Audio Engineering Society

Convention Paper

Presented at the 123rd Convention
2007 October 5–8 New York, NY

Music at Your Fingertips: An Electrotactile Fader

Jörn Loviscach[1]

[1]*Hochschule Bremen (University of Applied Sciences), 28199 Bremen, Germany*

Correspondence should be addressed to Jörn Loviscach (`jlovisca@informatik.hs-bremen.de`)

ABSTRACT

Tactile sensations can be invoked by applying short high-voltage low-current electrical pulses to the skin. This phenomenon has been researched into extensively to support visually or hearing impaired persons. However, it can also be applied to operate audio production tools in eyes-free mode and without acoustical interferences. The electrotactile fader presented in this paper is used to indicate markers or to "display" a track's short-time spectrum using five electrodes mounted on the lever. As opposed to mechanical solutions, which may for instance involve the fader's motor, the electrotactile display neither causes acoustic noise nor reduces the fader's input precision due to vibration.

1. INTRODUCTION

Three pathways offer themselves to build a user interface for audio mixing applications: the aural, visual, and tactile sensory channels, see Table 1. A usual mixing desk employs the aural channel to render the result. The faders, knobs, and switches act as haptic input devices; on top of that, they provide both visual and tactile feedback: One can see *and* feel the position. In principle, this remains true for motorized faders, the difference being that they can be used for total recall and time-variant automation.

A requirement of many everyday applications is that their user interface can be operated in eyes-free mode. For instance, a professional piano player doesn't look at his or her fingers; a car driver does not look at his or her feet when stepping on the brake. The same is true in audio production: Many sound engineers try to operate a mixing desk as blindly as possible. To improve on that, the user interface has to employ other channels of perception and action. If one rules out highly disruptive ideas such as foot controls or visual gestures, extending the use of the haptic/tactile channel seems to be the only solution.

Channel	Input	Output
Visual	*Gestures? Eye-tracking?*	See the positions of faders/knobs/switches *2D/3D visualization?*
Aural	*Voice control?*	Audio playback *Sonification?*
Haptic/Tactile	Set faders/knobs/switches	Feel the positions of faders/knobs/switches

Table 1: Three major sensory channels are used to operate a mixing desk. Some uncommon techniques may be used in addition (in italics).

There have been attempts to augment the output capabilities of motor faders through their contact to the user's fingertips: In their Q-Slider, Beamish et al. [1] add haptic feedback of the level. Andersen et al. [2] propose to control the lever's position or the force that it exerts through a sound recording's amplitude envelope.

The objective of this work was to create a tactile "display" for situations specific to audio production. The following two solutions emerged from this:

- To help in blindly setting a fader, markers are indicated at regular intervals along the track. This may be particularly useful to (pre-)set the level for a track which is momentarily silent, so that listening cannot help. The same technique may be employed to set a time position for playback or recording. In this case, the markers could for instance identify beats, measures, or parts of the song.
- To help in distinguishing one track from another, a sketch of the track's short-time spectrum is indicated in real time on its lever. Thus, to identify a track without watching its VU meter, it is no longer necessary to solo it or to raise its level shortly. In particular, in a live setting, it is assuring to know that one's fingers rest on the right faders.

These functions could be implemented through mechanical vibration. However, vibratory units that are small enough to mount several of them in a fader's lever are not easily available. On top of that, their activation would cause acoustic noise, which is objectionable in the sound studio. Actually, the author's first attempt to create a tactile fader was to use a fader's motor to generate vibration. (As in all other experiments in this work, an ALPS RSA0K11V was used.) This experiment showed that a noticeable vibration requires an oscillation frequency of 40 to 80 Hz, a range which results from the interplay of human perception and the fader's mechanics and leads to audible humming. On top of that, a vibration that is clearly noticeable leads to value changes of several tenths of a percent of the fader's readout range. To indicate virtual markers by briefly switching on the fader's motor also turned out to be no good idea because the mechanical shocks prevented a fluent motion of the lever.

These issues can be solved by presenting the tactile feedback not via actual mechanical vibration but through electrical pulses that are applied to the skin. Such electrotactile displays have been employed to support the blind and the deaf. An approach particularly inspiring for the solution proposed in this paper is SmartTouch [3], a system that lets the user feel visual information: The user can freely wipe a flat piece of electronics over an planar underground. A light sensor picks up black-and-white patterns on the underground and maps them to an array of electrodes on the top side of the electronics. A variety of other techniques have been researched into to create tactile displays [4, 5], but none can be implemented as easily as the electrotactile solution.

2. ELECTRONICS

The prototype of this work is created around an Arduino microcontroller that communicates with a host PC via a serial connection. To ensure a swift response and consistent timing, the microcontroller does not generate the pulses itself but only sets their frequency and their strength, see Figure 1. An

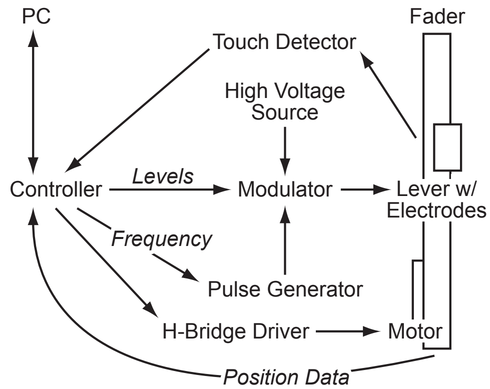

Fig. 1: The prototype is built around a microcontroller that interfaces with a PC on which the audio processing is done.

H-bridge that is controlled by the microprocessor drives the fader's motor. A small, unnoticeable voltage is always applied to the electrodes to sense if a finger is placed on the lever. On a standard motor fader, the change in capacitance would be measured instead, treating the lever as one single electrode. The touch detector allows the user to take over control while recorded fader motions are being played back; this is not used further in this work.

Five independent active electrodes are provided on the fader's lever, see Figure 2. They are distributed over 1.5 cm to cover the full length of a fingertip. The active electrodes are interleaved in a zigzag pattern with electrodes connected to ground. This results in a distance of 3.6 mm between the active electrodes, which is about the spatial resolution of the perception, which will be studied further in Section 3. The electrodes are not flat surfaces but pins to prevent grease or dust from shortcutting them.

The active electrodes are fed with positive voltage pulses. These cause a stronger sensation than negative pulses [6]. The sensation that is elicited is that of a tingling like the colloquial "pins and needles." The perceived vibration of the tingling is virtually independent of the pulse frequency that is applied. The perceived strength of the stimulus does neither depend strongly on the pulse width (1 ms in the prototype) nor on the frequency, with a slight maximum around 40 Hz. Thus, the prototype creates pulse at 40 Hz. The microcontroller can, however, switch the

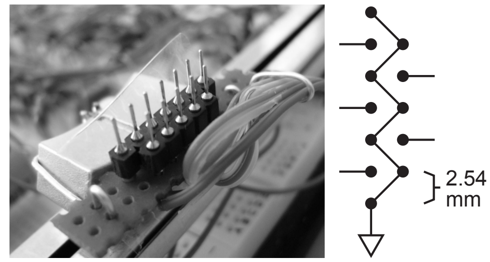

Fig. 2: In the prototype, the electrodes are affixed to a standard lever.

frequency to 120 Hz to improve the rendering of fast changes.

The intensity of the stimulus is controlled independently for each of the active electrodes. The microcontroller sets the current between 0 and 0.5 mA by applying a voltage of up to 270 V. The circuit employs current control, which is known to provide a more consistent sensation [6]. Nonetheless, there is some variation from finger to finger (see Section 3), with changing pressure, and in particular between different people.

To apply high voltages to the human body is potentially harmful. There is not yet much research on short-time or even long-time damages that may arise through the use of electrotactile displays. On top of that, a malfunctioning of the circuit may be dangerous. One basic safety measure is to ensure that the maximum output power of the circuit used to create the high voltage is reasonably low.

3. PERCEPTUAL RESOLUTION

The prototype can convey information to the user in three different ways: by the timing of the pulses, by their position, and by their level. Preliminary tests were done to examine the discriminative power.

To test the spatial resolution, pulse trains of 40 Hz with a duration of 0.5 s are presented on random electrodes at a fixed, comfortable level. For each of these stimuli, the user has to specify whether the last one occurred on an electrode that was either above or below the one before or at the same position. Figure 3 shows typical results that indicate that the just noticeable difference (JND) is on the order of the distance of two electrodes for stimuli of 0.5 s duration;

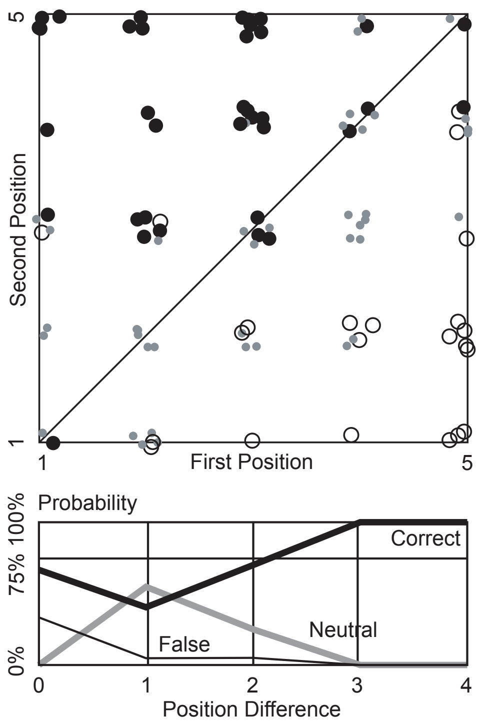

Fig. 3: The user has to specify whether the current stimulus occurred above (filled disk), below (hollow disk) or at the same place (gray disk). (Test data for a user's index finger on the dominant hand; pulse frequency 40 Hz, stimulus duration 0.5 s.)

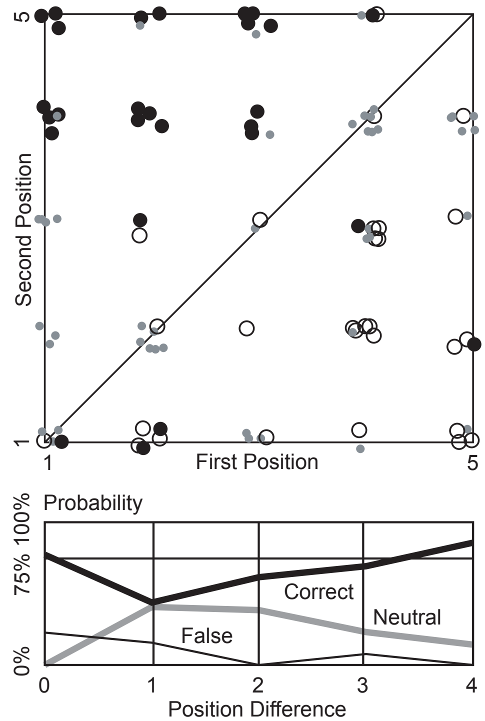

Fig. 4: The spatial discrimination deteriorates for shorter stimuli. (Test data for a user's index finger on the dominant hand; pulse frequency 120 Hz, stimulus duration 0.1 s.)

with stimuli of 0.1 s duration, the sensation begins to get ambiguous, see Figure 4. Thus, the number of five electrodes is on the safe side. The JND result for the spatial resolution is in accordance with the two-point discrimination and line-width discrimination probability found by Kajimoto et al. [3].

Note that the standard approach to determine JNDs is a forced choice between two alternatives [7]. Given the relatively low resolution of the system, this approach seemed to be overly sophisticated, however.

The psychophysical limits of the *spatial* resolution of electrotactile displays have been research into for decades [8]. However, the prototype presented here offers an additional output mode: the strength of the pulses. To learn about this, an approach similar to the one described before is employed: Stimuli of random strength are applied; the user has to specify whether the last one was stronger than, weaker than, or equal to the one before.

The user determines the range of the stimuli: He or she sets a minimum strength that is barely noticeable and a maximum strength at the border of what still feels comfortable, which regrettably but inevitably is a vague notion. For a given current, the voltage may shrink by about 25 % with increasing pressure of the finger against the electrodes. The difference between the minimum and maximum stimulus is rather small in terms of electrical quantities, see Table 2. The JND between stimulus levels suffices to discern about three steps, see Figures 5 and 6.

4. VIRTUAL MARKERS

On guitar amplifiers and studio effect equipment, it is customary to place eleven markers from 0 to 10 along the range of sliders and rotary controls. In its

Finger	Hand	Frequency	Minimum		Maximum	
		Hz	V	mA	V	mA
Index	Dominant	40	94	0.21	102	0.30
Index	Dominant	120	96	0.20	112	0.26
Little	Non-dominant	40	86	0.19	98	0.33

Table 2: The small fingers tend to offer a larger dynamic range (data from a test user).

virtual marker mode, the prototype created for this work indicates eleven markers through electrotactile stimuli. The spatial spread of the electrodes is employed to simulate the motion of a marker below the fingertip, see Figure 7. To also represent the markers well for a quick motion of the lever, the system switches from a pulse frequency of 40 Hz to one of 120 Hz and increases the strength of the stimulus when the lever's velocity surpasses 2 cm/s.

As opposed to a completely mechanical solution, the motion of the marker below the finger does not need to conform geometrically to the motion of the lever. In particular, it may be accelerated to seemingly spread the fader's extent. This idea is related to works that use non-linear distortion of a slider's range to allow the user to set near values with more accuracy than distant values. [9, 10] The prototype software allows to control the fraction $p \in (0, 1]$ of the fader's track that is covered with marker signals, see Figure 7. As eleven markers and thus ten regions are used on the track of 100 mm, the simulated length of touch is $p \cdot 10\,\text{mm}$. In essence, the virtual motion is always accelerated, since the electrodes are actually placed over a length of 15 mm.

Initial tests with a p of 0.25, 0.5, and 0.75 showed that a user can place the lever accurately at a marker by making sure that the middle electrode fires. Independent of p, this task is accomplished with an accuracy of plus or minus one electrode, which conforms to the JND results of Section 3. This accuracy corresponds to a fraction of the track of $p/(5 \cdot 10)$, that is 0.005, 0.01, and 0.015, respectively. Thus, with the setting of $p = 0.25$ one can achieve sub-percent accuracy. This comes at a price, however: For a rapid motion of the lever at for instance 10 cm/s, hardly a single pulse will be generated per electrode, even at the higher firing rate of 120 Hz. This renders the electrotactile stimulus hard to perceive and interpret as a directed motion.

The straightforward solution of this problem is to use $p = 0.5$ instead. To still achieve a high accuracy, however, the prototype's software was changed to no longer evenly assign a length's fraction of $0.5/5 = 0.1$ to every electrode, but to assign a fraction of 0.04 to the middle three electrodes and 0.19 to the two outer ones. This allows a high precision for slow motions and a clear perceptual result for rapid motions.

5. SPECTRUM DISPLAY

The second mode of the prototype concerns the identification of tracks. Here, the stimulus represents the current volume of the track plus its spectral distribution, which is determined in real time and mapped to the five electrodes.

To determine the spectrum, a 1024-sample FFT with a pre-emphasis of 3 dB/octave and 50 % frame overlap is done on the host PC. This is divided into five bands according to transition frequencies set by the user, see Figure 8. The FFT is only used to provide a detailed view of the tracks' spectra, which may not be needed in an actual application of the technique. Then simple IIR filters may be used instead of the FFT to save computational power.

There does not seem to be a meaningful way to map the complete dynamic range of human hearing onto the limited dynamic range of the electrotactile sensation. To compress the signal at the upper end, the cubic root is taken of each power spectrum value before computing the average of the five frequency bands. As the overall level may vary highly between tracks and even from one part to another part of the same track, an automatic gain control is applied to even out medium-term level differences. The result is transmitted to the microcontroller for every frame: 86 times per second for standard CD quality. To allow for quick variations, the rate of the pulses applied to the skin is set to 120 Hz.

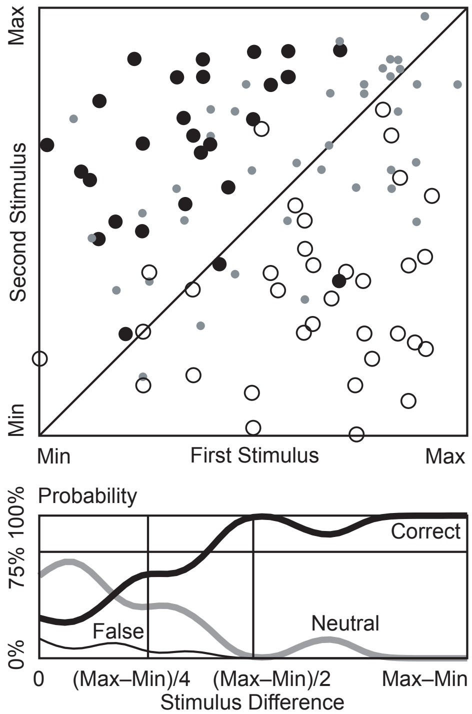

Fig. 5: The user has to specify whether the current stimulus appears stronger (filled disk) or weaker (hollow disk) than the one before or equal to it (gray disk). (Test data for a user's index finger on the dominant hand; pulse frequency 40 Hz, stimulus duration 0.5 s.)

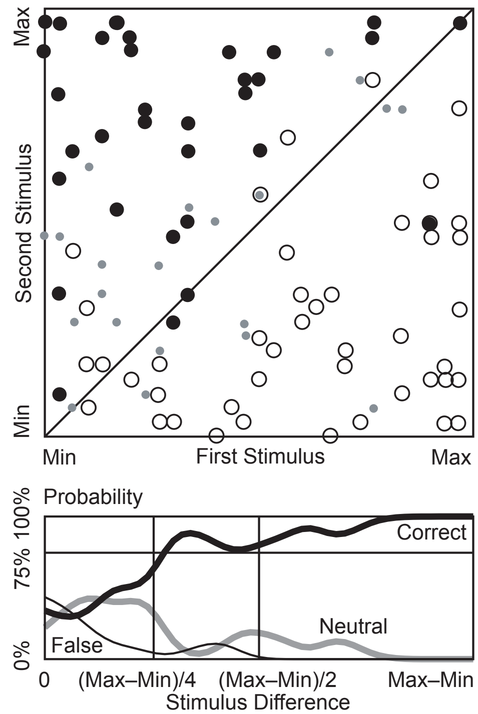

Fig. 6: Not only the voltage that has to applied but also the resolution of the strength of the stimulus changes strongly from finger to finger. (Test data for a user's little finger on the non-dominant hand; pulse frequency 40 Hz, stimulus duration 0.5 s.)

Most tracks of a typical recording session are readily distinguishable on the electrotactile display from their level profile alone. To look into the use of the multi-band/multi-electrode solution, tracks with the same rhythmic profile but different instruments such as drums, electric bass, and lead synthesizer were chosen. To set the transition frequencies in the range of the instruments' fundamental frequencies was vital to separate the tracks in terms of their spectrum, see Figure 8.

Nonetheless, the mapping of the five bands to the electrodes still poses two problems: First, due to the JND of the spatial resolution, there is a huge degree of perceptual cross-talk between the bands, in particular for short stimuli, see Figure 4. Second, due to the random placement of the nerve cells in the skin, the sensitivity varies from electrode to electrode. To overcome these issues, the software prototype offers several strategies from which the user can mix and match, see Figure 8:

- Spectral bands with a strong signal can reduce the power of neighboring bands, like a sharpening filter in image manipulation.
- A strong signal in one band can cause the levels of the neighboring bands to rise, similar to a blur filter.
- Quick changes can be boosted.
- The duration of every peak can be extended by up to 0.5 s.
- Only the electrode with the maximum signal fires.

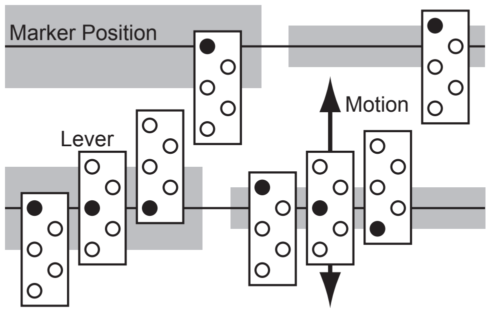

Fig. 7: The electrotactile stimulation does not conform to the geometric position of the virtual markers (left) but is accelerated instead (right). The fraction of the lever's track on which stimuli are generated is shaded.

- Only the electrode with the maximum signal and its top and bottom neighbor fire, all at the same level.

Preliminary user tests showed good perceptual results when the following options were switched on simultaneously: boost quick changes, extend every peak by 0.15 s, and let only the electrode with the maximum signal and its two neighbors fire. With this setting, it is for instance possible to tell apart the bass drum, the snare drum, and the hi-hat on a drum track.

6. CONCLUSION AND OUTLOOK

This work presented a perceptually enhanced interface for audio mixing. Addressing two basic problems in eyes-free interaction with a mixing desk, an electrotactile stimulus is applied either to signal the position and/or motion of a fader's lever or to help the user identify the audio track.

Tests of the prototype support the overall principle but also point out some issues. In particular, the strength and resolution of the sensation varies from finger to finger. As mixing is not a one-finger application but may be done with ten fingers on ten faders, the system needs some intelligence as to which finger is placed where. Heuristic strategies may help here: For instance, if the system finds two fingers that are placed on two faders far apart from each other, then these will almost certainly be the user's two index fingers.

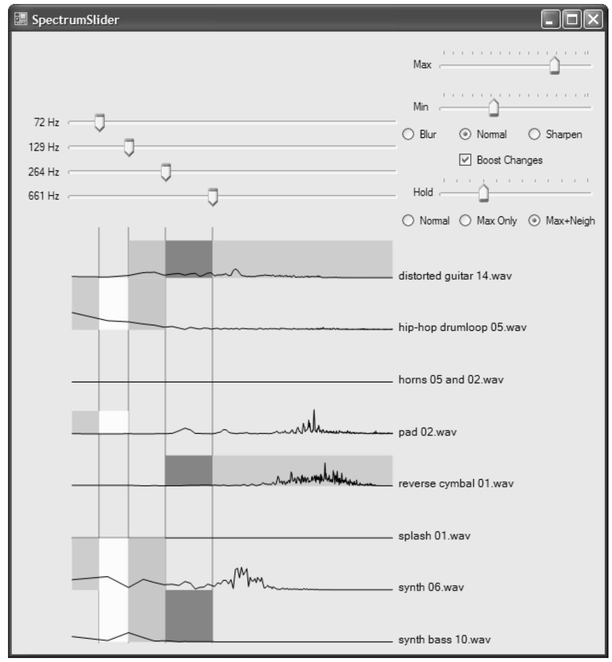

Fig. 8: The prototype's software offers to set the transition frequencies and to tweak the electrotactile rendering. The colored blocks indicate which electrodes are currently active on the different tracks.

To achieve a constant force of the electrodes against the fingertips, one could equip the electrodes with springs, see Figure 9. Future work may also research into how to set the transition frequencies for the electrotactile spectrum display automatically. Such a procedure could maximize the perceptual contrast between the tracks.

7. ACKNOWLEDGMENT

This work was funded by grant 1742B04 of the German Ministry of Education and Research (BMBF). The views and conclusions contained herein are those of the author.

8. REFERENCES

[1] T. Beamish, K. Maclean, S. Fels, "Manipulating Music: Multimodal Interaction for DJs," CHI Letters **6** (2004), no. 1, 327–334

[2] T. H. Andersen, R. Huber, A. Kretz, M. Fjeld, M. "Feel the Beat: Direct Manipulation of

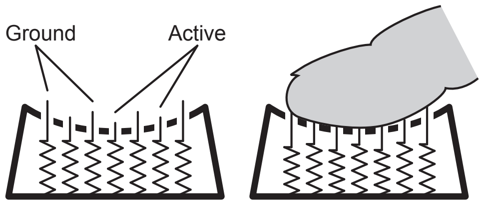

Fig. 9: Springs may be used to ensure a constant force between the electrodes and the skin. The electrodes that are connected to ground may protrude more to be touched first.

Sound during Playback," TableTop 2006: First IEEE International Workshop on Horizontal Interactive Human-Computer Systems. 123–126

[3] H. Kajimoto, N. Kawakami, S. Tachi, "Smart-Touch: Electric Skin to Touch the Untouchable," IEEE Computer Graphics and Applications **24** (2004), no. 1, 36–43

[4] V. G. Chouvardas, A. N. Miliou, M. K. Hatalis, "Tactile Displays: a Short Overview and Recent Developments," ICTA 2005: 5th International Conference on Technology and Automation, 246–251

[5] M. Hafez, "Tactile Interfaces: Technologies, Applications and Challenges," The Visual Computer **23** (2007), no. 4, 267–272

[6] K. A. Kaczmarek, M. E. Tyler, P. Bach-y-Rita, "Electrotactile Haptic Display on the Fingertips: Preliminary Results," 16th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (1994), 940–941

[7] R. Ulrich, J. Miller, "Threshold Estimation in Two-Alternative Forced-Choice (2AFC) Tasks: The Spearman-Kärber Method," Perception & Psychophysics **66** (2004), no. 3, 517–533

[8] O. E. Solomonow, L. Raplee, J. Lyman, "Electrotactile Two Point Discrimination as a Function of Frequency, Pulse Width and Pulse Time Delay," Annals of Biomedial Engineering **5** (1978), 117–125

[9] Y. Koike, A. Sugiura, Y. Koseki, "TimeSlider: An Interface to Specify Time Point," UIST 1997: 10th Annual ACM Symposium on User Interface Software and Technology, 43–44

[10] W. Hürst, G. Götz, P. Jarvers, "Advanced User Interfaces for Dynamic Video Browsing," MM 2004: 12th Annual ACM International Conference on Multimedia, 742–743