

Christopher Godley

May 4, 2018

CSCI 5602 - Information Visualization

Danielle Szafir

Project 4: Experimentation

Part 1: Designing an Experiment

In music production, a waveform representing sound level in decibels is used to display the contents of a track. This is great when tracking music, as it allows quick and efficient management of sound levels from the hardware so that all the tracks can maintain a similar sound floor. The problem comes when mixing these tracks, as the information is largely useless at this stage. Identifying when snare hits happen or when the singer starts singing can still be useful information, but most studio tracks are recorded to a click and quantized to remove variance. This effectively means that you shouldn't need this information after tracking is complete. While mixing, however, the goal is to balance not only the sound levels, but the sonic space of each track. Removing low frequencies from a female vocalist may clean up those frequencies in the mix where clarity is needed, and helps each track shine through the mix. This is almost entirely accomplished by ear, as visualizing frequency isn't always easy. That's where spectrograms come in. A spectrogram plots frequency along a y-axis as time moves in the x-axis, with a heat map applied for decibel levels. A comparison of these two visualizations can be seen below.

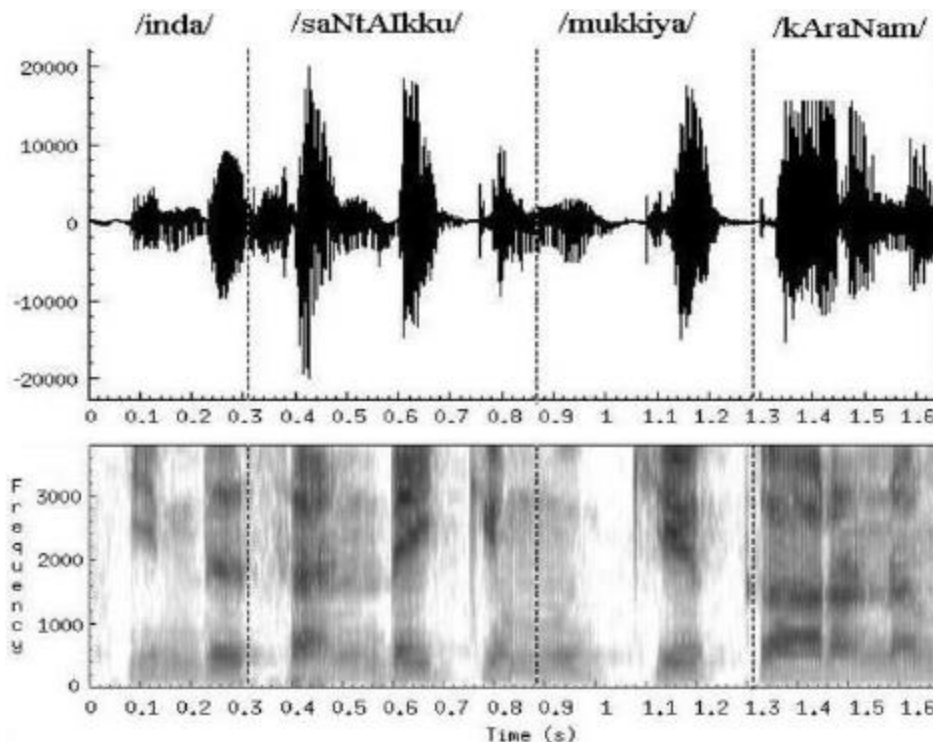


Figure 1: Thomas, Samuel & Nageshwara Rao, M & Murthy, Hema & S. Ramalingam, C. (2012). Natural sounding TTS based on syllable-like units.

The hypothesis I propose is that it is more efficient to mix audio using a spectrogram as the visual representation of tracks as opposed to waveforms. To test this hypothesis, a rather complex experiment must be designed. Effectively every piece of audio production software uses the waveform to represent recorded audio, so changing this will require modifying an open source DAW (digital audio workstation). The independent variable in this experiment will be whether waveforms or spectrograms are used. The dependent variable will be the time to generate a mix the user is satisfied with. This is our first control variable. There must be some kind of standardization amongst users of the expectation of how good the mix should sound. Since this is a measure of efficiency, time is the critical factor, but rushing to meet this measure will result in poor data. That is why only experience audio professionals are acceptable candidates for this experiment, as they will already have a certain quality they strive for instilled in their workflow. This is also why it'll be important to keep the time constraint secret from the participants. Users given spectrogram programs will be told they are trial running a new piece of software, while users given waveforms may be told the study is focused on the qualitative appeal of the music being mixed.

The modified DAW is a severely limiting aspect of this experiment. A major part of mixing and producing audio recordings lies in the workflow. Most mainstream DAWs have effective parity in their capabilities, while the workflow, how screens are laid out, how tracks are managed, and how plugins are loaded may vary greatly among them. Due to this, participants will also be difficult to find. A single DAW should be chosen for modification, as the modification process will not be trivial, and the rest of the DAW needs to maintain the critical workflow the participant is otherwise familiar with. This is why a remotely run experiment may be best. There are likely many forum users on various DAW help sites that would be willing to participate, but we will likely need to automate the deployment of the experiment.

To run this experiment remotely, it will be necessary to present the project as consistently as possible. A single recording project may be used by all participants, keeping musical variance away from the experiment. This project will need to have a timer running whenever it is open, so that total mix time can be measured without moderation. This will have to be told to participants, so that they do not leave it open while they sleep for example. I believe we can inform users of this without compromising the experiment by labeling it as more of a logistical necessity. Once modified, the DAW can be sent with the project to willing participants who will be randomly assigned either waveform or spectrogram visualization of the track view. Each participant will mix the track to the best of their ability and send the final mixdown back, along with the project metadata including the timing. Other metadata could be gathered as well, and actually implementing this experiment will surely reveal many more interesting metrics to gather. A qualitative review of those using the spectrogram representation will also provide testimonials to be evaluated for any significant data.

To analyze these results, simply plotting the mix time of those with spectrogram against those with the traditional waveform will provide an immediate answer to whether this saves time or not. More interestingly, a spectral analysis of the final mixdowns and individual project stems may reveal differences in how frequencies and sonic spaces were managed. Using various sound processing techniques, we can identify modifications to the frequency space of each individual track in the project and compare those with spectrograms vs waveforms to see if one group was more or less heavy handed with their equalizing. An 'optimal' solution, or multiple of them, may also be compared to see if either group was more 'accurate' in their EQ decisions.