# I HEAR NY4D: HYBRID ACOUSTIC AND AUGMENTED AUDITORY DISPLAY FOR URBAN SOUNDSCAPES

*Braxton Boren, Michael Musick, Jennifer Grossman, Agnieszka Roginska*

Music and Audio Research Laboratory
New York University
New York, NY
{bbb259, musick, jkg297, roginska}@nyu.edu

## ABSTRACT

This project, I Hear NY4D, presents a modular auditory display platform for layering recorded sound and sonified data into an immersive environment. Our specific use of the platform layers Ambisonic recordings of New York City and a palette of virtual sound events that correspond to various static and realtime data feeds based on the listener's location. This creates a virtual listening environment modeled on an augmented reality stream of sonified data in an existing acoustic soundscape, allowing for closer study of the interaction between real and virtual sound events and testing the limits of auditory comprehension.

## 1. INTRODUCTION

One advantage of a controlled audio reproduction environment is the ability to circumvent the natural tendency of the auditory system to normalize itself with respect to its current acoustic surroundings. Providing quick A/B comparisons between auditory scenes makes differences in loudness, texture, or clarity stand out to listeners more distinctly than if they visited each site in person.

This advantage has long been known and was the primary motivation for Schroeder's famous crosstalk cancellation system, which he used to provide pairwise comparisons of concert hall parameters for listeners in a controlled environment [1]. A similar desire underlay the I Hear NY3D project at NYU [2, 3], which allowed listeners to navigate to different Manhattan soundscapes instantly via Ambisonic recordings made at 14 locations throughout the island. SPL measurements taken during each recording allowed the reproduction system to be calibrated such that the same SPL was reproduced during playback. This allowed users to quickly hear changes not only in texture and spatial composition, but also the absolute dynamic range of the acoustic system that is New York City.

These extensive recordings and measurements provided a useful tool for interacting with the urban soundscapes around us, but depicted only a single timeframe at each location. As a next step, we began locating additional spatially oriented data feeds that could add spatial and temporal context to the soundscapes presented. This project, I Hear NY4D, adds 3D sonifications for static and dynamic data at each location, allowing users to hear changes in many different virtual layers as well as environmental sounds across Manhattan.



## 2. BACKGROUND

### 2.1. Soundscapes

R. Murray Schafer believed that modern urban life is 'schizophonic' because it is detrimental to look at a sound as distinct from its natural environment [4]. In contrast, Augoyard and Torgue [5] argue that every urban moment has a sound signature and no sound can be completely isolated from its environment because it both holds a set of particular DNA of its original space and time and also immediately takes on the qualities of whatever new space it encounters. Sound cannot be limited to the description of signals, but rather "Sound is a propagation and is therefore directly connected to circumstances linked to constructed environment and conditions of hearing and listening: filtration, anamorphosis, listener's location: space and sound integrally linked" [5]. Sounds are also shaped subjectively due to auditory capacity, listening culture, attention, and attitude. In the urban environment, there is already a soundscape occurring, but it is also shaped by how we perceive it. I Hear NY4D illustrates both of these ideas. Depending on the lens through which you observe it, be it through data, recordings, or interpretive layers of sound, the soundscape is complex and holds various layers of meaning [5].

### 2.2. Sonification

While Hermann [6] gives a more rigorous set of prerequisites for what constitutes a sonification, we will here only use the simpler definition given by Kramer et. al.: "the use of non-speech audio to convey information" [7]. More broadly, a 3D recording which seeks to accurately represent the locations of acoustic data in a real-world setting could be a type of sonification of the sources' spatial positions (with the caveat that such recordings will sometimes include speech from the location, but that this speech is included for the purpose of showing the acoustic and spatial attributes of a talker rather than the semantic information represented in the speech itself).

With this realization, the barrier between natural and virtual acoustic events in a sonification becomes more arbitrary, since other events with temporal and spatial data could also be included. Indeed, some work has already used this spatial parallel to use virtual sonification to create auditory maps for the blind and visually impaired [8]. By injecting virtual auditory icons into the same spatial layout as a real-world location, users can generate an accurate mental map of important landmarks with only auditory data [9]. Besides static landmarks, there is already significant infrastructure for the representation of dynamic real-time events through online data feeds of various types. One of the best examples of this is the Tweetscapes of Hermann et. al. [10], which sonified Twitter data from across Germany in a variety of different streams. This allowed some spatial designations of tweet origins, but the project mainly focused on the temporal representation of the fast, fleeting world of social media, in which events happen quickly and may be forgotten just as quickly or instead go viral. By combining the best aspects of these projects, we hoped to convey both spatial and temporal positions for static and dynamic sources.

### 2.3. Augmented Virtual Reality

This work also touches upon work in the field of augmented reality (AR) audio, which seeks to extend a real sound environment "with virtual auditory environments and communications scenarios" [11]. AR is distinguished from virtual reality (VR) in its emphasis on adding to a real environment rather than constructing a wholly synthetic sight or sound field that seeks to stand in for a real environment. Yet the distinction between the two is not always so simple, leading some to propose the term mixed reality (MR) to designate a continuum between the extremes of AR and VR technologies.

AR audio systems typically focus on a wearable audio headset that allows natural acoustic sound to reach the listener while providing additional temporal and spatial audio cues germane to the listener's current task. Numerous such tasks and augmentation systems have been proposed, and many have shown significant improvements using AR audio [12]. Previous AR systems have begun with a 'pseudoacoustic' environment in laboratory tests, which is meant to create a realistic real-world soundscape at the listener's ears onto which virtual cues may be superimposed [12]. In a similar way this project can be seen as pseudoacoustic environment well-suited to testing AR audio techniques and sonification approaches. Specific techniques that are found to work well within the laboratory environment may then be ported to an AR audio headset for use alone in real-world environments.

While AR audio systems are traditionally thought of in terms of virtual signals that augment the experience, the converse can be true as well - a purely synthetic sonification system can also be 'augmented' by the addition of real audio. This project is likewise difficult to designate as only AR or VR, as it involves a virtual laboratory listening environment divorced from live acoustic sources, yet it seeks to present an immersive audio rendering of a real auditory environment augmented with additional data in the form of spatialized sonifications of non-acoustic data. By augmenting an existing 3D audio portal of distinctive New York City soundscapes, this project might as accurately be termed "augmented virtual reality." The various data feeds being accessed and sonified are conversely augmented through the ability to listen to the real soundscape of a given location. The real audio helps ground each listening location by giving a sense of place to this data and these events.

## 3. METHOD

### 3.1. Theory

The I Hear NY3D project focused on the capture of acoustic events that were situated at specific points in time and space. But by opening up the project to other types of data, new spatial and temporal paradigms can be explored. We have organized these data by a 2x2 matrix representing spatial character (global vs. local) and temporal character (static vs. dynamic), as shown in Table 1.

Table 1: Classification of data by spatial and temporal character

| | **Local** | **Global** |
|---|---|---|
| **Dynamic** | Acoustic Recordings, Tweets | Weather Data |
| **Static** | Yelp Data | Not Considered |

Along the dimension of spatiality, data can be either local (i.e. present in a single location) or global (present at many or all locations represented). Since this project's scope is limited to the borough of Manhattan, features need not be literally global as long as they affect many different locations at once. It will be noted that a global characteristic undermines the ability to make instantaneous comparisons between geographic locations.

In the time domain, most data feeds are dynamic, meaning that they change over time. Since everything varies given enough time, it might be argued that all such features are thus dynamic. But if we constrain ourselves to a time window relative to a human lifespan, we can consider some data to be static in time and unchanging over the course of a single listening period or between multiple listenings. Thus temporally it may also be useful to think of these data feeds on a continuous scale based on how often they update themselves. Perfectly static data will have an update rate of 0 Hz, while slowly dynamic data will have rates far below 1 Hz, indicating that they will usually be static over a single listening period. Faster dynamic data, like sound or social media feeds, will have update rates much higher than 1 Hz.

#### *3.1.1. Local/Dynamic Data*

By limiting ourselves to two sets of binary classifiers along the twin dimensions of space and time, we have four collectively exhaustive categories of data types. As mentioned, most acoustic data recorded is in the Local/Dynamic category - changing over time and space. It is possible for some very loud acoustic events (such as thunderclaps) to be considered global over a set of locations close together, but at any rate the current Ambisonic recordings were not made simultaneously, and thus all acoustic recorded data is effectively local to a single position.

In addition to these acoustic events, there are many publicly available data streams that are both real-time and geo-tagged, meaning that they can be placed precisely in time and space. Much of this data is provided by users uploading information via smartphones, which can be seen as distributed sensors of various types of information. For the purposes of this prototype, we have focused on Twitter data in a similar fashion to [10] because of the large quantity and variety of Twitter data available. Other similar Local/Dynamic data feeds could include Instagram, FourSquare, and New York City Metropolitan Transit Authority (MTA) GPS data for taxi, bus, and train tracking.

The prototype Twitter sonification Layer uses live Twitter data taken from each location in NYC to control tuned bell sounds with various levels of granular synthesis whose parameters are based on distance from the tweet's geolocation. Covering three octaves of pentatonic tones, lower, more reverberant tones are associated with greater distance while higher pitched, less reverberant tones are associated with closer distance.

#### *3.1.2. Global/Dynamic Data*

Though it is rare for acoustic events to be global over any relatively large area, access to internet data feeds provides other features that change over time but are spatially uniform across the positions in our project. These Global/Dynamic features still change over large enough time scales, but do not directly differentiate specific locations from one another. These may give additional depth to repeated listenings, as a specific site may change over the course of a day or a period of many days. For our prototyping purposes, the chief Global/Dynamic feature considered was weather data, which has a large perceptual effect on our experience of a space but is often not directly audible.

#### *3.1.3. Local/Static Data*

Conversely, other available data is tied to a very specific location but does not change over time. This Local/Static data can help to further differentiate a location that has many buildings nearby that provide no auditory cues as to their function within their environment. This is similar to many of the mapped entities that were sonified by [9]. While certain static entities have a constant acoustic output (e.g. fountains), others are acoustically silent while still contributing to the perceptual experience of a given space. For this current version, Local/Static data is represented by data from Yelp!, which provides an API with information on restaurants, businesses, parks, and other static entities in New York neighborhoods.

#### *3.1.4. Global/Static Data*

Finally, the last corner of the diagram represents Global/Static data - features which are uniform over both space and time. While these types of data do exist at least within our universe (e.g. the immutable laws of physics), it is outside the scope of this project to represent them. For this reason, this category of data was ignored.

### 3.2. Layers

#### *3.2.1. Soundscape Layer*

As previously mentioned this sonification system builds on previous work by this group in the capture and reproduction of immersive urban soundscapes. Various locations around the borough of Manhattan were collected using Ambisonic techniques [2], which allow for the flexible immersive reproduction of these spaces on any periphonic or pantophonic speaker array. In a previous project, I Hear NY3D, these Ambisonic recordings were reproduced in a way that allowed participants to directly experience the soundscapes of Manhattan [3]. This was accomplished by using SPL data, collected at the same time as the Ambisonic captures, which was then used to ensure that the reproductions were recreated at the same level by measuring the output SPL with the same equipment. Experiencing the soundscape in the safety of a studio, absent of the fear of being hit by a taxicab or run into by a hurried New Yorker allows a participant to fully focus on the soundscape. During the past installations of this system, participants' primary comments were about "how intense the sound levels really are", which they were able to notice by comparison to other quiet locations. Another unique experience, that was mentioned to the installation developers, was the ability to track the position and movement of sound objects around the entire sound sphere. The final comment received from participants towards the installation of this system was about the experience of being able to quickly compare soundscapes, or to directly compare soundscapes through Ambisonic mixing and positioning techniques.

The current project builds on these benefits by retaining the Ambisonic reproduction of Manhattan soundscapes. These soundscapes serve as a canvas or rather point of reference about the quality of sound and sound data that exists around Manhattan. This allows participants in the current system the ability to create relationships between these qualities of Manhattan and the sonified data streams.

As with I Hear NY3D, this system will maintain the ability to reproduce these captured soundscapes through any periphonic or pantophonic speaker system because they have been maintained in the Ambisonic B format. Additionally, this system retains the capability to dynamically switch between soundscape locations based upon human interaction and location data pulled from the virtual layers. Finally, this system maintains the calibration necessary to allow for the reproduction of the soundscapes at their original sound pressure levels. By doing this, the system has the capability of demonstrating the immense amount of sound data that is thrust onto New Yorkers. These levels will only be approached during moments of intense activity within the virtual layers. Such a situation may occur during a particularly bad rush hour or events in which a lot of people are congregating.

#### *3.2.2. Virtual Layers*

The I Hear NY3D system gave participants an opportunity to hear and create relationships between the various soundscapes, through means of quick A/B comparisons, and direct spatial mixes of two or more soundscapes. This project furthers the participants' ability to observe new relationships by creating an infrastructure that allows for the quick comparison of and direct comparison between many data sources. This is accomplished by creating flexible layers, which influence each other, are dynamic in their reactions to user input and external data information, and can quickly be pulled up or muted. These layers need to be capable of becoming active or inactive quickly so that they can respond to user interaction or react to external data. This allows participants to drive the experience and create sonifications of the data streams they are most interested in.

For example, a user may be interested in hearing a representation of the rate of Tweets coming from or mentioning Central Park (dynamic/local), with the weather layer adjusting a high frequency shelf dampening filter (global/dynamic), against the actual soundscape of Central Park. Chordal drones that raise in pitch will represent the frequency of Tweets matching this criteria, and the position of the tones contributing to this drone could be place around the compose rose, thereby taking advantage of the size of Central Park. As Tweets count up around the busier south end and west ends of the park, the tones in those locations of the immersive environment would raise in pitch, and as the weather temperature drops over the course of an afternoon the brightness of these tones, and the soundscape, would be brought down with the previously mentioned filter. Additionally, Tweets that mention Central Park with a hashtag (#centralpark) or explicitly use its known name

[13], could be individually sonified and their geolocation would influence their placement around the soundfield of the sonification. By augmenting this with the actual immersive soundscape capture of Central Park, sense of place is offered to the participant and encourages their contemplation about this site. Whereas, a similar scenario, centered around Times Square would portray a more intense experience. The reproduced soundscape immediately heightens the aggressive energy of the experience through the multitude of horns, and ambient conversation occurring. Additionally, with the large numbers of people who regularly pack into the tiny area between 42nd and 49th streets, the Tweets have a much smaller geolocation area, therefore their position in the soundfield is more centralized. An additional layer that might add in would represent the poor Yelp reviews which tend to collect for establishments around Times Square (local/static). This could be used to push sounds into the soundfield, essentially shrinking the size of the sonic space in order to represent the negative reviews associated with Times Square.

## 4. IMPLEMENTATION

### 4.1. Multichannel

The Ambisonic recordings that comprised the I Hear NY3D project were processed in dual formats allowing multichannel loudspeaker playback or 2-channel binaural playback. Within our laboratory setting, the multichannel format was produced using the 16-channel loudspeaker system in NYU's Immersive Audio Research Laboratory (fig. 1).

Figure 1: NYU's Immersive Audio Laboratory

On top of the existing Ambisonic matrixing setup, any of the various AR layers may be added into the multichannel mix. First, the position data of each virtual layer is calculated in reference to the current listening location chosen by the user. Only data within a given cutoff radius is included, which is different for the various geolocations and AR layers based on the density of these respective characteristics. If a datum is close enough to be included, its relative position based on the virtual orientation of the listener is calculated. The source is then mixed into the Ambisonic output via vector based amplitude panning (VBAP) [14] and given atmospheric lowpass filtering and distance-based attenuation according to [15].

### 4.2. Binaural

Since the end goal for most AR systems is a portable 2-channel rendering, it is also necessary to create a binaural version that allows sonifications to be spatialized using only earphones. The existing Ambisonic mixes have also been converted to binaural versions using a KEMAR HRTF measurement from the U.C. Davis CIPIC dataset [16]. Similarly to the multichannel mixing, relative positions are calculated, faraway sources are excluded, and the source is positioned, this time using distance-based filtering and HRTF convolution instead of VBAP. This system can then double as a ready-for-deployment AR audio system by removing the Ambisonic audio data and ensuring that the earphones or headphones used are sufficiently open to allow good localization of external acoustic sources.

### 4.3. Participation

As stated, one of the goals for this type of system, is to afford the participant the opportunity to dynamically alter the state of the system at their whim, in order to maximize engagement. This also allows them the opportunity to experience connections between the various layers, and unique geolocations. These ideas maintain the interaction goals established by this team for the I Hear NY3D system. In order to facilitate this participants can quickly and smoothly alter their reference geolocation through a simple map type interface. When switching positions, data layers that represent global types of data remain constant. Likewise, layers that are local to the geolocation seamlessly cross-fade, taking on the characteristics of the new data feed. This allows a participant to experience the same type of data, but quickly compare various perspectives on this data, according to location.

The participant will also have the ability to mix in what layers are effecting the sonification occurring in the system. Almost like an audio mixer, the participant has the ability to bring in various data layers (such as Twitter, weather, public transportation, or soundscapes) and the amount that they contribute. This level of control is important for the participant, because they are the one wanting to find connections, and this type of mixing allows them to subtly adjust the scene to the elements that intrigue them.

Although participation is anticipated and encouraged, this system retains the ability to exist autonomously. After a timeout is reached, the system takes over the control of geolocation and layer mixing. This offers the development team the opportunity to present specific combinations or sequences that have been found to be particularly compelling during the creation process. This mode allows the participant the chance to experience combinations and relationships that the content curators have worked with extensively.

## 5. DISCUSSION

### 5.1. Listening Modes

I Hear NY4D maximizes the listener's physical and psychological experience of augmented virtual reality through multi-data layering. On one hand it displays a variety of existing locations throughout the city in a single isolated space, each location calibrated to match the sound environment of the listening room. The installation also integrates interpreted sonic layers that result from the real-time data mining of these environments. This allows for an open-ended experience of a very precisely crafted sonic image. Soundscapes are overlayed with more mutable, interpretive layers, perhaps, more accurately representing how we actually encounter spaces in today's smartphone 'enhanced' experiences of reality. Simultaneously, we can have similar objective observations or even physical effects, and internally subjective associations with the space that are then morphed by real time changes.

Certain psychoacoustic effects initiated by this sort of display enhance the notion of 'augmented virtual reality'. For example, delocalization, or the awareness of displacement in where a sound is coming from or desynchronization, a temporal decontextualization effect breaks the perception of an established sound structure [5]. Amnesis, the notion that sound triggers memory, might help the listener paint a certain mental picture of the environment, which will then be distorted by the sound layers triggered by the data. So, though the installation is portraying a specific 'past' sonic environment, the listener is able to experience a variety of augmented realities of that space in real time.

The timing, depending on the data stream will be variable. For example, Twitter data will likely trigger sounds more often and sporadically than weather data, which will be gradual and slow moving. We will use the characteristics of the actual data to dictate whether it will be triggering sounds or just sonic effect. Potential effects include shifts in LFO, EQ, delay, volume, compression etc. Sounds will vary from being literally reflective of the original environment or data stream, i.e. sounds of silverware for restaurant/bar location or Yelp data. Changes in traffic data could trigger increased compression or literal car sounds. In this case, the specific sound element is open-ended and will be determined as the installation progresses, but the platform could be applied to more than just sound.

The multi-dimensional, multi-purpose, and interactive nature of the sound environment allows for the options of true immersion and navigational agency within the soundscape. The listener can choose to be active or inactive and the resulting experiences will be different. The notion of situated listening or audio-positioning asserts that the level of interactivity and user mobility will change the meaning and material of the soundscape. By increasing the physical, experiential dimension of the soundscape, embodied listening is enhanced and active perception is encouraged.

### 5.2. Future Work

Additional sonification of dynamic layers we intend to execute in the final I Hear NY4D installation will include simple sine tone oscillators, various metallic, tuned bell sounds, and percussive cymbal strikes with reverberant trails to reflect the above mentioned data streams. Since at the local level sounds will be more greatly affected by geolocation, distance will play a greater role in the depiction of the sound through grain size and rate of granular synthesis. With less dynamic layers, spatial distinction of the sounds will be less prominent. For slowly updating (nearly-static) global data, we will adjust LFO filtering over the whole installation as to reflect the gradual changes. Chosen sounds will also represent the extent to which the data stream is reflecting a 'man-made' or 'natural' phenomenon, for example: sine tones will be used for Twitter data vs. cymbal strikes for Foursquare.

### 5.3. Conclusions

To summarize, this paper describes a modular system and theory for incorporating immersive virtual audio with spatial sonification data. The modules and examples given here are archetypal examples for each of the categories outlined. We are continuing to research existing real-time data feeds that may be useful as AVR audio layers. In addition, certain extant data feeds (such as GPS data for New York City taxis) only exist in non-realtime at the moment, but may incorporate realtime tracking in the future. While many scholars wish to make a strong demarcation between sonification and computer music [Hermann2008], sonified data may have aesthetic or compositional uses apart from conveying information. This project's union of immersive recorded soundscapes and spatial synthetic data may also be used as the basis for compositional exploration outside the principal informational goals stated. Indeed, it is hoped that as more layers are added the final output will have sufficient spatial and spectral density to be enjoyed aesthetically with or without the Ambisonic soundscape recordings. The ability to retain virtual spatial data while dropping in or out of the real acoustic space offers many exciting compositional possibilities. Likewise, the emerging field of Sound Art could most definitely utilize this platform for expanding creative options. The science and psychology communities may be able to conduct more accurate studies by replicating the bodily experience of music and sound through an immersive, interactive audio platform. Finally, as has been mentioned before, the current stage in this project is not seen as an end point but merely the beginning of a It also increases possibilities for AVR audio and its interactions with the acoustic soundscape of a real environment. It is hoped that the technologies and theories explored here will be useful for mobile deployment in real acoustic environments. While there is good evidence that substantial 'bandwidth' exists that is not being used by the auditory system for most acoustic environments, more work needs to be done to establish the point of sensory overload. The day may even come when users wear earphones that partially block outside sound to allow the auditory system to process more sonified data. This project, by allowing a strict control over the amount of acoustic and virtual cues, is the ideal incubator for creating the AVR systems of the future, representing the complex connections and relationships between all of the publicly available, geo-tagged data that this society lives within.

## 6. ACKNOWLEDGMENTS

Thanks to Areti Andreopoulou, Hariharan Mohanraj, Samuel Nacach, Michael Tierney, and Charles Deluga for their help in the development of this system.

## 7. REFERENCES

[1] M. R. Schroeder, D. Gottlob, and K. F. Siebrasse, "Comparative study of European concert halls: correlation of subjective preference with geometric and acoustic parameters," *Journal of the Acoustical Society of America*, vol. 56, no. 4, pp. 1195–1201, 1974.

[2] Boren, B., A. Andreopoulou, M. Musick, H. Mohanraj, and A. Roginska, "I Hear NY3D: Ambisonic Capture and Reproduction of an Urban Sound Environment," in *Proceedings of the 135th Audio Engineering Society Convention*, New York, NY, 2013.

[3] M. Musick, A. Andreopoulou, B. Boren, H. Mohanraj, and A. Roginska, "I Hear NY3D: an ambisonic installation reproducing NYC soundscapes," in *Proceedings of the 135th Audio Engineering Society Convention*, New York, NY, 2013.

[4] R. Schafer, *The Soundscape: Our Sonic Environment and the Tuning of the World*. Rochester, VT: Destiny Books, 1993.

[5] J. Augoyard and H. Torgue, *Sonic Experience: a Guide to Everyday Sounds*. Montreal and Kingston: McGill-Queen's University Press, 2005.

[6] T. Hermann, "TAXONOMY AND DEFINITIONS FOR SONIFICATION AND AUDITORY DISPLAY," in *Proceedings of the 14th International Conference on Auditory Display*, Paris, 2008, pp. 1–8.

[7] G. Kramer, B. Walker, T. Bonebright, P. Cook, J. Flowers, N. Miner, and J. Neuhoff, "Sonification report: Status of the field and research agenda," International Community for Auditory Display, Tech. Rep., 1999.

[8] E. Loeliger and T. Stockman, "Wayfinding without Visual Cues: Evaluation of an Interactive Audio Map System," *Interacting with Computers*, 2013.

[9] W. Heuten, D. Wichmann, and S. Boll, "Interactive 3D Sonification for the Exploration of City Maps," in *Proceedings of the 4th Nordic conference on Human computer interaction*, no. October, 2006, pp. 155–164.

[10] T. Hermann, A. V. Nehls, F. Eitel, T. Barri, and M. Gammel, "TWEETSCAPES REAL-TIME SONIFICATION OF TWITTER DATA STREAMS FOR RADIO BROADCASTING," in *Proceedings of the 18th International Conference on Auditory Display*, Atlanta, GA, 2012, pp. 113–120.

[11] A. Harma, J. Jakka, M. Tikander, M. Karjalainen, T. Lokki, J. Hiipakka, and G. Lorho, "Augmented Reality Audio for Mobile and Wearable Appliances," *Journal of the Audio Engineering Society*, vol. 52, no. 6, pp. 618–639, 2004.

[12] A. Harma, J. Jakka, M. Tikander, M. Karjalainen, T. Lokki, H. Nironen, and S. Vesa, "Techniques and applications of wearable augmented reality audio," in *Proceedings of the 114th Audio Engineering Society Convention*, Amsterdam, 2003.

[13] C. Li, J. Weng, Q. He, Y. Yao, A. Datta, A. Sun, and B.-s. Lee, "TwiNER : Named Entity Recognition in Targeted Twitter Stream," in *SIGIR '12*, Portland, OR, 2012.

[14] V. Pulkki, "Virtual Sound Source Positioning Using Vector Base Amplitude Panning," *Journal of the Audio Engineering Society*, vol. 45, no. 6, pp. 456–466, 1997.

[15] ANSI, "Method for Calculation of the Absorption of Sound by the Atmosphere," American National Standards Institute, Tech. Rep., 2009.

[16] V. R. Algazi, R. Duda, D. M. Thompson, U. C. Davis, and C. Avendano, "The CIPIC HRTF database," in *2001 IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*, no. October, New Paltz, NY, 2001, pp. 99–102.