Brian Hansen
MAT 200C
Stephen Travis Pope
Spring 2010

Modeling Spacialized Sensory Dissonance

Issues of consonant and dissonant sonorities in music have defined its compositional practices for centuries. Musical issues continue to arise as new compositional styles and musical vocabularies emerge, constantly forcing composers and listeners to reevaluate what is consonant, what is dissonant, and how are musical relationships formed between the two. Contributing to our understanding of consonant and dissonant sonorities is the quantification of sensory dissonance. There has been much research done in developing a method to quantify the dissonance between two tones. All methods consider the physical and psychoacoustical aspects of sonic perseption. However, these models are typically without dimension, as they do not consider sound as it occurs in space. This paper proposes a method for calculating the dissonance between sounding tones, taking into consideration their spacial relationship to the listener.

Throughout music history, there have been numerous perspectives on how to define consonance and dissonance. Pythagoreas had a mathematical approach, claiming tones with a ratio of small whole numbers were more consonant than those with a higher ratio. Helmholtz viewed the ear as a spectral analyzer, positing that the ear recongnizes when two tones share harmonics, and tones sharing more harmonics are more consonant than tones not sharing any at all. Terhardt claimed that consonance corresponds to fusion or pattern matching, where consonance is defined by how close two tones may fit into a psychoacoustically posited sub harmonic. Many disregard any scientific approach altogether claiming that the issue is completely contrived. They claim dissonance is a sound that simply sounds bad, so if one doesn't like a song or sound they simply label it as dissonant. Contributing to the debate about what consonance or dissonance really is, there has been extensive research on what is labelled as *sensory dissonance*. This type of dissonance is defined by the amount of roughness that is present in two sine tones separated at a given interval. In 1965, Plomp and Levelt explored this phenomena and showed how dissonance could quantifiably be related to the distance between two frequencies. This core discovery has been adopted by many and used in modeling the dissonance of sonority.

When calculating sensory dissonance, it is not enough to simply measure the distance between two tones. First of all, we can't merely consider the fundamentals of two tones. A given tone is not simply one sine tone, but a combination of sine tones with specific frequencies and intensities in the tone. Thus, we must consider the spectral components of two tones when calculating the sensory dissonance between them. In addition, due to psychoacoustics, our auditory mechanism skews how frequencies and intensities are heard. Thus, any measured frequencies and intensities must be recalibrated according to psychoacoustical properties before their sensory dissonance is measured. Considering these factors, sensory dissonance can be calculated by the following method:

1. Identify the sinusoidal components of a given sonority by computing a fast fourier transform of the sound. Use the fast fourier transform to estimate the true frequency of the sinusoidal components in the sound spectrum and their corresponding intensities.
2. As sensory consonance has psychoacoustical properties, all frequency components in the spectrum must be converted to the bark scale to compensate for the psychouacoustical effects of the critical bandwidth.
3. Again in accordance with psychoacoustics, all corresponding intensities must be converted to phons utilizing equal loudness contours.
4. Using Parncutt's approximation to the dissonance curve defined by Plomp & Levelt, calculate the dissonance between each sinusoidal component in the sound spectrum. The equation is shown below:

$$\text{Dissonance} = \{ 4 * (\text{interval}) * e^{[1 - 4 * (\text{interval})]} \}^{2}.$$

5. Assign a weight to each dissonance calculated based on the phon level present between the two sinusoidal components. In addtion, assign all tones masked by the intensity of other sine tones a weighting of 0, effectively excluding them from the calculation. This must be done because such sine tones will not be perceived and thus no dissonance will be heard.
6. Finally, the total dissonance present in the sonority will be the weighted sum of all dissonances present in the spectrum.

This model of dissonance calculation works when considering sound outside of spacial dimension. However, in reality we hear tones from different locations. Some are above us while some are below. Some are behind us and some are in front of us. Thus, we need to consider how the spacialized location of a sound is transformed by the auditory mechanism. The auditory mechanism essentially filters the sonic spectra that enters the inner ear depending on the sounds origin with respect to the listener. Since the sound is filtered, its spectrum is altered having consequences on the perceived frequencies and intensities present, which in turn affects sensory dissonance. This must first be considered before the sensory dissonance is calculated, and head-related transfer functions can be of aid.

Head-related transfer functions (HRTFs) are used to describe how the brain, inner ear, and the pinna work together to derive information about the spacial location of a sound source. It is posited that this collaboration may have developed in humans as an evolutionary necessity. Since the eyes can only focus on one direction, it is up to the ears to detect sounds continuously in all directions. This ability allows us to more completely sense our entire surroundings, and can alert us to potential dangers that are out of sight.

The HRTF can be used to model how a given sound wave is filtered by the diffraction and reflection properties of the head, pinna, and torso, before the sound reaches the machinery of the inner ear. The geometrical relationship of the sound source to the listener can be used as parameters to model how the HRTF filters an incoming sound. For example, the elevation of a sound, its distance from the source, and its horizontal position act as parameters for the HRTF filtering of the sound.

The main problem with the HRTF model is that each individual's is unique. This is because we all have uniquely shaped ears, that are uniquely positioned on our bodies. Thus, one person's HRTF is not equal to another's, and the HRTF can not perfectly model spacialized sound for everyone. Nevertheless, we can make generalizations. For example, we know everyone's ears at least point forward, and everyone has ears on their head and not on their feet. Such generalizations may sacrifice precision for the individual, but they can help us move closer to modeling the spacialization of sound. We can also adopt this notion when modeling sensory dissonance.

The sensory dissonance model detailed above can be extended to consider sound spacialized in a three dimensional environment. This can be accomplished by incorporating HRTFs into the model. Thus, to effectively model spacialized sensory dissonance, our procedure is as follows:

1. For a given sound, apply HRTF filtering to it based on geometrical parameters of its spacial locatoin with respect to the listener.
2. Identify the sinusoidal components of the sonority by computing a fast fourier transform. Use the FFT to estimate the true frequency of the sinusoidal components in the sound spectrum and their corresponding intensities.
3. Complete steps 2 through 5 as outlined above.

When comparing this revised approach to the original method, the dissonance value computed should be altered. This should be caused by the filtering parameters of the HRTF routine altering mainly the intensities of the sound spectrum. With the intensities of each sinusoidal component altered, the masking of some sinusoids may change, and a new weighting distribution will be applied to the weighted sum of dissonance in the spectrum.

The practice of computing spacialized sensory dissonace has the potential for numerous applications. It can aid music and sound design from both an analytical and compsoitional perspective. For example, composing contrapuntally is largely based on the level of consonance and dissonance in a sound and how such sounds relate to each other. The computation of dissonance could be very informative in analyzing the antiphonal choral works of the late Rennaisance Venetian School. in addition, there are numerous modern composers that have written works to be performed in a highly spacialized setting. Many of these works are recorded, but the experience is never the same when listening to a recorded vs. a live rendition. Computing the dissonance may point to certain traits of the sound that is more or less clear when spacialized vs. not. In addition, compositional sound modeling could benefit from the calculation. The calculation could be used to determine if and what levels of dissonance are more or less pleasing with respect to spacial location.

In summary, consonance vs. dissonance has always been a defining issue for compositional practices throughout music history. The calculation of sensory dissonance can be very informative and useful when applied to music composition, analysis, and sound design. However, up to this point the calculation has not considered spacialization of sound sources. Thus, incorporating HRTFs into the procedure of quantifying sensory dissonance allows for its extension into three dimensions. This may allow extended applications of composing, analyzing, or designing sound in three dimensions.

Bibliography

Cheng, C., Introduction to Head-Related Transfer Functions (HRTFs): Representations of HRTFs in Time, Frequency, and Space. Journal of Audio Engineering Society, Vol. 49, No. 4, April 2001.

Chon H. C., Quantifying the Consonance of Complex Tones With Missing Fundamentals. Thesis submitted to Department of Engineering, Stanford University, (2008)

Einbond, A., MacCallum, J. “Real-Time Analysis of Sensory Dissonance” Center for New Music and Audio Technologies (CNMAT), Berkely, CA

MacPherson, E., On the Role of Head-Related Transfer Function Spectral Notches in the Judgement of ound Source Elevation, Waisman Center, University of Wisconsin-Madison.

Mashinter, K.: Calculating Sensory Dissonance: Some Discrepancies Arising from the Models of Kameoka & Kuriyagawa, and Hutchinson & Knopoff. Empirical Musicology Review 1, 2 (2006)

Plomp, R., Levelt, W.J.M: Tonal Consonance and Crtical Bandwidth. Journal of the Acousstical Society of America 38 (1965).

Terhardt, E.: On the perception of periodic sound fluctuations (roughness). Acustica 30, 201–213 (1974)