Tactual Perception. In: *Handbook of Perception and Human Performance*, Chp. 31.
- Max/MSP website: www.cycling74.com
- McCullough, M. (1996). *Abstracting Craft, The Practised Digital Hand*. Cambridge MA: MIT Press.
- Minsky, M., Ouh-Young, M., Steel, O. and Brooks, F. P. (1990). Feeling and seeing: issues in force display. *Computer Graphics*, **24**, 235-243.
- Mithal, A.K. (1995). Using psychomotor models of movement in the analysis and design of computer pointing devices. In: *Proceedings of the CHI*, pp. 65 – 66
- Münch, S. and Dillmann, R. (1997). Haptic output in multimodal user interfaces. *Proceedings of the International Conference on Intelligent User Interfaces*, pp. 105-112.
- Nielsen, J. (1986). A virtual protocol model for computer-human interaction. *International Journal of Man-Machine Studies*, **24**, 301-312.
- Oakley, I., McGee, M. R., Brewster, S. A. and Gray, P. D. (2000). Putting the ‘feel’ into ‘look and feel’. In: *Proceedings of the CHI 2000*. ACM Press, pp. 415-422.
- Oakley, I., Brewster, S. A. and Gray, P. D. (2001). Solving multi-target haptic problems in menu interaction. In: *Proceedings of the CHI 2001*, ACM Press, pp. 357-358
- Oakley, I., Adams, A., Brewster, S. A. and Gray, P. D. (2002). Guidelines for the design of haptic widgets. In: *Proceedings of the HCI 2002 Conference*.
- Phillips, J. G and Triggs, T. J. (2001). Characteristics of Cursor Trajectories Controlled by the Computer Mouse. *Ergonomics*, **44/5**, 527 – 536.
- Poupyrev, I., Maruyama, S. and Rekimoto, J. (2002). Ambient touch: designing tactile interfaces for handheld devices. *Proceedings of the User Interface and Software Technology (UIST) Conference*, Paris.
- Rosengren, K. E. (2000). *Communication, An Introduction*. London: Sage.
- Schomaker, L., Münch, S., and Hartung, K., eds. (1995). *A Taxonomy of Multimodal Interaction in the Human Information Processing System*. Report of the ESPRIT project 8579: MIAMI.
- Sherrick, C. E. and Cholewiak, R. W. (1986). Cutaneous Sensitivity. In: *Handbook of Perception and Human Performance*, Chpt. 12.
- Taylor, M. M. (1988). Layered protocol for computer-human dialogue. I: Principles. *International Journal of Man-Machine Studies*, **28**, 175-218
- Verrillo, R.T. (1992). Vibration sensing in humans. *Music Perception*, **9/3**, pp. 281-302.
- Vertegaal, R. (1998). *Look Who's Talking to Whom*. Doctoral Thesis, Twente University, The Netherlands.
- Wensveen, S. A. G., Djajadiningrat, J. P. and Overbeeke, C. J. (2004). Interaction Frogger: a Design Framework to Couple Action and Function through Feedback and Feedforward. In: *Proceedings of the Designing Interactive Systems DIS'04 conference*.
- Wickens, C. D. (1992). *Engineering Psychology and Human Performance*. Second Edition, New York NY: Harper Collins.
- Wickens, C. D., & Carswell, M. (1995). The proximity compatibility principle: Its psychological foundation and relevance to display design. *Human Factors*, **37**, 473-494.