“Silooets” - Audiotactile Vision-Substitution Software

David Dewhurst
HFVE.org

daviddewhurst@HFVE.org

ABSTRACT
This poster describes "work-in-progress" on “Silooets” (Sensory Image Layouts and Object Outlines Exhibited via Touch and Sound) - experimental vision-substitution software that uses audiotactile methods to present certain features of visual images to blind people.

Keywords
Blindness, deafblindness, Silooets, sensory-substitution, vision-substitution, audiotactile, haptic, Morse code, braille.

INTRODUCTION
At the 1st HAID Workshop in 2006, the “HiFiVE” (Heard & Felt Vision Effects) vision-substitution system was presented at a poster session [1]. The system has been further developed since then.

The HiFiVE system highlights features of visual images that are normally perceived categorically, and substitutes with coded sound effects and their tactile equivalents. By smoothly changing the pitch and binaural positioning of the sounds, they can be made to appear to "move", whether following a systematic path or describing a specific shape. Such moving effects are referred to as “tracers”, and can be "shape-tracers", whose paths convey the shapes of items in an image; or "area-tracers", which systematically present the properties of parts of an image.

In the tactile modality, tracer location and movement are presented via force-feedback devices; and categorically-perceived features via braille, or via Morse code-like “tapping” effects.

Other work in the field includes tone-sound scanning methods that have been devised for presenting text [4], and for general images [5]; and software for presenting audiotactile descriptions of pixels in computer images [6]. Audio description is used to supplement television etc. (The merits of other approaches are not discussed in this poster.)

Potential applications
This project is not focussed on a specific application, but is trying various methods for presenting sequences of visual images via touch and sound. More straightforward material (such as simple shapes and diagrams) can also be presented.

Possible applications include:- presenting shapes and lines for instructional purposes; adding shape, colour and texture data to diagrams; providing ad-hoc information to users wishing to know the colour and shape of an item; and for specific tasks such as seeking distinctively-coloured items (for which corresponding sets of parameters are selected).

FEATURES OF THE SYSTEM
(This section recaps some of the previously-published features of the system [2,3], and describes enhancements.)

The Silooets software can present images as arrangements of pixels known as “layouts” Figure 1; and as the outlines of the objects in images Figure 2. In both cases the categorical properties of the areas are exhibited via groups of coded CV (Consonant-Vowel) speech sounds; or as braille dots; or via Morse code-like “tapping” effects.

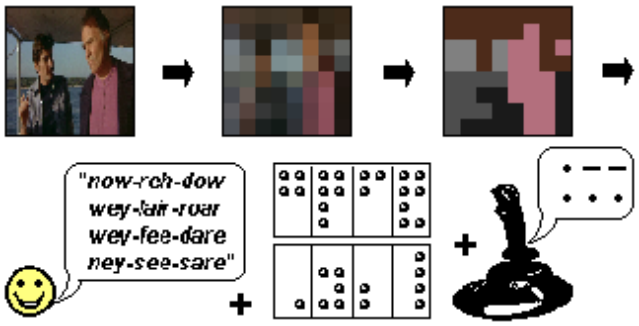

Figure 1. Diagram illustrating the conversion of the “layout” of an image into coded phonetics, braille, and “tap codes”.

Figure 1 shows an example of an image being reduced to two colour shades in each image quarter / “panel”, and some of the corresponding speech sounds, braille cells and tapping codes, which describe the 8 by 8 pixels shown. “Layouts” can be used to present any section of an image.

Alternatively, if the shapes of entities in an image can be determined (Figure 2), then audiotactile “shape-tracers” can exhibit those shapes and their corners, at the same time as the categorical properties are being presented.

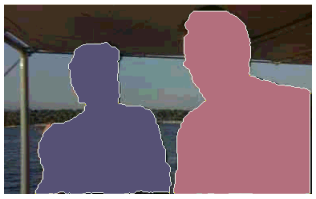

Figure 2. Identified entity shapes, whose outlines and corners can be presented via audiotactile “shape tracers”.

Colour mapping

The system normally uses two colour shades (e.g. "blue and red") when presenting an area or entity, allowing the shades to be given via a modest number of corresponding effects. The two-colour-shade approach can simply present the two associated shades, or have the effect of painting a picture with one colour on a differently-coloured background.

A 15-colour-palette format allows two colours to be easily presented via a single "CV" syllable; via a single 8-dot braille cell; or via "tap codes" of 8 or less pulses.

Speech mappings are available for English speakers using the full range of English phonemes; or an "International" format can be used, which uses the sounds found in most languages to present combinations of 24 shades (Table 1).

Vowel sound 2nd ↓	Consonant sound 1st →				
	S (w)	**R** (L)	**K** (g)	**N** (m)	**D** (b,p,t)
I	Lt.Purple	Lt.Brown	White	Cream/24	(Special)
E	Pink	Yellow	Light Grey	Lt.Green	Light Blue
A	Red	Orange	Mid Grey	GrnYell/23	Turquoise
O	Purple	Brown	Dark Grey	Green	Blue
U	Dk.Purple	Dk.Brown	Black	Dk.Green	Dark Blue

Table 1. An example of colour-to-speech mapping.

Tactile effects and user interaction

Programmable 8-dot braille cells are available commercially, and are an effective way of presenting categorical data to blind people who are able to read braille.

Alternatively, coded pulses can be induced on a standard force-feedback device, presenting "tap codes" to the user. It is found that short, evenly-spaced pulses of two force levels are more straightforward to interpret than conventional Morse-code timings. "Tap codes" are relatively slow when compared to speech or braille, but may be useful for deafblind users who cannot read braille.

A force-feedback joystick makes an effective control and pointing device, as it can also be moved by the system, pushing and pulling the user's hand and arm, tracing out any shapes (and highlighting corners) that are to be presented.

Corners

Corners (such as the vertices of an octagon) produce a considerable effect in giving the impression of a shape to sighted people, and the system emphasises them by introducing effects within shape tracers at the appropriate point in time. For example, the system can momentarily stop the movement of a "shape-tracer" (by stopping a moving force-feedback joystick); and in the audio modality the system can alter the volume of the sound.

RECENT DEVELOPMENTS

At the 1st HAID Workshop, the system was shown presenting standard demonstration shapes with corners; and speech-like sound representing the "layouts" of live images.

Since then, the following features have been progressed:-

Prepared media

Prepared media can be presented, with predetermined shapes etc. embedded in common image formats (e.g. bitmap, GIF or JPEG still images; or DVDs or AVI movie files), produced via a straightforward procedure. (Such media can be also be viewed on standard media players.)

Enhanced layouts and colour mappings

The Silooets software is designed so that "layout" configurations and colour shade mappings are flexible. New layout configurations include ones with greater resolution at the centre of the layout. Standard colour mappings have been simplified and enhanced.

Morse code-like tapping output

The tap-codes described above have been implemented, and can exhibit certain properties. Various timings and force levels have been tested, and are available as options.

Enhanced tactile effects

The system can now exhibit a notchy "grid" effect when a force-feedback joystick is used, so giving the user a tactile indication of where the joystick is currently located.

Two joysticks can be used. For example, the main joystick can be used as a pointer by the user, and by the system to indicate the location and size of the area being presented. The other device, for example a force-feedback mouse, can be used by the system to present any shapes and tap codes.

Activity-related processing

Users can rapidly switch between sets of parameters and options by selecting identifiers for particular activities.

SUMMARY

When completed, the Silooets software will allow visual features of images, ranging from basic properties, to object descriptions (if known), to be presented to blind people. Several features of the system can now be demonstrated.

REFERENCES

1. David Dewhurst, "An Audiotactile Vision-Substitution System" in *Proc. of First International Workshop on Haptic and Audio Interaction Design*, Vol.2 (2006) 17-20.
2. See website http://interactive-sonification.org /ISon2007/ISon2007Proceedings /papers/Dewhurst_ISon2007.pdf.
3. U.S. Patent Appl. No. US 2008/0058894 A1
4. E. E. Fournier d'Albe, "On a Type-Reading Optophone" in *Proc. of the Royal Society of London. Series A*, Vol. 90, No. 619 (Jul. 1, 1914), pp. 373-375.
5. See website http://www.seeingwithsound.com.
6. See website http://www.ifeelpixel.com.