Granular Model of Multidimensional Spatial Sonification

Muhammad Hafiz Wan Rosli[1], Andrés Cabrera[2], Matthew Wright[3], and Curtis Roads[4]

Media Arts and Technology,
University of California,
Santa Barbara CA 93106-6065 USA

Email: {[1]hafiz, [2]andres, [3]matt, [4]clang}@mat.ucsb.edu

ABSTRACT

Sonification is the use of sonic materials to represent information. The use of spatial sonification to represent spatial data, i.e., that which contains positional information, is inherent due to the nature of sound. However, perceptual issues such as the *Precedence Effect* and *Minimum Audible Angle* attenuate our ability to perceive directional stimuli. Furthermore, the mapping of multivariate datasets to synthesis engine parameters is non-trivial as a result of the vast information space. This paper presents a model for representing spatial datasets via spatial sonification through the use of granular synthesis.

1. INTRODUCTION

Sonification is the process of representing data through the sound domain. It has been proven that sonification is able to augment the visual display, and provide a platform for perceiving data for the visually impaired. Furthermore, a greater number of variables can be presented in one display, and some types of data are better suited for the sound domain, such as rapidly changing information.

However, the representation of data via visualization has been much more developed, as opposed to its auditory counterpart. In the case of spatial datasets (which contain location components along with other dependent variables), the mapping process from data to sound is inherently complex due to the dimensionality of the dataset.

In order to map the spatial data to spatial sound via the use of multiple loudspeakers (surround sound), perceptual attributes that pertain to the auditory system, such as the limitation of spatial acuity, has to be taken into consideration.

Through our research, we have developed a granular model for multimodal representation of spatial data via spatial sonification. Perceptual issues that pertain to spatial attributes (surround sound), and microsound are discussed, and taken into consideration in developing the system. Mapping strategies for granular synthesis (in the context of sonification) are also explored and described. Finally, we present how transformation of (micro) sounds, and interactivity could assist in the discovery of unknown patterns in a dataset.

2. MOTIVATION

One family of sounds is known as *microsound* [1]– sound particles in the range of 1 ms to 100 ms, that span the boundary between what can be perceived as an individual entity, and what has no distinct perceptual characteristics.

Motivated by the interest of this specific family of sounds, we started to search for natural occurrences that fall into the same category. One particular phenomena that interested us was the sound of thunder, as this is, in fact the result of a burst of electrical energy lasting for about approximately 20 µs. As the mass of energy is introduced, part of the energy is transferred into light (lightning) and sound (thunder).

On April 30th 2014, the *High Definition Earth Viewing* experiment was activated aboard the *International Space Station*, which presents viewers with images of Earth seen from outer space. One of the elements that was clearly shown was lightning occurrences. Guided by the interest in this particular phenomenon, we started to inquire if there were any correlations between lightning strikes in different geographic locations. This curiosity lead us to examine NASA's lightning dataset, and develop a research based on multimodal data representation, which resulted in a perceptual model of sonification– shown in the artwork *Point cloud* [2].

Sonification as a means to understand and display datasets, as well as the use in aesthetic explorations has been explored by various disciplines [3–8].

3. DATA SOURCE

The data for *Point Cloud* were gathered from the NASA Marshall Space Flight Center (MSFC) [9]. The data were generated by two of their space-borne optical sensors: the Optical Transient Detector (OTD) and the Lightning Imaging Sensor (LIS). The actual data span more than 16 years of lightning information, but for the purpose of this project, only a single year's worth of data was used.

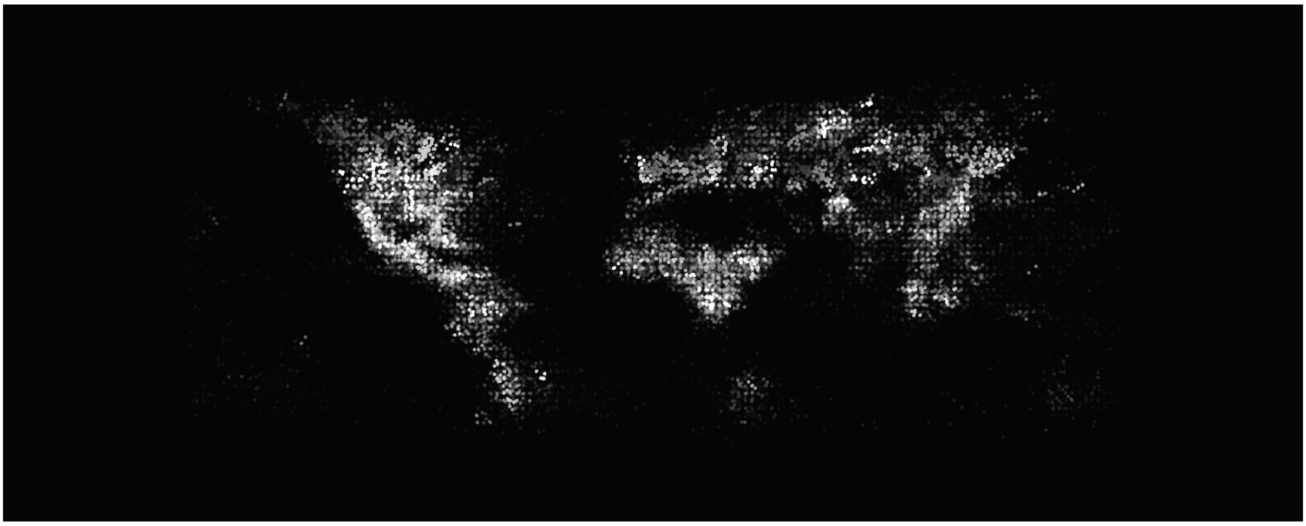

Figure 1. Visualization of *Point Cloud*: A map of the Earth showing lightning occurrences over the course of a single year (day 333).

The data take the form of an annual cycle of flash rates with a product dimension of 720 x 360 x 365 (bin size of 0.5 degrees × 0.5 degrees × one day) (Fig. 2). Each *data slice* (single day) contains a *flash rate* (number of occurrences per day) for each spatial *data point* (720 x 360).

The MSFC dataset is distributed in Hierarchical Data Format [10] and was parsed using the Python PyHDF library.

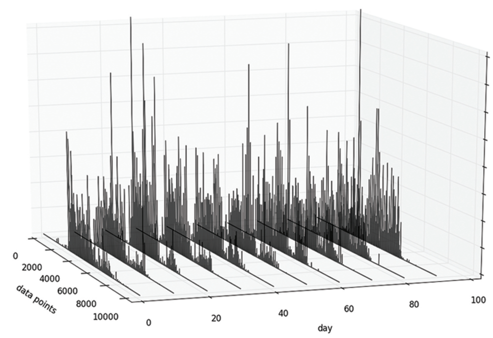

Figure 2. Lightning occurences for flattened [720 x 360] data points over series of days

4. SONIFICATION

Data sonification in auditory displays can provide information to complete, augment, or replace visual displays. As with most systems that present data, sonification techniques aim to provide the means for extracting information from a dataset to be parsed by the perceiver [3]. Additionally, sonification allows us to potentially extract new patterns and relationships, which are not necessarily perceived by simply analyzing the dataset.

The domain of data visualization faces similar issues, as it deals not only with the dataset being represented, but has to also consider the user as part of the process in parsing information.

4.1 Multimodality and Sonification

We typically take for granted how our senses parse information in our daily lives. When we interact with the real-world, we almost always receive feedback through various modalities, allowing us to understand our surrounding, and successfully navigate through the environment. Our perceptual system functions by correlating the information from these various senses to construct a mental image of our surrounding.

In order to design an effective sonification system, these well developed mechanisms of perception need to be involved. Auditory and visual stimuli needs to be coupled by the same mechanism which couples perceptual units in the real world to create a cohesive *environment* [3]. There have been efforts to simultaneously present data in both the auditory and the visual domain, in order to give the perceiver a more concise understanding of the dataset [3]. Such an example can be seen in *Point Cloud* (Fig. 1) [2].

4.2 Perceptual Issues

Our perception tends to fail at grasping the bigger picture when presented with a single viewpoint of information. This phenomena is exemplified, for instance, by the need to interactively change perspectives while viewing a 3-dimensional structure in the real world. When doing so, we allow ourselves to acquire different views, which then provide better sense of the object.

Sonification suffers from an analogous problem. When a complex data space is projected onto a linear audio signal, we are unable to gain different sonic views of the dataset, rendering the system less effective. One solution is to change the mapping parameters so that the perceiver is able to acquire a variety of sonic perspectives [3], discussed in Section 5.2. Additionally, the use of sound spatialization enables us to address this problem [11].

Sound spatialization is used in our system to assist a perceiver in gaining multiple perspectives of the dataset. In doing so, the complex data space is not collapsed into a single audio signal originating from one direction.

4.3 Spatial Sound

> "A cascading sequence of sound objects, each emanating from a different virtual space, provides the dimension of spatial depth to an otherwise flat perspective and articulates a varying topography" [12].

The use of multiple loudspeakers for spatial sound reproduction allows stimuli to be presented from different locations, preventing the complex data space from collapsing into a single audio signal originating from one direction. This, in turn allows a perceiver to interactively navigate around the acoustic environment in order to acquire different sonic views. Similarly, this mode of interaction is how humans localize acoustic energy in the physical world [8, 13]. Nasir and Roberts [11] explains that "location information can be used to enhance the sonification, or can be used to represent qualitative information."

Similarly, representation of spatial data via the means of visualization has had its share of exposure, dating back to the thorough dissection by semiologist Jacques Bertin [11, 14].

As we are dealing with a dataset that presents spatial information within specifically localized regions, it is only natural to include spatialization as a key aspect of the representational system. By correlating each spherical coordinate of the earth to the spatial position in the rendered (sound) field (longitude and latitude mapped to azimuth and elevation), we enable the auditory stimulus to be localized at its respective position. In other words, a perceiver would be "looking" at the dataset from a viewpoint inside the Earth.

One of the main objectives of the underlying research is to represent a single spatial dataset via multimodal stimuli. However, spatial acuity is much finer for vision than it is for hearing [3]. In order to allow users to distinguish stimuli coming from separate distinct locations, some consideration needs to be addressed.

4.4 Localization of sound

In order to effectively convey the perception of space, we have explored methods of sound localization, specifically those that pertain to sonification, as discussed in [11]. These methods are also thoroughly examined in various texts concerning sound spatialization, such as [3, 8, 15].

Non- spatial audible variables: These are the building blocks of sonification, which typically includes synthesis parameters such as pitch, loudness, and tempo. As discussed in Section 5.2, we have mapped the flash rate value of every data point to its corresponding granular stream's grain density, and grain amplitude.

Non- spatial motifs: These higher order components are intended to provide a better system for the perceiver to understand patterns in the dataset. Description of our implementation can be found in Section 6. Although these specific structures typically needs to be learned, we believe that the human brain is able to adapt, and find patterns in these higher-level dimensions– if there are patterns to be perceived.

IID & ITD: The Duplex Theory [16] states that we perceive directionality (and auditory space in general), through the use of *Interaural Time Difference* and *Interaural Intensity Difference*. The use of multiple loudspeakers allow us to successfully use this mechanism in placing a localized stimuli in a radial space. Due to the fact that this mechanism is a well developed component of our perception, the spatial data could be perceived without further training.

Time-based effects: Temporal factors provide excellent cues for sonic data exploration. The ability to traverse different time scales provide the means to understand various hidden structures in a dataset. We have explored temporal transformations to analyze *Microstructures* and *Macrostructures* in the dataset (Section 6.3).

5. SYNTHESIS ENGINE

Some of the various synthesis techniques used for sonification are more suitable than others, depending on the data that is analyzed (for example, in the case of multidimensional datasets). Undoubtedly, most of the "effective" sonification systems consider the dataset, and implements techniques that would best fit the data.

Our system is implemented using *Parameter Mapping Sonification* due to its effectiveness in displaying multivariate data [3]. This technique involves the mapping of data features onto parameters of sonic events, such as pitch, level, and onset time. Our model implements granular synthesis as the synthesis engine in order to render short, discrete events in the dataset (lightning occurrence). Furthermore, these short bursts of energy resembles the sound of thunder, which in turn enhances the effect of *Gestalt Principle of Past Experience*.

As the visualization algorithm displays a unit of occurrence for a brief period of time, the sonification engine renders a short burst of sonic energy. This allows the auditory and visual stream to be coupled together, as discussed in 4.1. We respect natural physical coherences by binding the visual and auditory events together temporally, giving the impression of causality [3, 14].

5.1 Granular Reverberation

Thunder is the result of a shock wave caused by a sudden thermal expansion as lightning passes through the air. A typical lightning bolt lasts for about approximately 20 μs. As the mass of energy is introduced into the cumulus cloud enclosure, its impulse response takes the shape of irregularly spaced delays as a result of spectral reflections in the cloud formation [17].

This effect can be synthesized using a technique known as granular reverberation. The foundation for granular reverberation is *Asynchronous Granular Synthesis*, which scatters grains statistically within a region defined on the time/frequency plane (Fig. 3) [18]. The number of grains in a particular region determines the density of sound particles (for each granular cloud).

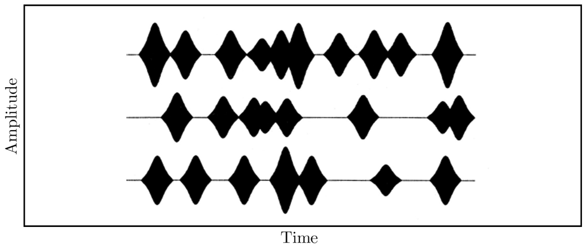

Figure 3. Varying density of 3 granular streams mapped to flash rate [19]

5.2 Mapping Strategies

The data of lightning occurrences is presented as a time series corresponding to the days in a single year. Every *data point* (Fig. 4) in the dataset holds a value corresponding to the amount of lightning (flash rate) for a particular geographic coordinate (longitude, latitude) of a given day. To simulate the individual lightning strikes (while retaining the ratio between each data point), we have chosen to map the flash rate values to the *density* of grains (discussed in Section 5.2.1).

Although this is not the "actual individual lightning occurrence," the triggering rate of the grains somewhat gives us a cue of how "dense" the occurrences are around a particular part of the world. The *maximum amplitude* and the *duration* of the grains in a particular stream are also correlated to the values of each data point, as a means to intensify the data mapping. Other synthesis parameters such as *grain triggering rate* and *grain length random deviation* are mapped to stochastic processes.

The use of synthetic grains with sharp *attack* and *decay*, and a lifespan between 10 ms to 50 ms enables us to evoke the idea of individual bursts of lightning bolts. The short barrage of energy also results in a very strong association with the rendered points in the visualization. This, in turn, results in the tendency for these two different stimuli to be grouped together as an interconnected event.

The nature of this technique allows for a multitude of low-level parameter manipulations. In contrast, the usage of granular synthesis in creative applications (as opposed to sonification practices) often requires thousands of control parameters per second, resulting in the need of higher-dimensional control parameters.

> "Granular synthesis requires a massive amount of control data. If *n* is the number of parameters per grain, and *d* is the density of grains per second, it takes *n* times *d* parameter values to specify one second of sound" [1].

On the other hand, multivariate datasets provide us with a wide range of parameters that could be assigned to granular synthesis' control data. The issue lies in fine tuning the synthesis engine to fit the dataset, and finding the best ways to *parameterize*, so as to allow changes in the data to be perceived by the user. This is where aesthetic features of the sonification plays an important role– to allow the dataset to be cognized by the user.

5.2.1 Parameterization

For every flash rate value in the dataset, we create a Gaussian function (1) centered on the *data point*. The normalized flash rate is then set to the height of the curve's peak (Fig. 5). In effect, sonic grains are statistically rendered around the data points, based on their actual location in the dataset (Fig. 6). Consequently, the density of grains in a particular area now gives a perceptual description of occurrences in that region.

$$f(x) = a \exp(-\frac{(x-b)^2}{2c^2}) \qquad (1)$$

where a, b and c are real constants

Additionally, this allows the algorithm to retain each data point's *relative weight* compared to other points on the same data slice, independent of the temporal scaling (discussed in Section 6.3). One could implement an algorithm that is set to render a non-statistical element at each data point, but the result would be a repeating cycle, akin to looping an audio file. Instead, the synthesis engine renders a sequence of grains for each data point on the grid. As such, the density of grains at a particular location in time retains its overall weight every cycle. However, it does not appear to be an exact repetition of the previous cycle as a result of statistically generating new grains every time the data point is updated.

5.3 Grain Density

The attempt to provide perceptually distinct cues for individual lightning occurrence also causes the data to be obscured. The number of grains (per data slice) at an instance becomes far too dense for the differences to be perceived [1]. We discuss the technique of focusing on individual streams in Section 6.2.

Our auditory system perceives events happening with intervals less than 20 Hz as distinct events. However, as these events are sped up to more than 20 Hz, they are perceived as a continuous stream. The visual equivalent of this can be seen in the phenomenon called *Persistence of Vision*. At 30 frames per second, our brain processes these visual stimuli as a continuous event.

In the case of the represented dataset, the speed at which the grains are generated would correspond to grain density per data point. When we take the whole data slice into account, we get a number of grains that is generated at a rate that temporally smears the grains into a continuous tone.

[1] https://soundcloud.com/muhammad-hafiz-wan-rosli/graindensity

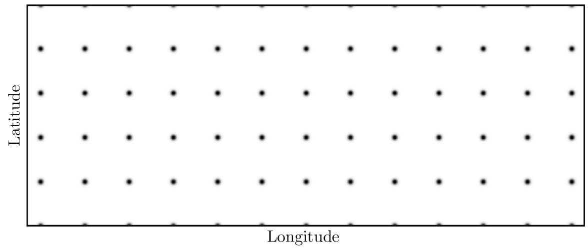

Figure 4. An array of data points for a single day

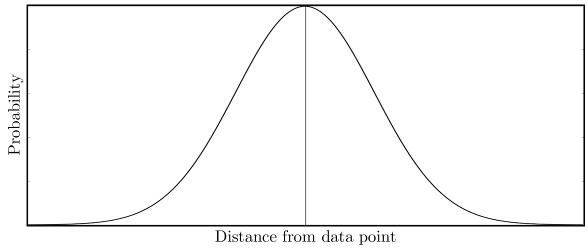

Figure 5. Probability curve of one data point in Figure 4

Suppose we have an average grain density of 100 grains per second for a single data point. The number of grains in one second of time would be:

[longitude resolution ∗ latitude resolution ∗ grain density]

$$720 \times 360 \times 100 = 25,920,000 \text{ grains per second} \quad (2)$$

Therefore, the problem of data representation through the means of granular synthesis is somewhat reduced to a perceptual and psychoacoustical problem. How do we pose a potential solution to parse this dense information space acoustically?

6. GRANULAR TRANSFORMATION

> "Xenakis observed how sound particles could be viewed as short vectors within a three dimensional space bounded by frequency, amplitude and time" [1].

6.1 Frequency Transformation

As is well known, the Cocktail Party Effect illustrates that our brain is capable of focusing auditory attention on a particular stimulus while filtering out a range of stimuli [8, 20] [2]. However, this effect is influenced not only by our ability to segregate sounds based on their spectral and temporal qualities, but also by the spatial relationships between the sounds.

Consider, for example, the ability to segregate multiple instruments in a recording, and to focus on a specific instrument. We are able to do so because we associate each unique instrument with a specific timbre (and melodic motives). Furthermore, our ability to identify unique instruments is also affected by the direction of which the sounds originate from [8, 13]. These cues help us localize sound

[2] http://sonification.de/handbook/media/chapter3/SHB-S3.1b.mp3 [20]

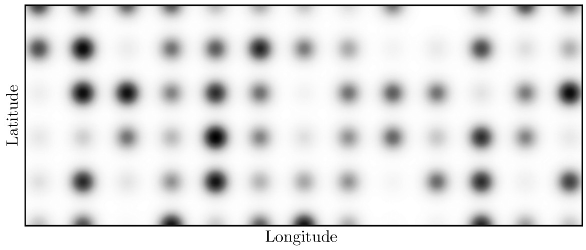

Figure 6. Grain distribution for data points in Figure 4

sources, which ultimately contributes to recognizing a stimulus as a continuous pattern.

6.1.1 Unique bands per quantized longitude

> "The ability to selectively attend to simultaneously sounding *auditory objects* is an ability that is not yet completely understood. Nonetheless it provides fertile ground for use by designers of auditory displays" [3].

Bandlimiting a set of grains allow us to theoretically differentiate between separate groups of stimuli, i.e "granular streams". As discussed in section 5, the synthesis engine exploits our perceptual ability by bridging the connection between what is seen and what is heard. However, the downside of using granular means for synthesizing the stimuli disables us from clearly identifying the separate grains, causing the mass of sounds to be perceived as a single evolving event. Although this effect is useful in parsing the overall macro pattern, the microstructure tends to lose its meaning through the dense cloud of sound particles.

By mapping the differences in azimuth (of the dataset) not only to its corresponding spatial position (in the rendered field), but also to a specific frequency band (Fig. 7), we allow the data to be segregated based on its spectral content and spatial location. However, the generation of synthetic (sinusoidal) grains is far too similar to one another, even with the assistance of spatial relationships.

This effect is further diminished due to the Minimum Audible Angle, which is defined as the *Just Noticable Difference* (in azimuth) for listeners [21]. One solution might be to fine tune the quantization of longitudinal space to fit the space where the model will be rendered in. The number of speakers used affects the ability to render directional stimuli, which, in turn, helps in segregating directional sources. We intend to further explore this possibility in the near future via the use of the *Allosphere* [22].

6.1.2 Granular clouds

The grouping of elements is further explored by segregating groups of grains to form what is known as granular clouds. If a set of grains are bounded by a pre-determined set of rules and parameters, then they would appear to morph "in unison". We implemented this technique for the different continents, which allowed us to analyze the trend of change per continent, and how one continent's flash rates relate to another's. In doing so, we now reduce the amount of concurrent events to “concentrate” on, enabling us to analyze the macroscale patterns. Here lies another example of how mapping a dataset to a higher-level representation could give rise to new meanings, and allow us to find patterns that were otherwise difficult to perceive (or even non-existent).

6.2 Amplitude Transformation

As discussed in [2], our visual system is able to focus on a group of stimulus in a specific position, while disregarding the other stimuli– akin to looking through a magnifying glass. Although our auditory system is able to segregate, and focus on different stimulus [20], the dense spatial dataset prevents a perceiver to tune in to specific areas of interest. To achieve a similar result (as the visual senses), we have implemented a means to “blur out” or “smear” the dataset– except for the area being viewed. This notion of *Interactive Data Selection* has been implemented, and discussed in the context of Parameter Mapping Sonification [3, 23, 24].

6.2.1 Acoustic focus

To focus on a specific area of the dataset, we *pass* the sonic grains through a conditional construct that checks if the generated grains are within a specific boundary. If these grains are within the boundary, then they are rendered. Otherwise, the grains are not rendered.

If we were to perform the conditional statement on a data point, instead of the rendered grains, we would not be able to render areas in between the data points. Instead, we are now able to seamlessly move the focal point around the dataset, while rendering grains that are only within the boundary.

Implementing this type of control not only enables us to focus on a specific data point, but also allows us to control the width of the scope, i.e the number of data points to be included in the focused region. Another parameter that is now at our disposal is the ability to control the loudness roll-off of the regions around the focused area.

6.2.2 Multiple focus

The number of focused regions could also be controlled (Fig. 8), so as to allow the perceiver to, for example, compare the data of several areas of interest. This control scheme reinforces the effect of temporal, and frequency transformations.

6.3 Temporal Transformation

Time domain transformation is well known in the realm of electronic music, discussed in depth by composers and musicians alike, including Stockhausen in his 1972 lecture entitled *Four Criteria of Electronic Music*. Speeding up a sequence of rhythmic events causes a transformation from distinct individual events perceived as rhythmic, into a continuous tone. Further increment of the speed creates an increase in pitch, whilst a decrease in speed results in a lowering of the pitch.

Temporal transformations allow us to traverse between time scales to perceive different relationships in the dataset’s temporal structure. In the case of our implementation, it allows us to perceive differences in *Microstructure* and *Macrostructure*.

6.3.1 Microstructure

The analysis of fluctuations in lightning occurrences for a particular location might not be a trivial task, as there are a multitude of concurrent granular streams. Coupled with the ability to segregate granular streams via amplitude transformations (Section 6.2), the relationship of one particular data point through time would be easier perceived if the temporal domain is stretched.

6.3.2 Macrostructure

On the other hand, if we were to compress the temporal domain, the distinct granular streams would be transformed into continuous tones [3]. A crucial point to note is that these manipulations do not effect the ratio between data point values (in a data slice). Therefore, the individual flash rates per time frame retains the same weight throughout the temporal transformation. What seemed to be a mass of micro-events resembling noise is now transformed into (720 x 360) pitched tones.

As a result, we can now compare data points over time by listening to the differences in pitches: The higher the pitch of a particular data point, the higher the lightning occurrence. Additionally, we can now analyze the data to extract higher level information, such as the ratios between tones (how the flash rate of a particular point relates to the flash rate of another point), the amount of frequency shift (glissando) corresponding to the changes in flash rates of a particular location (data point), and the rate of frequency shift in relation to another data point (rate of change for flash rates).

[3] https://soundcloud.com/muhammad-hafiz-wan-rosli/graintemporaltransformation

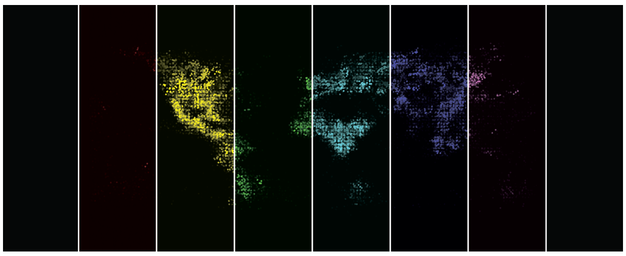

Figure 7. Bandlimiting grains per quantized longitude

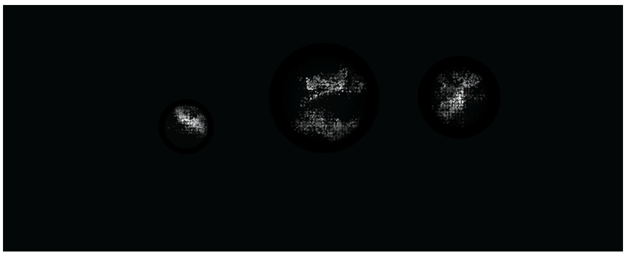

Figure 8. Multiple focus and loudness roll-off

If a granular stream does not change through time, then the number of occurrence for that data point does not fluctuate. In effect, we can hear the fluctuating changes of a particular location by listening to the granular streams that change from tone to rhythm, and vice versa.

7. INTERFACE FOR EXPLORATION

The mapping of spatial data to spatial sound was indeed one of the crucial components of this research. However, as the perceiver is an important component to the sonification, interface, naturally becomes an important factor as well.

We are currently in the process of designing custom hardware that would enable us to address spherical coordinates via a multi-touch spherical interface. However, we have explored (off-the-shelf) interfaces to navigate the system, which has produced satisfactory results.

Frequency and Amplitude transformations, as well as zooming capabilities are executed via the use of trackpad or *Wacom* tablet [25]. The *Griffin Powermate* [26] was used to achieve temporal transformations, whereas a *Graphical User Interface* was used for parameter control.

As a compositional aesthetic, we have also included a mode where the rate of reading the data slice is increased after every yearly iteration. In the "final" iteration, the whole year's worth of data would be presented as a single impulse, which contains the "energy contained in one year's worth of lightning/ thunder".

8. TOOLS

The initial version of the system was realized using a heterogeneous setup to allow fast prototyping. The data parsing, handling and processing was done using python, in particular the interactive ipython notebook. The sonification was done using Csound [27] within the ipython notebook, and the visuals were done using processing [28]. The synchronization and data interchange between applications was done using Open Sound Control [29].

9. CONCLUSION AND FUTURE WORK

We have explored, and described a granular model for multimodal representation of spatial data via spatial sonification. We have also discussed the perceptual issues that are taken into consideration, and propose solutions to overcome them. Interactivity, mapping strategies and transformations related to granular synthesis have also been examined to provide a platform for pattern discovery. These techniques can be used as an auditory display to complement a visualization component, as well as an interactive tool for sonic data exploration.

We are currently in the process of porting the system from its *surround-sound* version to a large immersive 3D space, the *Allosphere*. This would allow us to render sounds in a *periphonic* (full 360^o) environment. Explorations on spatial interfaces will also be carried out as we believe interactivity is a major component to sonification. The Allosphere is a 3-story facility that contains a 10 meter diameter sphere that provides 360^o realtime stereographic visualization using a cluster of servers driving 26 high-resolution active-stereo projectors. Audio is projected by 54 loudspeakers positioned along three rings of the sphere [22, 30].

Another approach to the granular paradigm is through the use of granulation, which divides a sample into short enveloped grains, and reproduces them in high densities (as opposed to generating *synthetic* grains via granular synthesis). Granulation of samples possesses a unique, *organic* aesthetic quality which could assist in unraveling the "poetics" of the dataset, which in turn could allow users to be more perceptually engaged. Unique spectral transformations could also be applied to selected areas in order to assist the user in data exploration– examples of these transformations range from user defined systems to algorithmic processes, such as *Dictionary- Based Methods* [12].

9.1 Perceptual validation

We plan to conduct several user studies to analyze the effectiveness of our system. The following are potential scenarios of a user-study.

The users are exposed to multiple 10 second segments of the sonified dataset, specifically those which contain correlation in the change of lightning occurrences over time between two data points. The excerpts would contain a combination of various segments (both sparse and dense) to be analyzed by the user. In each segment, the user is presented with the dataset, sonified using well-known spatial sonification techniques [11], followed by our system.

For every segment, the user is asked:

- **If there were any correlations in the data**. The user would be asked to determine the number of perceived distinct stimuli. They are are also asked to determine which data point contains more occurrences using the frequency and/ or density difference (Section 6.3).
- **To point towards the source of the incoming stimuli**. Users are allowed to navigate (Section 7) via discussed transformations (Section 6). The accuracy is measured based on radial distance from actual stimuli.
- **Which type of sonification is preferred**. This qualitative selection is compared to the quantitative result of the tests, and the correlation between aesthetics, and accuracy (function) is measured to determine their interdependence.

The result of this user-study would allow us to quantitatively measure our system, and show if multimodality in spatial data representation assists the accuracy of data perception. Furthermore, it would also show if the aesthetics of a system play an important role in the perception of data (in a data representational system). Additional tests would include different datasets, and different synthesis techniques.

10. ACKNOWLEDGMENTS

We would like to thank CREATE and the Allosphere Research Group for the use of facilities. This work was supported in part by the Graduate Program in Media, Arts and Technology at the University of California, Santa Barbara.

11. REFERENCES

[1] C. Roads, *Microsound*. MIT Press, 2004.

[2] M. H. W. Rosli and A. Cabrera, "Application of gestalt principles to multimodal data representation." IEEE VIS, 2014.

[3] T. Hermann, A. Hunt, and J. G. Neuhoff, Eds., *The Sonification Handbook*. Berlin, Germany: Logos Publishing House, 2011.

[4] J. A. Dribus, "The other ear: A musical sonification of eeg data." International Conference on Auditory Display, 2004.

[5] H. van Raaij, "Listening to the mind listening - a sonification/composition." International Conference on Auditory Display, 2004.

[6] E. Childs, "Icad 2006: Global music 8212;the world by ear," *Computer Music Journal*, vol. 31, no. 1, pp. 95–96, March 2007.

[7] M. H. W. Rosli and A. Cabrera, "Gestalt principles in multimodal data representation," *Computer Graphics and Applications, IEEE*, vol. 35, no. 2, pp. 80–87, Mar 2015.

[8] A. S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press, 1994.

[9] "Nasa marshall space flight center," http://www.nasa.gov/centers/marshall/, accessed: 2014-06-28.

[10] "The HDF group - information, support, and software," http://www.hdfgroup.org/, accessed: 2014-06-28.

[11] T. Nasir and J. C. Roberts, "Sonification of Spatial Data." International Conference on Auditory Display, June 2007, pp. 182–196.

[12] A. McLeran, C. Roads, B. L. Sturm, and J. J. Shynk, "Granular sound spatialization using dictionary-based methods," in *Proceedings of the 5th Sound and Music Computing Conference*, Berlin, Germany, 2008.

[13] L. Rayleigh, *The Theory of Sound*. Dover Publications, 1945.

[14] M. Chion, C. Gorbman, and W. Murch, *Audio-vision: Sound on Screen*, ser. Film and Culture. Columbia University Press, 1994.

[15] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*. MIT Press, 1997.

[16] L. Rayleigh, "On our perception of the direction of a source of sound," *Proceedings of the Musical Association*, vol. 2, pp. pp. 75–84, 1875.

[17] M. Uman, *Lightning*, ser. Dover Books on Physics. Dover Publications, 1969.

[18] C. Roads, *The Computer Music Tutorial*. MIT Press, Jan. 1996.

[19] ——, "Representations of musical signals," G. De Poli, A. Piccialli, and C. Roads, Eds. Cambridge, MA, USA: MIT Press, 1991, ch. Asynchronous Granular Synthesis, pp. 143–186.

[20] A. W. Bronkhorst, "The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions," *Acta Acustica united with Acustica*, vol. 86, no. 1, pp. 117–128, January 2000.

[21] A. W. Mills, "On the minimum audible angle," *The Journal of the Acoustical Society of America*, vol. 30, no. 4, pp. 237–246, 1958.

[22] X. Amatriain, J. Kuchera-Morin, T. Hollerer, and S. T. Pope, "The allosphere: Immersive multimedia for scientific discovery and artistic exploration," *IEEE Multimedia*, vol. 16, no. 2, pp. 64–75, 2009.

[23] D. Ó. Maidín and M. Fernström, "The best of two worlds: Retrieving and browsing," in *In Conference of Digital Audio Effects*, 2000.

[24] T. Hermann, T. Henning, and H. Ritter, "Gesture desk, an integrated multi-modal gestural workplace for sonification." Springer Berlin Heidelberg, 2004, vol. 2915, pp. 369–379.

[25] "Wacom pen tablet," http://www.wacom.com/en-us/products/pen-tablets, accessed: 2015-03-03.

[26] "Griffin powermate," https://store.griffintechnology.com/powermate, accessed: 2015-03-03.

[27] R. C. Boulanger, *The Csound book: perspectives in software synthesis, sound design, signal processing, and programming*. MIT press, 2000.

[28] "Processing.org," https://www.processing.org/, accessed: 2014-06-28.

[29] M. Wright and A. Freed, "Open sound control: A new protocol for communicating with sound synthesizers." Thessaloniki, Hellas: International Computer Music Association, 1997.

[30] J. Kuchera-Morin, M. Wright *et al.*, "Immersive full-surround multi-user system design," *Computers & Graphics*, vol. 40, pp. 10–21, May 2014.