I Hear NY3D: an ambisonic installation reproducing NYC soundscapes

Michael Musick,1 Areti Andreopoulou,1 Braxton Boren,1 Hariharan Mohanraj,1 and Agnieszka Roginska1

1 New York University, New York, NY 10012 USA
musick@nyu.edu, aa1510@nyu.edu, bbboren@gmail.com, hari.m@nyu.edu, roginska@nyu.edu

ABSTRACT
This paper describes the development of a reproduction installation for the "I Hear NY3D" project. This project’s aim is the capture and reproduction of immersive soundfields around Manhattan. A means of creating an engaging reproduction of these soundfields through the medium of an installation will also be discussed. The goal for this installation is an engaging, immersive experience that allows participants to create connections to the soundscapes and observe relationships between the soundscapes. This required the consideration of how to best capture and reproduce these recordings, the presentation of simultaneous multiple soundscapes, and a means of interaction with the material.

1. INTRODUCTION
Since recording technology became portable there has been interest among artists, naturalists, and scientists to capture, study, and reproduce soundscapes. The Immersive Audio Research Group at NYU’s Music and Audio Research Laboratory (MARL) has been participating in this practice over the past year through the “I Hear NY3D” project. The project has aimed to build on the history of soundscape practice by taking advantage of recent equipment and software that makes it easier to capture and reproduce high quality immersive recordings. These recordings have been used to create an installation piece that allows participants to experience, learn about, and directly compare the vibrancy of New York City’s varied soundscapes in a way not possible through typical transportation or documentation. This project captures the details of unique locations around the city and recreates the immersive experience in the comfort of an installation through the use of ambisonic techniques. This method also allows for the preservation of locations at this moment in history, with high quality audio that will be capable of reproduction on a variety of speaker arrays for many years. The methods used for this project also make it possible for the installation to be presented at other facilities, which will allow people who do not have the means or capabilities to travel an opportunity to experience New York City’s unique soundscapes.

This paper will focus on the installation portion of the “I Hear NY3D” project, how this builds on the history of
soundscape art, what ways ambisonic techniques were used in the capture and reproduction as specific to the installation, the composition process, and the intended participant experience. More information on technical details of the capture process and other methods for distribution, including an information website which will host binaural renderings of the soundscape recordings can be found in a companion paper for this project. [1]

2. BACKGROUND

2.1. Preserving and Capturing Soundscapes

The practice of capturing soundscapes came about shortly after recording technology became portable. Reasons for recording locations have varied. However, Pierre Schafer, Barry Truax and the World Soundscape Projects initial goals can serve as a reference point for a large portion of work that has and is happening. [2] [3] This group got started to (a) bring awareness to the issues of noise pollution; (b) highlight the rapid change of urban [Vancouver’s] soundscapes; (c) preserve, for historical purposes, the soundscapes as they were at that moment; (d) create an opportunity to experience unique locations for those without the means to visit them, and; (e) gain data and insight to make policy suggestions. The “Acoustic Ecology Institute” has a page on their website with links to many soundscape groups and related websites. [4] From this website alone one can navigate to many projects for capturing and presenting soundscapes that are inspired by these ideas.

Although much of the practice around capturing soundscapes is about preserving the temporal sonic characteristics of a particularly interesting location, there are many who use soundscapes as a medium for artistic expression. Artists such as Toshiya Tsunoda and Francisco Lopez strive to capture locations in ways that go beyond the typical of the soundscape tradition. [5] Mr. Tsunoda, for example may capture the soundscape as heard from a microphone inside a hollow tree, or the minuscule textures created as sand is blown across small pebbles on a beach. [6] There are also artists who use soundscapes as material for compositions, mixing multiple scenes together in a way that brings higher meaning to the whole. [7]

2.2. Data Capture and Measurement of Urban Environments

Capturing the total information of soundscapes goes beyond the simple field recording of a particular location. Studies are now being conducted that work to measure and quantify features of soundscapes. The most obvious feature collected is that of overall loudness, measured both as instantaneous and rms levels. This is used in studies to try and measure the “noisiness” of urban areas. [8] Other features such as spectral centroid, or acoustic event detection are also considered in recent studies. [9] This ambient noise data may be used to look for links with population trends, potential risks to residents’ health, [10] or as factors in determining real-estate values.

3. AMBISONIC AUDIO

Ambisonic audio techniques allow for the capture and reproduction of immersive audio. First order ambisonics allow for the localization of sources with a level of resolution equal to that of first order microphones, such as cardioid or figure-eight patterns. [11] Since first order ambisonic audio is kept in a four-channel format, portability and scalability are much greater than other immersive playback systems, such as wavefield synthesis or holophony. At the time of reproduction these four channels can be decoded to the appropriate number and location of each loudspeaker. This principle means that users of ambisonic techniques can perform work in any studio capable of periphonic ambisonic playback and know that this work will scale accordingly to any other location. [12]

Ambisonic audio is also much easier to capture than comparable immersive technologies. Whereas wavefield synthesis requires an equal number of microphones as the desired number of reproduction speakers, spaced over a large area in order to sample many locations of a wavefront, ambisonic capture techniques require a single point-source microphone, where the capsules are as close as possible together. [13] Since the array of microphone capsules need to be so close to one another, it is most practical to create a single unit housing all of the capsules. Even though such microphones have existed for a number of years, ambisonic techniques have not been widely utilized, partially because the prices for such microphones have been relatively high compared to other field recording solutions. Recently, companies have started producing quality ambisonic microphones that are more price
conscious, such as Core Sounds TetraMic soundfield microphone, which was released in 2007. [14] This is making ambisonic capture techniques more accessible to individuals and institutions for use in field recording situations.

3.1. The Choice to Use Ambisonic Techniques

Although Michael Gerzon at Oxford’s Mathematical Institute first described ambisonic ideas and techniques in the 1970’s, they have not been widely utilized for soundscape recording. Instead, soundscape recording has primarily been executed using a variety of stereo capture techniques. Cost effective ambisonic microphones, and high quality multi-channel field recorders have made it possible to use this technology out in the field. In order to capture soundscapes that will be useful into the future, and to allow listeners the opportunity to experience a realistic reproduction of the original space it is important to capture the immersive characteristics of locations. In addition ambisonic techniques are preferable to methods such as wavefield synthesis because, as previously mentioned, it only requires a single unit housing all of the necessary capsules.

The scalability of ambisonic audio in reproduction also makes it a practical method for capturing soundscape locations. This means that anyone with a system capable of decoding the four-channel ambisonic file into their pantophonic or periphonic speaker array can fully experience these locations. Ambisonic files can also be decoded for binaural or traditional stereo reproduction, allowing individuals to experience these locations over headphones or stereo speaker arrays anywhere they please. This projects use of binaural renderings is further discussed in the companion paper.

4. CAPTURE PROCESS FOR “I HEAR NY3D”

Soundscapes for this project were selected based on geographic location, sonic content, and popularity. At the time of this writing six locations have been successfully captured, processed and fully incorporated into the project. The project team is currently planning the next round of recordings to add to the database. These initial locations are a diverse representation of Manhattan. Each was recorded at a time best representing the locations’ unique qualities, which creates an engaging sonic palate from which to experience a slice of Manhattan’s soundscapes.

In addition to collecting the audio soundscape of each location, the team also captured noise levels using a sound level meter device. At a minimum this allows each location’s loudness to be accurately balanced relative to the other locations. For this project this data has also been used in analysis of the types and differences of sounds present at each location (see companion paper). [1] This data is also used to reproduce the exact sound pressure level of the locations for the installation reproduction.

Every effort was made to place the microphones in a location that would best capture source localization, and the unique character of the location. For example, the Greenwich Village capture took place on the iconic corner of MacDougal Street and Bleecker Street, with the equipment set-up as close to the actual street as was safely possible. The recording took place on a Friday evening, which usually sees the young NYU crowd mix with neighborhood residents, creating a wide variety of sources, types of people, and activities to be captured.

Figure 1 Greenwich Village Capture Set-Up – Left; Noise Level Microphone. Right; TetraMic

5. REPRODUCTION

This project is interested in capturing, cataloguing, and analyzing the soundscapes of New York City. An installation system was envisioned as a way of presenting this material for others to fully experience,
compare, and explore the soundscapes. The ability to experience soundscapes from the comfort of a studio or gallery allows participants the opportunity to engage more deeply with the material than is possible in the actual physical locations. This is also true when experiencing the soundscapes in their rendered binaural format over headphones, however, physically being in the studio allows participants an opportunity to engage more deeply with the material. Good results can be achieved when rendering the ambisonic files to the binaural format through the use of Head Related Impulse Responses (HRIR), however participants will be able to localize individual sources better in an ambisonic speaker reproduction through their own head movement, which helps illuminate front/back and up/down confusions. An installation also allows a participant to hear the exact sound pressure levels of the soundscapes due to the use of the noise measurement data that was collected at the same time as the original capture.

5.1. Goals for an Installation Reproduction

During initial playback tests the team was immediately engaged with the level of realism heard over the periphonic speaker array in NYU’s Spatial Audio Research Laboratory. It was possible to visualize the soundcape being reproduced from the material being heard. This auditory realism is the most important part of the installation and the primary quality that is always emphasized during the development of this system.

In addition to recreating these sonically realistic soundscapes two major goals were established for the installation. First, the system should create a safe space from which to experience and engage with the material. As mentioned previously, this goal is partially met by simply moving these soundscapes from the wild of the actual urban locations to a studio or gallery. This allows the participant to be free to focus on the sonic qualities, locations, and interactions occurring in the captured soundscape. This goal also required devising a method of user interaction that allowed the participant to remain in the sweet spot of the space, while providing the necessary means to select locations and receive any other non-sonic information that was deemed as useful.

The other goal for this space was to create a system that allowed participants to experience or study the relationships between soundscapes. This meant creating a system that could quickly and seamlessly switch to various soundscapes. This also meant creating a system that could simultaneously present multiple soundscapes.

5.2. Design Process

5.2.1. Playback

Originally the audio for this project was processed and matrixed from the B-format ambisonic files to the 16-channel speaker array of the Spatial Research Lab in MATLAB. These 16-channel files were then played using Max patches. This allowed for the quick development of a graphical user interface, and through Max’s standard set of objects it was easy to facilitate quick transitions between soundscapes. However, there were however major drawbacks to this workflow; chief among them was that it hindered the potential portability of the system to different spaces. This workflow also made it difficult to develop the system without being in the lab. Finally, because all processing was performed prior to reproduction there did not appear to be an elegant solution for the presentation of simultaneous soundscapes.

Through the process of looking for other solutions the “Ambisonic Toolkit” for SuperCollider was found and chosen as an ideal option for this project. Originally developed by Joseph Anderson, this toolkit is now maintained and developed by a community under the terms of the GNU General Public License. This toolkit allows for the real-time encoding and decoding of ambisonic audio. Since this toolkit is also now maintained as a public project, the code is available for further development and study. This allowed the team to verify the procedures being employed during the decoding process.

This toolkit easily facilitates the real-time decoding of ambisonic B-format files as long as the positions (ie. the azimuth and elevation angles) are known for each speaker in the array. This allows the project to be easily presented in a variety of spaces anywhere in the world that have a system capable of ambisonic reproduction. This toolkit also provides for real-time rendering of these same files to a number of additional formats, including binaural, which made development of the system possible away from the research lab.
5.2.2. Simultaneous Reproduction

As mentioned previously, it was decided that in order to create enhanced opportunities for participants to create connections to the soundscape material, the system needed to be capable of simultaneous playback. A number of solutions were auditioned to achieve this. The first and most obvious solution was to simply overlay soundscapes on top of each other. However, this does not allow a user to distinguish and compare the unique qualities of each soundscape. Instead, the two individual soundscapes blend into one indistinguishable space without creating the opportunity for critical comparison that is desired.

The next solution was inspired from the sonic cartography themes that this project explores. [1] The idea was to use single channel representations of the soundscapes that would be combined and place them in the speaker array according to their geographical spatial relationships. In order to accomplish this the omnidirectional channel was isolated from the ambisonic files and placed in the speaker array using vector-based amplitude panning (VBAP) techniques. Since the soundscapes were spatially separated it was far easier to recognize relationships in the content, texture and timbre of the soundscapes. It was also useful to aurally experience the geolocation relationships between soundscapes through this panning technique. However, the loss of spatialization data due to mono playback within the soundscapes themselves was not ideal. The consequence of this was a far less engaging experience.

The solution to facilitating simultaneous playback came in using the Ambisonic Toolkits “press” transformation. The object is described as a dominance related transform that distorts a soundfield towards a specified angular location. Aurally this moves a soundscape towards the specified location of the total soundfield, while maintaining elements of the original spatial characteristics. When multiple soundscapes are positioned and pressed around the field according to their relative geolocation it becomes possible for the participant to observe the same relationships that they could through the VBAP technique. The user is also offered an opportunity to observe relationships between the individual soundscapes spatial characteristics and the spatial movement of sources within each soundscape.

5.2.3. Interaction Considerations

For this installation the goal was to create a space and systems that encouraged focus on the soundscapes themselves, not on the technology, the interaction, or extraneous information about the project. The original iteration designed in Max allowed users to chose locations on a computer display with a traditional mouse. This unfortunately encouraged participants to stare at the screen, as they are accustomed to in their daily lives and keep their hand on the mouse while auditioning locations.

In order to break the reliance most users feel to stare at a screen for information, a physical controller was built. The controller is a large map of Manhattan with an LED light and push button installed at each captured location. Additional buttons were installed between locations in order to facilitate the choosing of pre-defined simultaneous reproductions. This map and interaction paradigm were inspired from similar devices that seem to exist in most United States National Park Service museums across the country.

![Figure 2 Physical Map Controller](image)

It is an interaction that people are familiar and comfortable with, and it can offer enough useful visual feedback to a participant without encouraging the learned behaviors of computer screens. The map’s components are connected to an Arduino board, which enables the state of each sensor button to be sent to SuperCollider and simultaneously allows the software to control which LEDs are illuminated. These illuminated...
LEDs allow participants to know which location/s are currently active.

Using this controller allows the computer display to be hidden from the view of participants, and creates a more aesthetically pleasing space by allowing for a dimly lit room without the glow of a computer screen. More importantly, this controller can be placed at a slight angle above a horizontal position, and placed directly in front of the system’s listening sweet spot. This placement of the controller encourages participants to stand in the ideal spot to experience the full immersive potential of these recordings.

Although the intention is for users to select the reproduction location from the controller, in the case that no input is received the software randomly selects locations. The active location/s are still shown through the LEDs.

6. DISCUSSION

Soundscape practices preserve unique locations at moments in time and afford those incapable of visiting them an opportunity to experience the sonic qualities. Equally important, soundscape practice allows listeners the ability to observe, study, and create relationships to specific locations and between locations that, in many instances, would be impossible in the actual physical location. Although one can sit calmly in the middle of Times Square with their eyes closed and focus on the interaction of the sounds and sources, this is not advised as a safe practice. Nor is it a reality for a listener to experience the same time period of a location multiple times to study every aspect of it. Also, although it is possible for an individual to travel around a city to compare soundscapes of multiple locations, the transit time, and unrelated noise between experiences may alter the ability to truly compare the qualities of multiple locations.

Soundscape practice makes all of these situations possible. The ability to safely observe a soundscape, listen to it multiple times, or compare it immediately to another soundscape allows listeners to create relationships, focus on the nuances, and engage more deeply with locations. These are strengths of soundscape practice. The “I Hear NY3D” project successfully incorporates these strengths and builds on them with the use of encoded spatialization data and noise measurements, not captured by traditional soundscape practice. The immersive presentation of these locations set this project apart and offers more material for participants to study. The participant hears the movement of objects, can locate background noises, or experience the feel of a loud truck passing. Additionally, the reproduction of these soundscapes at accurate levels allows participants an opportunity to understand how loud the actual locations are. Finally, this project offers a unique opportunity to compare locations by reproducing multiple soundscapes simultaneously, while maintaining qualities of their spatialization, with accurate geolocation positioning around the soundfield sphere.

This installation system allows both New Yorkers and those who have never visited the city an opportunity to experience these soundscapes in an engaging and informative manner. Since this system easily scales to any space with a speaker array capable of ambisonic reproduction, it will be possible for individuals to experience the sonic character of this city, study connections between locations, and develop relationships to these soundscapes without ever having to step foot in Manhattan.

6.1. Future Work

As this project grows it will become important to create a more robust interface that allows for the easy inclusion of new soundscape captures. There is also work occurring to create a new interface model, which will facilitate two goals. First, to take advantage of the real-time spatial mixing of the press object and allow for a user to hear what a mix of soundscapes from any location might be. Second, is to create an interface that will make it easier to distribute this project. Such an interface might take the form of a tablet application to control SuperCollider, that could be bundled with the full installation and allow users to pick a listening position from anywhere on Manhattan.

7. REFERENCES


