SonicTexting

Michal Rinott
Interaction Design Institute Ivrea
Via Montenavale 1, Ivrea, Italy
m.rinott@interaction-ivrea.it

ABSTRACT
SonicTexting is a system for inputting text – 'texting' – using gestures and sound. As in musical instruments and everyday mechanical objects, sound in SonicTexting is synchronous and responsive to actions. SonicTexting explores people's hand-ear coordination and demonstrates the use of informative digital sound. It proposes that through touch and sound, a functional activity like text entry can become an experience on the borders between performing a task, playing an instrument and playing a game.

Author Keywords
Auditory interaction, sound design, input devices, interactive experiences.

ACM Classification Keywords
H.5.2 [Information Interfaces and Presentation]: User Interfaces – Auditory (non-speech) feedback, Input devices and strategies.

INTRODUCTION
SonicTexting is an audio-tactile system for inputting text using continuous thumb gestures and synchronous digital sound. Sound provides feedback for the gestures, aiding orientation and navigation.

SonicTexting is related to existing solutions for gesture-based text entry on mobile devices. 'Quikwriting' is a stylus-based system with continuous gestures, in which the stylus is never lifted from the surface [3]. The gestural vocabulary of SonicTexting is based on this Quikwriting model. In the 'T-cube' text entry system auditory feedback is provided for stylus writing gestures [4].

SonicTexting is a predominantly auditory interface with which expert users can input text without the use of vision. It introduces a one-handed input device for performing gestures. Rather than maximizing word-per-minute efficiency, the emphasis in designing SonicTexting was placed on creating an engaging audio-tactile experience.

CHI 2005, April 2–7, 2005, Portland, Oregon, Usa.
ACM 1-59593-002-7/05/0004.

MOTIVATIONS

Audio-tactile interaction
Our interactions with physical objects in the world usually produce sound; we use this sound to gather information and adjust our actions [1]. Pressing the car gas pedal, shaking a bottle and tuning an analog radio are common examples. This tight connection between actions and resulting sound is often lost in digital products, in which the introduced sounds are more binary, on/off phenomena.

In additional to its informational value, the audio-tactile connection can be a source of pleasure; the addictive qualities of puncturing 'bubble-wrap' are probably an example. SonicTexting is an attempt to tap into the sources of audio-tactile gratification.

Texting
Texting on the mobile phone interface is an extremely widespread activity. There is actually a world championship in this skill! Teenagers especially excel at it: sociologist Sadie Plant finds that texting has had a profound effect on the way they use their thumbs. Some now point and even ring doorbells with their thumb instead of their forefinger.

SonicTexting is an interface for texting designed to require skill and to encourage mastery through an interaction that is challenging and playful. The thumb is used not for pressing but for performing continuous, rounded gestures.

THE SONIC TEXTING PROTOTYPE

Input Device
The Keybong is a one-handed input device that fits in the palm of the hand and is manipulated with the thumb (see Figure 1). The Keybong's joystick naturally supports the common gesture pattern of SonicTexting: moving from a central location, through a specific path, back to the center. The joystick is also a button: pressing it down clears the entered text. The Keybong contains a small eccentric motor that gives gentle vibration feedback in the writing process. This tactile layer accompanies and augments the sound layer. The Keybong form is designed to fit in the hand and to ensure that it is held in a fixed orientation.

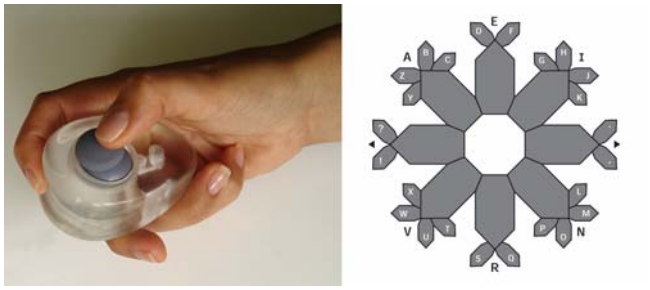

Figure 1. The Keybong and Gesture Map

Sound

In SonicTexting letters are located using sound. Sound provides continuous information during the navigation movement: an interactive sonification of the gesture. There are two main sound modes: Beginner Mode, which optimizes learning and memorization, and Expert Mode, which optimizes speed. Sound is also used after the gesture for a letter-by-letter 'readback' of entered words.

Beginner mode – In this mode continuous, looping letter phonemes (e.g., the sound "bh") are played in synchrony to movement. Loudness is a function of the distance between the controller and the eight equidistant points distributed on the periphery of the controller range. When the controller is located between two points, two phoneme sounds mix. When the controller is located close to a point, one phoneme sound is pure. Letter phonemes are spoken in different pitches according to their position: pitch ascends clockwise, completing one octave around the circle.

Expert Mode – In this mode discrete, synthetic percussive sounds are played in synchrony to movement. Sounds are played only when the controller reaches a perimeter location. As in Beginner Mode, each of the eight perimeter locations has a pitch. The velocity of the movement is mapped to the loudness of the initial part of the sound (attack), so that faster movements produce stronger sounds. Beginner and Expert sounds are optimized for different movement patterns: Beginner Mode is designed for slower movement and listening; Expert Mode for rapid movement.

Gesture Mapping

Two main letter mappings were considered: alphabetical and frequency-ordered (in which frequent characters require shorter gestures). The alphabetical mapping was chosen due to: 1) a prioritization of memorability over gesture length, given to the ease of performing all gestures, 2) the good fit between alphabetical order and the vocal rising pitch.

The mapping is presented and learned via the gesture map – a static visual representation of the gesture paths. The fractal-inspired design of the map, created by duplicating and rotating a basic graphic element, communicates the 'nested' nature of the gesture model.

The gesture map is read as follows: to write an 'axis' letter (i.e. the letters on the main axes, in black in Figure 1) the joystick is moved to that location and back to the center. Upon return to the center the letter is written. To write a nested letter (all other letters) the joystick is initially moved to the axis location, then around to one or two positions according to the letter location relative to the axis letter, and returned to the center.

Implementation

A BX24 microcontroller is used for two-way serial communication between the Keybong and Max/MSP software [2]. Normalized x, y coordinates and button data is sent from the Keybong to Max, and on/off commands are sent from Max to the Keybong motor. For every position received, Max calculates its distance from eight locations on the perimeter of the controller range. These distances, along with the accumulated path, determine the sounds played and their relative loudness.

EXHIBITION SETTING

The SonicTexting prototype has been tried out extensively in a design exhibition setting. The majority of visitors, both adults and children, reacted enthusiastically to the experience and were motivated to learn to 'SonicText'. Visitors' interpretation of SonicTexting varied: some saw it as a game, others as a kind of musical instrument and others as desirable mobile device feature.

Visitors tended to expect visual feedback on the gesture map. Instructing them to move in the direction of the letters using sound, and to seek the pure sound, helped increase their dependence on the audio feedback. Most visitors could successfully use SonicTexting after a few minutes of practice, a much shorter time than had been anticipated.

CONCLUSION

Future work on SonicTexting should include fine-tuning the design and conducting formal testing, also with long-term users and in a real world texting context.

SonicTexting explores how sound can be connected to actions in a sophisticated manner in a digital interaction, and how this connection can be a source of pleasure. These insights have wider implications for the use of sound and touch in interaction design.

ACKNOWLEDGMENTS

I thank Michael Kieslinger, Jan-Christoph Zoels, Edoardo Brambilla and Christian Palino for their advice and help.

REFERENCES

1. Gaver, W. "Auditory Interfaces". In: M. Helander. T Landauer, & P. Prabhu (Eds.), *Handbook of Human-Computer Interaction*, 2nd ed. (1997), pp. 1003–1042.
2. Max/Msp software, <http://www.cycling74.com>
3. Perlin, K. The Quikwriting Shorthand System, <http://mrl.nyu.edu/projects/quikwriting>
4. Venolia, D., Neiberg, F. "T-Cube: A Fast, Self-Disclosing Pen-Based Alphabet". *Proceedings of CHI '94* (Boston, April 24-28, 1994), ACM, pp. 265–270.