Acoustic Rendering as Support for Sustained Attention during Biomedical Procedures

Emil Jovanov, Dusan Starcevic

University of Belgrade
Belgrade, Yugoslavia

Kristen Wegner, Daniel Karron

Computer Aided Surgery, Inc.
New York, U.S.A.

Vlada Radivojevic

Institute of Mental Health
Belgrade, Yugoslavia

Abstract

Biomedical procedures of long duration cause mental fatigue and attention deficit. We investigated using sound as a means to support sustained attention during prolonged procedures and analysis. In this paper we present tactical audio as support for precise manual positioning of a surgical instrument, and introduce acoustic rendering as an additional information channel and/or warning signal in EEG analysis.

1 Introduction

The increased performance of present computer systems and commercially available high performance human-computer interfaces (HCI) have made possible the perceptualization of large data sets[1-3]. Contrary to previous generation of computer systems, the main limiting factor is often the characteristics of human perception. In addition to visual presentation, acoustic and other presentation modalities have come to be used increasingly to improve insight into complex biomedical phenomena and to decrease cognitive workload [4, 5].

Prolonged procedures and analysis induce fatigue and decrease attention. This disorder could be manifested as:

- Mental fatigue,
- Sensory-motor habituation,
- Distorted perception of time flow.

We investigate the possibility for improving attention during prolonged procedures and analysis through the use of multi-modal human-computer interfaces to create integrated and intuitive interface. Combined sensory workload allows optimal human resource utilization, and as a result we expect to have sustained attention and better performance. Typical examples are visuo-motor coordination during surgery and evaluation of long biomedical recordings (EEG, ECG, etc.).

We outline methods below to achieve sustained attention using acoustic rendering. The second section presents tactical audio as support for precise manual positioning of a surgical instrument. In the following section we introduce acoustic rendering in EEG analysis.

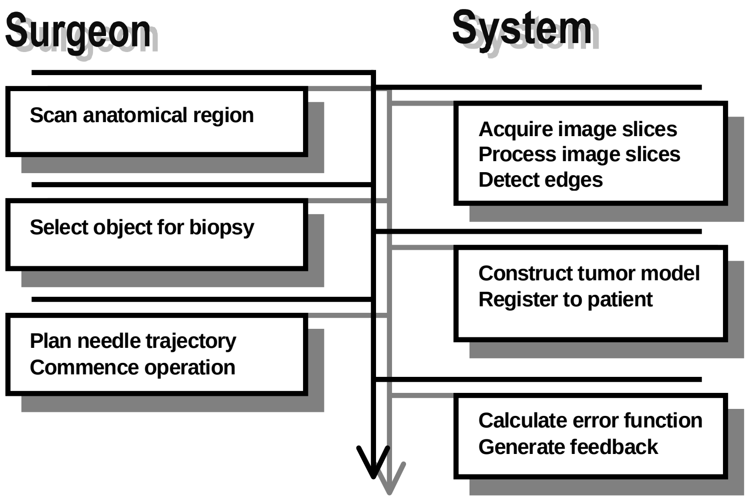
2 Tactical Audio

Tactical audio concerns the use of audio feedback for facilitating the precise and accurate positioning of an object with respect to another object. This has valuable application in the field of surgery. In the course of the typical diagnostic surgical procedure there are numerous needle placement errors, especially with regard to insertion depth, e.g., missing the tumor in a biopsy. Although ultrasound and other imaging modalities attempt to alleviate this problem, the nature and configuration of the equipment requires the surgeon to take his/her eyes off the patient. The use of tactical audio feedback enables the surgeon to effect a precise placement by enhancing his/her comprehension of the three-dimensional position of a surgical implement with respect to some predetermined desired position within the patient’s body. This system consists of sensors and software tools which allow the surgeon to acquire a target, and then navigate correctly in order to intercept that target. Navigation is facilitated by a novel integrated audio-visual feedback system. In addition to enabling more precise placement of surgical instruments, and therefore improved surgical outcomes, we believe this approach will yield substantial savings in time and cost. This will clearly benefit both patient and healthcare provider.

Current technology for image-guided needle insertion is severely limited in terms of its ease of manipulation for the surgeon, flexibility of application for many different procedures, and ability to resolve and intercept targets less than half a centimeter in diameter. Commercially available technology for this class of procedure either consist of many loose, non-integrated tools - which, though flexible in application for many different procedures, are difficult to manipulate and require an assistant - or consist of a single monolithic system which is useful for only a single kind of procedure. In both of these cases there are many tasks the operator must perform which easily could be automated. Furthermore, available systems have a poor facility for preoperative procedure planning. The operator either makes an educated guess about where the target is, and how the needle should be inserted (the insertion trajectory) based upon what he/she knows about human anatomical structure, and by looking frequently at an ultrasound monitor, or plans the procedure using stereotactic techniques. Stereotaxis is a primitive technique which is fundamentally limited by the fact that it uses static images as the basis of the plan because it is unable to take into account deformation of the patient's tissue as the needle is inserted, and is completely thrown off if the patient moves. All of these factors have significant effects upon the speed of execution of the procedure, and most importantly, the quality of care.

The function of this system, at the most fundamental level, is to facilitate ultrasound-guided biopsy procedures by providing real-time navigational guidance. The key element in enabling this guidance is the use of audio feedback.

This system may be most easily understood in terms of a division of tasks: those executed by the human operator, and those by the computer. The taskflow diagram in Fig 1 depicts the different tasks the system requires for successful operation, and by which entity (human or computer) they are performed.



Procedure Successfully Accomplished

Figure 1. System taskflow diagram.

The system requires four consecutive stages of operator input. The first three represent a non-time-critical planning session. The fourth represents the actual surgical execution of the plan, (which is time-critical).

In clinical practice this task-flow would take the following form: The surgeon performs an ultrasound pass on the region of interest in order to acquire images for planning the procedure. The system captures these images and processes them so that suspicious objects are rendered highly visible. The surgeon selects one of these objects as the candidate for biopsy.

The system reconstructs a three-dimensional model of the object using the ultrasound images. The system constructs a three-dimensional representation of the patient's skin surface with respect to the tumor, and registers the entire 3D anatomical model to the patient's actual anatomical position and orientation, which, along with the ultrasound transducer, and the biopsy needle, is being continuously and precisely tracked in three-dimensional space throughout the procedure. The surgeon plans a biopsy needle trajectory that will intercept the tumor using this anatomical scene model. While the surgeon is performing the procedure the system provides integrated graphical and audio feedback which allows him/her to reliably, precisely and accurately track the preplanned biopsy needle trajectory and intercept the tumor.

We are implementing tactical audio as an extension to the Sofamor Danek Group's Stealth Station system for frameless stereotactical neurosurgery. The Stealth Station provides a means for planning a surgical insertion path relative to a patient's anatomy. The system is then used in the operating room to track the position and orientation of the surgical instrument relative to the patient and the pre-planned insertion path and to provide real-time navigational guidance to the surgeon. The current version of the Stealth Station provides a three-dimensional graphical rendering of the operative scene - e.g. patient position and orientation, the pre-planned insertion path, and the instrument position and orientation. While the Stealth Station has gone a long way to improve the practice of stereotactical neurosurgery, the fact that the surgeon is obliged to repeatedly take his/her eyes away from the patient in order to keep track of the instrument placement evidences an error in the user-interface design.

Our proposed solution to this user-interface design problem is to provide audio feedback which will allow the surgeon to be continuously aware of the placement of the surgical instrument relative to the pre-planned trajectory, without ever needing to take his/her eyes off the patient. We plan to accomplish this by calculating and then sonifying the error vector of the current position of the instrument relative to the pre-planned insertion path. While we are still considering a number of approaches to the actual sonification of this error vector, our broad approach will be to employ some form of polyphonic consonance/dissonance function which will indicate the relative degree to which the instrument placement is in error. We are however of the philosophy that the most appropriate sonification method may only be determined through extensive usability testing. We plan to pursue this in the near future at Saint Louis University Medical Center.

3 Long EEG record analysis

In conventional clinical settings, analog EEG recording at 3 cm/s for a 24-hour period would require 2.6 km of paper [6]. Although a trained neurophisiologist can rapidly scan through long EEG recordings, two issues are critical: prolonged inspection induces mental fatigue, and some clinically important features are difficult to discern from simple visual inspection. Automatic feature extraction and at least warning for possibly significant sections are very important issues in everyday clinical practice and research.

Our multimodal interactive environment for biomedical data presentation uses VRML-based visualization and sonification. The Virtual Reality Modeling Language (VRML) is a file format for describing interactive 3D objects and worlds, applicable on the Internet, intranets, and local client systems [7]. VRML is capable of representing static and animated dynamic 3D and multimedia objects with hyperlinks to other media such as text, sounds, movies, and images. VRML browsers, as well as authoring tools for the creation of VRML files, are widely available for many different platforms. In our system the VRML world is controlled by Java applets.

Limited resources of previous generation information systems established the concept of optimal resource utilization, which implies non-redundancy. As a consequence, conventional applications still rely on this principle of using minimal resources to mediate the information. As a result of this poverty of resources, the presentation modality of interfaces was mostly uni-modal. Simultaneous presentation of the same information in different modalities was seen as a loss of resources. In contrast, our natural perception is based on redundancy. As an example, using mouse as pointing device we are not conscious of the additional sensory modalities which we use as feedback, such as cursor movement, perceived hand position, and the sound of the mouse friction upon the mouse pad, and the click of the mouse button.

Redundancy in human-computer interfaces should be accomplished using multi-modal presentation. The central concern in the design of a multi-modal representation is to determine an appropriate level of redundancy. A low level of redundancy increases the user's cognitive workload, while a high level of redundancy irritates the user. A user-specific metric of the appropriate degree of multi modal redundancy for a given application is necessary.

There exist two principle forms of multi-modal data representation. The simplest one is to signal state transitions or indicate certain states. This form is often used in implementing sound alarms. The second form is the presentation of current values as a data stream. Additional modes of presentation may be employed either as redundant modes of representation emphasizing certain data features or by introducing new data channels. Redundant presentation induces an artificial synesthetic perception of the observed phenomena [8]. Artificial *synesthesia* (*syn* = together, and *aisthesis* = perception in Greek) generates a sensory joining in which the real information of one sense is accompanied by a perception in another sense. Multi-sensory perception in this manner can improve understanding of complex phenomena by giving different cues or triggering different associations. In addition, the use of such an acoustic channel can permit the use of new information channels without information overload.

We implemented an environment for monitoring brain electrical activity. This environment consists of a 3D visualization system synchronized with a data sonification system driven by EEG data. The visualization is based on the use of topographic maps which are projected on the scalp of a 3D head model. The sonification system modulates natural sound patterns to reflect certain features of the processed EEG data, which creates a pleasant acoustic environment. This feature is particularly important for prolonged system use.

3.1 Hemisphere Activity Spatialization

The complicated interdependence of EEG channels makes it hard to perceive global patterns in the data as they evolve over time during visual analysis of EEG topographic maps. Moreover, a small visual memory capacity makes it difficult to remember patterns which have passed by rapidly. We have extended the analysis of animated topographic maps of brain electrical activity through sonifying derived parameters of global brain activity. This sonification system additionally provided 3D spatialisation of sound. Changes in the sound location correlated to EEG changes. This technique provided additional information to the examiner and an aid for localizing his attention. Although sound provides limited spatial distribution, it is more appropriate for attention focusing and localization [9], which is particularly important for sound alarms.

We applied sonification to left/right brain hemisphere EEG power symmetry. We sonified the index of symmetry (*IS*) which is calculated as:

$$IS = (P_1 - P_2) / (P_1 + P_2)$$

where P_1 and P_2 represent power over the left and right hemispheres or a pair of symmetrical EEG channels, like *O1* and *O2* for example.

The index of symmetry is sonified as the position of a sound source in space. This audio cursor shifts left when the left hemisphere dominates, and right when the right hemisphere dominates. A graph of the change of *IS* in an experiment with flash stimulation is given in Fig. 2.

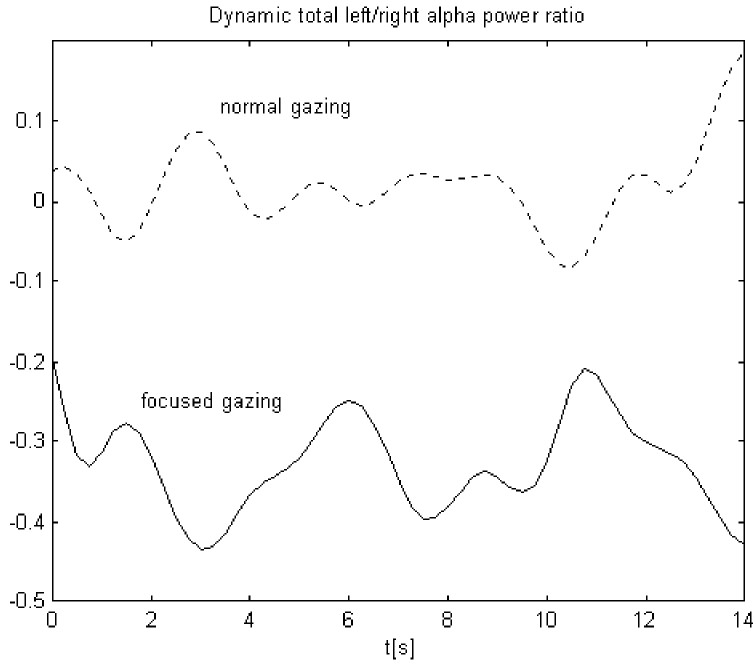


Figure 2: Variation of total alpha power index of symmetry over both left and right hemisphere during normal and focused gazing; Dominance of right hemisphere during focused gaze can be clearly seen; This parameter can be efficiently sonified as the position of a sound source (an audio cursor).

3.2 Audio Alarms

The auditory channel may also be effectively applied as an alert signal, either continuously changing in time, or as a discrete sound alarm played when certain conditions are satisfied. Mental fatigue during the evaluation of long EEG recordings increases the probability that short but important segments will be missed in the analysis.

We devised a sonification technique using a repeating sound phrase with appropriately changing pitch. This change in pitch corresponds to changes in the patient's EEG caused by drowsiness and functions to alert the EEG technician during such periods. We sonified the EEG index of Theta to Alpha frequency band power (ITA) as the classical correlate of drowsiness [10], depicted in Fig. 3.

4 Conclusions

In this paper we have shown how sound may be used to support sustained attention during prolonged procedures and analysis. Tactical audio facilitates precise manual positioning of a surgical instrument, even beyond the capability of visual sense. Acoustic rendering provides an additional sensory channel that carries information obtained by data reduction, which could not be otherwise perceived. This information can be simultaneously processed with primary visual data without increasing mental workload. It also decreases the need for sustained visual attention during the detection of transient events.

We are currently implementing tactical audio technology as an extension to the Stealth Station system for frameless stereotactical neurosurgery. This project is being executed by Computer Aided Surgery, Inc. in conjunction with Richard Bucholz, M.D. of Saint Louis University Medical Center Department of Neurosurgery, and is funded by DARPA.

Sonification support during analysis of long EEG records was tested at Institute of Mental Health in Belgrade. Results of the pilot test in clinical settings indicate reduced mental fatigue of neurophysiologists during analysis of long EEG records, and better insights into global brain electrical activity. In the next phase we plan to implement real time vigilance assessment as acoustic alert."

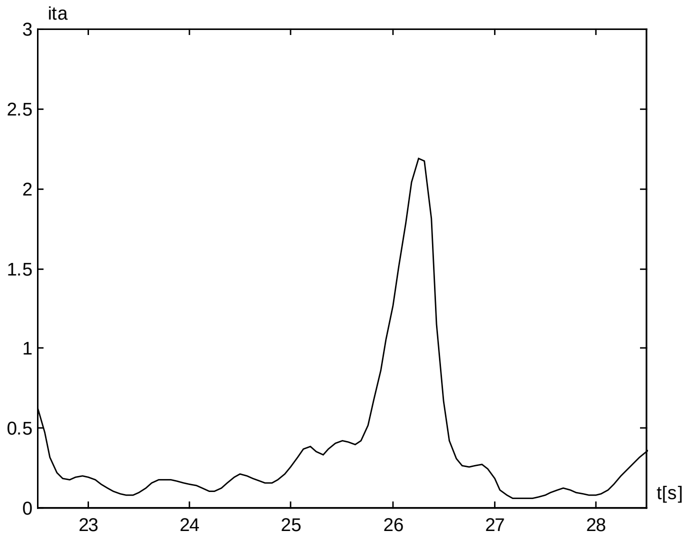


Figure 3: Increase of average ITA index for the right hemisphere during the drowsy period; this parameter could be effectively sonified as changing pitch of repeating sound phrase.

5 References

1. Kramer G Ed. Auditory Display, Sonification, Audification and Auditory Interfaces. Addison Wesley, 1994.

2. Burdea GC. Force and Touch Feedback for Virtual Reality, John Wiley & Sons, Inc., New York, 1996.

3. Begault DR. 3D Sound for Virtual Reality and Multimedia, Academic Press, Inc., Boston, 1994.

4. Bernsen NO, Foundations Of Multimodal Representations: A taxonomy of representational modalities, Interacting with Computers 1994; 6:347-371.

5. Bernsen NO, A Toolbox of Output Modalities, http://www.mrc-apu.cam.ac.uk/amodeus/

6. van Gils M, Rosenfalck A, White S, Prior P, Gade J, et al: Signal Processing in Prolonged EEG Recordings During Intensive Care. IEEE EMBS 1997, 16(6): 56-63.

7. The Virtual Reality Modeling Language, http://www.vrml.org/Specifications/VRML97

8. Cytowic RE: Synesthesia: Phenomenology And Neuropsychology A Review of Current Knowledge. Psyche 1995, 2(10).

9. Barnard PJ, May J (Eds): Computers, Communication and Usability: Design Issues, research and methods for integrated services. North Holland Series in Tele-communication, Amsterdam, Elsevier, 1993.

10. Santamaria J, Chiappa KH, The EEG of Drowsiness in Normal Adults, J. Clin. Neurophysiol. 1987; 4(4):327-382.