SPATIAL UTILIZATION OF SENSORY DISSONANCE AND THE CREATION OF SONIC SCULPTURE

Brian Hansen
University of California at Santa Barbara
Media Arts & Technology Program
brian.hansen78@gmail.com

ABSTRACT

Issues of musically consonant and dissonant sonorities have defined compositional practices for centuries. 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 sensory dissonance between two tones. All methods consider the physical and psychoacoustical aspects of sonic perception. However, these models typically ignore the dimension of physical space.

This paper aims to develop a model for representing sensory dissonance in three-dimensional space. In doing so, the proposed method accounts for factors that impact the spatialization of sound and, in turn, sensory dissonance. These factors include the inverse-square law, atmospheric absorption, and phase. The implementation of these factors will be discussed in detail, ultimately resulting in a method to model the sensory dissonance of sound in space.

Once the method is established, dissonance fields will be calculated, displaying the contours of dissonance that occur in a given space with multiple sound sources. It will then present how such dissonance fields and contours can be utilized to create atmospheric sculptures resulting from the sonic arrangement of a given space.

1. INTRODUCTION

Sensory dissonance is of a highly physical nature, as at its core it is caused by the activity of actual atmospheric vibrations occurring at a particular point in space. If we imagine all the points in a particular space being shaped by the forces generated from sound sources, it is as if the atmosphere itself has been sculpted by the sound. Furthermore, if we consider an instant of music, it is precisely a snapshot of a particular state of the expansion/rarefaction occurring in the air. An entire piece of music then, listened to from beginning to end, is a sequence of these snapshots, a film yielding a dynamic sculpture of the atmosphere. If the atmosphere is essentially a sonic sculpture, what does it look like? Is there a meaningful visual representation of it, or is it limited to an auditory and psychoacoustic experience?

A work highly influential to this line of thought is La Monte Young's Dream House. Dream House, located in lower Manhattan, is a sound and light installation, where the sound was designed by Young and the lighting by artist Marian Zazeela. The installation is set as a monochromatic room, sounding a single chord that has remained unchanged since 1993. The sustained chord encompasses a large portion of the acoustic spectrum ranging from extremely high to extremely low frequencies. The chord is spatially separated, as the different sinusoids of the chord emanate from multiple positions throughout the room. The overall effect is a completely immersive drone the listener can enter and explore. As the visitor navigates the room, the mix of sound changes based on their position and orientation. Certain tones are reinforced or attenuated, causing a morphology of harmony and melody for the visitor. The visitor is thus allowed to shape their own experience and form their own composition by how they choose to explore the space, or by how they subtly shift their orientation, simply tilting their head from side to side.

Works like Young's Dream House inspire questions about how a listener experiences spatialized sound. Although the space taken in its entirety consists of only the sound of a droning chord with each component remaining constant in frequency and amplitude, the listener has a dramatically different experience when entering the space. As the visitor navigates Dream House, they experience the reinforcement and attenuation of particular tones, causing the formation of unique sonic spectra particular to the visitor's perception. Such spectra are continuously changing as the visitor moves about, and each one yields a unique level of sensory dissonance. The visitor is immersed in a sonic field, where the listener experiences its contour largely by sensory dissonance.

We can use sensory dissonance as a tool to help generate atmospheres like Young's Dream House. Sensory dissonance could reveal to us where the various locations are more or less rough in a given space. However, in order to obtain this perspective, we need to extend the core calculation of sensory dissonance to incorporate spatialization factors. These factors would include the locations of each sound source and how the sound occupies and traverses a given space. Once we incorporate such factors into a sensory dissonance model, we have the potential to design immersive sonic environments like Dream House.

2. CORE CALCULATION

There have been essentially two methods of approach in quantifying sensory dissonance. The first was put forth by Kameoka and Kuriyagawa in 1969 which acknowledged the role of the critical band but did not utilize it. The second was implemented by Hutchinson and Knopoff in 1978 which fully utilized results from Plomp and Levelt's tonal consonance and critical bandwith research reported in 1965 [15]. Their calculation is thus based on the distance measured in terms of critical band (barks) between tones. A bark refers to the bark scale, a psychoacoustic-based scale ranging from 1 to 24 which corresponds to the first 24 critical bands of hearing.

Between the two approaches, the Hutchinson Knopoff approach has been more widely utilized, as it has been shown to yield comparatively better results. The Hutchinson & Knopoff model calculates the dissonance of all combinations of dyads according to the Plomp & Levelt curve. Upon calculating the dissonance, it weights each dyad's contribution to a given spectrum's dissonance by computing the product of the two amplitudes of the dyad and dividing it by the sum squared of all amplitudes.

$$D = \frac{\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} A_i A_j g_{ij}}{\sum_{i=1}^{N} A_i^2} \quad (1)$$

where N is the number of partials, A_i represents the amplitude of a partial, and g_{ij} is the sensory dissonance of a given dyad based on the critical bandwidth between their frequencies.

The critical bandwidth dissonance input into the Hutchinson & Knopoff approach is calculated here by using Richard Parncutt's approximation of Plomp and Levelt's dissonance curve, yielding the dissonance for the dyad [9].

$$g(b) = \left(4b \cdot e^{-4b}\right)^2 \quad (2)$$

where b is the critical bandwidth distance as determined by the distance in barks between two tones.

3. DISSONANCE IN PHYSICAL SPACE

In developing a spatialization method, we begin by utlizing the sensory dissonance calculation from Hutchinson & Knopoff as our foundation. We then build on this foundation by accounting for multiple sound sources emanating from multiple locations. When only one sound source is considered outside of space or time, there is only one spectrum that results with a unique level of sensory dissonance. However, when multiple sound sources exist, at each point in the space there is a unique sound spectrum present. Each spectrum, with its unique pairings of frequencies and corresponding amplitudes yields a unique level of sensory dissonance. Because of this, we can compute a unique value of sensory dissonance at any point in a given space. As a result, we will produce a dissonance field, where there will be different levels of dissonance dispersed throughout the space.

To accomplish this, we need to consider the proximity of a listener with relation to sound sources. Proximity factors include the inverse-square law, atmospheric absorption, and phase. In addition to this, psychoacoustic factors will be considered to more accurately represent a listener's perception of sound. These factors include masking, equal-loudness contours, and the critical bandwidth. Each of these factors, their impact on sensory dissonance, and their incorporation into the model will be detailed below.

3.1 Inverse-Square Law

The first factor our model accounts for is the sound's decrease in energy per unit area as a result of the distance it travels. For this, the inverse-square law is applied to adjust loudness levels in the spectrum of each sound source in the space.

As the sound travels radially from a point source in a free field, its energy is dispersed in a spherical pattern. The inverse-square law is an idealization because it assumes that sound propagates equally in all directions. In reality, there are reflective surfaces and structures that, depending on how the sound encounters such objects, have additive and subtractive effects to sound intensity. Nevertheless, our model assumes an anechoic environment, so a direct application of the inverse-square law is applied.

$$I_2 = I_1 \cdot \left(\frac{r_1}{r_2}\right)^2 \quad (3)$$

where I_1 and I_2 are the sound source intensities and r_1 and r_2 are the sound source distances.

3.2 Atmospheric Absorption

After adjusting the loudness level in the spectrum of each sound source for the inverse-square law, we make further adjustments for the effects of atmospheric absorption. Essentially, the atmosphere acts as a low-pass filter, since high frequencies become more dissipated than low ones as the sound travels through the air. In order to quantify this effect, ISO standard 9613-1: 1993 was used. This ISO standard gives an analytical method for calculating the attenuation of sound pressure given certain atmospheric conditions. The main inputs required for atmospheric conditions are temperature in degrees Kelvin, atmospheric pressure in kilopascals, and the percentage of relative humidity. The method works for pure-tone sounds ranging from 50 Hz to 10 kHz, with temperatures between -20C and 50C, relative humidity 10% to 100%, and the pressure of air equal to 101.325 kPa.

Because our model assumes an anechoic setting, the effects of atmospheric absorption apply, since like the inverse-square law, reflective surfaces nullify its effects indoors. Our model also will assume atmospheric conditions with a temperature of 20 degrees Celsius, relative humidity of 50%, and an ambient atmospheric pressure of 101.325 kilopascals. Given these assumptions, we can directly input the frequencies present in a given sound source's spectrum and the distance from that source. What results is the attenuated spectrum of our sound source after undergoing effects of atmospheric absorption.

3.3 Phase

We must also consider the phase differences in the sounds emanating from various source locations. If the sound sources are at different locations, then the frequency is out of phase. This must be accounted for when combining amplitudes for our dissonance calculation.

In prior works utilizing the sensory dissonance model of Hutchinson & Knopoff, it has been assumed that all partials of a spectrum are in phase. Thus, when computing the dissonance between two complex tones, it requires only simple arithmetic to combine the amplitudes of two coinciding partials. Because our model is considering spatialization, we can no longer make the assumption of a zero phase spectrum.

Our model does make the assumption that all partials emanating from the same sound source have a relative phase shift of zero. However, with multiple sound sources, since each sound source traverses a unique distance to reach the listener's location, we must consider the phase perceived by the listener at that point. For most cases, our approach holds because phase only impacts the combination of amplitudes when the difference in frequencies is extremely small. Further, in our case, we do not consider the time domain and its effects of phase on very small differences in frequency. Thus, for our model, we only consider the effects of phase when combining partials of the same frequency.

When combining the amplitudes of two equal frequencies, we need to know the relative phase between two partials before the combined amplitude can be calculated. First, calculate the distance between each sound source and the listener. Given this information and the frequency, we can calculate the phase shift present in each sinusoid. Then, simply subtract the phase shifts present in each sinusoid to get the relative phase between the sinusoids. Finally, knowing the relative phase between the partials, the formula below is utilized to determine the combined amplitude of the partials.

$$A_{combined} = \sqrt{A_1^2 + A_2^2 + 2A_1A_2\cos(\phi)} \qquad (4)$$

where A_1 and A_2 are the amplitudes for two partials with a given frequency, and ϕ is the relative phase between the two frequencies.

3.4 Auditory Masking

After accounting for proximity factors of spatialization, the next step is to adjust loudness levels for psychoacoustic properties of the listener. The first factor accounted for in this regard is auditory masking.

There are essentially two types of auditory masking, simultaneous masking and temporal masking. Simultaneous masking occurs when both sounds are present at the same time. In addition to simultaneous masking, temporal masking occurs when the masking effect extends outside the period when the masker is present. This type of masking can occur prior to and after the presence of the masker.

Auditory masking is very important to the calculation of sensory dissonance. Any given spectrum could have loud partials that drown out the sound of softer ones. If a given partial is masked, and thus not perceptible to the listener, then we assume it cannot contribute dissonance to the spectrum. Without the consideration of masking, our calculations would depict a dissonance level higher than is actually perceived, and the results could easily be skewed.

Our model assumes a continuous emanation of tones, eliminating the need to consider time. Thus, our model accounts only for simultaneous masking. The masking effect is modeled by utilizing a triangular spread function [2]. The spread function is written in terms of the bark scale difference between the maskee and masker frequencies:

$$d_{brk} = brk_{maskee} - brk_{masker} \qquad (5)$$

The bark difference is then input into our triangle function and a masking threshold T is calculated

$$T = L_M - \left(-27 + 0.37MAX\{L_M - 40, 0\}\theta(d_{brk})\right)|d_{brk}| \qquad (6)$$

where L_M is the maskers sound pressure level and $\theta(d_{brk})$ is the step function equal to zero for negative values of d_{brk} and one for positive values. If the sound pressure level of a given partial is less than the masking threshold, as computed by the triangle function above, then that particular partial is considered masked and is eliminated from the dissonance calculation.

3.5 Equal-Loudness Contours

After applying the spectral adjustments of auditory masking, the sound-pressure level of each partial is converted to sones in order to account for perceived loudness. When calculating sensory dissonance, prior models rarely take into account psychoacoustic effects of perceived loudness when weighting dyad amplitudes within a spectrum. This approach can lead to inaccurate results because there are drastic differences between a given frequency's sound pressure level and how its loudness is perceived. Thus, when weighting together the dissonance of dyads in a spectrum, our model follows the approach of Clarence Barlow by accounting for perceived loudness via representing amplitude loudness in sones rather than decibels [1]. This is accomplished by utilizing equal-loudness contours.

The equal-loudness contours used in our model are from the current international standard ISO226:2003 [12]. Given a frequency and its sound pressure level, our model uses the equal-loudness contours to convert sound-pressure level to phons. The phons are then converted to sones, which is a linearized unit of measurement for perceived loudness. The sones are then used when weighting together the dissonance of dyads in a spectrum. Using sones rather than sound-pressure level is a more accurate depiction of how the listener perceives partials in the spectrum. The sone based weighting reflects which partials are perceptually more prominent and in turn contribute the most to the sensory dissonance of a spectrum.

Converting decibel levels to sones marks the final adjustment required to calculate the sensory dissonance of a spectrum with consideration for the spatialization of sound sources. Thus, after completing the conversion from decibels to sones, we calculate the sensory dissonance of the spectrum utilizing the modified Hutchinson & Knopoff approach explained above.

4. CREATING AND VISUALIZING A DISSONANCE FIELD

Upon implementing all physical and psychoacoustic impacts into our spatialized dissonance model, we are able to produce a dissonance field. The dissonance field gives us a "topographical" representation of where different levels of dissonance occur in a given space. This is a very powerful result, as it gives us a vivid perspective on how sensory dissonance can occupy a space in the presence of multiple sound sources.

To construct the field, we first devise a spectrum and three-dimensional location for each sound source. With the sources in place, we then calculate the sensory dissonance at an equally distributed grid of locations in the space. The calculated dissonance field is then visualized via the technique of isosurfacing. This technique allows us to scan the dissonance field, revealing the contours and concentrations of dissonance throughout the space.

In constructing the dissonance field we make some key assumptions. First, we assume the space is an anechoic environment. Thus, we do not account for any impacts reflective surfaces may have on the sound. Our model then has its closest practical application in an outdoors setting. Secondly, we assume that the emissions of all sound sources are omnidirectional. Finally, we ignore the impacts of head-related transfer functions because we want to reveal a more objective perspective of the dissonance field.

Following suit with the modeling above, we can simulate an example of a dissonance field. To begin, we construct four tones consisting of band-limited saw tooth waves. The tones are then uniquely positioned in a virtual space of 40 cubic meters. Table 6 below displays the tones constructed and their positions in the space.

Fundamental Frequency (Hz)	Scale Degree	Sound Source Position (m) (x, y, z)
440	Tonic	(20, 20, 20)
493	Major 2nd	(0, 40, 20)
554	Major 3rd	(40, 0, 20)
660	Perfect 5th	(0, 0, 20)

Figure 1. Notes constituting the dissonance field

The tones were constructed and positioned based on their musical implications. A fundamental of 440 Hz was placed in the center of the space because of its foundational relationship as tonic to the other tones. Relative to this, the third and fifth scale degrees were placed on the floor because they form a major triad with the fundamental. Placing them here allows for the illumination of the dissonance relationship in the triad. The second scale degree was placed at the ceiling.

With the tones constructed and positioned, we calculated the sensory dissonance throughout the space at increments of 2.5 meters in all directions. Thus, we generated a 16 cubed matrix housing a total number of 4,096 measurements of dissonance. The image below in figure 2 displays our results.

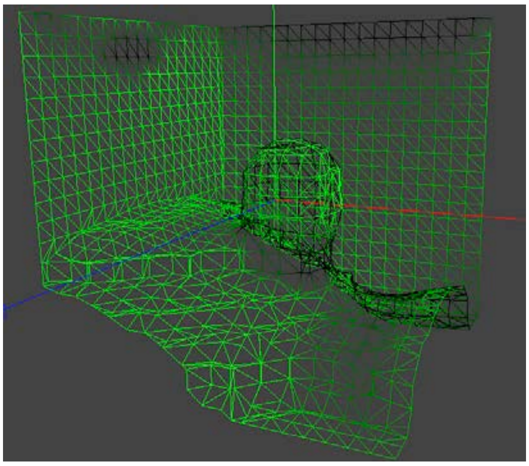

Figure 2. A dissonance field in virtual space as represented by isosurfacing.

The image displays the topography of the dissonance formed by the sound sources and its placement in the space. The contour displayed is a snapshot of where a particular level of sensory dissonance is present in the space. This representation gives us a unique insight into the relationships between the tones present and how the sonic field permeates the space.

5. IMMERSIVE ENVIRONMENTS AND SONIC SCULPTURE

The construction of a dissonance field is an informative result in modeling the spatialization of sensory dissonance. The field provides a unique perspective on how different sound sources relate to each other and how they are experienced in a space. Exploring the contours of different sonic arrangements in a space can be not only of practical use, but it also yields enormous artistic potential.

Recalling Lamonte Young's Dream House, the calculation of a dissonance field allows us to achieve a similar result. We can design immersive sonic environments yielding dissonance contours lush with sonic sensations that a visitor can explore. However, as opposed to Dream House, we have the added element of visualization, allowing us to more vividly design the sound space and visitor experience.

In addition, considering the purely physical nature of sound as the vibration of air, we are essentially creating sonic sculptures. We are using sound to position the vibration of air in highly specific ways. With this idea, the physicality of the sound itself is experienced as the artistic focus, rather than the sound having secondary significance as a metaphorical representation. This approach is analogous to the work of James Turrell, whose work involves creating immersive environments and sculptures utilizing light. Turrell creates visual art by flooding spaces or sculpting objects with light. In the environments he constructs, visitors are able to experience the purity of light's affect. In addition, Turrell "sculpts" objects that appear concretely physical in nature, but in actuality are comprised entirely of light. With his approach, it is not the illumination of a particular object that is the focus, but the light itself.

Utilizing our sensory dissonance spatialization model, we were able to simulate the sonic sculptures displayed in Figure 6 below, Curl and Arch. Curl is a sonic sculpture based on a justly tuned major triad with fundamental frequency of 100 Hz. The root of the triad is placed in the lower left of the space, while the third and fifth of the chord emanate from sound sources located above and to each side. The overall effect of the sonority is that the dissonance closes in on the fundamental in the curl like shape displayed.

Arch was simulated with six sound sources, where each sound source was placed in the center of one of the six sides constituting the border of the cubical space. Each sound source emanates a sine tone, and each tone is separated in frequency by 0.25 barks (the point of maximum roughness, i. e. sensory dissonance, in a given critical band). The setup results in the arch like dissonance field caused by the convergence of the sounding tones.

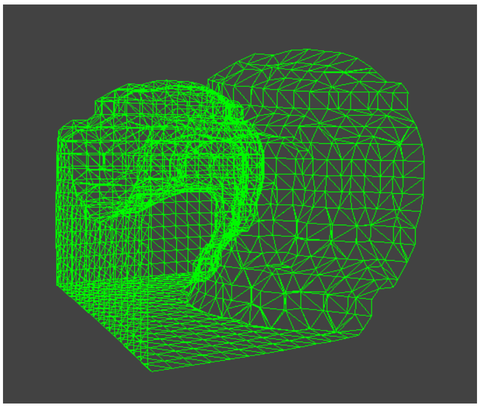

Figure 3. Sonic sculpture: Curl.

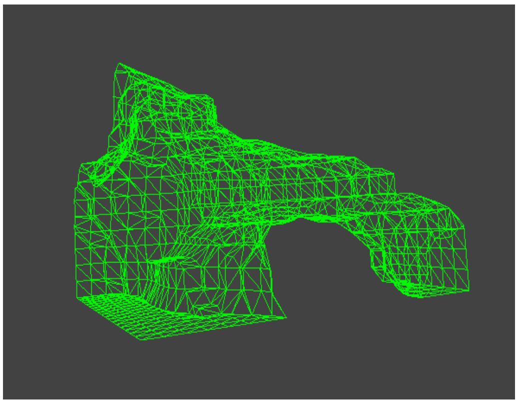

Figure 4. Sonic sculpture: Arch.

6. CONCLUSION

The method outlined above gives us enormous compositional potential for utilizing the concept and calculation of sensory dissonance. By calculating sonic roughness at various points throughout a three-dimensional space, dissonance fields are revealed to us through isosurfacing. We can use these fields to help design the sonic permeation of a given space, and they can help us shape the sound into sonic sculptures that visitors can explore with their ears rather than eyes.

Acknowledgments

I would like to thank Clarence Barlow, Curtis Roads, and Mathew Wright for their guidance in conducting this research.

7. REFERENCES

[1] Barlow, Clarence (1980). Bus Journey to Parametron. Feedback Papers 21-23. Feedback Studios Cologne.

[2] Bosi, M. (2003). Audio Coding: Basic Principles and Recent Developments. Paper presented at the HC International Conference.

[3] "Calculation Method of Absorption of Sound by the Atmosphere Air Damping Dissipation Absorbtion - Attenuation of Sound during Propagation Outdoors Outdoor - Sengpielaudio Sengpiel Berlin." Web. 20 Nov. 2011. <http://www.sengpielaudio.com/calculator-air.htm>.

[4] "Calculation of the Damping of Air Formula Atmosphere Acoustics Noise Absorption by Air Attenuation of Sound Air Damping - Sengpielaudio." Web. 20 Nov. 2011. <http://www.sengpielaudio.com/AirdampingFormula.htm>.

[5] Helmholtz, H. von. (1877). Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik. 1877, 6th ed., Braunschweig: Vieweg, 1913; trans. by A.J. Ellis as On the

sensations of tone as a physiological basis for the theory of music" (1885). Reprinted New York: Dover, 1954.

[6] Huron, D. (1994). Interval-class content in equally-tempered pitch-class sets: Common scales exhibit optimum tonal consonance. Music Perception, Vol. 11, No. 3, pp. 289-305.

[7] Hutchinson, W. & Knopoff, L. (1978). The Acoustic Component of Western Consonance. Interface, Vol. 7, No. 1, pp. 1-29.

[8] "Inverse-square Law." Wikipedia, the Free Encyclopedia. Web. 20 Nov. 2011. <http://en.wikipedia.org/wiki/Inverse-square_law>.

[9] Jacobsson, B. & Jerkert, J. (2000). Consonance of non-harmonic complex tones: Testing the limits of the theory of beats. Unpublished project report. Department of Speech, Music and Hearing (TMH), Royal Institute of Technology.

[10] Kameoka, A. & Kuriyagawa, M. (1969a). Consonace Theory, Part I: Consonace of Dyads. Journal of the Acoustical Society of America, Vol. 45, No. 6, pp. 1451-1459.

[11] Kameoka, A. & Kuriyagawa, M. (1969b). Consonance Theory, Part II: Consonance of Complex Tones and its Computation Method. Journal of the Acoustical Society of America, Vol. 45, No. 6, pp. 1460-1469.

[12] Mashinter, K. (1995). Discrepancies in theories of sensory dissonance arising from the models of Kameoka & Kuriyagawa and Hutchinson & Knopoff. Bachelor of Applied Mathematics and Bachelor of Music joint thesis, University of Waterloo.

[13] MacCallum, J., Einbond, A. (2007) Real-Time Analysis of Sensory Dissonance Center for New Music and Audio Technologies (CNMAT), Berkely, CA.

[14] Parncutt, R. (2006). Commentary of Keith Mashinter’s “Calculating Sensory Dissonance: Some Discrepancies Arising from the Models of Kameoka & Kuriyagawa, and Hutchinson & Knopoff.” Empirical Musicology Review. Vol. 1, No. 4, 2006

[15] Plomp, R. & Levelt, W.J.M. (1965). Tonal consonance and critical bandwidth. Journal of the Acoustical Society of America, Vol. 38, pp. 548-560.

[16] Suzuki, Yoiti. (2003) Precise and Full-Range Determination of Two-dimensional Equal Loudness Contours.

[17] Tenney, J. (1988). A History of "Consonance" and "Dissonance." White Plains, NY: Excelsior, 1988; New York: Gordon and Breach, 1988.