



2003-06-01

An investigation into non-verbal sound-based modes of human-to-computer communication with rehabilitation applications

Ted Burke

Dublin Institute of Technology, ted.burke@dit.ie

Follow this and additional works at: <http://arrow.dit.ie/teapotcon>

 Part of the [Biomedical Commons](#), and the [Computer Engineering Commons](#)

Recommended Citation

Burke, Ted: An investigation into non-verbal sound-based modes of human-to-computer communication with rehabilitation applications. Adjunct Proceedings of HCI International, 2003, Crete, Greece, 22-27 June, 2003, pp. 241-242.

This Article is brought to you for free and open access by the tPOT: People Oriented Technology at ARROW@DIT. It has been accepted for inclusion in Conference Papers by an authorized administrator of ARROW@DIT.

For more information, please contact yvonne.desmond@dit.ie, arrow.admin@dit.ie, brian.widdis@dit.ie.



This work is licensed under a [Creative Commons Attribution-NonCommercial-Share Alike 3.0 License](#)



An Investigation into Non-Verbal Sound-Based Modes of Human-to-Computer Communication with Rehabilitation Applications.

Edward Burke, Yvonne Nolan and Annraoi de Paor

Department of Electronic and Electrical Engineering,
National University of Ireland, Dublin,
Belfield, Dublin 4, Ireland.

and

National Rehabilitation Hospital,
Rochestown Avenue, Dun Laoghaire, Co. Dublin, Ireland.
tedburke@iol.ie, ynolan@ee.ucd.ie, annraoi.depaor@ucd.ie

Abstract

During the course of our work in the National Rehabilitation Hospital in Ireland we have encountered patients who, while unable to speak, are capable of making reproducible utterances. We present techniques used to harness such utterances, in addition to whistling, as a means of communication and control. A simple technique for identifying the phonemes /o/ and /s/ (in single-symbol ARPAbet notation) is presented with applications. The use of pitch variation as a means of controlling a continuously variable parameter is described with two applications - a microcontroller based light dimmer switch and a computer program which facilitates mouse pointer control. Finally, a technique for the recognition of short note sequences is presented. A program is described which allows arbitrary commands to be executed in response to tunes either sung or whistled by the user. These commands may be used to switch on or off electrical appliances in the home.

1 Simple Phoneme Recognition

Phonemic, non-verbal input has previously been used as means of communication in rehabilitation (Igarashi and Hughes, 2001). Here, phoneme recognition based on the timing of positive-going zero-crossings (PGZCs) in the audio signal is used. The criterion for recognition of the phoneme, /o/, requires the time interval between successive PGZCs to remain roughly constant, and to be greater than a threshold value. The criterion for recognition of the phoneme, /s/, requires that the average rate of PGZCs be greater than a threshold value, typically set at or above 2000 PGZC/s.

Using this technique, a microcontroller-based phoneme recognition device was developed to control any device operable by two switches, and in particular, a reading machine designed for users with severe physical disabilities. The signal from a microphone is amplified and infinitely clipped. Two PIC16F84 microcontrollers - one for each phoneme - take this rectangular wave as input. Each actuates a relay switch at its output when the appropriate phoneme is detected.

A C++ class, called *audio_widget*, was also developed implementing this phoneme recognition technique. The *audio_widget* facilitates the integration of phoneme recognition into many programs, including one which provides a graphical menu for a patient, using icons drawn by a

therapist, each of which may be associated with an arbitrary command. The user may scroll through menu items by making an /s/ sound and may select a menu item by making an /o/ sound.

2 Continuous Pitch Control

A microcontroller-based light dimmer switch was developed which is controlled using whistles (or vocalisations) of varying pitch. The instantaneous pitch of the controlling sound is calculated from the time which elapsed between the two most recent PGZCs of the signal. The controlling sound is required to have only a single PGZC per pitch period, a criterion typically satisfied by both whistling and the phoneme /o/. While the instantaneous pitch is varying only gradually (e.g. during a whistle), these changes in pitch are mirrored by changes in the light intensity.

The *audio_widget* features a pitch-tracking mode, which uses a power spectrum based pitch estimation algorithm called Harmonic Series Identification (Burke, 2002). Using this, a computer program for controlling movement of the mouse pointer and simulating clicks was developed. Horizontal and vertical mouse movement are controlled alternately using whistles (or vocalisations) of varying pitch in a manner similar to that used in the pitch controlled dimmer switch described above. Any whistle shorter than a threshold duration is interpreted as a click.

3 Tune Based Control

Another computer program incorporating the *audio_widget* allows a number of user-defined commands to be triggered by whistling or singing the appropriate tune. A subsidiary application, called *X10action*, facilitates the control of household appliances. The initial pitch of each new note is appended to an array of recent note pitches maintained by the program. If a specified timeout period elapses after the end of one note, without a new note beginning, a (non-unique) tune signature is generated from the stored sequence of note pitches, before clearing the array. If the signature matches that of one of the user-defined tunes, then the corresponding command is executed. Each signature is a sequence of binary digits, each representing the change of pitch in a pair of successive notes: '1' represents an increase in pitch. A '0' represents a decrease in pitch. Note transitions involving no significant change in pitch are ignored. This method of encoding note sequences is similar to the Parsons code used by Prechelt and Typke (2001).

4 Conclusion

As shown here, non-verbal sound-based modes of human-to-computer communication have several useful applications in rehabilitation, in particular where other non-contact modalities, such as speech recognition, may be unsuitable, impractical or uneconomical.

References

- Burke, E. (2002). The Harmonic Series Identification Method of Pitch Estimation. In *Rehabilitation Applications of Musical Human-Computer Interfaces* (pp. 84-91). MEngSc Thesis, National University of Ireland, Dublin.
- Igarashi, T. and Hughes, J. (2001). Voice as Sound: Using Non-verbal Voice Input for Interactive Control. *14th Annual Symposium on User Interface Software and Technology, ACM UIST'01*, Orlando, FL.
- Prechelt, L. and Typke, R. (2001). An Interface for Melody Input. *ACM Transactions on Computer-Human Interaction, Vol. 8, No. 2*, 133-149.