Perception of Dynamic Audiotactile Feedback to Gesture Input

Teemu Ahmaniemi Vuokko Lantz Juha Marila

Nokia Research Center
P.O. Box 407
FI-00045 Nokia Group, Finland

{ teemu.ahmaniemi, vuokko.lantz, juha.marila } @nokia.com

ABSTRACT

In this paper we present results of a study where perception of dynamic audiotactile feedback to gesture input was examined. Our main motivation was to investigate how users' active input and different modality conditions effect the perception of the feedback. The experimental prototype in the study was a handheld sensor-actuator device that responds dynamically to user's hand movements creating an impression of a virtual texture. The feedback was designed so that the amplitude and frequency of texture were proportional to the overall angular velocity of the device. We used four different textures with different velocity responses. The feedback was presented to the user by the tactile actuator in the device, by audio through headphones, or by both. During the experiments, textures were switched in random intervals and the task of the user was to detect the changes while moving the device freely. The performances of the users with audio or audiotactile feedback were quite equal while tactile feedback alone yielded poorer performance. The texture design did not influence the movement velocity or periodicity but tactile feedback induced most and audio feedback the least energetic motion. In addition, significantly better performance was achieved with slower motion. We also found that significant learning happened over time; detection accuracy increased significantly during and between the experiments. The masking noise used in tactile modality condition did not significantly influence the detection accuracy when compared to acoustic blocking but it increased the average detection time.

Categories and Subject Descriptors

H.5.2 User Interfaces: Haptic I/O; H.5.2 User Interfaces: Auditory (nonspeech) feedback

General Terms

Documentation, Experimentation, Human Factors

Keywords

Gesture interaction, audio, haptics.

ICMI'08, October 20–22, 2008, Chania, Crete, Greece.

1. INTRODUCTION

Dynamic audiotactile feedback is natural to any hand motion. In fact, almost all of our manual interactions with the physical world produce tactile sensations and audible cues which are proportional to the motion itself. When we touch a table, explore a textile, or hit a tennis ball with a racquet, the immediate feedback seems obvious but at the same time it tells us essential information about the physical properties of the objects and interaction between them. If the natural feedback is missing or artificially modified it is easy to notice the change in interaction experience. Much more difficult is to specify the particular features which have changed in the feedback.

1.1 Gestural Interaction

In human-to-human communication hand motions, gestures are an essential part of interaction and their interpretation is learned together with the language and culture of communication [4]. Using gestures in Human-Computer Interaction (HCI) is a new concept and gestural interfaces are challenge for both the human and the computer. In handheld devices, the motion and touch sensing technologies have developed rapidly together with increased processing capacity. Thus, the main challenge is not necessarily the gesture recognition technology itself but rather the way how the technology is presented to the user. Because of the short evolution of gestural interfaces, users do not have expectations on what kind of gestural input the system is able to understand. From the user interface design perspective the challenge is to add proper cues, affordances, to indicate the available gestural action possibilities [10].

1.2 Dynamic Feedback

If the system responds immediately to user's hand motion, it already gives a clue of the action possibilities. A simple example is a visual mouse pointer: usually user has to shake the mouse in order to locate the pointer on the screen. Small random input is needed to sense how the system responds. When using handheld devices, audio and tactile are practical feedback modalities for gesture interaction. Firstly, the visual feedback on the screen can be difficult to follow since the device is moving together with the motion. Secondly, tactile actuation is surely sensed since the device is in user's hand. But at the same time, designing the feedback for gesturing is challenging due to the different spatial characteristics of input and output. Motion input has always direction and magnitude which are rather easy to visualize. With audio and tactile feedback, the user needs to interpret the mapping between the spatial input and non-spatial output. Of course, audio can be represented spatially as well as tactile feedback [5] but the spatial resolution of those is much worse than that of visual feedback.

Williamson and Murray-Smith [15] introduced a gesture recognition system where state of the classifier was presented to the user by real-time audio feedback. Similarly, in the study by Rath and Rohs [12], a real-time audio feedback was coupled to a rolling ball movement controlled by tilting. The audio feedback shortened the time spent on tilting task. Another interesting finding of the study was that the best learning result was achieved with an abstract sound. The other feedback designs were a sound that reflects the visual feedback of the rolling ball and no sound at all. Furthermore, Mononen [9] showed that real-time visual or auditory feedback during the aiming of rifle shooting improves the learning and task performance. The above studies show the potential of the real-time feedback but lack the evidence of how the feedback is perceived and do not provide insight into how an optimal feedback should be designed.

1.3 Perception of Audiotactile Feedback

The perception of tactile feedback in human computer interfaces has been investigated in several studies. Brown *et al.* [2] presented results of a study where the effectiveness of tactile patterns was examined. Rhythm and roughness of the vibrotactile pulses were considered the best design parameters yielding the best result in a discrimination task. Salminen *et al.* [13] studied the tactile perception with a horizontally rotating fingertip stimulator. Again, altering the temporal characteristics of the stimulus appeared to lead to the best performance in discrimination task. Hoggan and Brewster investigated the transfer of stimulus recognition skills between audio and tactile modalities [5]. Subjects recognized the stimuli with similar temporal characteristics quite well even they were trained with a different modality. Jousmäki and Hari [7] presented an interesting experiment where amplification and attenuation of perceived audio signal influenced the subjective sensation of touch (skin roughness) when subjects were rubbing their hands together. The finding suggests that audio feedback modifications can be used for modifying the tactile sensation.

Lederman *et al.* [8] studied user's performance in texture identification task. Subjects explored different textures with a probe and were asked to identify the textures based on tactile cues, auditory cues or both. The best identification performance was obtained with tactile and audiotactile cues, while audio alone resulted in worst performance. In the study, the experiment setup did not enable the elimination of the tactile modality, which caused modality conditions to be unequal; in the audio only modality condition, the audio feedback was generated not from subject's active exploration but from experimenter's motions. This may have had strong impact on the results.

1.4 Previous Work

We presented recently a study where a sensor-actuator prototype was introduced [1]. The prototype was designed for investigating the closed-loop interaction with motion input and real-time audiotactile output. We designed four simple audiotactile feedback textures and tested how well our test subjects were able to make difference between them in different feedback modality conditions through active perception. In our preliminary analysis we found that the audio feedback is dominating the perception. Tactile feedback led to decent results but the modalities together did not yield better performance than the audio alone.

In this paper we analyze the test results in more detail. Special attention is paid to users' active role in perception by examining how the scale and periodicity of motion influence in change detection performance. Also, the effects of masking noise used with tactile feedback, and subjects' learning are taken into consideration.

2. EXPERIMENTS AND METHODS

We conducted two experiments to assess active audiotactile perception. In both experiments, the task of the subject was exactly the same: indicate by a button press as promptly as possible whenever some characteristics of the system's response to their movements have changed independently of subject's own actions. The information content in audio and tactile feedback modalities was exactly the same and synchronized. Both the audio and tactile feedback were generated from the overall instantaneous angular velocity of the subject's hand movement. We experimented with two feedback design principles: 1) regular or irregular, 2) and slow, moderate, or fast response to overall angular velocity (see section 2.3). In both experiments, subjects were asked to detect changes between the four different textures: slow, moderate, and fast versions of regular feedback, and a slow version of irregular feedback.

The main motivation for the first experiment (Experiment 1) was to study active perception of virtual textures by comparing the effect of gesturing measures, the different modality conditions and texture designs on change detection performance. These analyses provide us insight into how to set the design parameters in order to create distinguishable virtual textures (see section 3.1.1), roles of different feedback modalities in the change detection (3.1.3), skill transfer between different modality conditions (3.1.6), what kind of gesturing patterns the dynamic feedback elicits (3.1.2, 3.1.4), relationships between the change detection performance and gesturing measures (3.1.5).

The second experiment (Experiment 2) was conducted in order to clarify the role of masking noise in tactile only condition (3.2.1). It also enabled us to study how well the subjects can maintain their skills to detect texture changes without regular exposure to them (3.2.2). Since the audio noise masking was not present in the other modality conditions of Experiment 1, it was unclear whether it disturbed the change detection performance by increasing the cognitive load of the user. Generally, the effects of noise on performance are variable: it may enhance or decrease performance, or have no effect at all [3]. Adding masking noise to the modality condition with audio feedback only was not considered as a viable approach as it would have reduced the signal-to-noise ratio of the feedback and made it incomparable with other modality conditions. Instead, we decided to use heavy acoustic blocking to attenuate the leakage sound of the tactile actuator.

2.1 Test Subjects

Experiment 1 was carried out with 29 subjects of whom 9 were females. The test was performed with the test device in the dominant or preferred hand. Majority, that is 27 subjects, used right hand (26 right-handed, 1 ambidextrous) for holding and moving the test device. Experiment 2 was performed by 12 subjects (5 females) who had participated also in Experiment 1. One of the subjects was left-handed and preferred to hold the test device in the left hand. In both tests, the other hand was used for reporting the texture change detections in the feedback with button presses on a small keypad. The ages of the subjects varied between 23 and 44 years, the average being 33 years in Experiment 1, and 30 years in Experiment 2.

2.2 Test Equipment

The test equipment consisted of handheld device made of commercial off-the-shelf components fit in a plastic chassis (Figure 1). It was connected to a PC laptop (Acer TravelMate C110) with a wire. The movements of the test device were captured with a commercial MT9 motion sensor by xSens (www.xsens.com). The sensor signal was used for the immediate synthesis of dynamic audiotactile feedback and saved for the later analysis of the gesturing behavior of the test subjects. The audio and tactile feedback signals were generated with Pure Data real-time audio synthesizer software (PD, http://puredata.info/) running on the laptop. The audio- and tactile signals were channeled to the standard stereo audio output of the PC. Left channel was used for tactile and right channel for audio output. The tactile signal was fed to the actuator (C2 Tactor by Engineering Acoustics, Inc. http://www.eaiinfo.com/) housed in the device while the audio signal was fed to standard headphones.

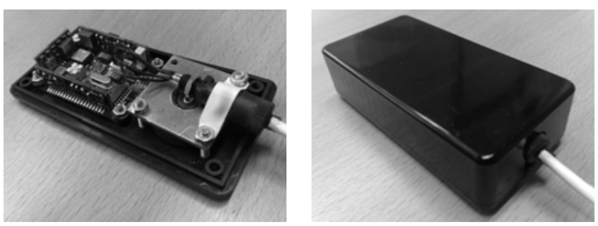

Figure 1. Handheld prototype consisting of a motion sensor and tactile actuator.

In test conditions with only tactile modality (Experiment 2), acoustic blocking of the tactile actuator's sound leakage was realized with Peltor Alert M2RX7A hearing protector headset equipped with an FM radio receiver (http://www.peltor.se/) and with Bilson 303S/303L disposable earplugs. The hearing protector headset provided 29 dB(A) attenuation for frequencies below 250Hz (Peltor specs) while the earplugs attenuation was 24 dB(A) [14]. The sound pressure levels measured in A-weighted decibels dB(A) have been adjusted for the human ear's varying sensitivity to different frequencies. The mitten and acoustic board which were used in Experiment 2 provided additional 3 dB(A) attenuation according to our own measurements (TES 1352A Sound Level Meter).

The noise and signal levels were measured with an ear and cheek simulator system provided by GRAS [6]. The masking noise from the hearing protector used in the both experiments was 75 dB(A) radio noise from the radio receiver with bandwidth of 0-6000Hz. The sound pressure level of the leakage sound of tactile actuator depended on the gesture input (see section 2.3). The measured level of the sound with a typical gesture input was 37dB(A). So, both the acoustic attenuation (Σ=56dB) and noise masking (75dB vs. 37 dB(A)) can be considered sufficient enough for protecting the subject from auditory sensation when only tactile feedback was present. In the test blocks with audio feedback, the sound pressure level of the audio signal with a typical gesture input was 75 dB(A). The audio feedback was played through AKG K240 Monitor headphones (Experiment 1) or Nokia stereo-HF HS-45 earphones (Experiment 2) worn under the hearing protector headset.

2.3 Feedback Synthesis

The sensor hardware consisted of 3D-accelerometer, -gyroscope, and -magnetometer. The sampling frequency was 100 Hz. Sensor data acquisition, initial preprocessing, and recording were realized using the Windows software provided by the sensor manufacturer. The magnetometer signal was disregarded in the sensor data fusion stage because of the electromagnetic interference with the tactile actuator. The movements of the test device were represented by Euler angles with a sampling rate of 100 Hz. The orientation angles were derivated numerically into angular velocities. The angular velocities were sent to PD for feedback synthesis. Tactile actuator was utilized in its optimal range; the recommended drive for the C2 was 250 Hz sinusoidal signal. The tactile feedback was generated by modulating the sinusoidal signal with an envelope signal [1]. The frequency (texture density F) and amplitude (texture intensity I) of the envelope signal was proportional to the overall angular velocity (Ω) of the device motion. The audio signal was generated using the same envelope signal but instead of sinusoid, noise was used as the modulated signal. Figure 2 presents the audio and tactile output of the feedback synthesis with a simulated input signal corresponding to a constant acceleration of the device.

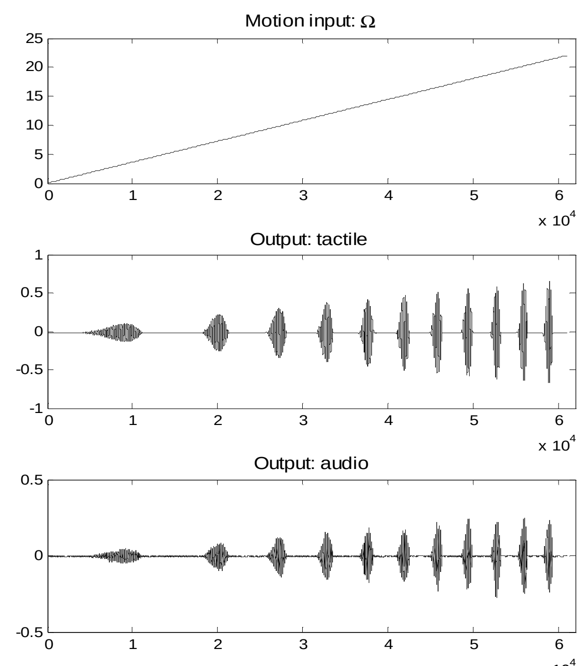

Figure 2. Tactile (middle) and audio (bottom) output of the feedback synthesis with a simulated linear sensor motion Ω (top).

We created four different textures. In the first three cases, the texture density was linearly proportional to the sum of absolute values of angular velocities ($\Omega=|\omega_x|+|\omega_y|+|\omega_z|$) of the handheld device with three different slopes, slow ($k_1=b\Omega$), medium ($k_2=2b\Omega$) and fast ($k_3=3b\Omega$). Coefficient b was hand tuned to make the textures pleasant within a wide frequency range. The fourth texture was otherwise similar to the slow one but noise was added to the input signal (Ω) creating an impression of irregular texture. The noise was generated by sampling an even distribution with maximum value inversely proportional to the angular velocity Ω. In all textures, the amplitude response was the same and linearly proportional to Ω. The actual displacement of the device was assessed with a laser measurement device (Keyence LK G152). The measured displacement of all four textures is presented in Figure 3. The simulated input for the measurement was corresponding to a linearly accelerating movement of the device.

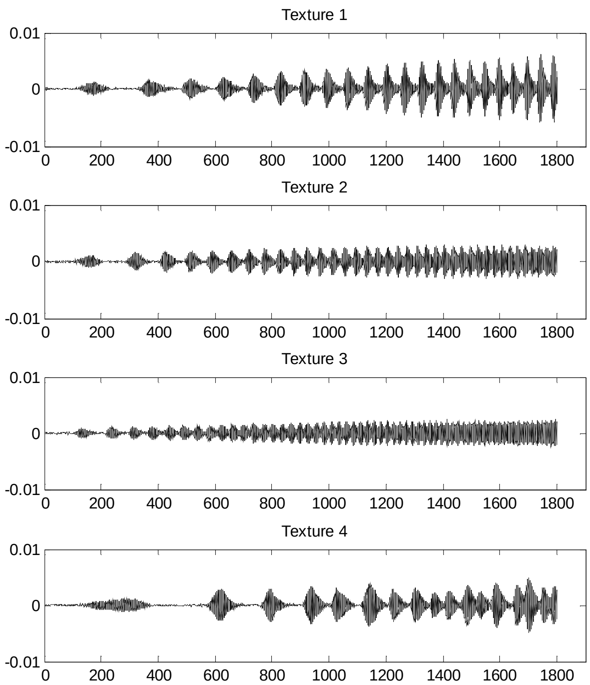

Figure 3. Measured displacement of the device with simulated input signal corresponding to linearly accelerating movement. The unit of displacement is millimeters and horizontal axis represents time in samples (1/1000 s).

2.4 Test Procedure

During both experiments, subject were seated on an office chair with the response keypad next to them on an adjacent office chair (Experiment 1) or in subject's lap (Experiment 1 and 2). Subjects were asked to sustain one finger poised for a button press in order to minimize the response times to texture change detections.

In the Experiment 1, the subjects could move the test device freely but were instructed to keep the magnitude of the movement modest and to maintain visual contact to the device. Majority of the subjects performed small- to moderate-sized gesturing movements above their laps. In Experiment 2, subjects were seated next to an acoustic board. They moved the test device underneath it and wore a mitten over the hand and test device for maximal attenuation of sound leakage from tactile actuator. Naturally, the acoustic board limited the movements in vertical direction and the size of gesturing was relatively small in all the cases. The setup of Experiment 2 is illustrated in Figure 4.

Figure 4. An example of the subject's position in Experiment 2

In the Experiment 1 we had three feedback conditions: tactile feedback with auditory noise masking (T+n), audio feedback (A), and combined tactile and audio feedback (T+A). Subjects performed one test block with the T+A condition and another test block with either T+n or A condition. The order of test blocks was balanced. The 4 different textures were presented 24 times each in a pseudo random order which was the same for all the subjects. All the 12 different types of changes from one texture to another one occurred 8 times. In total this makes 96 texture changes. Each texture was applied for 5 seconds plus a random, uniformly distributed time interval from 0 to 3 seconds. To avoid striking discontinuities at the texture changes, the transition from texture to another was done smoothly. This was realized by changing the slope parameter k_i of one texture (section 2.3.) to another k_j gradually within 100ms.

In the Experiment 2, all subjects performed the following three test blocks: tactile feedback with acoustic sound blocking (T), tactile feedback with audio noise masking (T+n, as in Experiment 1), and audio feedback (A, as Experiment 1). The procedure was similar to Experiment 1; only the length of the test blocks was halved into 48 texture changes.

The Experiment 1 was video recorded for a later qualitative analysis of the subject's gesturing behavior. Subjects were also interviewed afterwards. They were asked to estimate the number of different textures and to describe the differences between the textures. The additional questions concerned the users' impressions of the relationship between the motion and feedback, their gesturing strategies, learning process, and the roles and usefulness of each feedback modality in the change detection task. The subjects of the Experiment 2 were not interviewed afterwards as they had participated in Experiment 1 and were familiar with the feedback design principles.

2.5 Sensor Data Post Processing

The motion data of each test session was stored for later analysis. Hand movements were analyzed with respect to the energy, periodicity and period time. These measures were estimated over the time window between the two consecutive texture changes (5-8 s) (see Figure 5).

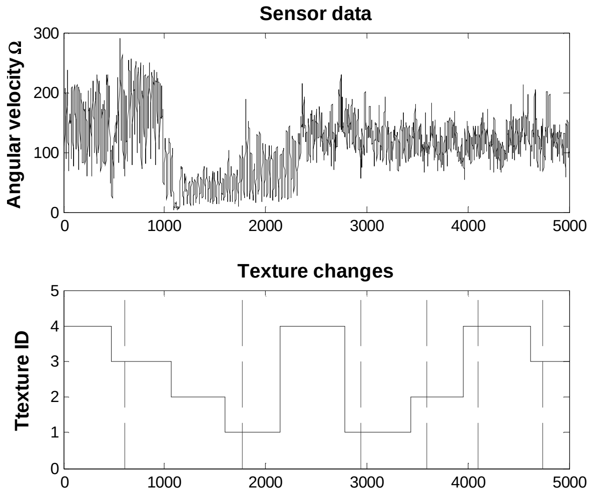

Figure 5. An example of collected data from a test session. Upper plot: overall angular velocity (Ω); lower plot: detected texture changes at dashed lines.

The energy of the movement was determined by the mean of the squared overall angular velocity Ω. The periodicity of the gesturing movement was estimated with a normalized auto-correlation function. First, the overall angular velocity was filtered with first-order Butterworth high- and low-pass IIR filters with cutoff frequencies at 0.4 Hz and 2 Hz, respectively, as the observed gesture durations were approximately 1 second. The height and location of the first positive peak of the autocorrelation function above 0.4 Hz were used as periodicity and period time estimates. If the subject gestured in a periodic manner and did not alter her/his gesturing pattern within texture, the periodicity estimate is close to one. Otherwise, the periodicity estimate is close to zero, even if the subject experimented with various periodic gestures. Therefore, a low periodicity estimate value can be an indication of either completely nonrecurring gesturing behavior or intermittent and frequent variations between multiple periodical gesturing patterns.

3. RESULTS

3.1 Experiment 1

3.1.1 Detection performance for different texture change types

The texture frequency affected most the detection accuracy and detection times. The changes between the most similar texture pairs were detected with worst accuracy and highest average detection times. The average rate of correct detections was the lowest for the changes between regular and irregular slow textures. Also, the changes between fast and medium regular textures were difficult to detect. Decent detection performance (>70%) was found for the changes between slow (irregular or regular) and fast and medium frequency textures. The average detection times followed the suite. Table 1 shows the correct detection rates and average detections times for all the different types of texture changes. One-way ANOVA between textures and detection rate as well as detection time shows the differences are significant ($p<0.001$).

3.1.2 Gesturing measures for different textures

Textures did not have a significant effect on gesturing behavior measures, either averaged over all feedback conditions or inside them. One-way ANOVAs show the differences in the energy, periodicity, and period time to be independent of texture type. Thus, the following analyses will not make difference between the textures or texture change types - the focus will be on the effects of the feedback modality condition on the change detection performance and gesturing behavior.

Table 1. Correct detection rate (%) and detection time (seconds) for the different types of texture change, results are averaged over all feedback modality conditions and subjects.

From \ *to*	***Slow regular***	***Medium regular***	***Fast regular***	***Slow irregular***
Slow regular	-	60.5% 1.86s	77.7% 1.60s	46.3% 2.55s
Medium regular	71.0% 1.94s	-	43.9% 2.58s	79.9% 1.87s
Fast regular	84.6% 1.56s	35.4% 2.68s	-	87.5% 1.66s
Slow irregular	28.4% 2.64s	68.5% 1.79s	74.0% 1.68s	-

3.1.3 Detection performance in different modality conditions

Audio feedback enabled better detection of texture changes than tactile feedback on its own. The correct detection rate was 68% for A, 54% for T+n, and 66% for T+A condition. The difference between the detection rates observed for A and T+n conditions approaches statistical significance (independent samples t-test: $t=-2.08$, $df=27$, $p<0.05$).

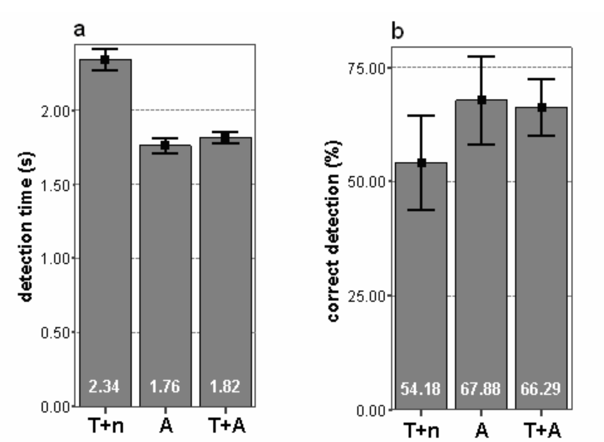

Figure 6. Detection performance as a) detection time and b) correct texture change detection rate, averaged over feedback design change types and show with the 95% confidence intervals.

Detection times differed along the same lines: the average time to correctly detect a texture change was 1.76 seconds for A, 2.34 s for T+n, and 1.82 s for T+A condition. The differences in detection times are significant between A and T+n (independent t=3.24, df=27, $p<0.01$) and T+n and T+A conditions (paired samples t-test: t=4.32, df=14, $p<0.01$). These results are illustrated in Figure 6. The number of false detections, i.e. subject indicated a change in the feedback design when there was none, was low in all feedback modality conditions. The percentage of false detections of all subject indications ranged from 6% to 10% in the first test block and from 3% to 8% in the second test block. From this part of analysis, data from subject 8 was omitted due to an abnormally high rate of false detections, 15-fold compared to average.

3.1.4 Gesturing measures in different modality conditions

The feedback modality conditions differed significantly in respect to the energy invested in the hand and device movements. The energy of the movement was higher in the T condition than in either A or T+A condition, see Figure 7.

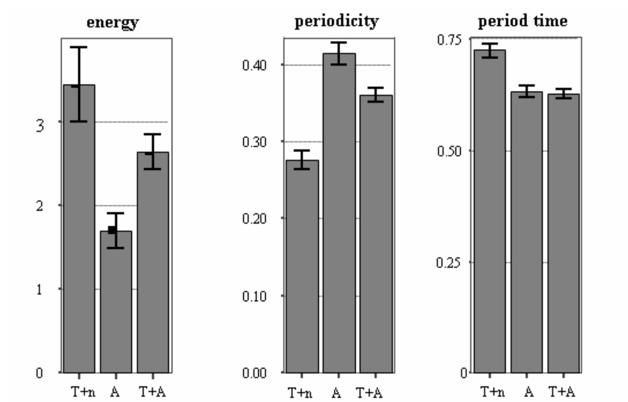

Figure 7. Means and 95% confidence intervals for energy ($10^5 \cdot (deg/sec)^2$), periodicity, and period time (seconds) of gesturing movements in different feedback modality conditions.

The movement was less periodic in the T condition than in A or T+A conditions. In one-way ANOVA, all the differences were found to be statistically significant ($p<0.001$). Tukey LSD post-hoc test reveals significant pair-wise differences for all variables between all conditions except for period time between T & T+A and T & A.

Table 2. The effects of periodic movement and energy on change detections, averaged over all subjects and modality conditions. All differences are significant at $p<0.01$ (equal variance assumed).

Gesturing measure	*Missed detect.*	*Correct detect.*	*Mean differ.*	*t*
Period time (sec)	0.67s	0.64s	0.03	4.46
Periodicity	0.31	0.37	-0.06	- 8.13
Energy (deg/sec)²	298422	240658	57763	3.31

3.1.5 Effect of movement characteristics on change detection performance

Correct and missed detections also differed in respect to the characteristics of the preceding gesturing movements. Periodicity, period time, and energy of the movement affected the change detection performance significantly. Correct detection was aided by shorter period time, higher periodicity, and lower energy of the gesturing movements (See Table 2). Two subjects were removed from this analysis as their energy measures were over two units of standard deviation above the mean in all feedback modality conditions.

3.1.6 Learning and skill transfer effects

Detection rate did clearly improve between the test blocks, regardless of the order of feedback modality conditions, see Table 3. Percentages of missed and false detections were significantly lower in the second block in all feedback modality conditions (chi^2=16.68, 130.68, and 39.21, df=2, $p<0.01$ for T+n, A, and T+A conditions, respectively). The relative improvement of the rate of missed detections was most prominent when T+n condition was followed by T+A (33.3%). Relative improvement was least prominent when T+A was followed by A (3.6%). The relative improvement was 18.2% and 9.6% for the cases where A was followed by T+A and T+A was followed by T+n, respectively. Table 3 shows missed and false detection rates and average detection times for correct response by feedback modality condition and testing order.

Table 3. Texture change misses and false detections, and correct detection times.

Order	*Cond. in block 1 or 2*	*Missed (%)*	*False (%)*	*Detect. time (s)*
1	**T+n**	45	8	2.46
	T+A	30	5	1.70
2	**A**	44	6	2.04
	T+A	36	3	1.97
3	**T+A**	52	10	1.97
	T+n	47	4	2.22
4	**T+A**	28	8	1.69
	A	27	8	1.55
Block 1 mean		42	8	2.01
Block 2 mean		35	5	1.82

3.1.7 Qualitative findings of the interview and video analysis

According to the video analysis, almost all test subjects (25 out of 29) established a stable movement pattern by the end of the first block of Experiment 1. Typical gesturing patterns were circles, repetitions of left-and-right movements and ∞-shapes. In the interviews, majority of the subjects (20) were able to articulate the relationship between the velocity of the device and frequency of the feedback. Almost as many (19) subjects could make a difference between regular and irregular texture designs. However, only 11 subjects understood both of the design principles. 7 subjects were able to define the number of presented textures (4) correctly. 6 subjects estimated it too low (<4) and 15 subjects too high (>4). Four subjects claimed that there were more than 10 textures. Majority (20) reported that the second block was easier than the first one, independently on the order of modality conditions. 5 subjects did not see any difference between the test blocks. 4 subjects felt that the second block was more difficult. Out of those 4 subjects, 3 subjects performed the test blocks in T+A followed by T+n order and one subject in T+A followed by A order.

3.2 Experiment 2

3.2.1 *Effects of audio masking noise on change detection performance*

The results of Experiment 2 were in line with those of Experiment 1: change detection performance with tactile feedback was lower than with audio feedback. One-way ANOVAs between feedback modality condition (3 levels) and detection rate, and condition and detection time, approach statistical significance (F=3.275 and 3.928 with significance levels 0.038 and 0.020, respectively). Tukey LSD post hoc tests reveal that the detection rates differed significantly between A and T conditions. See Table 4 for details.

Table 4. Tukey LSD comparisons between feedback modality conditions for correct detection rates and detection times. Significance at *p* <0.01 is marked with double asterisk (**) and *p*<0.05 with a single asterisk (*).

Cond. I	*Cond. J*	*detection rate, mean diff.(I-J)*	*detection time, mean diff. (I-J)*
A	T	**7.00 %** **	-0.01
T+n	T	3.40 %	**0.24** *
T+n	A	-3.70 %	**0.25** **

Detection times differed significantly between A and T+n as well as between T and T+n. Detection times were longer in the T+n feedback condition than in T or A conditions. The noise mask thus increased response times.

3.2.2 *Skill transfer between Experiment 1 and 2*

Subjects performed clearly better in Experiment 2 than in Experiment 1. This implies that the learning that had occurred during Experiment 1 had persisted for the three months that passed between the experiments.

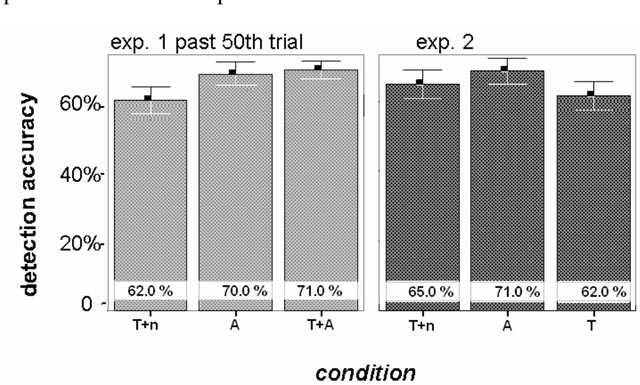

Figure 8. Detection accuracy in the latter part of Experiment 1 vs. Experiment 2. Means (bars) with 95% CI.

A comparison between Experiment 2 and Experiment 1, excluding the 50 first texture changes of both blocks, shows that subjects' detection rate in Experiment 2 was as high or better than at the end of Experiment 1. The detection rates were 70% and 71% for feedback condition A and 62% and 65% for T+n condition in Experiment 1 and Experiment 2, respectively (see Figure 8).

4. SUMMARY AND CONCLUSIONS

In this paper we presented the results of two experiments where perception of dynamic audiotactile feedback was studied with a handheld sensor-actuator device. In the first experiment we assessed the discrimination performance with respect to the modality conditions, texture designs and gesturing behavior. The experiment consisted of two test blocks in which the modality configurations were tactile, audio, or audiotactile. In the second experiment we investigated the role of noise used for masking the leakage sound of the tactile actuator in the first experiment and bring out possible learning effects. There were two test blocks of the following modality conditions: audio, tactile with noise masking, and tactile with acoustic blocking.

Our results allow us to answer the questions we set forth in chapter 2. Considering feedback design, it appeared to be the most difficult to discriminate the slow irregular feedback design from the slow regular one, almost 75% of such texture changes were not detected. The three regular designs with different velocity responses were easier to distinguish from each other, on the average 60% of texture changes between them were detected. So, *temporal density* seems to be a more discriminating design parameter than regularity. The decent detection rates show that the range of the parameter values was set reasonably: the differences between texture designs were neither trivial nor overly difficult to perceive.

The feedback design principles did not have effect on how the subjects were gesturing with the device i.e. energy and periodicity of the movement. This may have been caused by the lack of particular task or target in the experiment or constraints posed by the environment. The subjects were advised just to detect the changes in the feedback independent on the gesturing and therefore preferred to keep the movements as steady as possible.

With respect to modality condition, the audio alone led to best detection performance. Audio and tactile modalities together did not improve the performance, even though tactile feedback alone yielded satisfactory results. This finding indicates that audio dominates perception when audio and tactile modalities contain the same information. Gesturing behavior was also affected by the modality condition. Tactile feedback induced subjects to do more energetic movements while audio feedback led to slower gesturing. This can be due to the dynamic nature of the texture intensities and differences in degrees of amplification of tactile and audio feedbacks. In our feedback designs, the amplitude of the feedback was linearly proportional to the angular velocity, so more rapid motion caused higher amplitude. The tactile feedback alone appeared to yield poorest performance in perception. The subjects may have tried to compensate it by gesturing more rapidly to cause more intense feedback.

Significant learning happened during the experiments. In the first experiment, overall detection accuracies were much better in the second test block than in the first block. This effect persisted over the second experiment which took place 3 months after the first one. However, the differences between the modality conditions remained similar; audio feedback yielded better performance than tactile feedback. Overall, when ignoring the texture type and modality condition, *slower motion* enabled significantly better detection accuracy. The finding can be specific to our feedback design; the mapping between the overall angular velocity and texture density were hand tuned and may have supported less energetic movements. If overall response had been slower (smaller slope *k* in section 2.3.), detection accuracy could have been better with more energetic movements. Another explanation could be the interference of rapid hand motion with the audiotactile perception.

In the second experiment we found that masking noise did not decrease the detection accuracy but might have actually improved it when compared to acoustic blocking. Although the difference was found not to be significant, it proves that our findings in the first experiment are valid. The significant increase in detection time in noise masking condition compared to that of acoustic blocking condition was also an interesting finding. The effects of masking noise follow the other findings of noise effect research; it can either deteriorate or benefit performance, depending on task and type and level of noise [3]. Masking noise obviously has some influence in performance but detailed understanding of it requires more studies. Furthermore, even though our acoustic attenuation used in the Experiment 2 can be considered satisfactory, we cannot be sure that all auditory sensation was blocked. Bone conductivity is another channel for auditory information to travel to ear drum. This cannot be fully eliminated with external blocking.

Our results are conflicting with the findings of Lederman *et al.* [8]. In their study, tactile modality appeared to be most accurate in texture identification task while audio modality was the worst. We suggest following explanations for the conflict: First, in our study, both of the feedback modalities were generated similar manner being proportional to the user's active input while in the study by Lederman *et al*., the audio was generated from experimenter's movements. Second, in our setup the feedback signals were delivered over a large surface area in palm and fingers where as in the study by Lederman *et al*. the tactile cues were delivered to a narrow area, due to the probe's pencil-like grab. Third, in our study the users were not guided for any particular motion path, but allowed to move the device freely. In the test procedure of Lederman *et al.* users were instructed to follow the track of a dot on a computer screen with the probe. And finally, in our setup the feedback signals were artificially generated by tactile and audio actuators while in the other study the audio and tactile cues were induced by the physical interaction with real texture. The users may be more familiar with real textures and thus modality preferences may be different with virtual ones. All of these explanations still lack evidence and raise new research questions for further studies.

However, our prototype and test setup turned out to be useful for both demonstrating the principle of dynamic audiotactile feedback and assessing the active perception. In further studies attention also should be paid on how the perception accuracy changes when the experiments are conducted in non-laboratory environment. Audio and tactile interference as well as attention shift to a secondary task are likely to influence the performance [11].

5. REFERENCES

[1] Ahmaniemi, T., Lantz V., Marila J.,2008. Dynamic Audiotactile Feedback in Gesture Interaction. Proceedings of the 10th International Conference on Human-Computer Interaction with Mobile Devices and Services, Amsterdam, the Netherlands.

[2] Brown, L., Brewster, S., A., Purchase, H., C. 2005. A First Investigation into the Effectiveness of Tactons. Proceedings of WorldHaptics, Pisa, Italy.

[3] Dalton, B., Behm, D., 2007. Effects of noise and music on human and task performance: A systematic review Occupational Ergonomics, 7(3), 143-152.

[4] Goldin-Meadow, S. 2006. The Hand's Role in Talking and Thinking, Handbook pp. 336-369.

[5] Hoggan, E., and Brewster, S. Designing Audio and Tactile Crossmodal Icons for Mobile Devices, 2006. Proceedings of ACM International Conference on Multimodal Interfaces, Nagoya, Japan. ACM Press, 162-169.

[6] http://www.gras.dk/00012/00058/00166/00344/, 13.5.2008.

[7] Jousmäki, V., and Hari, R. 1998, Parchment-skin illusion: sound-biased touch. Current Biology, 1998, 8(6) 190-191.

[8] Lederman, S.J., Klatzky, R.L., Martin, A. & Tong, C. 2003. Relative performance using haptic and/or touch-produced auditory cues in a remote absolute texture identification task. IEEE VR'03, 11th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 151-158.

[9] Mononen, K. 2007. The Effects of Augmented Feedback on Motor Skill Learning in Shooting, A Feedback Training Intervention among Inexperienced Rifle Shooters, Doctoral Thesis, University of Jyväskylä.

[10] Norman, D., A. 1990. The Design of Everyday Things. Currency/Doubleday, New York. 9-12, 87-104.

[11] Oakley, I., Park J., 2008. Did you feel something? Distracter tasks and recognition of vibrotactile cues. Interacting with Computers, 20(2008), 354-363.

[12] Rath, M., and Rohs, M. 2006. Explorations in Sound for Tilting-based Interfaces. Proceedings of the 8th international conference on Multimodal interfaces, Banff, Alberta, Canada.

[13] Salminen, K., Surakka, V., Lylykangas, J., Raisamo, J., Saarinen, R., Raisamo, R., Rantala, J., and Evreinov, G. 2008. Emotional and Behavioral Responses to Haptic Stimulation. Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems. Florence, Italy. ACM Press, NY, USA, 1555-1562.

[14] Toivonen M, Rauno P., Savolainen S., Lehtomäki K. 2002. Noise Attenuation and Proper Insertion of Earplugs into Ear Canals. Ann. Occup. Hyg. 46 (6), 527-530, Oxford University Press.

[15] Williamson, J. and Murray-Smith, 2002. R. Audio feedback for gesture recognition. Technical Report, Department of Computing Science, University of Glasgow, Glasgow.