Modeling Sensory Dissonance in Space: Revelations in Sonic Sculpture.

Brian Hansen
Department of Media Arts and Technology
University of California Santa Barbara
Winter Quarter 2012

Table of Contents

1. Introduction

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 musical relationships are 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 sensory dissonance between two tones. All methods consider the physical and psychoacoustical aspects of sonic perception. However, these models typically do not consider the effects of physical space.

The spatial consideration of sound gained ever increasing attention over the 20th century, as the medium for which music is composed and presented largely expanded. Particularly with the advent of electronic music, spatialization of sound holds particular importance. Sound can be easily placed and dynamically distributed throughout a given space. For example, at the 1970 World's Fair in Osaka, a hemispherical auditorium was built according to designs drawn up by Karlheinz Stockhausen. Fifty speakers were distributed throughout the auditorium and thirty-five light sources were intuitively controlled, allowing the sounds to move spatially throughout the auditorium.

Another example highly influential to this work 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 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.

Figure 1. Inside Lamonte Young's Dream House (left). Occupants exploring the space (right).

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.

Sensory dissonance is of a highly physical nature, as at its core it is 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 by 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 can it only be limited to an auditory and psychoacoustic experience?

This project aims to answer these questions by utilizing the calculation of sensory dissonance. First of all, a method is proposed for calculating the dissonance between sounding tones, taking into consideration their spatial relationship to the listener. The method, described below, accounts for numerous factors that impact the spatialization of sound, where the key is the spatial impact on the loudness of components in a given spectrum. Factors considered will include the inverse-square law, head-related transfer functions, atmospheric absorption, phase, and masking. The paper will discuss the implementation of these factors in detail, ultimately resulting in a method to model the sensory dissonance of sound occurring in space. Once the method is established, various dissonance fields will be calculated, displaying atmospheric sculptures resulting from various sound sources and their locations.

2. Quantifying Sensory Dissonance

2.1 A History of Quantifying Dissonance

Historically, there have been numerous perspectives on how to define consonance and dissonance. Such perspectives span multiple contexts of culture, history, and science, and have yet to achieve universal agreement. As indicated by Richard Parncutt [22], it has always been fundamentally problematic to distinguish the cultural, learned, or "top-down" vs. the innate, sensory, or "bottom-up" aspects of what consonance is. He claims the boundary between the sensory and cultural is not clearly defined and depends on its operational definition. It is difficult to determine the boundaries of culture because of a lack of study and data related to the issue. In turn, this may impact the scientific or "bottom-up," in that the quantifiable component of dissonance in a musical sonority, sensory dissonance or "roughness," is potentially contaminated by the experiences and preferences of the listeners.

In addition to the philosophical questions surrounding consonance vs. dissonance, there have been historical attempts to quantify it. Pythagoras has been credited with the first attempt, claiming tones related by simple frequency ratios are more consonant. Around 1638, Galileo Galilei made claims about the eardrum's response to sound. In support of the Pythagorean view, he claimed that simple ratios produce regular vibrations on the eardrum, and more complex ratios produce irregular motions, yielding more or less pleasant sounds respectively. In 1863, Hermann von Helmholtz's *On the Sensations of Tone* presented the ear as a spectral analyzer, positing that the ear recognizes when two tones share harmonics, and tones sharing more harmonics are more consonant than tones not sharing any at all. More recently, in 1965 Plomp and Levelt more deeply explored the ideas set forth by Helmholtz and showed how dissonance could quantifiably be related to the distance between two frequencies. Their results show that dissonance is a function of critical bandwidth, and they provide a response curve for the perceptual dissonance in a dyad. They further show that when applied to complex tones, they could produce dissonance curves that bear some similarity with those produced by Helmholtz [23].

This type of dissonance, defined by the amount of beating or roughness present between two tones, has since been termed *sensory dissonance*. Through numerous studies, sensory dissonance has been shown to be highly correlative with Western common practices of music theory, and it has been successful in meaningfully modeling the phenomenon of dissonance in sonority. It is this type of dissonance, sensory dissonance, which will serve as the foundation for our model.

2.2 Choosing a Sensory Dissonance Calculation

Of the contemporary approaches to modeling sensory dissonance, there have been essentially two methods of approach in its quantification. 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. Each approach presents strengths and problem areas.

Kameoka and Kuriyagawa present their approach to quantifying sensory dissonance in a two-part paper [16][17]. Their first step, as detailed in the first paper, was to quantify the dissonance in a dyad. Then in the second paper, they proposed a method for aggregating the dyadic dissonance in order to quantify the dissonance between two complex tones. The Kameoka and Kuriyagawa method of quantifying sensory dissonance was based on direct observation of a test population, rather than building on the research of Plomp and Levelt. With a test population of sufficient size, they had enough data to observe their own relationships, and thus develop their own dissonance formula.

Kameoka & Kuriyagawa acknowledge the importance of Plomp & Levelt, but also cite their deficiencies. One area of criticism is in accounting for sound pressure level. The dissonance curve constructed by Plomp & Levelt is independent of sound pressure level, as it assumes that maximum dissonance is at the same critical band distance regardless of sound pressure level. Kameoka & Kuriyagawa claim this is not the case, explaining that the interval of maximum dissonance as indicated on the Plomp & Levelt curve actually increases as sound pressure increases. Consequently, Kameoka & Kuriyagawa's method accounts for various sound pressure levels when calculating the roughness present between two sinusoids.

In addition to accounting for various sound pressure levels of sinusoids, the Kameoka & Kuriyagawa model accounts for spectral masking. They claim that a difference in amplitudes exceeding 25dB results in perfect masking. Any difference less than this is accounted for with their formula by adjusting dyad dissonances with different sound pressure level.

Although the sound pressure level and masking considerations are advantageous in Kameoka & Kuriyagawa's model, their treatment of weighting dyad dissonance is problematic when considering complex tones. Their model does not accurately consider the contribution of each dyad's dissonance to the spectrum of two complex tones. It simply adds dyad dissonance without the consideration of how strong the dyad's amplitudes are with respect to entire spectral intensity. Because of this, as more partials are considered in a spectrum, the dissonance increases monotonically.

In addition to this, each dyad's dissonance is calculated with the consideration of the sinusoids' amplitudes. The decibel level of each partial is converted to microbars, where the ratio between to the microbars of each partial is raised to a pre-determined exponent. This is problematic for our model because we are considering psychoacoustic effects on perceived loudness. Because psychoacoustic conversions are non-linear, we cannot convert decibels to sones and apply the same scheme utilizing the exponent given to us by Kameoka & Kuriyagawa. Thus, if we were to use the Kameoka & Kuriyagawa approach we would either have to perform our calculations without psychoacoustic considerations or assume equal loudness for all combinations of dyads.

The Hutchinson & Knopoff model, proposed in 1979, is a direct application of the Plomp & Levelt model [13]. This approach 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. Thus, as apposed to the Kameoka & Kuriyagama method, a dyad's dissonance contribution is accounted for via this weighting.

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

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.

An area of concern with the Hutchinson & Knopoff approach is that it treats all partials of a complex tone as equal-tempered. When evaluating partials of a complex tone, the model snaps all frequencies of the tone's partials to nearest well-tempered frequency. This causes large deviations, particularly when higher frequencies are shifted to the nearest equal tempered pitch.

This problem is easily remedied however by using the actual, unshifted frequencies of a complex tone's partials. First the critical bandwidth between the frequencies must be determined. Then, the calculated CBW can be input into Parncutt's approximation of Plomp and Levelt's dissonance curve, yielding the dissonance for the dyad [19].

$$g(f_{cb}) = \left(\frac{f_{cb}}{0.25} \cdot e^{(-x/0.25)}\right)^2 \tag{2}$$

where f_{cb} is the critical bandwidth distance as determined by the distance in barks between two tones.

For our model we have decided to implement a modified version of the Hutchinson & Knopoff approach. Although the Kameoka & Kuriyagawa method has some intriguing benefits, it is simply unusable because it does not accommodate our psychoacoustic considerations for sound intensities, and it does properly consider the weighted contribution of a given dyad to the overall spectrum under consideration. In addition, shortcomings of the Hutchinson & Knopoff approach can be corrected with our modifications. First of all, we do not round the frequencies of partials to the nearest well-tempered value. Instead, we preserve the actual frequency and apply the Parncutt approximation to Plomp & Levelt's dissonance curve. In doing this, all frequencies are converted to the bark scale, in order to account for the critical bandwidth. It is then that the bark scale difference serves as input to the Parncutt equation. Our method will also account for the masking of frequencies via the implementation of a masking screen (explained in detail below) before calculating the dissonance of a spectrum.

Although our adopted approach does not answer Kameoka and Kuriyagawa's criticism of critical band dissonance being independent of sound pressure level, Hutchinson & Knopoff acknowledge this phenomena claiming that critical band width dissonance only becomes subjective to sound pressure at very high levels. Since we are not considering extremes in sound pressure level, critical bandwidth dissonance will suit our needs.

Finally, the Hutchinson & Knopoff approach has been more widely implemented and has yielded comparatively better results. For example, Vos [26] found that the Hutchinson & Knopoff approach was more accurate than the Kamaoka & Kuriyagawa method when comparing intervallic purity of 3rds and 5ths. In addition, Song Hui Chon [6] displayed successful results when utilizing a modified Hutchinson & Knopoff approach to consider the case of a missing fundamental in the calculation of sensory dissonance.

2.3 Effects of Spatialized Sound on Sensory Dissonance

Nearly all prior explorations of sensory dissonance have been limited to the dimension of frequency. There have been very few extensions of the application beyond this, yet there is enormous potential. For example, MacCallum and Einbond [20] consider the additional dimension of time by measuring how levels of sensory dissonance change over the duration of a composition. In addition, there has not been an exploration of sensory dissonance where the sound is considered in a spatial setting. In reality, a listener hears sounds from different locations. Some are above, while some are below. Some are behind, and some are in front. Thus, we can consider how a listener's orientation with respect to a sound source transforms the sonic image. In addition, if multiple sound sources are present, the listener's proximity to one vs. the others may impact their perception of the sound. Sensory dissonance relies on the perception of frequency and intensity. Essentially, each combination of partials in a given spectrum yields a certain level of sensory dissonance. This level of dissonance is dependent on the frequency difference between the two partials. The amount of dissonance a particular pair of partials contributes to the spectrum's total sensory dissonance is dependent on the intensity of the two partials. Of the two components, frequency and intensity, the spatial location of sound relative to the listener has the greatest impact on intensity. Thus, it is the intensity between two sound sources that impacts the sensory dissonance between them. The only situation where frequency is affected is if the listener and/or sound source is moving, in which case the Doppler effect is present.

When considering effects of intensity on sensory dissonance, a sound's intensity is dependent on the listener's orientation and proximity relative to the sound source. With respect to orientation, the auditory mechanism essentially filters the sonic spectra that enters the inner ear depending on the elevation and azimuth of the on the sound source with respect to the listener. Since the sound is filtered, its spectral mix is altered which impacts how much a particular frequency contributes to the overall dissonance of the spectrum. This aspect can be modeled using head-related transfer functions, which will be discussed in detail below.

In addition to orientation, proximity has a major impact on sound intensity. As sound travels through air, its energy is dispersed and absorbed into the atmosphere. Thus, the intensity of a sound source experienced by a listener is greatly determined by their proximity to it. For example, consider two sound sources separated by a distance such that only one source can be heard at each end, and the two can only be heard together when at the midpoint between them. Further, assume the intensity of each sound to be equal, and the spectra of each sound are independently consonant, but highly dissonant when heard together. At each end, the listener experiences a consonant sound because the space traversed by the conflicting sound is too large. The energy of the conflicting sound has been dispersed and absorbed into the atmosphere such that the listener does not perceive it. Clearly, if it cannot be heard it cannot impact the spectra perceived by the listener. Thus, they do not perceive dissonance between the two sources. Conversely, if we consider the listener's location at the midpoint between the sound sources, the listener hears the two sounds combined at equal intensity levels. Thus, they perceive all partials and the resulting roughness from dyad combinations throughout the spectrum. This example can be modeled by utilizing the inverse-square law and atmospheric absorption, and will be an integral part of our model.

The model detailed below constitutes a theoretical foundation for the spatialization of sensory dissonance. To realize this model, we will utilize prior research in calculating sensory dissonance as our foundation. However, we will 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 intensities 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. Thus, we will produce a dissonance field, where there will be different levels of dissonance dispersed throughout the space.

To accomplish this, we will mainly consider the orientation and proximity of a listener with relation to sound sources. Proximity factors include the inverse-square law, atmospheric absorption, and phase. For orientation, we will utilize head-related-transfer-functions. 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.

2.4 Method of Implementation

2.4.1 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.

$$
\begin{aligned}
&I \propto \frac{1}{r^2} \\
&\frac{I_1}{I_2} = \frac{r_2^2}{r_1^2} \\
&I_2 = I_1 \cdot \left(\frac{r_1}{r_2}\right)^2
\end{aligned}
\tag{3}
$$

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

For implementation into our model, intensities are represented on a decibel scale. Thus, we used the following representation of the inverse-square law for our calculation:

$$dB_d = dB_0 - 20\log_{10}(1/d) \tag{4}$$

where dB_0 is the original sound pressure level of the sound source, d is the distance from the sound source, and dB_d is the sound pressure level experienced at a distance d from the source location.

2.4.2 Atmospheric Absorption

After adjusting the loudness level in each sound source spectrum 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 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 assumes 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.

2.4.3 Head-Related Transfer Functions

After adjusting loudness levels for the inverse-square law and atmospheric absorption, we assume the sound has reached the body and outer ear of the listener. At this point, we must make further adjustments for the orientation of the listener with respect to each sound source. Thus, a head-related-transfer-function routine is applied next.

The auditory mechanism essentially filters the sonic spectrum that enters the inner ear depending on the sound's origin with respect to the listener. Since the sound is filtered, its spectrum is altered affecting the amplitudes and phases in the spectrum, and in turn having consequences for sensory dissonance. Our model considers this before the sensory dissonance via the use of head-related transfer functions (HRTF).

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 inner ear. The azimuth and elevation of the sound source relative to the listener are used as parameters to select a HRTF for filtering an incoming sound. It is posited that this mechanism may have developed in humans as an evolutionary necessity [8]. Since the eyes can only focus on one general 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 it can alert us to potential dangers that are out of sight.

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 individual’s HRTF is not equal to another's, and the HRTF cannot perfectly model spatialized 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 spatialization of sensory dissonance.

For implementation into our model, we are using head-related impulse responses (HRIR) from the archives of IRCAM [21]. We processed the 512 sample point HRIRs via a fast Fourier transform to obtain the frequency response of each azimuth and elevation. Each frequency response was then stored in a database that was referenced for our HRTF process. When accessing the database, our listener's azimuth and elevation relative to the sound source is used to select a corresponding HRIR frequency response. If an exact match in elevation and azimuth does not exist, we perform a linear interpolation between the nearest HRIRs to obtain an approximation. The selected HRIR amplitudes are then multiplied by the amplitudes of our sound source spectrum in correspondence to frequency. The result of this operation is the transformation of the sound source spectrum to match the listener's perception of it, based on their orientation with respect to the point source location. Because the amplitudes of each partial have been adjusted, the overall weighting of dyad dissonance will change, thus impacting the calculated sensory dissonance.

2.4.4 Phase

At this point, it is assumed all sound sources have now funneled into the ear and must be combined into one spectrum. To accomplish this, the next step must account for phase impacts. In doing so, if two spectra share the same frequency, then the amplitudes must be combined. If the sound sources are at different locations, then the frequency is out of phase, and this must be accounted for when combining the amplitudes.

In prior works utilizing the sensory dissonance model of Plomp & Levelt, 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, 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.

$$A_{combined} = \sqrt{A_1^2 + A_2^2 + 2A_1A_2\cos(\phi)} \tag{5}$$

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.

2.4.5 Auditory Masking

With the sound having traversed physical distance, funneled its way into the listener's inner ear, and combined with all sources to form one perceivable spectrum, the sound has undergone all transformations resulting from physical properties. The next step then is to adjust loudness levels for psychological properties of the listener. Thus, the adjustments based on psychoacoustics are performed. The first factor accounted for in this regard is auditory masking.

Auditory masking occurs when a given sound, a "maskee", is made inaudible by another sound present, a "masker". For example, people who are talking cannot hear their conversation when a noisy bus drives by because the noise of the bus drowns out the sound of the conversation. In this case the bus noise is the masker, and the sound of the conversation is the maskee.

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:

$$d_{brk} = brk_{maskee} - brk_{masker} \tag{6}$$

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

$$T = L_M - \left(-27 + 0.37 MAX\left\{L_M - 40{,}0\right\}\theta\left(d_{brk}\right)\right)\left|d_{brk}\right| \qquad (7)$$

where L_M is the maskers sound pressure level and $\theta\left(d_{brk}\right)$ 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. Since partials are eliminated from the dissonance calculation our model receives the additional benefit of decreased computation time. Thus, accounting for the phenomenon of auditory masking not only improves our accuracy of perceived sensory dissonance, but it also improves the computation time of our model.

2.4.6 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 Balrow 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 international standard

ISO226:2003 [24]. These contours have replaced the original "Fletcher-Munson" curves, named from the earliest researchers of this phenomenon. They are the current standard in accuracy utilizing several studies by researchers from Japan, Germany, Denmark, the UK, and the USA.

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 Huthinson & Knopoff approach explained in section 2.2 above.

2.5 Assumptions and Omissions

It is important to note assumptions and omissions for our model of calculating sensory dissonance in space. First of all, our model assumes an anechoic environment. Because of this we do not account for any reinforcement or attenuation in sound sources due to reflective objects in the space. Essentially, our model has its closest practical application in an outdoors setting. Secondly, we assume the sound is traveling through air with atmospheric conditions of a temperature of 20C, relative humidity of 50%, and barometric pressure of 101.325kPa. This set of assumptions depicts a comfortable environment for the average listener, and it has particular importance when accounting for atmospheric absorption. Third, we assume all sound sources are omnidirectional. This assumption has its greatest impact on calculating a dissonance field, as directional sound can focus dissonance on certain areas. Fourth, a listener agent has directional perception. This impacts the mix of a sound spectrum when applying HRTFs as explained above. Fifth, sounds emitted in our environment are only sinusoidal tones with zero phase, and there is no noise present. Noise could have an impact on auditory masking, and it in itself has an intrinsic dissonance. Although, the Kameoka & Kuriyagawa approach accounts for noise, there is speculation as to its accuracy and effectiveness in their dissonance results (CTN Mashinter). In addition, since we are assuming the Hutchinson & Knopoff approach, we do not account for it. Sixth, our model does not account for the psychoacoustic phenomenon of a missing fundamental. Song Hui Chon (BIB CTN miss fund thesis) proposes that the missing fundamental phenomenon can be incorporated into the dissonance equation. This has potential impact for a listener agent in our environment, and may be implemented in the future. Finally, our model assumes a static time frame. Time can have interesting impacts on our calculation, particularly when considering phase. For example, if two sound sources are very close in frequency and out of phase, one would expect to experience undulations in the sensory dissonance present. In addition, sound sources can change over time with respect to their spectrum and location. Modeling this would allow us to visualize the fluidity of dissonant sonorities in a space.

In addition, it is important to note that we follow all the steps above when calculating spatialized sensory dissonance as perceived by a listener. However, when calculating a dissonance field, we do not assume a listener. In this case, our aim is to achieve a more objective point of view of the dissonance “topography” in the atmosphere. Thus, we remove impacts that would be subjective to one particular listener. Eliminating effects of listener orientation accomplishes this, and so the routine for HRTF impacts is bypassed.

3. Architecture and Implementation

3.1 Architecture of *spatialdiss* external.

The dissonance spatialization method described above was coded in C++ and realized in a Max/MSP external object called *spatialdiss* (see appendix A for the source code). The object acts as a hub for calculating the dissonance among all sound sources in a given three-dimensional space. The spectrum and location data of each sound source is sent to the *spatialdiss* object. In addition, the location of a "listener" agent is sent. The object then computes and returns the sensory dissonance present at the location of the listener, taking into account the physical and psychoacoustic factors outlined above.

There are two main data structures used in the object, a note structure and a dissonance spectrum structure. The note structure houses the sonic and location information for each emission agent. It contains an array to store the location of the agent in Cartesian coordinates. It also contains frequency and associated amplitude vectors for the agent’s sound spectrum, where the frequency is in Hz and amplitude is in decibels.

The second data structure is a dissonance spectrum. This is the main structure that gets populated by the routine, and it is used for the sensory dissonance calculation. Here the note data from each individual sound source is consolidated to one master spectrum. Thus, it contains a frequency vector and associated amplitude vector. Vectors for barks and sones are also contained in the dissonance spectrum struct, since frequency and amplitude must be converted to barks and sones respectively. In addition, the structure contains a vector for the phase of each frequency component. This is calculated based on the frequency present and the distance between each sound source and listener. Moreover, it contains two two-dimensional vectors; one for the calculated sensory dissonance between each pair of sinusoids, and one for the amplitude weighting assigned to the dissonance of each dyad. Finally, it does not contain a location vector, as this is unnecessary since all amplitudes are calibrated based on sound source location before being stored in this data structure.

There are three main steps that occur in *spatialdiss* in order to calculate our desired sensory dissonance. First of all, the spectral information and location of each sound source must be stored in the note structure. Second, each agent's sonic information is consolidated into a master spectrum. Finally, the sensory dissonance of the master spectrum is calculated and returned.

Each sound source is assigned a unique note structure when received by *spatialdiss*, thus forming an array of notes. A user can change a sound source by creating a new note or modifying an existing one. Once the program has all relevant agent note data, the routine getSpatialDiss() is called. This routine sets up a structure for our master dissonance spectrum. It calls the routine generateListenerSpect() which consolidates our note data and modifies it based on our factors for sound perception. Then, it passes the formulated spectrum into calcDissonace() which calculates the sensory dissonance of the spectrum. Finally, it returns the dissonance value via diss_out().

The generateListenerSpect() routine is the main method that prepares the data for the dissonance calculation. Before the dissonance can be calculated this routine consolidates all sound sources to one sound spectrum. The consolidation is accomplished by accounting for all the factors outlined above. First of all, the routine calculates the distance between the listener's location and a given sound source. Using this distance, and its position relative to the listener, it then determines the azimuth and elevation between the two and returns an appropriate HRTF to coincide with their relative positions. Once the distance and HRTF are obtained, the method proceeds to make adjustments to the spectrum of the incoming sound sources. These modifications coincide with the sequence of factors outlined in section 2.4 above.

3.2 *spatialdiss* in Max/MSP/Jitter

The external object, *spatialdiss*, requires data from sound sources and a listener. Sound sources are recognized by the object as agents. Each agent is uniquely indicated and sends its information into the left inlet of *spatialdiss* by interleaving its spectral makeup of frequency and amplitudes and appending its location on the end of the spectrum (see Figure 2 below). The listener only needs to specify its location, and this is sent into the right inlet of spatialdiss.

The object *spatialdiss* also allows for enabling or disabling certain spatial effects on the dissonance spectrum formation. One can independently enable or disable impacts of the auditory masking, inverse-square, atmospheric absorption, or head-related transfer functions. This allows the user to isolate the impacts each individual calculation has on spatializing sensory dissonance.

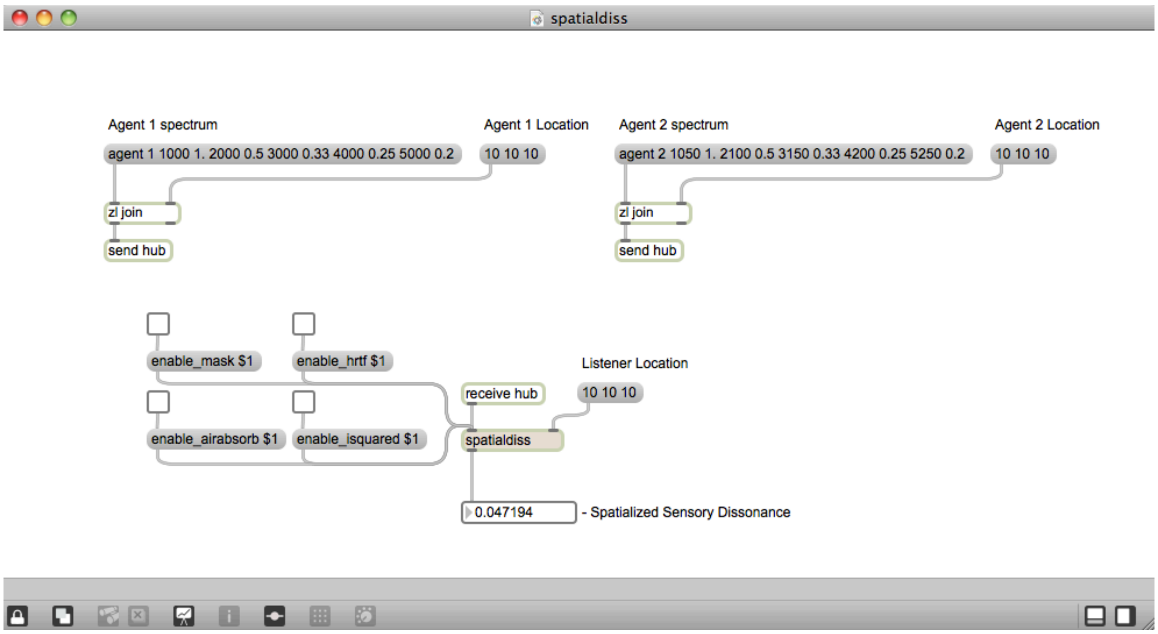

Figure 2. Max/MSP implementation of the spatialdiss object.

There are numerous ways in which to utilize *spatialdiss* with its implementation in the Max/MSP/Jitter environment. For example, we can calculate the dissonance field of a space. To do this we simply set up our sound sources. Then, we calculate the dissonance equidistantly distributed throughout the space and store it in a jitter matrix. We can then utilize the volumetric visualization technique of isosurfacing in jitter to illuminate the contours of the dissonance field.

In addition, we can utilize the Cosm extension package developed by Graham Wakefield and Wesley Smith at the University of California Santa Barbara with *spatialdiss* [7]. With this we can color code our listener and have a clear visualization as to the level of dissonance it is experiencing at different locations in the space. Further, we can create a dissonance network among our sound sources. To accomplish this, we make a copy of *spatialdiss* for each sound source and substitute their location for the listener's. Each sound source can be color coded as well, yielding a dissonance relationship among sound sources. There are presumably countless other ways the *spatialdiss* object could be utilized in Max/MSP, where more exploration on this topic will be presented when considering artistic explorations below.

4. Results

4.1 First Simulation: A listener Between Two Sound Sources.

Our first simulation for showing spatial impacts on sensory dissonance was to determine how a listener situated between two sound sources perceives the dissonance between them. For this test, two virtual sound sources were separated by a distance of 100 meters. Then, according to the method described above, the sensory dissonance was calculated at our listener's location in 10 meter increments between the sound sources. Our hypothesis was that dissonance would be lowest when the listener is nearest one of the sound sources, and the dissonance would reach it's maximum when the listener is located at some point between the two.

The sound sources were modeled as a band-limited saw tooth wave consisting of five partials. One sound source acted as a base note, while the other separated by a distance of 100 meters, was always a 12-tet minor 2nd interval higher than its base. The base notes tested were 100, 250, 500, 750, 1000, and 1250 Hz. The table below displays the results as our listener location was incremented between the two sources:

Listener Position (m)	Base Frequency Fundamental (Hz) 100	250	500	750	1000	1250
0	0.0317	0.0163	0.0162	0.0199	0.0176	0.0188
10	0.0766	0.0559	0.0537	0.0529	0.0544	0.0556
20	0.0912	0.0720	0.0686	0.0679	0.0704	0.0720
30	0.0976	0.0803	0.0765	0.0760	0.0793	0.0812
40	0.1003	0.0845	0.0805	0.0803	0.0841	0.0861
50	**0.1008**	**0.0857**	**0.0818**	**0.0817**	**0.0859**	**0.0878**
60	0.0992	0.0842	0.0805	0.0806	0.0849	0.0864
70	0.0954	0.0798	0.0765	0.0766	0.0807	0.0818
80	0.0879	0.0714	0.0687	0.0688	0.0724	0.0729
90	0.0723	0.0553	0.0538	0.0540	0.0566	0.0565
100	0.0289	0.0171	0.0178	0.0182	0.0190	0.0190

Table 1. Dissonance values of listener between two sound sources with their position from base note source indicated in the left column.

Table 1 confirms our original hypothesis. In essentially every case, the perceived spectral dissonance reached its maximum when the listener was located directly in between the two sound sources. Conversely, when the listener was close to either source, their experienced dissonance was drastically lower. In addition, the results show that the impacts of distance are independent of frequency. This is illuminated by the rate of change in dissonance as the listener moved from one source location to the next. The rate of change is very similar for every base frequency tested.

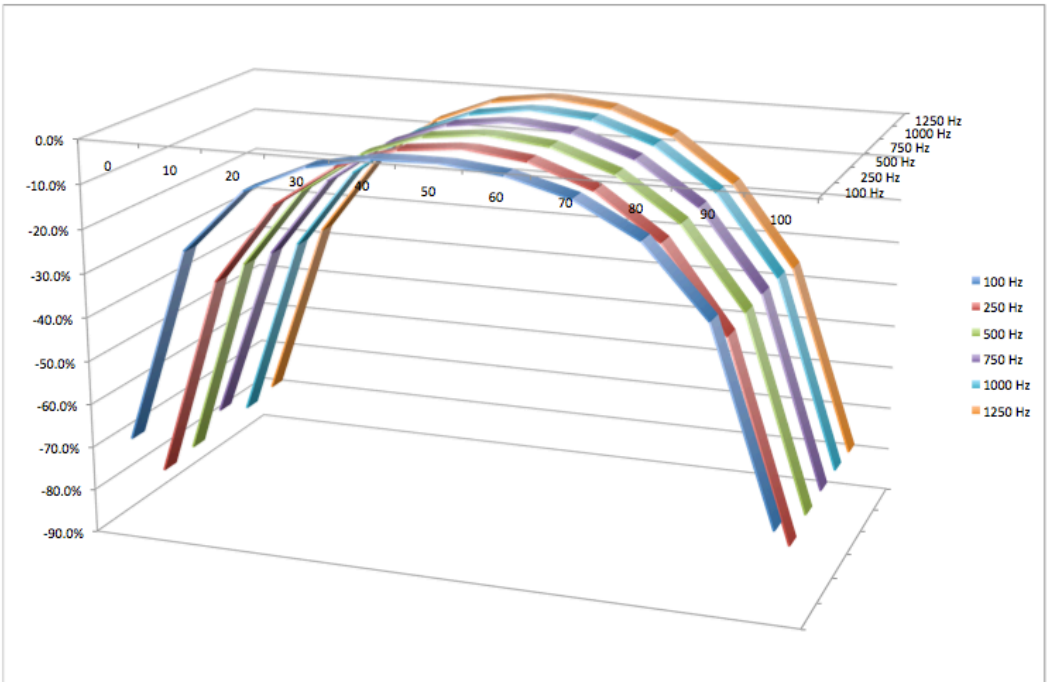

Figure 3. Graph depicting the rate of change in perceived dissonance as the listener changes position between two sound sources.

This first test was conducted without the influence of auditory masking. A second pass was conducted with the same parameters in place, however with the addition of auditory masking. The overarching conclusion was the same as with our first test. The listener experienced a maximum dissonance level half way in between the sound sources and the lowest dissonance nearest each source. The major difference is the dramatic effect our auditory masking implementation has on the perceived dissonance. The dissonant tone would be nearly completely filtered by the masker up until the listener was at least 30 meters from both sources. The filtering due to masking causes the dissonance to be zero within 30 feet of either source. Outside of this, the dissonance drastically jumps in value, as the listener can now perceive the two tones combined, and thus experience their dissonance. In addition to this, it is noted that incorporating the effects of masking introduces more dependency on frequency for our dissonance calculation.

Listener Position (m)	Base Frequency Fundamental (Hz)					
	100	**250**	**500**	**750**	**1000**	**1250**
0	0.0138	0.0000	0.0000	0.0000	0.0000	0.0005
10	0.0145	0.0000	0.0000	0.0000	0.0000	0.0005
20	0.0145	0.0000	0.0000	0.0000	0.0000	0.0004
30	0.0145	0.0677	0.0644	0.0749	0.0744	0.0771
40	0.1246	0.0839	0.0790	0.0791	0.0801	0.0849
50	**0.1110**	**0.0851**	**0.0802**	**0.0804**	**0.0818**	**0.0870**
60	0.1219	0.0836	0.0788	0.0789	0.0806	0.0857
70	0.0114	0.0663	0.0637	0.0745	0.0754	0.0802
80	0.0114	0.0000	0.0000	0.0000	0.0000	0.0004
90	0.0114	0.0000	0.0000	0.0000	0.0000	0.0005
100	0.0107	0.0000	0.0000	0.0000	0.0000	0.0005

Table 2. Dissonance values of a listener between two sound sources with the effects of auditory masking.

4.2 Isolating Spatialization Effects

4.2.1 Inverse-square law

The effects on our model that relate purely to spatialization are the inverse-square law, atmospheric absorption, and head-related transfer functions. With the introduction of these three factors into the quantification of sensory dissonance, we wanted to isolate the impacts of each. This will give insight into the ultimate behavior of our model.

When isolating the impact of the inverse-square law, we followed a similar approach as with our initial test explained above. This time, we started with two sound sources and the listener, all at the same location. Then, one of our sound sources was relocated by increments of 10 meters away from its original location. As before, we tested a saw tooth wave for various base tones with fundamentals of 100, 250, 500, 750, 1000, and 2000 Hz, each against a 12-tet minor second interval. We expected the sensory dissonance to follow in suit with the inverse-square law, and display an exponential decay in dissonance as the distance between the sound sources increased.

Source Separation (m)	Base Frequency Fundamental (Hz)					
	100	250	500	750	1000	2000
0	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
10	-34.9%	-39.7%	-38.0%	-36.6%	-37.3%	-37.0%
20	-49.1%	-55.3%	-53.4%	-51.9%	-52.4%	-52.2%
30	-56.0%	-63.2%	-61.3%	-59.8%	-60.2%	-60.0%
40	-60.2%	-68.1%	-66.2%	-64.8%	-65.2%	-64.9%
50	-63.1%	-71.5%	-69.7%	-68.4%	-68.7%	-68.4%
60	-65.3%	-74.1%	-72.4%	-71.0%	-71.3%	-71.0%
70	-66.9%	-76.1%	-74.4%	-73.2%	-73.4%	-73.1%
80	-68.3%	-77.7%	-76.1%	-74.9%	-75.1%	-74.8%
90	-69.4%	-79.0%	-77.5%	-76.3%	-76.6%	-76.2%
100	-70.3%	-80.2%	-78.7%	-77.5%	-77.8%	-77.4%

Table 3. Percentage change in dissonance from original position for separation between two sound sources.

Table 3 displays what we expected. It displays the percentage change in sensory dissonance as the separation between sound sources increases. Looking across all tested registers, it is evident that at each increment of separation, the percentage change is very similar, and overall the same rate of exponential decay is experienced.

4.2.2 Atmospheric Absorption

The same setup used in isolating inverse-square impacts was used for atomspheric absorption. However, atmospheric absorption has minimal impact over short distances. Thus, the repositioned sound sources were incremented by 100 meters over a total distance of 1000 meters in order to more clearly assess its effects.

The effects of atmospheric absorption are similar to that of a low pass filter. Knowing this, it was our expectation that the dissonance between low tones should be largely unaffected by atmospheric absorption, and we would notice an increasing impact with higher tones. This expectation proved to be the case as displayed by Table 4 below.

Distance (m)	Base Frequency Fundamental (Hz) 50	100	500	1000	2000	4000
0	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
100	0.0%	-0.1%	-0.2%	0.2%	4.4%	-7.1%
200	0.0%	-0.3%	-0.6%	-0.5%	3.0%	-10.7%
300	0.0%	-0.4%	-1.1%	-1.9%	-1.0%	-14.1%
400	0.0%	-0.6%	-1.8%	-3.8%	-5.9%	-19.7%
500	0.0%	-0.8%	-2.6%	-6.1%	-11.3%	-25.4%
600	0.0%	-0.9%	-3.5%	-8.6%	-16.7%	-30.5%
700	0.0%	-1.1%	-4.5%	-11.3%	-21.9%	-35.0%
800	0.0%	-1.3%	-5.6%	-14.0%	-26.9%	-38.8%
900	0.0%	-1.5%	-6.8%	-16.8%	-31.7%	-42.0%
1000	0.0%	-1.8%	-8.1%	-19.5%	-36.1%	-44.6%

Table 4. Percentage change in dissonance from original position over distance as a result of atmospheric absorption.

The table shows that atmospheric absorption has a clear frequency dependent impact on sensory dissonance. The higher the tones, the more they are absorbed by the atmosphere. Clearly, if a tone is largely absorbed into the atmosphere, it cannot be perceived in conjunction with another tone, and thus it cannot contribute to sensory dissonance.

4.2.3 Head-Related Transfer Functions

Again following the same procedure as above, we isolated the impacts of head-related transfer functions by independently testing impacts of azimuth and elevation. For azimuth impacts, sensory dissonance was measured at 45 degree increments of a full 360 degree circumference around the listener. Elevation was measured from -45 degrees to positive 45 degrees in increments of 15 degrees. The base tones tested were at 100, 250, 500, 1000, 2000, and 4000 Hz.

Our results show that, indeed, head-related transfer functions have a material impact on sensory dissonance. The calculated sensory dissonance can vary up to 20%, depending on the frequency and orientation of the listener with respect to the sound source. With respect to azimuth, frequencies in the mid range (250 - 1000 Hz) tended to yield a maximum spread in sensory dissonance at an average of about 5.5%. With respect to elevation, the average was about 4.7%. Table 5 below displays the results.

Azimuth (degrees)	Base Frequency Fundamental (Hz) 100	250	500	1000	2000	4000
0	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
45	-1.7%	-1.0%	0.9%	-3.4%	-4.2%	-4.5%
90	-5.4%	-1.8%	1.3%	2.4%	-10.0%	-6.0%
135	-1.3%	2.8%	2.2%	2.3%	-6.5%	-5.3%
180	1.7%	1.3%	4.4%	4.2%	-5.3%	8.8%
225	-3.0%	2.1%	0.9%	2.6%	-4.4%	-9.2%
270	-4.4%	-1.3%	0.3%	1.6%	-8.8%	-2.2%
315	-0.1%	0.3%	1.7%	-3.0%	-1.3%	6.2%

Elevation (degrees)	Base Frequency Fundamental (Hz) 100	250	500	1000	2000	4000
-45	2.5%	0.3%	3.3%	0.9%	-2.2%	-8.3%
-30	4.5%	0.1%	4.1%	0.5%	-3.4%	-8.4%
-15	4.7%	-1.7%	4.5%	4.8%	-6.3%	-14.1%
0	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
15	-3.3%	0.8%	4.7%	1.0%	-4.9%	-13.9%
30	-0.4%	-0.5%	1.2%	-0.9%	-1.7%	-2.1%
45	2.2%	-0.2%	-0.4%	-1.6%	-2.4%	-5.5%

Table 5. Percentage change in dissonance from original position with respect to changes in azimuth (top) and elevation (bottom). Both tables present changes with respect to an initial orientation of zero degrees.

With respect to azimuth, there tends to be an increase in dissonance when sound sources are at 180 degrees (directly behind) of the listener's orientation. This may be because sounds placed behind a listener undergo the effects of a lowpass filter, as the pinna not being oriented towards the sound source reduces much of the energy of high frequencies. Since dissonance of complex tones tends to be stronger for lower tones (verified in [13]), the low pass effect may accentuate this trait, thus causing the slight increase in perceived dissonance. In addition, when the sources are placed just to the left or right of the listener, the perceived dissonance tends to decrease. This may be due once again to the pinna's effects on high frequencies. Sounds directly in front of a listener need to be reflected off the pinna before entering the inner ear, where sounds to the left or right of the listener have a direct route into the ear canal. Frequencies with a more direct path do not lose energy from reflecting off the pinna, and therefore maintain their energy in the overall spectral mix. Thus the dissonance weighting among energy of higher frequencies may be greater, causing a decrease in the aggregate perceived dissonance.

With respect to elevation, there does not appear to be a common trend impacting sensory dissonance among all frequencies. Nevertheless, the table shows that elevation has the greatest impact in the highest frequencies. This seems reasonable, as the highest frequencies are the most directional, and thus would be the most sensitive to changes in orientation.

It is important to note that at this point, the concluding remarks above are only conjectural. Because of the nature of head-related transfer functions, it is difficult to conclude universal details about how listener orientation impacts sensory dissonance. Each HRTF is different for each individual, and there can be drastically different frequency responses for each individual. In turn, this property creates a vast variety in how each individual may perceive sensory dissonance. In order to understand the impacts of HRTFs and thus listener orientation on sensory dissonance in greater detail, a much more concerted effort must be conducted.

4.3 Creating and Visualizing a Dissonance Field.

After 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. In addition, we make the assumption that the emission and reception of all sounds are omnidirectional. Thus, we do not utilize head-related transfer functions for this representation.

The dissonance field is visualized utilizing the technique of isosurfacing. Each location's calculated sensory dissonance is stored into a matrix, where then the marching cubes algorithm is performed on the data matrix. Our implementation in the max/msp/jitter environment utilizes the jit.isosurf object. This object is very powerful, as it allows us to dynamically change isosurfacing levels. This lets us scan the dissonance field, revealing the contours and concentrations of dissonance throughout the space.

Following suit with the modeling above, the sound sources were constructed as band-limited saw tooth waves. Four tones were placed in a virtual space of 40 cubic meters. Of the four tones constructed, the first tone has a fundamental frequency of 440 Hz. The remaining tones were constructed on top of this as pitches lying in a 12-tet major scale. Table 6 below displays the tones constructed and their positions in the space.

Fundamental Frequency (Hz)	**Scale Degree**	**Source Location (x, y, z)**
440	Tonic	(8, 8, 8)
493	Major 2nd	(0, 16, 8)
554	Major 3rd	(16, 0, 8)
660	Perfect 5th	(0, 0, 8)

Table 6. Notes constituting the dissonance field

The position of the tones was selected based on their musical implications. The fundamental of 440 Hz was placed in the center of the space because of its foundational relationship as tonic to the other tones. The third and fifth were placed on the floor because they form a major triad with the fundamental of 440 Hz. Placing them here allows for the illumination of the dissonance relationship in the triad. The second scale degree was placed at the ceiling. As these are the most dissonant tones in the scale (particularly the second), we expect this portion of the room to be more dissonant.

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 images below in figure 4 display our results:

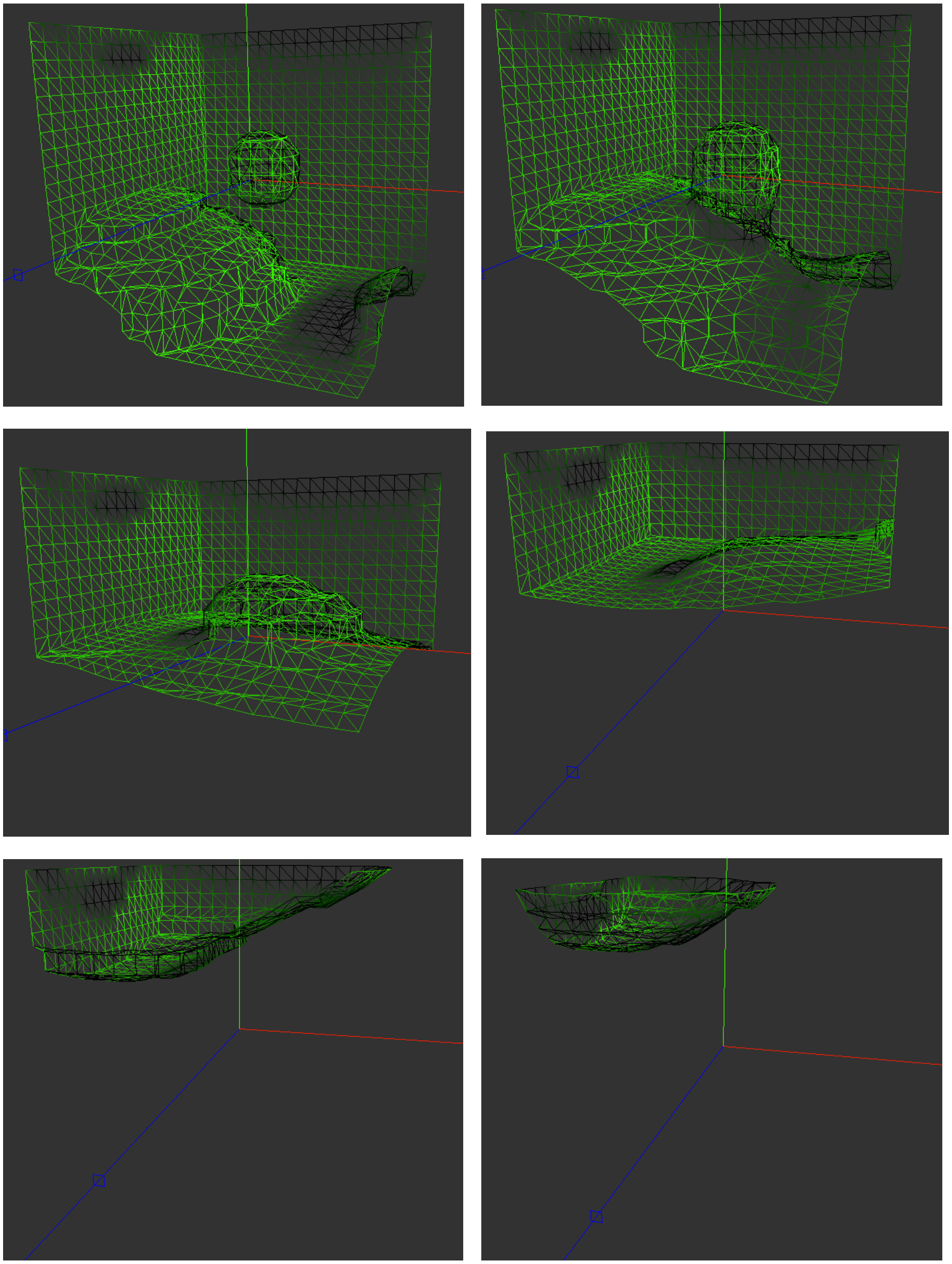

Figure 4. Dissonance field in the virtual space as represented by isosurfacing. Lower levels of dissonance are pictured in the upper left, progressing to the highest levels in the lower right.

The images clearly display the topography of the dissonance formed by the sound sources and their placement in the space. The images display snapshots of the different levels of dissonance present ranging from the lowest levels (upper left) to the highest levels (lower right). This representation gives us a unique insight into the relationships between the tones. It is not surprising that the lowest levels of dissonance are experienced about the placement of the fundamental and the fifth interval (upper left image). Next to this, as the dissonance increases, the topographical surface reveals the presence of the major third interval, and we have a clear visual sense about where the sensation of the common triad is perceived in the space. As the dissonance levels increase further, the topographical surfaces move towards the upper region of the space. It is in this region where the major 2nd interval was placed. As this forms the most dissonant interval against the root tone of 440 Hz, it is clear that this area of the space would have the most sonic roughness.

The construction of a dissonance field is our seminal 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. Some such ideas will be presented below.

5. Applications and Future Work

5.1 Immersive Environments and Sonic Sculpture

Recalling Lamonte Young's *Dream House* presented in the introduction, 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 Dreamhouse, 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.

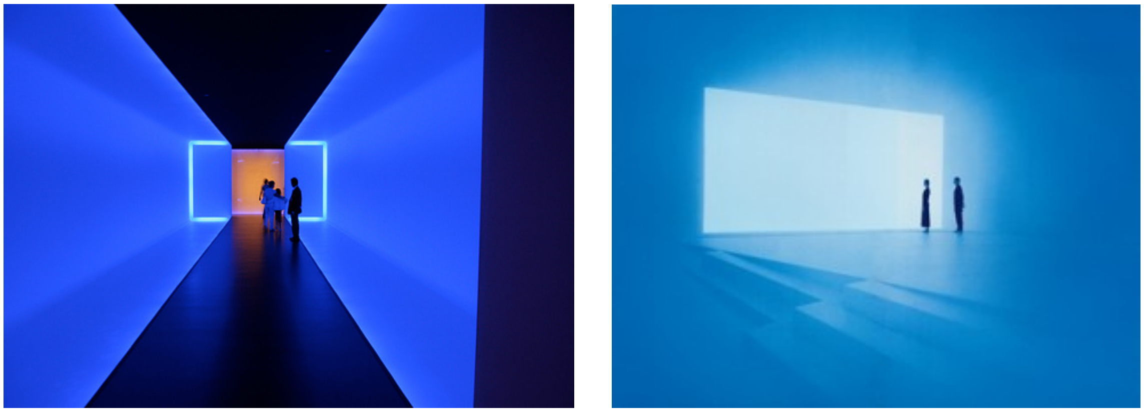

Figure 5. Works by James Turrell: The Light Inside (left) and Wide Out (right).

Utilizing our sensory dissonace spatialization model, we were able to construct 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 constructed with six sound sources, where each sound source was placed in the center of one of the six sides constituting the border of the space. Each sound source emanates a sine tone, and each tone is separated in frequency by 0.25 barks (the point of maximum roughness in a given critical band). The setup results in the arch like dissonance field caused by the convergence of the sounding tones.

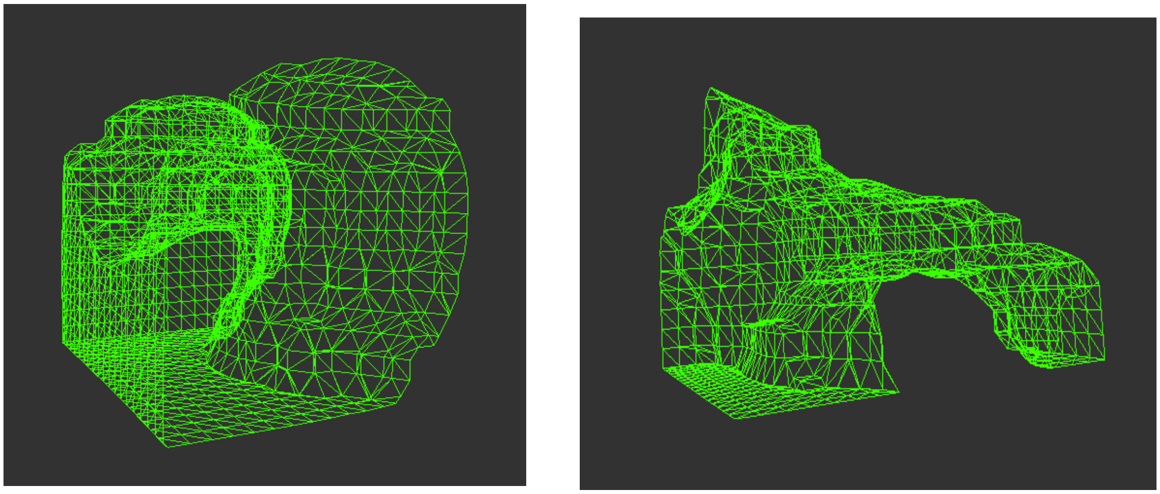

Figure 6. Two sonic sculptures: Curl (left) and Arch (right).

5.2 Future Work

Further enhancements can be incorporated into our model to more accurately construct sonic environments and refine sonic sculptures. As our model assumes an anechoic environment, it currently would not portray the dissonance field experienced by a listener indoors. Thus, the physical modeling of reflective surfaces must be considered for modeling the sound space. In addition, our sources are assumed to emit sound omnidirectionally, where in reality, sound is emitted directionally from speakers and other sources. Further, it is possible to generate highly directional sound or to create unique spatialization effects using wavefield synthesis or vector based amplitude panning (VBAP). Incorporating any of these strategies into our model would increase the potential and precision in our ability to sculpt the sound of a space. Finally, our sculptures are currently the result of static sonic structures. If we were to allow for dynamic panning or for the spectrum of sound sources to change over time, our dissonance environments would become highly dynamic. The result would be dissonance fields that ebb and flow with fluidity.

Bibliography

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. Cheng, C. (2001). Introduction to Head-Related Transfer Functions (HRTFs): Representations of HRTFs in Time, Frequency, and Space. *Journal of Audio Engineering Society*, Vol. 49, No. 4.
6. Chon H. C. (2008). *Quantifying the Consonance of Complex Tones With Missing Fundamentals*. Thesis submitted to Department of Engineering, Stanford University.
7. "Cosm for Max/MSP/Jitter" Web. 20 Nov. 2011. <http://www.allosphere.ucsb.edu/cosm/>.
8. "Head-related transfer function." *Wikipedia, the Free Encyclopedia*. Web. 20 Nov. 2011. <http://en.wikipedia.org/wiki/Head-related_transfer_function>.
9. 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.
10. Huron, D. (1991a). Tonal consonance versus tonal fusion in polyphonic sonorities. *Music Perception,* Vol. 9, No. 2, pp. 135-154.
11. Huron, D. (1993). A derivation of the rules of voice-leading from perceptual principles. *Journal of the Acoustical Society of America* Vol. 93, No. 4, p. S2362.
12. 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.
13. Hutchinson, W. & Knopoff, L. (1978). The Acoustic Component of Western Consonance. *Interface*, Vol. 7, No. 1, pp. 1-29.
14. "Inverse-square Law." *Wikipedia, the Free Encyclopedia*. Web. 20 Nov. 2011. <http://en.wikipedia.org/wiki/Inverse-square_law>.
15. 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.
16. 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.

17. 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.
18. MacPherson, E. (1994). *On the Role of Head-Related Transfer Function Spectral Notches in the Judgement of Sound Source Elevation*, Waisman Center, University of Wisconsin-Madison.
19. 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.
20. MacCallum, J., Einbond, A. (2007) *Real-Time Analysis of Sensory Dissonance* Center for New Music and Audio Technologies (CNMAT), Berkely, CA.
21. “Listen HRTF Database – WWW IRCAM. Web 9 Dec. 2011. <www.ircam.fr/equipes/salles/listen/>.
22. 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
23. Plomp, R. & Levelt, W.J.M. (1965). Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America,* Vol. 38, pp. 548-560.
24. Suzuki, Yoiti. (2003) *Precise and Full-Range Determination of Two-dimensional Equal Loudness Contours*.
25. Tenney, J. (1988). *A History of "Consonance" and "Dissonance."* White Plains, NY: Excelsior, 1988; New York: Gordon and Breach, 1988.
26. Vos, J. (1986). Purity Ratings of Tempered Fifths and Major Thirds. *Music Perception*. Vol. 3, No. 3, pp. 221–258.