

THE TRIPLE TONE SONIFICATION METHOD TO ENHANCE THE DIAGNOSIS OF ALZHEIMER'S DEMENTIA

*Letizia Gionfrida*¹, *Agnieszka Roginska*², *James Keary*², *Hariharan Mohanraj*², *Kent P. Friedman*³

¹School of Computer Engineering
Nanyang Technological University
Singapore
lgionfrida@ntu.edu.sg

²Music and Audio Research Lab
Music Technology
New York University
New York, NY, USA
{roginska, jpk349, hari.m}@nyu.edu

³Division of Nuclear Medicine
Department of Radiology
NYU School of Medicine
New York, NY, USA
kent.friedman@nyumc.org

ABSTRACT

For the current diagnosis of Alzheimer's dementia (AD), physicians and neuroscientists primarily call upon visual and statistical analysis methods of large, multi-dimensional positron emission tomography (PET) brain scan data sets. As these data sets are complex in nature, the assessment of disease severity proves challenging, and is susceptible to cognitive and perceptual errors causing intra and inter-reader variability among doctors. The Triple-Tone Sonification method, first presented and evaluated by Roginska et al., invites an audible element to the diagnosis process, offering doctors another tool to gain certainty and clarification of disease stages. Audible beating patterns resulting from three interacting frequencies extracted from PET brain scan data, the Triple-Tone method underwent a second round of subjective listening test and evaluation, this time on radiologists from NYU Langone Medical Center. Results show the method is effective at evaluation PET scan brain data.

1. INTRODUCTION

Medical imaging techniques such as X-rays, magnetic resonance imaging (MRI), and positron emission tomography (PET) have drastically aided the diagnosis of medical conditions. Such techniques provide physicians with multi-dimensional and time-varying data sets that continue to increase in precision and reso-

lution. These techniques rely heavily on visual displays and visual based analysis for the principal method of evaluation and subsequent diagnosis. The techniques listed above convey large amounts of complex data via medical imaging. Although diagnostic accuracy is higher than ever, clinicians continue to struggle when the visual analysis of the information provided leads to imperceptible differences between health and disease. It is a possibility that for the diagnosis of AD, visual imaging techniques of such complex data need supplementary information to aid diagnosis. The Triple Tone method aims to enhance clarity in the diagnosis process, particularly in regards to diagnostic accuracy and inter-observer variability.

2. BACKGROUND

The Triple Tone method proposed by Roginska et al. (2013) generates a sonification represented by three separate tones that correspond to three regions of the brain: the sensorimotor cortex (SMC), frontal lobe (FL), and parietal lobe (PL). Metabolic activity of each region maps to frequencies that create the tone. The SMC has a fixed tone of 440 Hz while the FL and PL tones are the relative standard deviation with respect to the SMC in the positive and negative directions. This difference in tones creates a beating pattern, which would be more perceptible for the listener to hear. Resulting sonification models were initially tested for accuracy on listeners without medical backgrounds. Participants categorized the sonification samples with both coarse and fine grained levels, which represented the severity of the different cases of AD. For instance, coarse categorization, on one side, involved four levels that ranged from severities (normal, mild, moderate, and severe), on the other, fine categorization involved twice as many levels by



This work is licensed under Creative Commons Attribution Non Commercial 4.0 International License. The full terms of the License are available at <http://creativecommons.org/licenses/by-nc/4.0>

splitting each of the coarse levels in half. This test showed that listeners could more accurately categorize the correct diagnosis with finer gradation.

2.1. Current issues in the diagnosis of Alzheimer's Dementia

Alzheimer's disease (AD) is a neurodegenerative disease that affects the frontal and parietal regions of the brain. Recently, a multiplicity of imaging techniques, including positron emission tomography (PET) studies of cerebral metabolism have shown characteristic changes in the brain of patients with AD. Particularly, when a PET scan is used on a brain affected by AD, a decrease in metabolic activity of the above cited regions can be found, as shown in Figure 1. PET scan datasets are displayed visually with MIM software, throughout which a radiologist can compare the brightness of the affected regions to the sensorimotor cortex, a region unaffected by the disease. The greater is the difference in brightness the more severe will be the diagnosis. To be fair, this is an oversimplification of the process; however, such a technique does rely heavily on the visual process. The downsides of this can be seen in the intra and inter reader variability when it comes to spotting the early stages of the disease's progression.

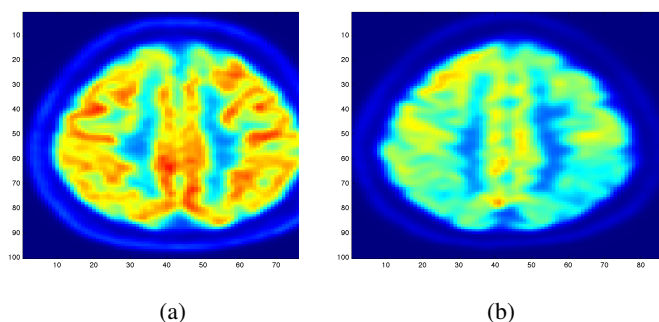


Figure 1: Example of orthogonal PET scan brain images of (a) a brain not affected by AD and (b) a brain severely affected by AD.

2.2. A review of Triple-tone sonification

The Triple-Tone Sonification method leverages the ear over the eye. For example, if presented with two shades of red one frequency apart on the color spectrum, the human eye will have a hard time determining the subtle difference between them. However, the human ear has a distinct advantage in that it hears two close frequencies beat against each other in a phenomenon known as beating. As thoroughly outlined by Roginska et al. (2014), complete with mathematical equations, the triple tone sonification method takes PET scan data-sets and assigns a single audible frequency value to the frontal, sensorimotor cortex and parietal regions. The sensorimotor cortex is given a base frequency of 440 Hz, and the other two regions are deviations from that frequency. These deviation values are determined by how far the overall metabolic regions activity deviates from the sensorimotor cortex. When these 3 frequencies are sonified, the differences between the regions are audibly heard as rhythmic beating patterns.

3. METHODOLOGY

3.1. Research structure and goals

As shown by Roginska et al. (2014) preexisting works, the previous round of testing validated the technique, proving its capability to provide accurate categorizations of the beating patterns. However, that testing round was done on the trained musical ears of audio professionals. The goal of this second round of testing was to determine how the Triple-Tone sonification technique affected the diagnoses process of those who would theoretically use it: physicians in the field of radiology. Thus, the methodology by which we tested the radiologists was informed accordingly. The goals of this round of testing were to determine the following:

- The accuracy and consistency of diagnosis by musically untrained ears (physicians).
- The extent of influence of the sonification in the diagnosis of a brain scan, for both experienced and inexperienced physicians.

The learning curve associated with this sonification technique on untrained ears, for both experienced and physicians with less experience.

3.2. Dataset preparation

The datasets used were the same as those in the first round of testing: 32 de-identified PET/CT scans of human brains diagnosed with varying stages of AD. Obtained from the Radiology Department of New York University Langone Medical Center, these 32 brains scans consisted of 8 brain scans in each of the four categories of diagnosis of AD; Normal, Mild AD, Moderate AD, and Severe AD. Only one slice per brain scan data set was used for sonification. The datasets were prepared in the same manner as the first round of testing as well. All datasets were spatially warped to a standard brain model using the MIM software package (www.mimsoftware.com). The spatial normalization process was necessary so that all 32 datasets used in testing were spatially consistent. This made it possible to sonify the same two dimensional subset of data points within each three dimensional dataset by simply choosing the same lateral slice each time for sonification. The next step was lobe segmentation. The MIM software was used to determine the boundary coordinates of the lobes within the dataset. The datasets were then output from the MIM software as DICOM standard RTSTRUCT file format, and input into SoniScan, our primary data analysis tool developed for the purpose of this research. SoniScan chose the 30th lateral slice of the spatially normalized datasets, segmented the lobes into the three regions to be sonified based on the given boundary coordinate points, and removed all irrelevant data points from the sub datasets. The 30th slice was chosen because it passes through representative regions of the frontal lobe, parietal lobe, and sensorimotor cortex. Thereafter, the 3 regions were mapped to a single frequency value each. The full frequency mapping process is outlined in detail by Roginska et al., after this process, SoniScan output 9 frequency values, 3 for the 3 regions in the left hemisphere, 3 for the 3 regions in the right, and 3 for both left and right together. AD can be asymmetrical in where it acts, it can happen in the left, right or both hemisphere. Thus it was necessary to dissect these regions by hemisphere, and choose the 3 values that demonstrated the most deviation, drawing attention to the abnormalities. Otherwise it was

possible that, for example, an asymmetrical severe case would be heard as mild or moderate when the left and right hemispheres were sonified together. A sonification software tool was developed for this research project using the sound synthesis engine and programming language SuperCollider. The 9 frequency values were input into this program for interactive testing of the sonifications and output audio recordings. The final results were 32 30 second audio files, one for each PET scan.

4. EVALUATION

4.1. General procedures

The Triple-Tone sonification technique was tested by one highly experienced and one less experienced radiologist from the New York University Langone Medical Center. Testing took place over four sessions, each session included diagnoses of the same 32 de-identified PET scan cases. The order of presentation of the sonifications was consistently randomized across diagnoses sessions to minimize recall bias. The sessions were set up to mimic the diagnosis process that these radiologist would normally undertake in which the radiologists would diagnose the left hemisphere's frontal lobe, the left's parietal lobe, the right hemisphere's frontal and the right's parietal. The radiologists were given no information about the test case patient's gender, age or medical history. For each session, the diagnoses were recorded as well as the time it took for the full diagnosis of each patient. The first session was a basic visual diagnosis where the radiologists used the MIM software to search through the 2D slices of each test case. The second session was the advanced visual diagnosis session, which saw the addition of the 3D projection views of the brain. 3D projection views are generated by the MIM software. The data points in the 3D projection view represent the average intensity of voxels along the line of view from an external perspective, directly left of and directly right of the brain, to a depth of three millimeters. Physicians utilize this view of the brain to assess the overall picture regarding the diagnosis of a brain. The third session was the basic visual diagnosis with sonification, and the fourth session was the advanced visual diagnosis with sonification. All sonifications were played through Sennheiser HD650 headphones.

4.2. Sonification Training and Testing

The sessions with sonifications saw the addition of a training session. The training was exactly the same for both sessions. The first step of the training was to verbally introduce the radiologists to the general idea of sonification. The next step was to present them with specific examples of PET scan sonifications. To do this, 2 training case sonifications were played for each of the 4 AD diagnosis categories, 2 normal, 2 mild, 2 moderate and 2 severe. These sonifications were played for 30 seconds while the radiologist looked at the lateral slice related to that particular sonification. This was done to introduce them to the sound of each AD diagnosis category and to closely match the form of the test itself. The radiologists were informed that the sounds they were hearing were sonifications of a 2D lateral slice of the region of the brain where the greatest deviation of metabolic activity occurred. The next part of the training was the introduction of a graphical user interface tool that the radiologists could use during the diagnosis process. Created in Matlab, the tool consisted of an interface of 40 buttons divided into 4 columns; Normal sonifications, Mild AD

Questions	Physician 1 (High Exp.)	Physician 2 (Low Exp.)
The sonification was tiring to listen to and induced fatigue.	2	1
The sonification was pleasant.	2	2
The sonification provided additional information that the visual display did not.	3	5
The sonification was helpful in discerning between:	-	-
a) normal and mild cases	5	4
b) mild and moderate cases	5	4
c) mild and severe cases	4	4
The sonification made me more certain about my diagnosis.	4	4

Table 1: Physician questionnaire results

sonifications, Moderate AD sonifications and Severe AD sonifications. Each column had 10 buttons that would each play a sonification from that column's category. Ten sonifications were generated for each category, resulting in a total of 40 training cases. The 40 training cases were generated from the spectra of the deviation values of the 32 test cases. Specifically and for example, the 8 severe test cases had deviation values that went across a range. The 10 severe training cases were generated on average from that range. The ranges of the categories were displayed in order through the Matlab tool, where the smallest deviation in that category was at the top of the column and the largest was at the bottom. The radiologists were given a few minutes to test out the interface in order to familiarize themselves with more sonifications from each category as well as familiarize themselves with the tool they would be using in the diagnosis process. Directly after the training session, the radiologists were told to go about their diagnosis process as they would normally do, searching through the 2D slices and 3D projections visually. Only this time, the 30 second sonification would play along with the visuals. The radiologist could control the playback and volume of the sonification. They were also given the ability to play the training cases, by clicking the buttons of the Matlab tool to do AB comparisons with the brain scan sonification.

5. RESULTS

5.1. Physician questionnaire

After the final session, the radiologists were given a brief survey. They were asked a short series of questions on if they found the sonification helpful in the diagnoses process. Their answers were given on a scale from 1 to 5, where 1 = strongly disagree, 2 = disagree, 3 = somewhere in the middle, 4 = agree, and 5 = strongly agree. As shown in Table 1, the radiologists found the sonification

Physician 1 (High Experienced)

Sess.	LF	LP	RF	RP	W
0 Basic Visual	0.64***	0.84***	0.56**	0.8***	0.88***
1 Adv. Visual	0.39*	0.76***	0.37*	0.74***	0.81***
2 Basic Visual + Sonif.	0.69***	0.94***	0.66***	0.94***	0.98***
3 Adv. Visual + Sonif.	0.68***	0.79***	0.67***	0.81***	0.86***

Physician 2 (Low Experienced)

Sess.	LF	LP	RF	RP	W
0 Basic Visual	0.58***	0.70***	0.54**	0.64***	0.87***
1 Adv. Visual	-0.15	0.00	-0.01	0.12	0.03
2 Basic Visual + Sonif.	0.61***	0.89***	0.51**	0.83***	0.93***
3 Adv. Visual + Sonif.	0.58**	0.85***	0.48*	0.79**	0.95***

Table 2: Pearson’s correlation for association with the ground truth, ***significant at $p < 0.001$, ** significant at $p < 0.005$, * significant at $p < 0.5$

technique, although not pleasant, helpful in the diagnosis process, especially when it came to discerning between degrees of AD and in solidifying the certainty of their overall diagnosis.

5.2. Statistical Analysis results

During each evaluation, both physicians took part in four distinct sessions, basic and advanced visual diagnosis with and without sonification. For each one of the 32 unidentified brain scans of patient, eight cases per each severity level (normal, mild, moderate, severe), physicians analyzed four part of the brain: left frontal (LF), left parietal (LP), right frontal (RF), right parietal (RP). The results were then compared to the ground truth provided by Dr. Kent P. Friedman.

5.2.1. Regional Scores

For each combination of two physicians and four sessions there is a significant positive correlation between the ground and the

Physician 1 (High Experienced)

Sess.	LF	LP	RF	RP
Basic Visual	0.68	0.86	0.64	0.84
Advanced Visual	0.57	0.97	0.55	0.91
Basic Visual + Sonification	0.74	0.99	0.71	0.96
Advanced Visual + Sonification	0.87	0.96	0.85	0.92

Physician 2 (Low Experienced)

Sess.	LF	LP	RF	RP
Basic Visual	0.64	0.86	0.59	0.75
Advanced Visual	0.57	0.97	0.55	0.91
Basic Visual + Sonification	0.61	0.96	0.48	0.84
Advanced Visual + Sonification	0.52	0.85	0.44	0.77

Table 3: Pearson’s correlation for association with worst scores

frontal, parietal and worst regions. This is excepted for the scores reported by the less experienced physician during the advanced visual diagnosis (shown in Table 2), where there is no significant correlation between the ground truth provided and the regional or worst. Comparing basic and advanced inspections with and without sonification a significant increase of the correlation is shown. Since the highest significant correlation was always between the ground truth and the worst scores, accuracy was assessed in terms of concordance between the truth and worst results. Therefore

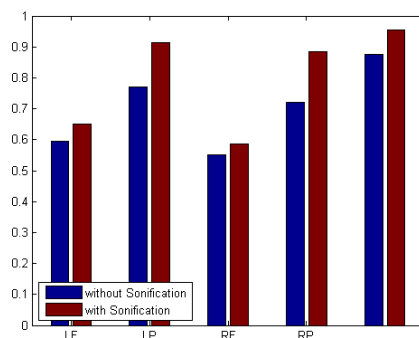


Figure 2: Basic visual with and without sonification for the two physicians.

the correlation for the association of the regional scores with worst scores is presented in Table 3. All correlations were significant ($p < 0.01$). The worst scores were most highly correlated with the

Sessions :	<i>Phys.1</i>	<i>High Exp.</i>	<i>Phys.2</i>	<i>Low Exp.</i>
	Lower	Upper	Lower	Upper
Basic Visual	41.9%	74.7%	41.9%	74.7%
Adv. Visual	41.9 %	74.7%	13.0%	42.6%
Basic Visual + Sonification	78.2%	98.4%	64.1%	91.3%
Adv. Visual + Sonification	64.1%	91.3%	67.5%	93.3%

Table 4: Concordance percentage of ground truth and worst scores and left parietal scores

scores from left parietal region, highlighted in bold. According to these results, looking at the parietal elements highlights more efficient scores. Thus, this result suggest that the worst scores may be approximated using only parietal region scores.

5.2.2. Accuracy Percentage

Logistic-Regression was used to characterize and compare sessions in terms of the accuracy of the worst scores for each subject. Accuracy was defined as the percentage of times the worst scores were concordant with the ground truth. Generalized estimating equations (GEE) is generally used to estimate the parameters of a generalized linear model with a possible, unknown, correlation between outcomes and it is based on binary logic regression. In this work it has been used as a measure of agreement between the ground truth and the worst and left parietal region. The lack of statistical independence among results from the same subject (i.e., results for the 4 sessions) was accounted for by assuming results symmetrically correlated when acquired from the same subject and independent when derived for different subjects. Output included a 95% confidence interval (CI), an estimate that calculate interval- of possible values of unknown parameters, for the percentage of times each session produced worst scores that were concordant with the ground truth. The percentage % and number (#) of times the worst scores and the left parietal scores were concordant with the ground truth, is shown in Table 4.

6. DISCUSSION

A sonification method throughout which medical researchers have access an additional way of interpreting data has been presented. This method can assist more physicians detecting something that may be missed with traditional visual inspection. To evaluate this technique, the correlation between each evaluation method has been proposed in this study and the correct diagnosis based on the ground truth has been computed to measure whether sonification helped improve accuracy in the diagnosis. Results from the statistical analysis show a higher correlation between tests using sonification and the ground truth. During the basic visual testing, one physician achieved an accuracy of 56-88% while the other performed at an average of 50-55%. For each session, the physicians performed more accurately with an average improvement of about 10% using the added sonification method. For the more severe

cases, the improvement in accuracy was as high as 30% with sonification added. The results of the evaluation show that subjects with untrained ears are able to categorize the sonifications generated by the triple-tone technique with accuracy.

7. FUTURE WORK

Future purpose for the Triple-tone sonification method will extend its data scope to spatial representation of the audio. It is theorized that by sonify by one slice to a larger subset of the brain, a more accurate description of the brain will be reached. Furthermore, by extending the spatial scene it may provide more information from the data. Each DICOM data set is a 3 dimensional matrix that contains 60 slices portraying the full brain. In the current triple tone sonification method one subset of each dataset was chosen for sonification: the 30th lateral slice (from the top) of the spatially normalized dataset. This slice was segmented into three lobes of interest (frontal and parietal) and the reference lobe (sensorimotor cortex). After data was extracted from each relevant lobe and frequencies per lobe were allocated, physicians were presented with a mono audio file, representative of the 30th slice of DICOM data, sonified by three frequencies. The beating patterns created by the three frequencies reveal the severity of the case. To extract more slices from the DIACOM data set, modifications to the SoniScan program were implemented. SoniScan is the C++ program used to perform slice selection, lobe segmentation and aid calculations for frequency allocations. Due to the physical makeup of the brain, coordinates for lobe segmentation will be individual to each slice. Segmentation information was performed with the aid of the MIM software. Each slice and its relevant coordinates were placed into a 2D matrix for the SoniScan program to read as per user defined slice number. The contours were approximated to straight lines connecting the coordinates as in the original program. Currently, physicians are presented with a mono audio file containing the three frequencies. To better represent this information, binaural renditions of case studies have been created. The binaural examples allow for multi-slice playback, and individual lobe isolation, all of which are user defined via a GUI. The frontal lobe is spatially placed at azimuth: -5 to +5 and parietal lobe at azimuth 175 to -175. Elevation of each slice is determined by how many slices there are during playback. The more slices rendered, the closer each slice will be to one another in elevation. Distance of each sound source is 1 meter. The Head Related Impulse Responses used to for the convolution process were extracted from the MARL-NYU file format for storing HRIRs.

8. CONCLUSIONS

Initial validation for the the triple-tone sonification method has been presented in this paper. The promising results obtained from this study highlight more possibilities in helping AD diagnosis. In fact, through spatialization analysis of the resulting slices and their sonification using this technique, radiologists can get more information to investigate a larger region of the brain. Given a 3D representation aurally also presents an opportunity to create a virtual space for the brain with which the listener can interact. The next steps involve possibly working with 3D audio and enhancing the analysis and/or experience of diagnosing AD. Furthermore this sonification technique can be enlarged to existing visualization imaging techniques, such as MRI, in order to bring considerable improvement to medical diagnosis area. Therefore, the proposed

method is broadly ready and now applicable to a number of different methods with a promising future.

9. ACKNOWLEDGMENT

This work was supported by the NYU CTSI Award Number NIH/NATS UL1 TR000038.

10. REFERENCES

- [1] G. Baier, T. Hermann, The Sonification of Rhythms in Human Electroencephalogram, in Proc. of the 10th Int. Auditory Display (ICAD 2004), Sydney, Australia, July 2004.
- [2] G. Baier, T. Hermann, U. Stephani, Multi-channel sonification of human EEG, in Proc. of the 10th Int. Auditory Display (ICAD 2007), Montreal, Canada 2007.
- [3] R. S. Frackowiak, C. Pozzilli, N. J. Legg, G. H. Du Boulay, J. Marshall, G. L. Lenzi, Regional Cerebral Oxygen Supply and Utilization in Dementia, *Brain* (104), 753-778.
- [4] J. W. Piper, Validation of an Artificial Neural Network Computer-Aided Diagnosis Scheme for Alzheimer's Disease using FDG-PET, Wake Forest University, Department of Computer Science, 2007.
- [5] J. Oh, Y. Wang, A. Apte, J. Deasy, A Statistical and Machine Learning-Based Tool for Modeling and Visualization of Radiotherapy Treatment Outcomes, *Medical Physics* (2012), 39 (6), 3763.
- [6] A. Roginska, K. P. Friedman, H. Mohanraj, Exploring Sonification for Augmenting Brain Scan Data, in Proc. of the 19th Int. Conf. on Auditory Display (ICAD 2013), Lodz, Poland, July 2013.
- [7] T. Kagawa, S. Tanoue, H. Kiyosue, H. Nishino, A Sonification Method for Medical Images to Support Diagnosis Imaging, in Proc. of the 7th Int. Conf. on Complex, Intelligent, and Software Intensive Systems (CISIS), Taichung, China, July 2013.
- [8] A. Roginska, H. Mohanraj, M. Ballora, K. P. Friedman, Immersive sonification for displaying brain scan data, in HEALTHINF 2013, 6th International Conference on Biomedical Electronics and Device, Barcelona, Spain, February 2013.
- [9] A. Roginska, H. Mohanraj, J. Keary, K. P. Friedman, Sonification Method to Enhance the Diagnosis of Dementia, in Proc. of the 20th Int. Conf. on Auditory Display (ICAD 2014), New York, New York, June 2014