LOCATIVE SONIFICATION: PLAYING THE WORLD THROUGH CITYGRAM

Tae Hong Park, Johnathan Turner, Christopher Jacoby, Alex Marse, Michael Musick
Music and Audio Research Laboratory (MARL)
New York University, New York, USA
{tae.hong.park, cbj238, jmt508, aem585}@nyu.edu

Ajay Kapur, Jingyin He
California Institute of the Arts
Valencia, California, USA
{akapur, jingyinhe}@calarts.edu

ABSTRACT

This paper presents an exploration platform for locative sonification based on audio feature vectors extracted from urban spaces. Our locative sonification research is part of a larger project called Citygram[17]. Citygram focuses on geospatial research that is concerned with automatically collecting, visualizing, analyzing, and mapping non-ocular energies from urban environments. The acoustic data is captured via off-the-shelf poly-sensory Android-based remote sensing devices (RSD). Audio feature vectors are streamed to and stored in the Citygram database which can then be used for sonification and visualization. The first iteration, Citygram One, concentrates on urban acoustic energies, rendering spatio-acoustic feature vectors with the aim of better understanding our environment, large cities in particular. This paper focuses on using the Citygram framework for creative practice via locative sonification.

1. INTRODUCTION

The advent of GPS and digital mapping systems have revolutionised how users navigate and find their way to a destination, whether in real-time or on the computer at home. Today, smart-phones with embedded GPS are ubiquitous. Mobile devices come installed with free apps like Google Maps that can inform the user about traffic patterns and help direct them home on the safest and fastest route. What is next in mapping technology? Imagine if we had similar data about crowds of people in a city. Imagine if we could decode affective attributes of spaces, flux in humidity patterns, and link cyber-physical spatial data through dynamic, multi-layered and multimodal mapping systems. These are some of the features that we seek to realise as part of the Citygram project by setting up a network of sensors to gather additional meaningful and dynamic information about our urban environments. This paper looks at creative application possibilities embracing big data science while focusing on a method we call locative sonification.

Sonification is an art form and type of auditory display technique used to convey information or perceptualize data through audio signals. The idea of locative sonification through Citygram was motivated by our interest in: (1) exploring musical and creative applications of Citygram and its collection and streaming mechanism of non-ocular spatial features and (2) providing a locative sonification system for the computer music community and the general public. Our sonification system provides a hub for exploring and engaging in musical projects driven by non-ocular spatial data accessible through our open databases. We envision that our sonification system will facilitate engagement in location-specific sound projects in a number of ways. For example, it can potentially be used in the context of dynamic sound and art installations that reflect changes of environments. Citygram could also be useful in cyber-physical based performance situations including simultaneous and multi-location telematic “flash-mobs.” Other examples are concert performances using tempo-spatial variables. Data could also be mapped onto musical parameters and structures as typically used in traditional sonification practices.

1.1. Related Work

There are a number of cartographic sound projects including BBC’s Save Our Sounds[7], NoiseTube, WideNoise, and Locus Stream[6]. Save Our Sounds is perhaps the simplest one as it gathers crowd-sourced audio snapshots, archives them, and makes “endangered sounds” soundscapes that will go extinct if not saved available on the Internet. NoiseTube is a project that began in 2008 at the Sony Computer Science Lab in Paris, France where personal cell-phones are used to crowd-source and measure geolocalized dB(A) measurements. WideNoise is similar to NoiseTube. However, WideNoise appears to have a larger user-community which utilizes social media features and similarly measures dB(A) levels to assess noise levels. The Locus-stream SoundMap project is based on a concept called networked “open mic” audio streaming. In this scenario, site-specific, unmodified audio is broadcast on the Internet through a mapping interface by “streamers” – persons who install the Locustream boxes in their apartments and share “non-spectacular or non-event based quality of the streams.”
1.2. Paper Overview

This paper describes a framework for creating locative sonification. Section 2 begins by discussing the Citygram project. Subsection 2.1 discusses Citygram’s System Overview including sensor network and feature extraction. Section 3 introduces locative sonification using the Citygram framework. The final section is a summary of accomplishments and outlines future work.

2. THE CITYGRAM PROJECT

Current topological mapping paradigms are typically static and focus on addressing layouts based on fixed objects such as buildings, streets, and bridges. This should not be surprising as visible objects that define spaces (eg. rivers and roads), and consequently the maps that represent them, generally remain unchanged for substantial periods of time. Imagery in Google Earth, for example, is updated every one to three years[1]. These traditional topographic maps are suitable in coding complex 3D spaces into 2D/quasi-3D visualization formats that best address the static nature of visual objects. However, spaces are not defined only by static objects, nor are all defining characteristics visible. Advances in affordable and powerful computing devices, electronic sensors, and wireless technologies are now facilitating the process of capturing dynamic and invisible dimensions of urban spaces.

Citygram is a large-scale, iterative, and interdisciplinary project that focuses on interactive mapping research based on big-data paradigms. The project is divided into a number of functional modules that renders a real-time visualization/mapping system with a focus on non-ocular energies through scale-accurate, non-intrusive, and data-driven strategies. Two fundamental shortcomings in modern digital mapping practice motivated us to engage in this research: (1) the virtual absence of real-time technologies, and (2) a neglect of addressing non-ocular energies.

2.1. System Overview

This section describes the system overview of Citygram. As shown in Figure 2, audio features are computed on RSDs and streamed via WiFi to the Citygram server and stored in an SQL database. Both visualization and sonification algorithms are created by streaming data from the database to create new music, sound, and visualizations.

2.2. Fixed and Mobile RSDs in Citygram

We are using two types of RSDs: fixed and mobile. The fixed RSDs are permanently installed and continually stream feature vectors to the Citygram server. The mobile RSDs (eg. smartphones) allow for additional, temporary spatial readings that can be especially useful in capturing meaningful events that are outside existing RSD networks. We tested many different scenarios before settling on the current RSD system. We researched a number of possible RSD candidates, including Alix system.
boards\textsuperscript{2}, Raspberry Pi\textsuperscript{3}, Arduino\textsuperscript{4}/Basic Stamp\textsuperscript{5}, smartphones, and others. At first, alix3d boards seemed to be suitable for our needs as it ran a scaled-down Linux OS, had multiple I/O capabilities, and came with optional rugged, weather resistant metal cases. Alix also included on-board audio I/O, compact flash sockets for adding WiFi cards (along with its built-in Ethernet ports), and USB. The problem with Alix was the size and need for WiFi dongles which in turn affected cost. Raspberry Pi was also a potential candidate as it was much smaller and cheaper than the Alix board. However, like the Alix board, the Raspberry Pi also required optional external hardware for audio I/O and WiFi.

Figure 3. RSD: The UHost Android mini PC

The Arduino board, like the Parallax Basic Stamp solution, was also considered, as it allowed for add-on WiFi features through a drop-on plug-and-play “WiFi Shield” option. However, it also required additional circuitry and setup for audio input as well as for any additional sensor types. The ultimate solution was found in ubiquitous mobile computing devices: Android-based hardware. Android hardware offered the best solution as it provided: (1) on-board WiFi, (2) on-board poly-sensory capabilities, including audio input, (3) rugged casing, (4) efficient power management, and (5) flexibility related OS. The drawback to using a regular touchscreen-based device was cost. This led us to eventually choose Android-based Mini PCs. The Mini Android PC UHost1, for example, features a built-in microphone, USB ports, HDMI for visual interfacing, and WiFi (Figure 3). It comes in a durable case and costs approximately 30 USD. Furthermore, Android OS is a very good fit for our purposes as it allows auto-updating for software applications, addressing the complex issue of RSD maintenance and manual software updates.

2.3. Feature Vectors, Database, and Visualizations

All feature extraction tasks are executed locally on each RSD. For example, when computing the spatial acoustic energy levels, each RSD computes its own \( dB_{RMS} \) values and streams the feature vector to the Citygram server where it is stored in a mySQL database. The aim is to extract and include a large number of feature sets for two main purposes: (1) to provide features for sonification exploration, (2) to investigate feature vector saliency in delivering effective features to our machine learning modules. The feature set is continually growing and currently includes RMS energy, spectral centroid, spectral flux, novelty/event detection, and noisiness level. Figure 5 shows the feature extraction process on the RSDs.

Figure 4. Indoor visualizations on CalArts campus

Figure 5. Feature Extraction Process on the RSD

3. LOCATIVE SONIFICATION

3.1. Locative Media

Locative media, coined by Karlis Kalnins\textsuperscript{12}, and locative arts\textsuperscript{13} focus on cyber-physical spaces for situated and spatially associated media production, utilization, and experiences. Essential to locative media practices are location-specific information such as GPS data. Examples of locative media include map-based browsing in-
murmur provides auditory display of location-based oral history recordings to render a reflection of urban spaces, where the recordings themselves are provided by participants. Another example is Sonic City. In Sonic City[8], participants put on wearable computers that receive data from sensors to monitor environmental conditions including light and temperature, as well as physical properties of participants, such as heart rate and directionality. The measured data is used to produce modulated “environmental sounds,” creating a continuously evolving soundscape that participants listen to over headphones as they move through the city. Other examples include the GPS Beatmap project where an automobile coupled with GPS and a Max/MSP patch provide a way to mix audio files and embraces the “Planet as Control Surface” idea[10]. A final example is Jason Freeman’s UrbanRemix[9], which has a key interest in exploring “soundscape composition” paradigms to reveal the inner sonic life of a particular place and frame of time. UrbanRemix utilizes mobile-device applications and web-based tools to facilitate collaborative field recording, sound exploration, and creation of soundscapes. The project allows mobile phone users to “record and share geo-tagged sounds and images captured from their environment” and provides web-based tools to “enable anyone to browse, remix and share the sounds through an intuitive map-based interface.”

3.2. Locative Sonification and Citygram

Citygram provides a system for locative sonification through simple data polling from popular computer music software like Max/MSP, Pd, ChucK, and Processing. For example, for Max/MSP we provide custom patches for easy access to spatial feature vectors stored in the Citygram database. The Max patches, which will be available at citygram.calarts.edu, allows users to either receive (1) a constant stream of selected feature vectors, or (2) receive feature vectors from a user-selectable window of time. The former enables the utilization of feature vector streams that characterize specific locations in real-time. These feature vectors can then be used to drive custom sonification designs implemented in Max/MSP or Pd.

3.3. RSD Deployment and Streaming Strategies

For Citygram’s locative sonification module to be effective, we need a large number of RSDs. “Massive” deploy-
ment of RSDs is without a doubt a concern. We have, however, devised a number of strategies to address this issue as follows: (1) provide Max/MSP and Pd patches to allow anyone with personal computers, a microphone, and network connection to serve as a “streamer,” (2) create free Android apps so that users with smartphones can also contribute, (3) continue with efforts in working with cities to deploy RSDs through existing infrastructures such as payphones, which will become, or already are, functionally irrelevant. New York City, for example, is planning to repurpose the public payphone system. Of the currently active 11,412 public payphones in the city (Figure 9 shows some the locations in the city), 49 have recently been converted to include free WiFi. This infrastructure is ideal for Citygram RSDs as it would quickly lead to availability of 11,412 nodes throughout the city.

![Figure 9. Current payphone locations in Manhattan](image)

Contributing to Citygram as a “streamer” is also straightforward. We again provide Max/MSP and Pd patches to simply push data onto the Citygram database. Once added to the database, a dynamically changing node will appear in the Citygram web interface according to the GPS coordinates.

3.4. Sonification Possibilities via Citygram

There are many locative sonification possibilities in Citygram without the need for complicated hardware setup. The simplest setup would be to use a single feature vector (eg. $dB_{RMS}$) to drive a sound synthesizer or some kind of generative music system in real-time. This is akin to Charles Dodge’s Earth Magnetic Field where annual magnetic field measurements are used for temporal, pitch, harmonic, and timbral structures. In Dodge’s work, the sonification rendered a fixed composition rather than a continually changing one as is facilitated by Citygram feature streams. Other possibilities include spatial telematic music. One scenario could entail groups of performers contributing to the soundscapes in multiple locations which will in turn affect their respective RSD measurements. The resulting output stream would immediately be available on the Citygram server, which could be used compositionally at a concert hall, living room, or gallery setting. Furthermore, a composer could distribute custom software to users around the world to produce sonic art on a global scale – a type of global performance.

4. SUMMARY AND DISCUSSION

4.1. Future Work

Citygram is still at the beginning stages of development when considered within the larger context of its potential to better understand our environment. We will look at connections between Citygram data and available crime stats, Twitter feeds, municipal/census data, and weather patterns. For sonification purposes, however, we plan to provide improved interfaces and additional data streams from machine learning modules that will enable monitoring and visualization of traffic patterns, noise pollution, and spatial mood/emotion. Other plans include enabling users to interact with Citygram maps to hear the “texture” of spaces. We are testing a modified granular synthesis blurring technique to allow the monitoring of spatial sound objects including speech without compromising what is being said. For example, this would enable users to recognize a dog barking or a car passing by without compromising the contents of conversations in public spaces.

The massive deployment of RSDs remains the most challenging problem and the success of Citygram will be dependent on the widespread deployment of such devices. Our crowd-sourcing model is already functional and in place, as anyone with a basic computer, microphone, and network connection can become a streamer. There remains, however, much work to be done in creating a framework where we have reliable and consistent data streams to our server. For example, we are investigating methods of auto-calibration using impulse response, filtering, and deconvolution. For Android apps, we are planning to include software auto-update features to streamline deployment of fixed RSDs.

We are also researching mesh networking technologies where RSDs simultaneously become receivers, transmitters, and relayers of data stream, expanding the RSD network WiFi boundaries. We will also freely distribute mobile apps for crowd-sourcing additional measurements as part of our community outreach efforts to inspire citizen/science-collaboration, support information accessibility/openness, and create a hyper-dimensional city exploration hub.
Another type of information that can potentially be inferred from sound is environmental emotion/mood. Although this interdisciplinary research is still in its nascent stages, automatic mood detection has found increasing interest in the field of music [15], speech analysis [18], face-recognition [16], and natural language processing [19]. Sentiment analysis is concerned with inferring a subject’s emotional state and is increasingly applied in political and marketing contexts to assess population level affinities towards candidates and brands. In lieu of research objectives, practical application examples include finding the closest “quiet” park ... now, restaurants in the “moodiest” areas, and visualizing noise pollution via measuring traffic noise, construction activities, and high-level acoustic events. Future iterations will include exploring other energies that can be captured by our RSDs (humidity, magnetic field, brightness, color, etc.).

Ultimately, we hope Citygram will contribute to future music-making possibilities for the computer music community.

5. ACKNOWLEDGEMENTS

We would like to thank Google for supporting Citygram research.

6. REFERENCES

[1] [Online]. Available: https://sites.google.com/site/earthhowdoi/Home/ageandclarityofimagery