New Scenic Subjects: Explorations of a System of Autonomous On-Stage Observers

Ludvig Elblaus
KTH Royal Institute of Technology
10044 Stockholm, Sweden
elblaus@kth.se

Åsa Unander-Scharin
Luleå University of Technology
94128 Piteå, Sweden
asa.scharin@ltu.se

Carl Unander-Scharin
Karlstad University
65188 Karlstad, Sweden
carl.unander-scharin@mhi.kau.se

CHI'16 Extended Abstracts, May 7–12, 2016, San Jose, CA, USA.
ACM 978-1-4503-4082-3/16/05.
http://dx.doi.org/10.1145/2851581.2889470

Abstract
This paper describes a full body interactive performance system where the users interact with a set of non-human observers. An observer is a sculptural scenic presence, vaguely anthropomorphic, equipped with a camera, that analyses the movements of the users, and creates a soundscape. The observers are movable but not mobile, meaning that their placement on stage defines a set of overlapping layers of activation for the users to interact with.

This work aims to answer questions on what scenic subjects emerge from interacting with the observers, what new identities and relationships are established, and how we can use emotional and scenic components of these relationships to find new expressive qualities in stage productions.

Author Keywords
Sound in Interaction; New Musical Instruments; Opera.

ACM Classification Keywords
H.5.5 [Sound and Music Computing]; H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellanious

Introduction
This paper describes a full body interactive performance system where the users' movement qualities are translated

Figure 1: Maria Callas in the role of Medea in Cherubini's opera Medea.

Figure 2: Early experimental setup using laptops on speaker stands.

into sound in real-time. This is achieved by positioning *observers* in the performance space that view and analyze the movements of the users. An observer is a sculptural presence, movable but not mobile, vaguely anthropomorphic, that is equipped with a digital video camera. The observers analyze what they see, and communicate their results to a central hub, wirelessly, so that a soundscape can be produced. The system can accommodate any number of observers and users. While the current iteration uses three observers, future plans include increasing this number substantially.

Background

This development process is a part of a project called Re:Callas. In Re:Callas, the extreme vocal expression of the iconic opera singer Maria Callas (1923-1977) in general, and her performances as Medea in particular (Figure 1), are used as an inspiration in exploring new ways of approaching vocal performance. The Re:Callas project is still ongoing at the time of this writing and will culminate in a full scale operatic performance in the Croatian National Theatre Ivan Zajc in Rijeka, Croatia, in 2017.

The development process started by defining an imagined ideal end goal: the notion of full body interaction as control over a particular voice. Or, put in another way, the ability to dance the voice of Maria Callas. With this in mind, a wide range of techniques were explored to try to connect with Callas' recorded vocal material. Many different synthesis techniques were explored, e.g. granular synthesis, concatenative synthesis, convolution techniques, but the experiments tended to either be able to offer interesting auditive results but be very hard to control, or respond well to interaction but somehow lose the qualities of Callas' voice that were the focus of the exploration to begin with. In a similar fashion, the control method has been arrived at through the exploration of many different versions, most of them revolving around assigning different positions in space to different positions in time in the source material.

In Bogue's book *Deleuze on Cinema*, he writes on notation that "we tend to spatialize, and hence distort melody" [1], and indeed, also in many of the authors' previous projects, space had been used as a primary metaphor for time [4], and this was also the case in the early experiments focused on trying to revive and control the voice of Callas. In response to this, another strategy was devised, leaving behind the axiomatic ideas of interactive control where the same controlling gesture would always produce the same response, the same output. Instead of laying a recorded or composed section of time out in the spatial domain to be played by moving in or manipulating that space, it was decided that the primary form of interaction would be to *push time forward* with full-body activity. The amount of activity would correspond to the speed of the passage of time, meaning that the same gesture would create a similar change each time it was performed, but the resulting auditive output might be very different.

Another important conceptual shift was to move from a fixed setup to a modular autonomous system. Instead of fixing sensors to a body like in *The Throat* [3], or fixing strings to a stage like *The Vocal Chorder* [5], or defining a fixed area on stage as interactive like in *The Charged Room* [4], the new system introduced movable sensing observers. Furthermore, as the metaphor for their sensing was architectural and scenic in itself, it didn't need to be staged. Big movements create big reactions, small movements create small reactions, and most importantly, a big movement looks small if performed from a large distance. This opened up a new conceptual world of *hiding*, of being *seen* or *detected*, a complex spatial and dynamic activation

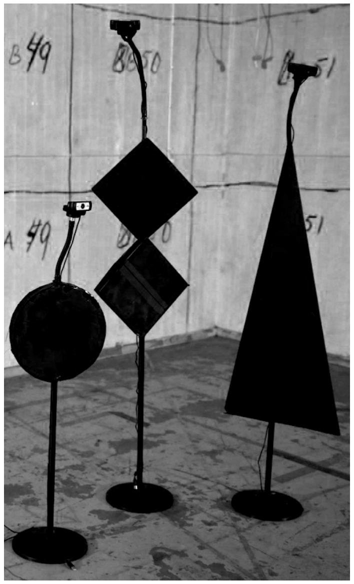

Figure 3: The observers. An ensemble of vaguely antropomorphic sculptural scenic presences, each equipped with a camera.

of the scenic space that responded intuitively to on-stage factors like scenography and lighting.

System Overview

The system consists of a set of observers, a physical gestalt that houses a computer and a camera, and a central hub computer. The observers and the hub are running custom software and the hardware is all off the shelf. The observers and the hub communicate over WiFi. The hub outputs the result of the interaction as audio through loudspeakers. The system was prototyped using laptop computers on speaker stands (as can be seen in Figure 2), but are now housed in the abstract but distinct sculptures in Figure 3.

The Observer Software

To allow the observers to make sense of what they see, they are running a computer vision system designed to interpret body movement detected by a video camera and to generate control data from those movements.

Much computer vision works by comparing the pixels of an incoming video to a static reference image. A difference between the two pictures indicates that something new is present in the video, and if several different pixels are close to each other, one can start to make assumptions about them being one object. This is often referred to as background subtraction. The observer software also does this, but with an important addition: Each new frame in the incoming video stream is very subtly fused with the reference image, causing it to gradually adapt to changes over time. This means that instead of comparing the video stream to a static reference image, it is compared to a somewhat blurred memory of how things looked a few seconds ago. If the scene that is viewed is still, the reference image will catch up to the present and become very similar to the live video stream, allowing the system to detect an absence of motion, or to be more precise, an amount of motion below the threshold of the system. The details of the video synthesis that blend the old and new images can be controlled to yield many different styles of analysis. This is especially powerful when combined with the different possible parameter setting of the blob-detection. Some settings detect slow movements and others fast ones, some are keen to pick up small objects while others pick up large ones. While the system as such is still in development, the potential for expressive control is immense. From a practical point of view, this adapting analysis engine will mitigate any changes in the environment that would cause a system that used a static frame of reference to need re-calibration. Thus, the observer software allows for changes in scenery, dynamic lighting, and moving the observers during performance.

The program is written in C++ using the openFrameworks framework, utilizing the openCV computer vision library [2]. While much more could be said about the observer software than the scope of this paper permits, its most important feature is the way in which it adapts to its environment over time.

The Hub Software

The central sound producing hub of the system runs code written in the SuperCollider programming language [6]. Each observer has an assigned layer in the audio engine and each layer consists of a data-file and a synthesizer. The data-file contains spectral analysis data that is derived from an audio-file. The synthesizer uses a technique called phase vocoding to produce a re-synthesis of the original audio file, using the data-file, with the added feature that the playback speed is separated from the pitch of the sound. Usually when you play an audio file, the pitch of the result is proportional to the rate of playback, e.g. playing a sound

Figure 4: A montage of frames from the video that accompanies this paper.

faster than its nominal playback speed will cause the resulting sound to be higher in pitch. This, however, does no longer have to be the case using the phase vocoder. Here instead, like conducting an orchestra, going faster or slower doesn't mean going out of tune. Being free to move through the musical material at any pace, each layer is given a separate playback position. That position corresponds to the point in the data file from which the synthesizer should take spectral information to create its sound. When the observers detect movement, the playback position of the layer assigned to the observer is pushed forward proportionally to the amount of movement, while the volume of that layer is turned up for the duration of the movement. Currently, the source material is divided into three layers that are assigned to an observer, as per the description above. One layer contains the voice of Maria Callas and the other two contain orchestral accompaniment of two different styles, one more energetic, fast, and loud, the other softer, more pensive, and slow. Some qualities in the synthesis have been fine tuned in the three layers to accommodate the different source material.

New Scenic Subjects

The main question that this work asks of its audience, as well as of its performers, is what scenic subjects emerge from interacting with the observers. In the upcoming workshops and stage production in the Re:Callas project, it will become clear what new identities form and what new types of interplay they are they engaged in. Our explorations so far (see for instance Figure 4) show a heightened presence and an increased sense of activation of the performance space, compared to other more rigid interactive techniques. Hopefully, the observers will also help explorations of how we can understand the relationship between space and scenography, light and shadow, performer and spectator, human and non-human, and how we can use the emotional and scenic components of these relationships to find new expressive qualities in stage productions. Finally, the ultimate goal is for the haunting spectre of Maria Callas to be summoned, controlled, and re-contextualised in a way that is musically appropriate using the custom digital tools that we develop, among them, the observers. Of course, the observers as such could be used to control an infinite number of processes, and the concept of a multi-camera motion tracking system that estimates magnitude of movement in an ecological way might be useful in many other situations, artistic and otherwise.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 645553.

References

[1] Ronald Bogue. 2013. *Deleuze on cinema*. Routledge.
[2] Gary Bradski. 2000. The opencv library. *Doctor Dobbs Journal* 25, 11 (2000), 120–126.
[3] Ludvig Elblaus, Kjetil Falkenberg Hansen, and Carl Unander-Scharin. 2012. Artistically directed prototyping in development and in practice. *Journal of New Music Research* 41, 4 (2012), 377–387.
[4] Ludvig Elblaus, Carl Unander-Scharin, and Åsa Unander-Scharin. 2014. Singing Interaction: Embodied Instruments for Musical Expression in Opera. *Leonardo music journal* 24 (2014), 7–12.
[5] Carl Unander-Scharin, Åsa Unander-Scharin, Kristina Höök, and Ludvig Elblaus. 2014. Interacting with the vocal chorder: re-empowering the opera diva. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. ACM, 603–606.
[6] Scott Wilson, David Cottle, and Nick Collins. 2011. *The SuperCollider Book*. The MIT Press.