Immersive Sonification for Displaying Brain Scan Data

Agnieszka Roginska[1], Hariharan Mohanraj[2], Mark Ballora[3] and Kent Friedman[4]
[1,2]Music and Audio Research Lab (MARL), New York University, 35 West 4th St, New York, NY, U.S.A.
[3]Penn State School of Music/School of Theatre, Music Building I, University Park, PA, U.S.A.
[4]Division of Nuclear Medicine, Department of Radiology, NYU School of Medicine, 560 First Ave, New York, NY, U.S.A.

Keywords: Sonification, Data Display, Data Visualization.

Abstract: Scans of brains result in data that can be challenging to display due to its complexity, multi-dimensionality, and range. Visual representations of such data are limited due to the nature of the display, the number of possible dimensions that can be represented visually, and the capacity of our visual system to perceive and interpret visual data. This paper describes the use of sonification to interpret brain scans and use sound as a complementary tool to view, analyze, and diagnose. The sonification tool SoniScan is described and evaluated as a method to augment visual brain data display.

1 INTRODUCTION

Modern medicine relies heavily on human perception as a means to detect and monitor disease. Techniques based on (1) verbal communication and (2) physical exam evolved over hundreds of years, long before the development of effective therapies.

A variety of human sensory systems and cognitive pathways are put to use by the physician when evaluating a patient. Verbal communication involves not only listening to words spoken but also an assessment of more subtle clues including a patient's tone of voice and gestures. Physical exam requires the clinician to visually inspect, listen to and touch the patient. Even a doctor’s sense of smell can aid in diagnosis.

In more recent times, technology has brought forth a third paradigm for disease detection: diagnostic testing. This methodology reduces the demands placed on the physical senses but does not dispose of them altogether. Laboratory tests generating simple numeric values are viewed once and processed mainly by cognitive portions of the brain, with very little required in the way of "perception." More complex testing places greater sensory demands on the diagnostician. For example, electroencephalograms (EEGs) depicting brain electrical activity generate complex arrays of time-variant data that are presented on large displays. Electrocardiograms generate similar visual data. Ultrasound devices detect the reflection of sound waves off the body and generate not only time-variant images but produce audio signals that provide information about the flow of fluids throughout the body. Stethoscopes present clinicians with detailed information about transit of blood in vessels and air in lungs.

It is clear from the above that physicians employ all the major physical senses in their pursuit of information regarding their patients. What has become apparent over the years however is that traditional methods of perceiving information do not always provide sufficient detail, and the methods by which data are presented to the clinican do not always allow appropriate analysis, even with advanced diagnostic testing. Without new innovations, the capacity to prevent or effectively treat many complex conditions will remain elusive.

2 CHALLENGES OF IMAGING

A broad category of diagnostic testing, which has revolutionized diagnosis and management of disease, is medical imaging. A variety of techniques based on x-rays, mechanical vibration, fluorescence, rotation of atoms and radioactive decay have been developed and produce multi-dimensional and time-varying arrays of spatial information.

Current methodologies used to "perceive" and interpret medical image data are largely based on the human visual system, in some cases enhanced by use

Roginska A., Mohanraj H., Ballora M. and Friedman K..
Immersive Sonification for Displaying Brain Scan Data.
DOI: 10.5220/0004202900240033
In *Proceedings of the International Conference on Health Informatics* (HEALTHINF-2013), pages 24-33
ISBN: 978-989-8565-37-2

of simple graphs or tables depicting numeric data. Such visual and traditional quantitative analysis methods have led to great advancements in nuclear medicine, radiology and other fields.

Despite these innovations, many limitations remain with respect to the medical community's ability to perceive and analyze the vast amounts of data now being generated by CT scanners, PET scanners, MRI machines and newer combined devices (including PET/CT and PET/MRI). Diagnostic accuracy is higher than ever, but clinicians still are unable to detect certain conditions when the information provided by visual analysis or basic quantization does not uncover perceptible differences between disease and health.

There are two possible causes for the existing limitations in the diagnosis of disease using medical image data. One possibility is simply that the techniques in use do not provide enough information to allow absolute differentiation between disease and health, no matter how sophisticated our methods of analysis might be or might become. Another possibility is that the information needed to diagnose disease is "hiding" within these images and is not perceptible because the means by which we have chosen to examine the data are insufficient for detection of complex patterns associated with disease. This second (and far more likely) possibility suggests that the medical community needs to find new ways to process and understand the data that is being acquired by advanced scanners and other testing equipment.

The concept of searching for new methods to understand complex medical image data is not without precedent. In recent years, researchers performing brain scans on patients found that basic visual analysis of the distribution of information was at times insufficient to detect subtle diseases (epilepsy being one example). In order to overcome the limitations of visual interpretation, new methodologies were developed whereby brain image data was segmented and compared to a database containing information about normal individuals. Such techniques have proven successful and provide useful complementary information when combined with visual inspection. (Ferrie et al., 1997; Minoshima et al., 1995; Kono et al., 2007).

Quantitative analysis techniques for examining medical image data represent a significant step forward beyond the traditional visual processing system of the human brain. However, they raise the question of whether or not even newer and perhaps less conventional methods for image data analysis might be able to improve our ability to diagnosis and monitor disease even further.

Advanced computer processing techniques generating increasingly complex visual reports may allow doctors to find new ways to detect disease. Multi-parametric predictive modeling (Najafi et al., 2012**)** and machine learning (Oh et al., 2012) are examples of two evolving areas in medical image analysis.

When considering possible diagnostic techniques of the future, it is important to look back to the history of medicine and remember that visual analysis has not been the only way for a human to detect disease. It is worth exploring other ways to perceive patterns buried deep within complex data.

One potential alternate method for interrogating medical image data is by means of translation of information in to a format that can be processed by an alternate pathway: the human auditory system. Already regarded by experts as superior to vision in the domains of frequency and time, the human ear and associated auditory cortices present a compelling alternative system for perception of medical scan data. Independent of considerations of resolution, the complex neurological pathways of hearing offer a new perspective for understanding spatial, frequency and temporal data. Fortunately, a nascent field exists to allow for such a line of inquiry: sonification.

The central hypothesis of this paper is that auditory processing of medical image data will identify patterns associated with disease that cannot be detected by traditional means.

Using a model of molecular brain imaging whereby detection of small molecules is accomplished through the use of radioactive decay of injected "tracers" (PET imaging), we propose that patterns, when sonified, will emerge from the data and show information that was not previously detected by visual and pre-existing computer analysis techniques. These patterns, when heard and processed by the human brain, might one day allow the medical community to detect diseases that are presently invisible by currently existing methodologies. To our knowledge, this is the first report of the use of sonification in the analysis of three-dimensional medical image data.

3 SONIFICATION BACKGROUND

In a report prepared for the NSF (Kramer, et. al., 1999), sonification is defined as “the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation." Sonification involves the translation and integration of quantitative data through mapping to a sound model, and enables recognition of patterns in data by their auditory signatures.

While the term "sonification" dates back to the 1990s (Kramer, 1994), we and other colleagues observe anecdotally that it has appeared with increasing frequency over the last two years or so in the parlances of the sciences and informatics. This increase in awareness is perhaps somewhat surprising, yet has for some time seemed inevitable to those who have worked in the field, for a number of reasons (Ballora, 2010). Since both the eyes and ears aid us in providing complementary and supplementary information as we navigate through the world, it seems unreasonable to expect that, as information sources increase, visualization would remain the sole means of representing abstract information. As researchers and analysts currently face datasets of higher dimensions, or multiple simultaneous situation reports from different areas, there is a commonly acknowledged problem of information overload.

Information presentation is critical for both research and education. Scientists frequently rely on highly developed visualization techniques for their own understanding, as well as an aid in presenting material to lay audiences. Science museums and television programs such as *Nova* commonly engage audiences with eye-popping visuals; these visuals become the cultural basis of a generation of students who are inspired to go into the sciences and push the boundaries of a field of knowledge. The inclusion of sonic demonstrations along with visualizations is becoming a conceptual priority in science education and even in some popular music (Hart and Smoot, 2012; Hart, 2012).

Effective use of sound hinges on perceptual understanding, and what types of tasks we use the eyes and ears for. Visualizations are strongly synoptic, that is, an entire image can be seen at once. The eyes provide summary information of features such as shape, size, and texture.

Many organizing principles of visual cognition also apply to auditory perception. Like the eyes, the ears create auditory gestalts that aid understanding the nature of events, and make estimates when presented with incomplete information (Bregman, 1991).

In contrast to visualization, a sonification, like a piece of music, exists in time. It cannot be "listened to all at once." Being time-based, the ears give us a strong sense of dynamic elements of our environment. The auditory system is also highly adapted for following multiple streams of information (Fitch and Kramer, 1994). That is, listeners can readily apprehend a number of simultaneous melodies if they are presented effectively. Thus, sonification is an effective way to display a multitude of signal processing operations simultaneously, with each being represented as a line of counterpoint, a series of chords, or a succession of musical instruments. The auditory system is also extremely adept at pattern recognition, a capability that allows us to recognize melodies in spite of transpositions or variations.

The auditory system is most sensitive to dynamic changes involving periodicities: small changes in pitch or tremolo rate are perceptible to untrained listeners. Beyond this, other dimensions that may be represented in an auditory display include changes in loudness, instrument, stereo position, spectrum, transient time, duration, and distance. Through considered use of tools used in music synthesis, information can be presented in an engaging and informative fashion – in addition to boggling the eyes, science should also dazzle the ears.

The study of fMRI images is opening new avenues of understanding correlated activity among brain regions when performing certain tasks. As researchers advance in their understanding of this information, new pattern types will be recognized. We are exploring the use of sound to represent this information. On the one hand, this information may be static, and therefore not seem consistent with the time-specific strengths of the auditory system discussed above. On the other hand, by capitalizing on the sensitivities of the auditory system as we translate the information to sound, we expect that the pattern recognition capability of the auditory system will reveal new recognizable patterns.

3.1 Sonification Applications

In addition to the 1999 NSF summary referenced earlier, the most authoritative summary of work to date in the field is *The Sonification Handbook* (Hermann et al., 2011). Sonification has been demonstrated to offer advantages in representing multivariate data from a variety of domains, including the financial markets, quantum physics, and meteorology, as well as various situational monitoring implementations. Of particular interest to this project is prior work in medical informatics.

Sonifications of heart rate variability (Ballora et al., 2004) have been shown to have diagnostic potential. And there is a strong history of sonifying EEG readings of brain activity (Hermann et. al., 2002; Baier and Hermann, 2004; Baier et. al., 2007; de Campo et. al., 2007; Hermann et. al., 2008). Other relevant work includes EMG (Pauletto and Hunt, 2006) and cell culture data (Edwards, 2008). These sonifications have subsequently proved useful as general introductions to physiological health, making distinctions between healthy and diseased states easy for uneducated listeners (Hong, 2007). Sound has also been effective in combination with haptics as a surgical aid or training tool (Jovanov et. al., 1998; Müller-Tomefelde, 2004), and sound has functioned as an effective component in assistive rehabilitation therapy for stroke victims (Wallis et. al., 2007). There has also been work done in sound rendering of tissue biopsy slides for diagnostic purposes (Cassidy et. al., 2004).

3.2 Immersive Sonification

Displaying 3D data on visual displays is particularly challenging because of the many dimensions needed to be presented concurrently. Our eyes are limited with their "field of view", and cannot see behind objects that obstruct the view. Our hearing, however, is omni-directional, and we can hear sounds emanating from around us, and in distance.

We refer to immersive sonification as a method of sonifying data in such a way as to place the viewer/listener of this data inside of an immersive 3D environment so that she may be able to navigate through the environment. The sonified data is rendered constantly. However, the listener will be closer/farther to a region of the data depending on their virtual location in the immersive data environment. For example, we can think of an immersive sonification of 3D brain data as putting the listener inside of the scanned brain. Through the use of spatial sound, we are able to create an immersive environment in order to present a spatially distributed sonification.

3.2.1 Spatial Sound

This study emphasizes in particular the advantages of synthesizing spatial location. Unlike the visual system, the auditory system doesn't have a "field of view." Sounds can be heard and perceived anywhere around a listener. A listener does not have to be facing a sound in order to perceive it. Regardless of whether a sound is located in front, behind, above, or below, it can be detected by the listener (Wakefield et al., 2012).

Sounds representing objects on computer screens, or other auditory displays have used spatial location to represent the absolute location of objects or events. For example, sound representing an object on a graphic monitor can be played to be perceived at the same location as it appears on the screen – if a graphical icon is located in the lower right corner of the screen, the sound may be presented so that it appears to be emanating from the lower right corner.

In addition to providing location-specific information, when multiple sounds are presented concurrently in spatially disparate locations, they can be better segregated into individual streams (e.g. Barreto et al., 2007; Marston et al., 2006; Shilling et al., 2000). Without spatial separation, the multiple sounds will have a greater tendency to fuse together, making them more difficult to understand.

3.2.2 Human Localization

Our ability to localize sounds in a three dimensional world has been studied by researchers for almost 100 years, ever since Lord Rayleigh's duplex theory was introduced in 1911. Since then, many studies have resulted in a more comprehensive understanding of localization, human perception of 3D sound.

Spatial hearing refers to the ability of human listeners to judge the direction and distance of environmental sound sources. To determine the direction of a sound, the auditory system relies on various physical cues. Sound waves emanating from a source travel in all directions away from the source. Some waves travel to the listener using the most direct path (direct sound) while others reflect off walls and objects before reaching the listener's ears (indirect sound). The direct sound carries information about the location of the source relative to the listener. Indirect sound informs the listener about the space, and the relation of the source location to that space.

The Duplex Theory of Sound Localization (Rayleigh, 1911) states the two primary cues used in sound localization are time and level differences between the two ears. Because of the ears' spatial disparity and the mass between them, they each receive a different version of the arriving sound. The ear that is closest to the sound (ipsilateral ear) will receive the sound earlier and at a greater intensity or level than the ear farther away from the source (contralateral ear). The differences in time of arrival and in level are referred to as the Interaural Time Difference (ITD) and the Interaural Level Difference (ILD) respectively.

Although the ITD and ILD cues are good indicators for determining the location of sources along the interaural axis, they provide an insufficient basis for judging whether a sound is located above, below, in front or in back. For sources located at an equal distance on a conical surface extending from the listener's ear, ITD and ILD cues are virtually identical producing what is referred to as the "cones of confusion" (Woodworth, 1954). Along a cone of confusion, where the ITD and ILD cues along a cone of confusion for a source located in the front, back, up and down are equivalent, a listener can have difficulty determining the difference in location, which can lead to front/back or up/down confusion.

There is an additional acoustic cue that helps to resolve the position along a cone of confusion. Before reaching the listener's ears, the acoustic waves emitted by a source are filtered by the interaction with the listener's head, torso and the pinnae (outer ear), resulting in a directionally dependent spectral coloration of the sound. This systematic "distortion" of a sound's spectral composition acts as a unique fingerprint defining the location of a source. The auditory system uses this mapping between spectral coloration and physical location to disambiguate the points along a cone of confusion, leading to a more accurate localization of a sound source. The composite of the ITD, ILD and the spectral coloration characteristics are captured in Head-Related Transfer Functions (HRTF).

In creating an immersive sonification tool, we use the time, intensity and spectral cues contained in HRTFs in order to simulate spatial sound around a listener. The next section describes SoniScan – an immersive sonification tool.

4 SoniScan

SoniScan is a sonification tool developed in the Matlab technical computing software. Matlab is a flexible and versatile computing package that facilitates the manipulation of data, and synthesis and processing of sound. SoniScan provides a graphical interface through which a user can load, manipulate, and sonify Digital Imaging and Communications in Medicine (DICOM) data - a standard format for viewing and distributing medical imaging data. The program was constructed with a modular approach in mind at both a macro and micro level, in order to allow maximum flexibility for exploring sonification methods that are conducive to brain data display. The data and signal flow are illustrated in Figure 1.

The following four modules have been designed to map the data to sound: control, path, Sound synthesis, mappings, and audio controls. DICOM data is read into SoniScan, and data manipulation is performed in the Controls module. After passing through the Sonification Path module, which specifies the data path the sonification will follow, the Sound Synthesis module defines the details of the sound synthesis mechanism and passes to the Mappings module. Sound is generated using these parameters and the resulting audio is played using the Audio Controls module.

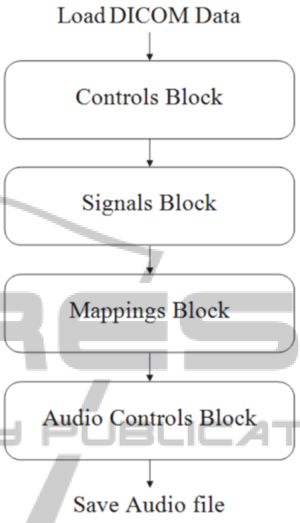


Figure 1: Schematic of SoniScan.

4.1 Control Module

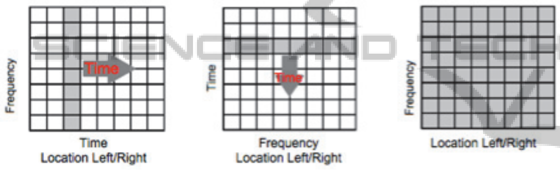
The Control module reads in and prepares the DICOM data to be sonified. This module has five main functionalities: data read, data selection and zoom, sonification duration, and data adjustment.

SoniScan loads DICOM-formatted files. A 3D matrix of the scan data is constructed and used for the sonification and the basic visual display provided. A single frame of data may be selected for sonification, a subsection of one frame, or a full three-dimensional scan. Using HRTF processing, the data selected is spatialized, a process by which directly correlates the location of each data pixel to an apparent sound location heard by the listener.

Transposing the DICOM data values that are suitable for auditory presentations requires that all data be easily re-scalable. Additional controls of range, shift and volume allow the mapping ranges to be adjusted. The range adjusts the data range. It allows a lesser or greater differentiation between the largest and smallest values in the data. The shift operation transposes pitches to higher or lower part of the musical scale (i.e. like shifting to higher or lower notes on a piano).

4.2 Sonification Path Module

The sonification of the data stored in a matrix can be done by sonifying one data voxel at a time, by row, by column, or sonifying the full 2D or 3D matrix simultaneously. The Sonification Path module allows the user to define the exact path through which to scan the selected data. SoniScan contains three preset paths: from left to right, top to bottom, or all data simultaneously, as seen in Figure 2. If the left-to-right path is selected, each data column is sonified and played concurrently, followed by the next column of data. If the right-to-left path is selected, each row is sonified, and rows are played concurrently. If simultaneous is chosen, effectively the path is removed: the sum of each row is taken, resulting in a single column, with each value representing the total values of the corresponding row. We refer to these three mapping trajectories as *conventional* paths.



(a) (b) (c)

Figure 2: Sonification paths: (a) left-to-right; (b) top-to-bottom; and (c) all data.

In addition to the three conventional paths, we also allow sonification along a *split-path*. The selected data frame is split along a specified line (split line) and sonified as two halves. The split-line is specified by the Cartesian x and y coordinates of any two points on the line. The motivation behind the split path is to facilitate detection of asymmetries between the two halves of the data, which correspond to the two brain hemispheres.

There are two types of split paths. The first is a *hard left/right*, where the selected data is split into two halves, the left side is sonified and played to the left ear, and the right side is sonified and played to the right ear. The second type is the *difference* scan, where the mirror of the right side of the data is subtracted from the left side. This can be thought of as folding the halves onto each other and taking a difference between the two halves. Resulting asymmetries between the two sides of the data are the only audio signals that are audible, since symmetric data is cancelled out. In both types of split paths, the sonifications generated are based on the mappings and scan paths specified, as described above.

4.3 Sound Synthesis Module

The Sound Synthesis module specifies the base signals to use for the sonification. It is modular in structure and can easily support additional signal types in the future. Any combination of signals contained in this module can be used simultaneously in the sonification. In our current implementation, the sonification can use band-passed noise, triangle wave tone, or a plucked string model. For example, when band-passed noise is selected, a noise signal is used for the sonification that is band-passed at the centre frequency that corresponds to a data value.

The parameters of these signals, namely amplitude, frequency, spatial location, and time of occurrence, are governed by the data to be sonified. The manner of mappings is specified in the Mappings block (next section). Any number of base signals can be simultaneously selected for sonification, and doing so will layer the signals on top of each other.

4.4 Mappings Module

The Mappings module defines how the image data is mapped to different parameters of the audio. Three different mapping techniques have been used: amplitude mapping, frequency mapping, and spatialization. These are explained in detail in the following sections.

4.4.1 Amplitude Mapping – Fourier Synthesis

In the case of amplitude mapping, each pixel's intensity is mapped to the physical amplitude of the audio signal corresponding to that pixel. The greater the data value, the greater the amplitude of the audio signal resulting in a louder sound. Therefore, highly active regions of the data result in louder regions.

When the mapping is performed, different frequencies are assigned to different pixels based on their location in the image. Frequencies are distributed between 500Hz and 5kHz. How the frequencies are distributed depends on the sonification paths. For Path 1 (left to right), frequency varies from top to bottom, with the highest frequency assigned to the topmost rows in the data (Figure 2a). For Path 2 (top to bottom), frequency varies from left to right, with highest frequency assigned to the rightmost column (Figure 2b). Path 3 (all data) follows the same frequency assignment as Path 1 (Figure 2c). While Paths 1 and 2 present each row or column sequentially, Path 3 renders an integrated spectrum of the entire image, since the same sets of frequencies for every column are presented simultaneously. A total sum of values is taken for each row/frequency, and this sum controls the amplitude of each frequency component.

4.4.2 Frequency Mapping

In the case of frequency mapping, pixels are assigned frequency based on their intensity, with all amplitude values being kept at a constant level. The assignable frequencies are quantized to integer multiples of 500Hz to a maximum of 5kHz, so all generated audio contains only harmonic content.

The motivation behind performing frequency mapping is to judge the intensity distribution of a particular dataset through timbre. Datasets containing more high-intensity pixels will have more high-frequency harmonic content. This may in some cases result in a harsh timbre, if there is a large range of values being presented and many frequencies are sounding, or a ringing, if most values are high, and thus each pixel is rendering the same frequency. Hence, through frequency mapping, the user can gain a quick idea of the composite intensity at each step along the selected scan path..

4.4.3 Spatialization

The Spatialization module distributes sonified sound spatially around the listener. Assuming users would employ off-the-shelf headphones of reasonable fidelity, a great range of apparent stereo locations may be synthesized. Spreading the sound around the listener can result in a better differentiation of sonified regions, and thus a better distinction of features pertaining to sections of the data.

There are four spatial mapping methods employed in SoniScan: intensity panning, vertical spatialization, horizontal spatialization, and full 3D spatialization.

The intensity panning method uses interaural intensity differences between the two ears to pan sounds from left to right. When the sounds sent to the left and right ears are of equal level, the virtual auditory image appears to be located in the center. As the sound level of one of the ears increases, the location of the source appears to be originating from the side with the greater sound level. Intensity panning is effective for creating changes along the horizontal plane, but not the vertical plane. Using the intensity panning method, the audio corresponding to each pixel is panned to an apparent position based on the Cartesian x-coordinate of the pixel. Pixels that are in the center of the image will be panned and perceived to be coming from the center, those that are on the right side of the image will be panned to the right, and so on. A left-to-right image path will result in the sound containing all the vertical image data moving from left to right, while in a top-to-bottom image path the sound constantly surrounds the listener from left to right, and the vertical data is presented consecutively.

Vertical HRTF spatialization utilizes head-related transfer functions to map the selected dataset onto a two-dimensional (up-down, left-right) vertical aural image space directly in front of the user. The spatialized audio correlates with the spatial distribution of the visual image. For example, data on the top left corner of the image is sonified and perceived to be coming from a high elevation, the left data that is in the middle of the image is presented at ear level, and data in the lower part of the image is spatialized below ear level (Figure 3a).

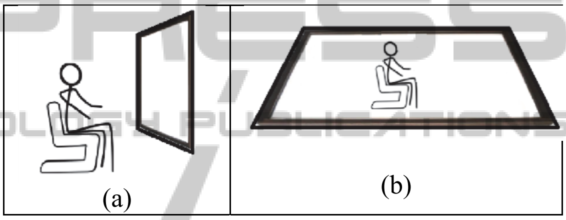
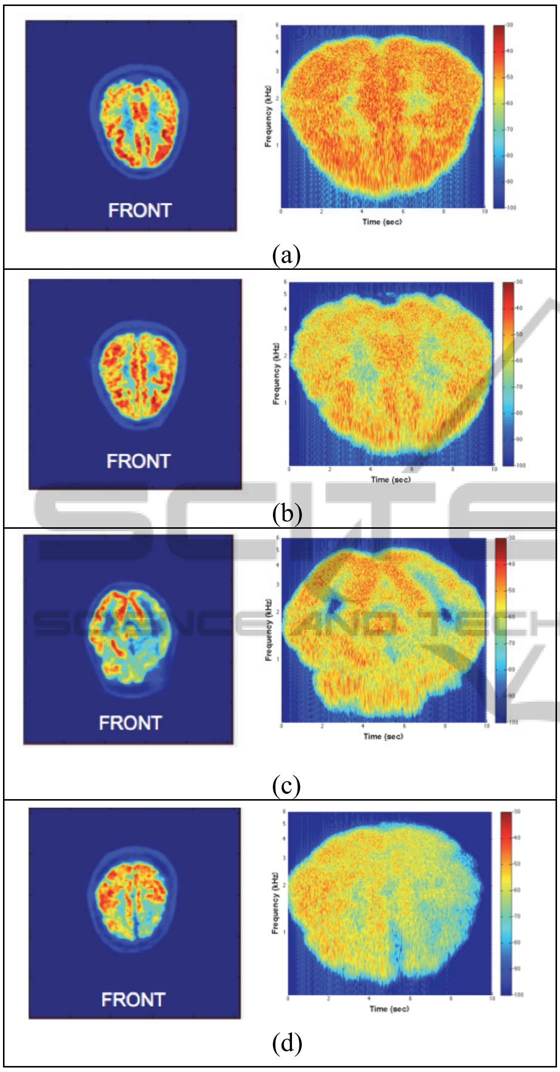

(a)
(b)

Figure 3: Representation of vertical (a) and horizontal (b) HRTF spatialization.

Horizontal HRTF Spatialization utilizes head-related transfer functions to map the selected dataset onto a two-dimensional horizontal aural image space (front-back, left-right) that places the user in the center. Effectively, the image is laid flat on the horizontal plane, and the listener is placed in the center of the image. The audio corresponding to the sonified dataset is spatially placed all around the user on the horizontal plane (Figure 3b). For example, data that is in the upper left corner of the image would be sonified and presented in the left front of the audio image.

The full 3D spatialization method can be used for data that is three-dimensional. This would be a VR version of the data, with a series of horizontal or vertical scans all active simultaneously. HRTFs and distance mapping are used to position the data points around the listener – front, back, left, right, up, down. In this case, all the data can be sonified concurrently and spatialized to reflect the position of each data point relative to the listener. The listener can place herself in the middle of the data, or listen to the data from another perspective. For example, the listener can be center of a 3D brain scan and listen to all the data in the scan concurrently. By selecting different listening locations, the listener can effectively “walk through” the data in a fully immersive manner.



(a)

(b)

(c)

(d)

Figure 4: Images of brain scans (left column) and spectrograms of sonifications (right column) of healthy brains (a and b), and unhealthy brains with Alzheimer's dementia (c) and frontotemporal dementia (d).

4.5 Audio Control Module

The Audio Control module manages the sonification process and handles the generated audio files. Once the relevant sonification and data parameters are set, the sonification is performed, and the rendered sonification can be played or stored to disc.

There is an additional A/B playback functionality, which facilitates the serial playback of two different sonifications as an A/B comparison. This is useful when comparing two sets of data that may have been taken at different points in time, or comparing two cross-sections of the same dataset in order to perceive the difference between the two. This capability for A/B comparison through sonification function is particularly important since the perceptual auditory system is more acute than the visual system at detecting temporal, spectral, and spatial changes. When small changes in data occur, they may not be immediately noticeable on the visual display, but may be more easily observed using sonification.

4.6 Examples

Due to the limitations of presenting audio examples in a written paper, we are using spectrograms (time-varying representations of the spectrum of an audio signal) to show the correlation between the PET scans and their sonifications. Figure 4 contains four examples of the spectrograms of sonifications that were created from normal and abnormal brain data. The left side of the figure contains single slices from three-dimensional PET scans depicting sugar utilization in the brain. On the right side are the spectrograms of the sonifications. In these examples, the x-axis of the spectrogram represents time, while the y-axis shows the frequencies from low (bottom part of the graph) to hi (top part of the graph). The intensity of the color represents the amplitude of the spectral content at each point in time – with blue indicating low amplitude, and red high amplitude.

The top two examples – Figure 4(a) and Figure 4(b) - are of healthy brains exhibiting a homogenous and symmetric pattern of sugar metabolism. As can be seen in the spectrograms (and heard in the sonifications), the symmetry of the spectral content as well as the full bandwidth reflects normal glucose uptake in the brain. Conversely, in Figure 4(c), there is asymmetric and decreased signal intensity in the brain of a patient with Alzheimer's dementia. A lack of low frequency content is heard in the sonification, and is visible in the second part of the spectrogram, starting at approximately 13 seconds. This region on the spectrogram corresponds to the most severely diseased portion of the brain, resulting in an audible hole in the frequency spectrum. Likewise in Figure 4(d), an image of a brain scan of a patient with frontotemporal dementia is presented and demonstrates a lack of low frequency spectral content corresponding to the frontal lobes of the brain. This leads to an unusual sonic representation of that hemisphere.

5 DISCUSSION & CONCLUSIONS

Brains scans contain complex, highly variable data and present a challenge to the interpreting physician. Imaging experts spend years learning to properly read such studies, yet detection of subtle disease remains difficult. Compounding matters, many disease processes remain invisible even to the best observers, either due to lack of meaningful information or undiscovered means by which to identify the relevant data. Visual quantitative techniques have improved matters but there is room for further improvement. It is suspected that as-yet-undiscovered information exists within these images and has diagnostic and therapeutic relevance.

Development of new ways to understand complex, multi-dimensional data from brain and other body scans is important for advancement of the medical imaging sciences. Sonification seems to be an appropriate target for further studies in this area, harnessing the auditory system's spatial acuteness and omni-directional hearing in order to identify patterns in data that may not be apparent by other means. The examples presented above confirm that even preliminary work in sonification can begin to differentiate disease from health. Increasingly complex auditory modelling might one day reveal sophisticated and relevant medical information.

The tool described in this paper, SoniScan, is a research system developed for exploring the benefits of sonification for medical data and researching sonic parameters that are promising in guiding physicians as they diagnose disease and monitor treatment. With the sonification mappings we have explored thus far, we see a significant and audible correlation between the image data and the sonification. These correlations serve as a foundation upon which refinements can be explored to pick out increasingly subtle variations in image data that might one day carry diagnostic relevance and ultimately impact patient care.

We are still in early stages of evaluating our renderings and exploring new mappings between PET data and audio parameters. The parameter mappings will significantly affect the audibility of patterns and possibility of extending sonification as providing value in monitoring and/or diagnosing disease. Initial reactions among the authors included that of an immediate, pleasurable sense of recognition when hearing the simultaneous ascending/descending pitch patterns, which clearly corresponded to the outline of the skull in the image and in/active areas of the brain. While not of any diagnostic value, this still made the "first handshake" with the environment a pleasant and engaging one. Areas of diseased brain are clearly audible as distinct sounds contrasting with the sonifications of healthy patient scans.

During a cursory comparison of normal and abnormal brain scans, it was immediately apparent that there was much more scattered activity with the dementia scans than with the normal scans, and there was a certain sonic quality that might be described as more "strident" in the central area of the scan. By zooming in on the central area of greatest activity, it became apparent that the dementia scans consisted of activity farther forward in the brain, due to the later starting time of audible activity in the L-R scan.

Next steps for this work includes subjective studies with junior and senior-level nuclear medicine physicians to systematically compare the sonification performance with traditional techniques and determine if it reveals additional information that may improve upon existing analysis methods.

As understanding of MRI, PET and other imaging technology increases, novel means of data presentation and analysis can only be welcome. Imaging scientists are constantly searching for new ways to examine their massive data archives, and sonification is an intriguing line of inquiry that complements other investigations in image analysis. Non-invasive imaging coupled with new perceptual methods such as sonification hold great promise for a future in which society can better detect and treat disease. Such advancements will be welcomed by doctors and, most importantly, patients.

REFERENCES

Baier, B., Hermann, T., & Stephani, U., 2007. Multi-Channel Sonification of Human EEG. *Proceedings of the 13th International Conference on Auditory Display*, p. 491-496.

Ballora, M., Pennycook, B., Ivanov, P. C., Glass, L. & Goldberger, A. L., 2004. Heart rate sonification: A new approach to medical diagnosis. *LEONARDO,* 37, p. 41-46.

Ballora, M., 2010. Beyond Visualization – Sonification Invited chapter in Hall, D. L. & Jordan, J. M., *Human-Centered Information Fusion*. Artech House Publishers.

Barreto, A., Jacko, J. A., & Hugh, P. 2007. Impact of spatial auditory feedback on the efficiency of iconic human-computer interfaces under conditions of visual impairment. *Computers in Human Behavior*, 23(3), p. 1211-1231.

Bregman, A. S., 1990. *Auditory Scene Analysis*. Cambridge, MA: MIT Press.

Brown, M. H., 1992 An Introduction to Zeus: Audiovisualization of some elementary sequential and parallel sorting algorithms. *Proceedings of the CHI '92 Conference*, p. 663-664.

Cassidy, R. J., Berger, J., Lee, K., Maggioni, M., & Coifman, R. R., 2004. Auditory display of hyperspectral colon tissue images using vocal synthesis models. *Proceedings of the 10th International Conference on Auditory Display*.

Edwards, A. D. N., Hines, G., Hunt, A. 2008. Segmentation of biolocial cell images for sonification. *Proceedings of the 2008 Congress on Image and Signal Processing*, Vol. 2, p. 128-32.

Ferrie, C. D., Marsden, P. K., Maisey, M. N., Robinson, R. O. 1997. Visual and semiquantitative analysis of cortical FDG-PET scans in childhood epileptic encephalopathies. *J Nucl Med*, 38(12), p. 1891-1894.

Fitch, T. & Kramer, G. Sonifying the Body Electric: Superiority of an Auditory over a Visual Display in a Complex, Multivariate System. In Kramer, G., ed., 1994. *Auditory Display: Sonification, Audification, and Auditory Interfaces*. Santa Fe Institute Studies in the Sciences of Complexity. Addison Wesley. Reading, MA.

Hart, M., & Smoot, G. S., 2012. Rhythms of the Universe, a multi-media production. Lawrence Berkeley Labs.

Hart, M., et. al., 2012. The Mickey Hart Band: Mysterium Tremendum. 360° Productions, Inc., Sepastopol, CA.

Hermann, T., Hunt, A., & Neuhoff, J. G., 2011. *The Sonification Handbook*. Logos Publishing House, Berlin.

Hermann, T., Meinicke, P., Bekel, H., Ritter, H, Muller, H.M., Weiss, S. 2002. Sonifications for EEG data analysis. *Proceedings of the 8th International Conference on Auditory Display*.

Hong, S. L., 2007. *Entropy Compensation in Human Motor Adaptation.*Ph.D thesis in Kinesiology, Penn State University.

Jovanov, E., Starcevic, D., Wegner, K., Karron, D., & Radivojevic, V., 1998. Acoustic rendering as support for sustained attention during biomedical procedures. *Proceedings of the 5th International Conference on Auditory Display*, p. 1–4.

Kramer, G., ed., 1994. *Auditory Display: Sonification, Audification, and Auditory Interfaces*. Santa Fe Institute Studies in the Sciences of Complexity, Addison Wesley. Reading, MA.

Kramer, G., Walker, B., Bonebright, T., Cook, P., Flowers, J., Miner, N.; Neuhoff, J., Bargar, R., Barrass, S., Berger, J., Evreinov, G., Fitch, W., Gröhn, M., Handel, S., Kaper, H., Levkowitz, H., Lodha, S., Shinn-Cunningham, B., Simoni, M., Tipei, S., 1999. The Sonification Report: Status of the Field and Research Agenda. Report prepared for the NSF by members of the International Community for Auditory Display. ICAD, Santa Fe, NM.

Kono, A.K., Ishii, K., Sofue, K., Miyamoto, N., Sakamoto, S., Mori, E. 2007. Fully automatic differential diagnosis system for dementia with Lewy bodies and Alzheimer's disease using FDG-PET and 3D-SSP. *Eur J Nucl Med Mol Imaging* 34, p. 1490-1497.

Marston, J.R., Loomis, J. M., Klatzky, R. L., Golledge, R. G., & Smith, E. L., 2006. Evaluation of Spatial Displays for Navigation Without Sight. *ACM Transactions on Applied Perception*, 3(2), p. 110-124.

Mauney, L. M., & Walker, B. N., 2007. Individual Differences and the Field of Auditory Display: Past Research, a Present Study, and an Agenda for the Future. *Proceedings of the 13th International Conference on Auditory Display*, p. 386-90.

Minoshima, S., Frey, K.A., Koeppe, R.A., Foster, N.L., Kuhl, D.E. 1995. A diagnostic approach in alzheimer's disease using three-dimensional stereotactic surface projections of fluorine-18-FDG PET. *J Nucl Med* 36(7), p. 1238-48.

Müller-Tomefelde, C., 2004. Interaction sound feedback in a haptic virtual environment to improve motor skill acquisition. *Proceedings of the 10th International Conference on Auditory Display*, p. 1–4.

Najafi, M., Soltanian-Zadeh, H., Jafari-Khouzani, K., Scarpace, L. Mikkelsen, T. 2012. Prediction of glioblastoma multiform response to bevacizumab treatment using multi-parametric MRI. *Plos One* 7(1), p. 1-7.

Oh J, Wang Y, Apte A, Deasy J. 2012. SU-E-T-259: A Statistical and Machine Learning-Based Tool for Modeling and Visualization of Radiotherapy Treatment Outcomes. *Med Phys*. 39(6), p. 3763.

Pauletto, S., Hunt, A. 2006. The sonification of EMG data. *Proceedings of the 12th International Conference on Auditory Display*. p. 152-157.

Rayleigh, Lord [Strutt, J.W.], 1907. On our perception of sound direction. *Philosophical Magazine*. 13, p. 214:232.

Shilling, R.D., Letowski, T., & Storms R., 2000. Spatial Auditory Displays for Use within Attack Rotary Wing Aircraft, P *Proceedings of the 6th International Conference on Auditory Display*, Atlanta, GA.

Wakefield, G.H., Roginska, A. & Santoro, T.S. 2012. Auditory detection of infrapitch signals for several spatial configurations of pink noise maskers, *Proceedings of the 41st International Congress on Noise Control Engineering, Inter-Noise 2012*, New York, NY.

Walker, B. N., 2002. Magnitude Estimation of Conceptual Data Dimensions for Use in Sonification. *Journal of Experimental Psychology: Applied*, 8(4): p. 211–221.

Wallis, I., Ingalls, T., Rikakis, T., Olsen, L., Chen, Y., Xu, W & Sundaram, H., 2007. Real-Time Sonification of Movement for an Immersive Stroke Rehabilitation Environment. *Proceedings of the 13th International Conference on Auditory Display*, pp. 497-503.

Woodworth, R. S. 1954. *Experimental Psychology*, revised edition. New York: Rinehart & Winston.