Sonification Method to Enhance the Diagnosis of Dementia

Agnieszka Roginska – New York University Steinhardt School

Hariharan Mohanraj – New York University Steinhardt School

James Keary – New York University Steinhardt School

Kent Friedman – New York University School of Medicine

Hello everyone, my name is James Keary, and I am part of the Medical Imaging Sonification team here at New York University. I am presenting on our paper, Sonification Method to Enhance the Diagnosis of Dementia. This is a joint project between the Music Technology department at Steinhardt, headed by Prof. Agneizska Roginska, and the Department of Radiology at the Langone Medical Center, at the School of Medicine. This interdisciplinary study combines sonification with the field of medical imaging infomatics.

Perception in the diagnosis and management of disease

- Physical exam
  - Visual
  - Touch
  - Smell
  - Sound
    - Augmented by the stethoscope

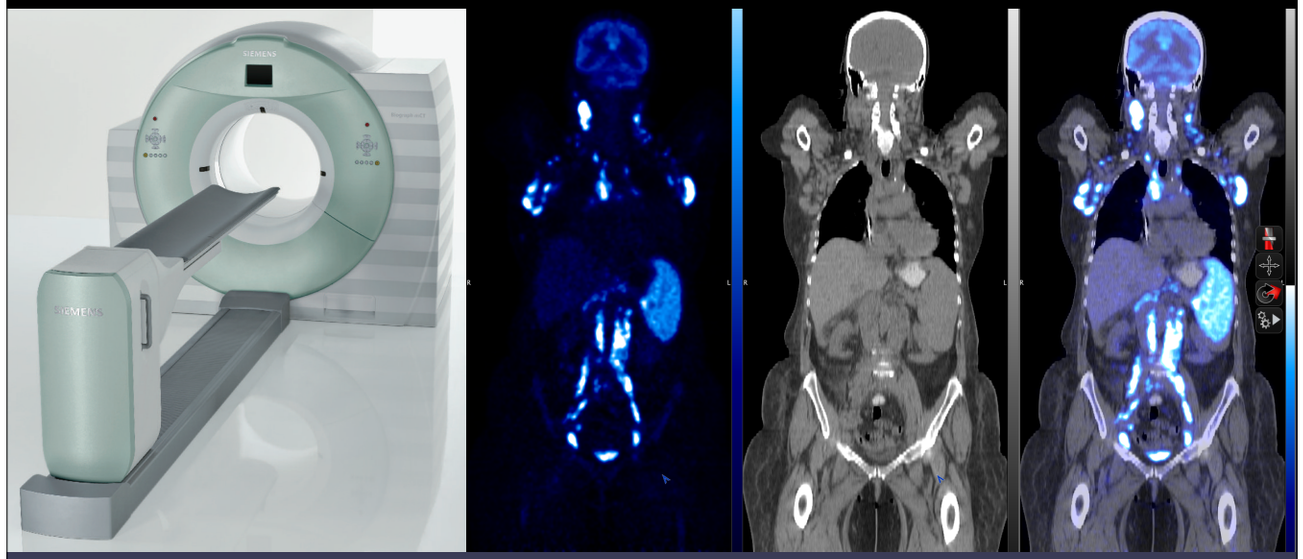
Sound is not necessarily a new or novel approach in the field of medicine. Ever since the stethoscope, doctors have attempted to use all their senses in determining diseases. And because there is a huge medical component to the ICAD conference this year, I will not go into much of the applications of sonification in medicine; many of those people are sitting in this very room.

Alzheimer's Dementia

- Affects more than 5M Americans
- Difficult to diagnose early stages
- Can be slowed down if diagnosed early
- Medical Imaging Techniques help us diagnose:
  - X-Rays
  - Medical Resonance Imaging (MRI)
  - Positron Emission Tomography (PET)

I will instead delve into the area that *we* focused on which is Alzheimer's Disease. Alzheimer's disease (AD) is a form of dementia that is common in the elderly. It is an incurable and it is difficult to diagnose. The up side is we do know that if we can diagnose it early then we have a better chance of slowing it down. So how do we diagnose it early on? We use medical imaging techniques X-rays, Medical Resonance Imaging (MRI), and Positron Emission what (PET) imaging techniques.

PET/CT Scanner



SUGAR METABOLISM (NUCLEAR - PET)

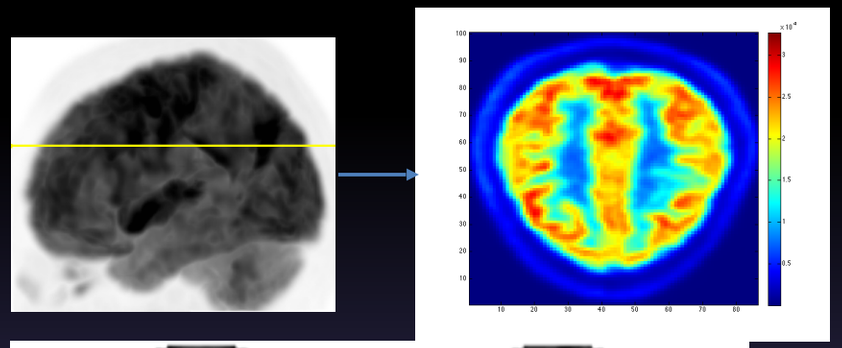
TISSUE DENSITY (CT SCAN)

BOTH (PET/CT SCAN)

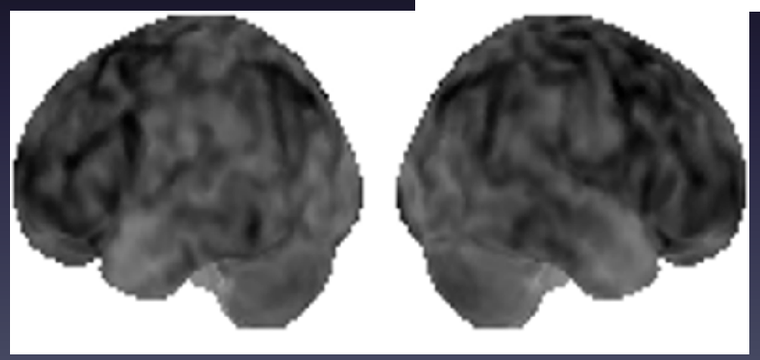
So, what we have here on the slide is what is known as a Positron Emission Tomography / Computed Tomography scanner. As well as some scans demonstrating the types of things that can be explored with such scanners. Now, its not entirely necessary to delve into the sub atomic chemistry behind how these machines work, but it is important for us to know what we are working with on a basic level. So, in broad strokes, the PET scanner is picking up radiation measurements. There is not a whole lot of radioactive radiation coursing through our bodies at all times, so a tracer element is injected into the patient before hand. That tracer element causes a low level chemical metabolic reaction that releases radiation.

The Data Sets

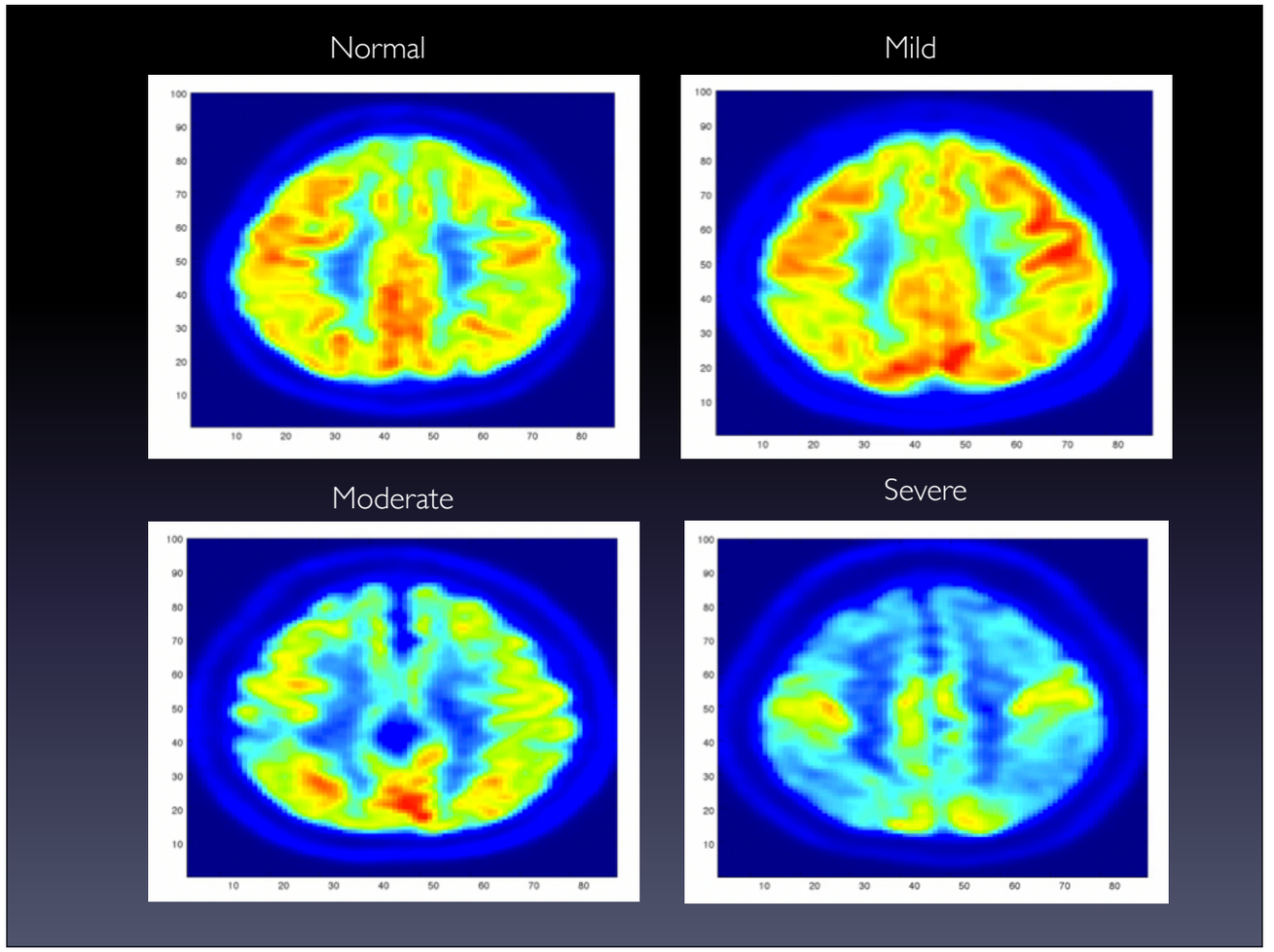
- Lateral



- 3D Projection



These radiation values picked up by the scanner can be reconstructed into a 3D data set of radiation intensity values. Normally in radiology infomatics, this gets mapped to a 3D or 2D color image, where red is high intensity and blue is low. Using medical imaging software, such as OsirX or MIM, doctors can take what is known as a *lateral slice,* a 2D sub section of the 3D data set pertaining to the median, or horizontal, plane of the body, and look at the intensity values. Doctors can also change the color palette and hue in these programs to examine the intensity values further.



What I am showing here are the four categories or stages of AD as those in the radiology informatics field have defined. In just a quick visual analysis of these 4 categories, the more important question is not what does severe AD look like, but how do we tell *when* AD starts, specifically when it goes from Normal to Mild. Or what is mildly abnormal? What is moderate, even? These are the important questions.

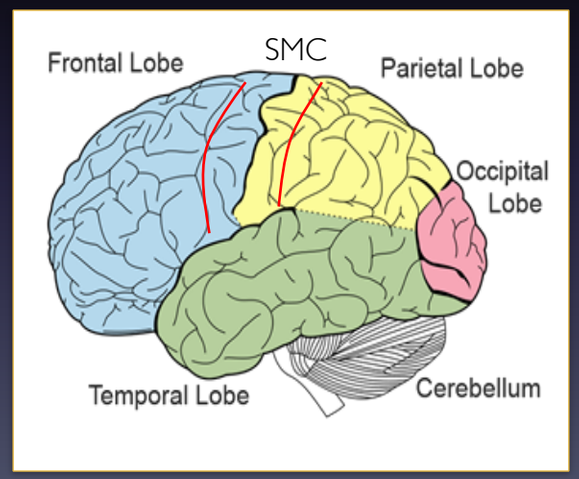
Challenges in Visual Interpretation

- Abnormalities may not be seen at all or not recognized
  - Perceptual error (“missed” findings)
  - Cognitive error (misinterpreted findings)
- What is normal?
- What is mildly abnormal?
- What is severe disease?
  - Subjective assessment of disease severity is variable

Despite all these technological advancements, the assessment of disease severity is still a subjective measurement and variable among doctors. The difference in doctor’s experience, and the intra reader variability among doctors is still an issue causing perceptual and cognitive errors. With the advancements in computers, the images have become more and more resolute with bigger and bigger data sets, and with more data, means more types of analysis besides just visual, quantitative analysis is performed, and of course this leads us to sonification.

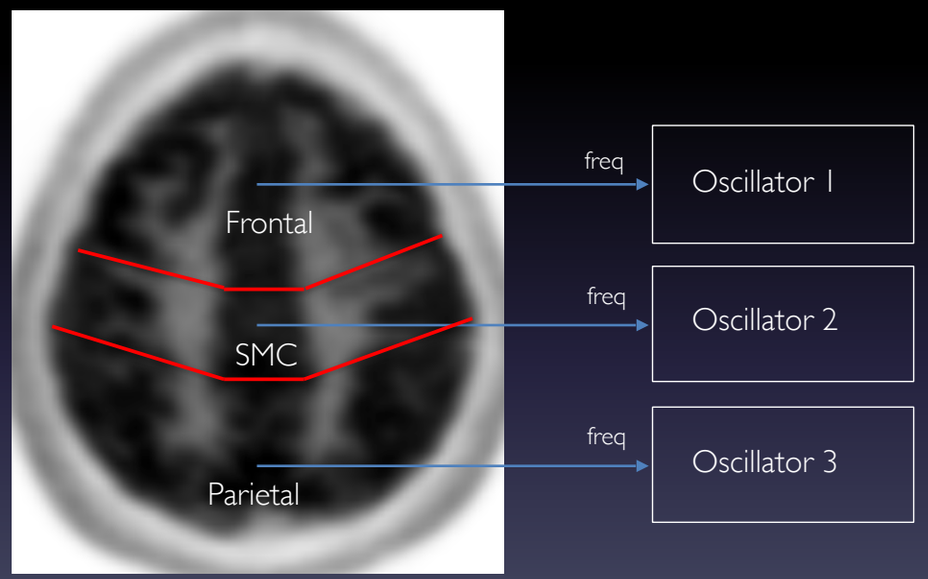
Alzheimer's Dementia Characteristics

- Decreased hypometabolism: parietal, temporal and frontal.
- Unaffected: sensorimotor cortex



We tested many other sonification methods, and what we settled on for this stage of research was to look at how doctors were making diagnosis already to build upon that. Alzheimer's Disease is characterized by decreased metabolic activity primarily in the parietal lobe and frontal lobe, while generally leaving the sensorimotor cortex and occipital lobe. In order for a sonification technique to directly target the diagnosis of Alzheimer's disease, this difference in metabolic activity between the lobes of interest (frontal and parietal) and the reference lobe (sensorimotor cortex) should be mapped to an obvious auditory parameter.

Triple-Tone Sonification



We call this technique the triple-tone sonification because we were sonify-ing 3 lobes, the frontal, the smc, and the parietal lobe of one lateral 2D slice, by assigning one oscillator to each lobe. These oscillators are assigned the same base frequency by default. Then, the frequency of each oscillator is “detuned” from default according to the metabolic activity of the corresponding lobe with respect to that of the reference lobe. The auditory parameter that is mapped to is the rhythmic beating pattern produced by three oscillators at slightly different frequencies. This leverages humans aural sensitivity to beating patterns for the purpose of diagnosis.

Frequency mapping

$$\Delta f_{FL} = \left| f_{default} * DF * \frac{av_{FL} - av_{SMC}}{av_{SMC}} \right|$$

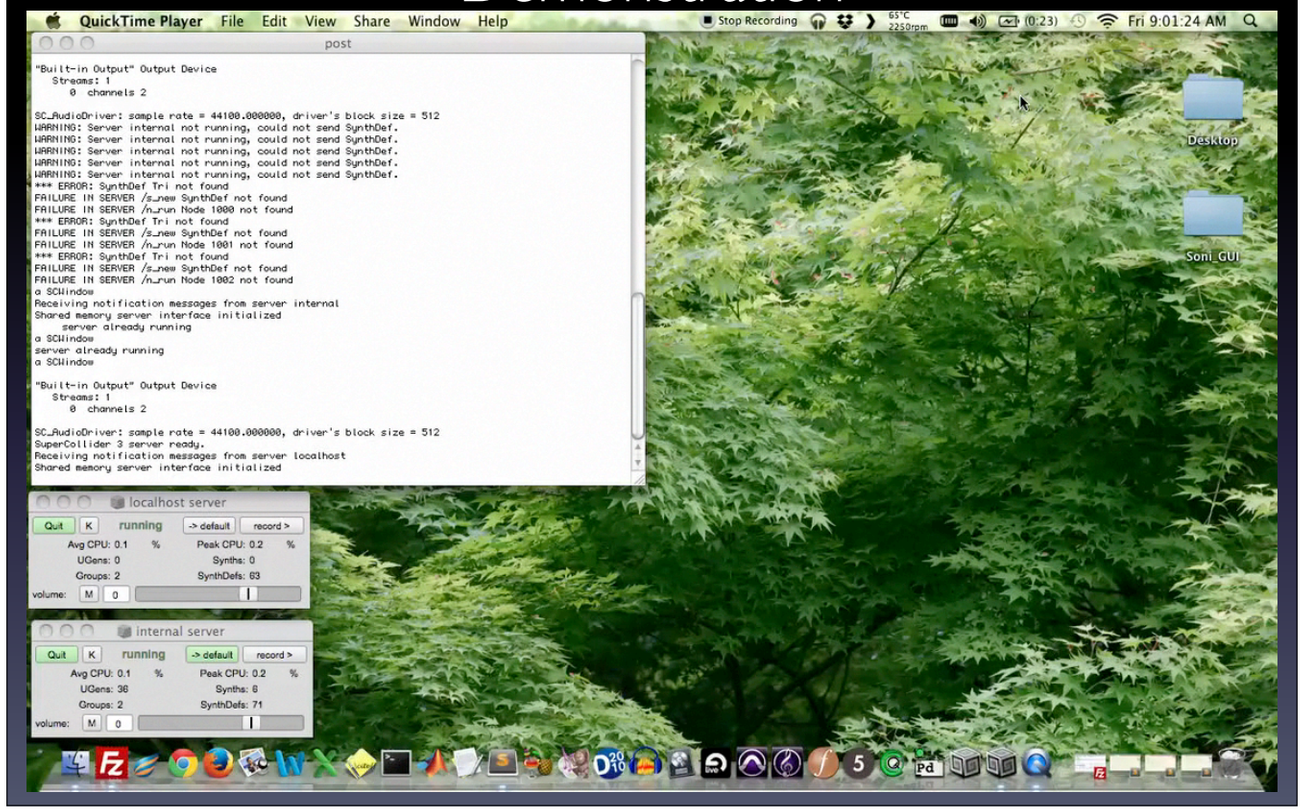
$$\Delta f_{PL} = - \left| f_{default} * DF * \frac{av_{PL} - av_{SMC}}{av_{SMC}} \right|$$

The metabolic activity of a lobe is calculated by computing the average intensity of the voxels of the PET scan lying inside the region that defines that lobe.

Relative deviation of average intensity with respect to the average intensity of the sensorimotor cortex (SMC) was linearly mapped to the relative deviation of the oscillator frequency with respect to the oscillator's base frequency through the detune factor coefficient.

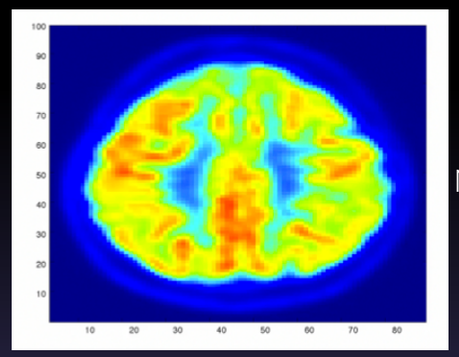
The detune factor was used to control the range of frequency deviations given the range of voxel average deviations. The goal of this parameter was to find different levels of beatings to indicate the varying degrees of AD: For example finding the level of beating that would generate split tones indicating severe cases of AD. A higher detune factor would result in a larger frequency deviation, and hence faster beating, for a set of voxel intensity averages.

Interactive Sonification with Supercollider Demonstration

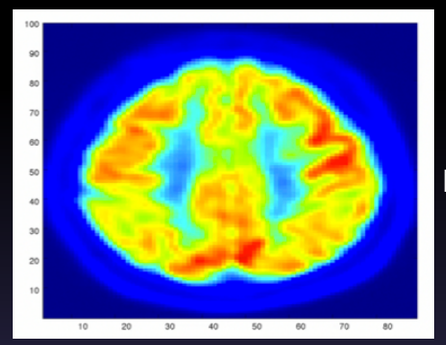


Here is a video demo of the interactive prototyping tool we developed in the Supercollider programming language. It needed to be interactive because the process was akin to defining colors of a crayon box, what is red? what is fire truck red versus brick red? All the parameters of what defined the categories of AD, normal, mild, moderate, and severe, had to be defined through this prototyping process.

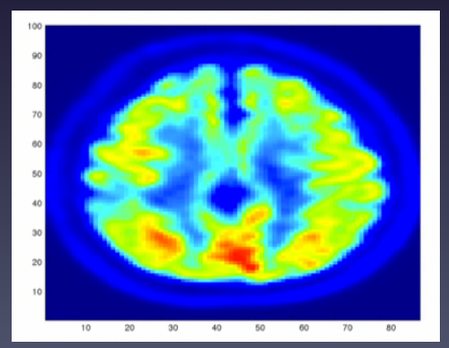
Diagnosis



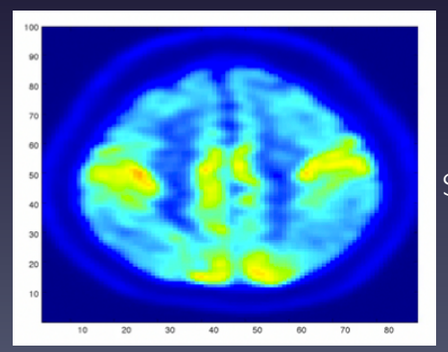
Normal



Mild



Moderate



Severe

A base frequency of 440Hz was chosen, and the oscillators were triangle waves. So here we have some examples of how we defined the auditory definitions per category. With the normal, non AD brain, we have a regular pattern. Mild - irregular arrhythmic, Moderate - the space ship quality as the tones start to split from each other, and Severe we have full on dissonance as the notes have fully diverged.

Phase 1: Evaluation

- 8 x 4 categories = 32 sonifications
- 10 x 4 categories = 40 training sonifications – Statistically generated
- 5 subjects – Music Tech graduate students and faculty
- 2 sessions:
  - Session 1 – Coarse Categorization: Categorize as 1, 2, 3, 4
  - Session 2 – Fine Categorization: Categorize as 1, 1.5, 2, 2.5, 3, 3.5, 4
- Evaluation goals:
  - Accuracy of categorization
  - Intra-reader consistency of categorization
  - Resolution of sonification

The evaluation had 5 subjects, all graduate students and faculty members of the Music Technology program, Steinhardt, New York University. This control group was chosen to first evaluate the triple-tone sonification technique with trained ears as a validation, before proceeding to Phase 2, evaluation by physicians. In Phase 1, all medical aspects of this sonification technique were withheld from the subjects, and their evaluation of this technique is influence solely by auditory parameters. We had 32 sonification files, 8 from each category, given to us by the Langone Medical Center, the ground truth was already established for those cases. There was a training session of 10 sonifications generated for each category, resulting in a total of 40 training sonifications. The listeners were allowed to use the training cases as reference throughout the process of categorization. There were 2 rounds of categorization, 1 where they were told to rate categories 1 through 4, and then a round of finer categorization where they had 7 ratings they could make, 1, 1.5, 2, 2.5, 3, 3.5, and 4.

The first goal of the evaluation is to evaluate the effectiveness of the triple-tone sonification technique in accurate distinguishability between brains of different levels of Alzheimer's disease. The second goal of the evaluation is to evaluate the intra-reader consistency of diagnosis. The third goal is to evaluate whether the sonification technique provides a finer gradation that improves diagnostic accuracy.

Results: Accuracy

Coarse Categorization

	Category 1	Category 2	Category 3	Category 4	All Categories
Subject 1	100%	87%	81%	87%	**89%**
Subject 2	87%	93%	37%	93%	**78%**
Subject 3	93%	100%	81%	100%	**93%**
Subject 4	100%	93%	75%	100%	**92%**
Subject 5	93%	93%	43%	100%	**82%**
All Subjects	**95%**	**93%**	**63%**	**96%**	**87%**

Fine Categorization

	Category 1	Category 2	Category 3	Category 4	All Categories
Subject 1	87%	87%	81%	87%	**85%**
Subject 2	100%	100%	56%	100%	**89%**
Subject 3	100%	100%	100%	100%	**100%**
Subject 4	93%	100%	87%	100%	**95%**
Subject 5	100%	100%	68%	100%	**92%**
All Subjects	**96%**	**97%**	**78%**	**97%**	**92%**

The accuracy of response against ground truth was computed for each ground truth category for each subject.

1 - categorization seems accurate just based on the audio

2 - finer quantization improves the categorization

These results show that the subjects were able to categorize the sonifications with a very high degree of accuracy. This shows that the sonification technique is successful in being differentiable by trained listeners. This suggests that in the case that untrained listeners are unable to perform the categorization with a high degree of accuracy, they can be trained to improve their accuracy to high levels.

Results: Intra-reader consistency

	Coarse Categorization	**Fine Categorization**
Subject 1	84%	88%
Subject 2	84%	100%
Subject 3	94%	100%
Subject 4	97%	88%
Subject 5	84%	88%
All Subjects	**89%**	**93%**

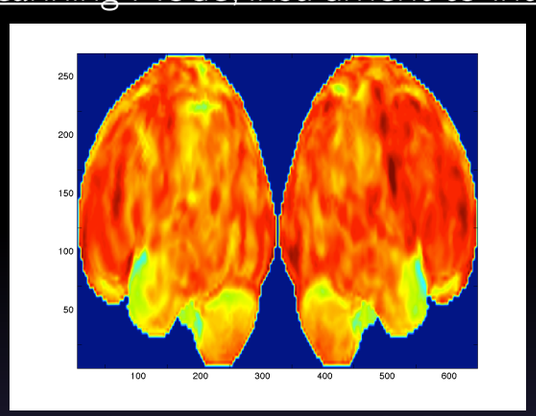
And intra-reader consistency was up as well with finer gradation.

Phase 2: Physician evaluation

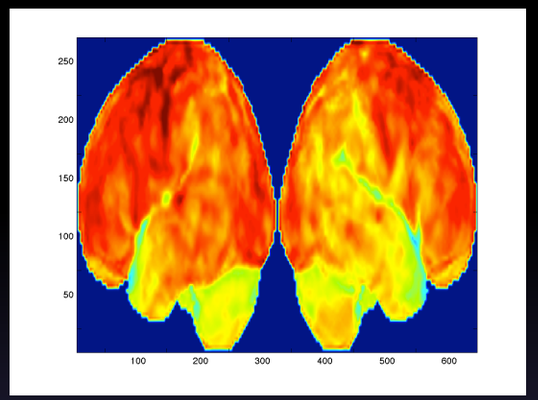
- 2 subjects – differing experience
- 4 sessions per subject
  - Basic visual
  - Advanced visual – with 3D projection map
  - Basic visual + sonification
  - Advanced visual + sonification
- Currently analyzing results
- (Looking very promising!)

Now, the next phase of the testing is Physician evaluation, and this work has been done on 2 subjects, radiologists from the Langone Medical Center of differing experience level. For these testings, there were 4 different sessions. The first session consisted of 40 cases to be diagnosed through the basic visual standard methodology that is currently implemented with the MIM software. Doctors flip through a bunch of 2D slices and give a diagnosis. The second session was like the first but this time with the addition of the 3D projection maps. Then, in the next two rounds the sonifications were added. Physicians were asked to diagnose and the time it took to diagnose was recorded. Currently these results are being analyzed by a medical statistician.

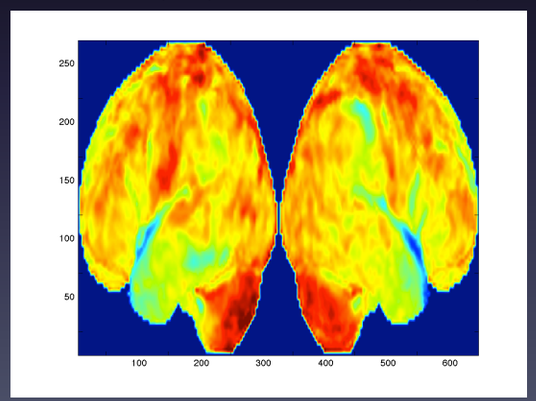
Scanning Mode, Instrument to Intensity Mapping, Perspective Slices



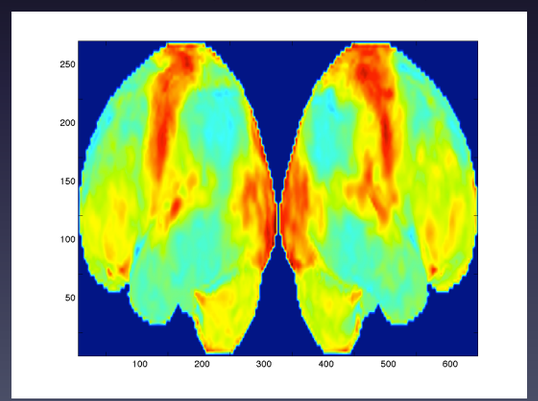
nor_KF01_perspec_Inst2Int_CbPnpFltHrp



mil_DO01_perspec_inst2int_CbPnoSynFHrn



mod_PM01_perspec_inst2int_SynPnoVbsHrp



sev_HB01_perspec_Inst2Int_PnoSynClntVln

If time

Here are some examples of former modes of sonification that were discarded. They were interesting, yet did not really contain any information that we deemed helpful for diagnosis.

Acknowledgements

- This work was supported by the CTSI Award Number NIH/ NCATS UL1 TR000038

http://marl.smusic.nyu.edu/projects/sonification

Thank You