SOUNDCATCHER: EXPLORATIONS IN AUDIO-LOOPING AND TIME-FREEZING USING AN OPEN-AIR GESTURAL CONTROLLER

Gabriel Vigliensoni, Marcelo M. Wanderley

McGill University
Music Technology Area, Montréal, Canada
gabriel@music.mcgill.ca, marcelo.wanderley@mcgill.ca

ABSTRACT

SoundCatcher is an open-air gestural controller designed to control a looper and time-freezing sound patch. It makes use of ultrasonic sensors to measure the distance of the performer's hands to the device located in a microphone stand. Tactile and visual feedback using a pair of vibrating motors and LEDs are provided to inform the performer when she is inside the sensed space. In addition, the rotational speed of the motors is scaled according to each hand distance to the microphone stand to provide tactile cues about hand position.

1. INTRODUCTION

Contemporary vocal performance is almost always associated with the use of microphones and amplification. Furthermore, vocal performances in music concerts and recordings are one of the most processed music signals. However, it is very common that the sound engineer is in charge of the audio processing, not allowing singers to *augment*, process, and control their vocal performance [4]. Hence, SoundCatcher is designed as an open-air gestural controller for singers that allows them to sample their performance, loop and process it in real-time, creating new possibilities for performance and composition in live, rehearsal, and recording contexts.

2. RELATED WORK

Melodic and rhythmic looping is an extended performance and compositional practice based on phrase repetition. Its use began in the sixties with the works of Terry Riley and Steve Reich, among other minimalists composers, who used tape loops in analog recording machines to play repeatedly audio portions with the possibility of recording overdubs and manipulate them. Analog delays appeared in the seventies, giving new creation possibilities to composers such as Brian Eno and Robert Fripp in creating imaginary landscapes for their own pieces or for other musicians such as David Bowie. Digital devices in the eighties and nineties provided musicians with new ways to perform and compose with the loop technique without the constraints of tape machines and analog devices [11].

Different audio software and hardware platforms allow musicians to design and use patches to loop, control, and process their performance in many different ways. While Electro Harmonix, Lexicon, Roland, TC Electronics and Electrix have offered out-of-the-box solutions for musicians that want to use hardware loopers, INA/GRM and more or less open software frameworks such as Reaktor, Pure Data and Max/MSP have loopers and time-freezing patches for software based performers.

On the gestural side, Hewitt and Stevenson [4] developed in 2003 the e-Mic, a gestural controller designed for singers to capture their movements and performance through the microphone and its stand, using the acquired data to derive control signals into a computer software's sound engine, allowing the control and processing of different vocal effects without staying behind a computer screen. Vocal performers gestures were studied in order to find the best places in the microphone stand to sense their movements, different sensing technologies were tested for the gestural acquisition, and the mapping strategies allowed the audience to relate the performer movements with what is been controlled and heard.

Michel Waisvisz invented and mastered The Hands [13], a music performance interface to perform live composing. A set of sensors were used to translate hand, arm, and finger movements into sounds. Through its different development stages since 1985, it has always been considered an expanded range controller and partially immersive, which the performer can escape to make movements without musical consequence [8]. Moreover, the performer can move and walk, making music without being tied to much analog electronica, but cable-wired. The same principles using different hardware technology are used by Alex Nowitz in his works for amplified and live electronics with a gestural controller (although both projects use STEIM's LiSa as the live sampling software)[9]. In addition, some new interesting devices have been developed to control voice and sound through gestures, such as Elena Jessop's *Vocal Augmentation and Manipulation Prosthesis* (VAMP)[6], and Emilie

Simon's *Brissot arm controller*, which allows her to remotely control, modulate and transform her voice.

Using non-contact technologies, open-air controllers give performers the possibility of generating or controlling sound without the necessity of holding, touching, and manipulating an instrument. However, by its nature, one of key feedback channel to sense the response and accuracy of the instrument is lost: the haptic channel. Thus, the performer must rely mainly on aural and visual feedback, as well as proprioception and egolocation [1]. Although there is a number of studies and experiments in restoring haptic/tactile feedback to open-air controllers, such as Rovan and Hayward's *VR/TX Vibrotactile Stimulator Project* [10], no commercial devices incorporate its use in their design.

3. TECHNICAL DETAILS

SoundCatcher is a music performance and compositional controller that allows a musician to sample and process his/her voice or other audio live input in real-time. In its most basic setup, the gestural interface allows the performer to record and overdub a sound engine's buffer at any moment, controlling its playback behavior by changing its loop start and end points related to his/her hand positions to a reference point. In order to provide more than aural feedback, tactile and visual feedback using a pair of vibrating motors and LEDs are provided to inform the performer when she is inside the sensed space. Furthermore, each motor's speed is scaled for each hand, in relation to its distance to the microphone stand, to provide tactile cues about hand position.

Our approach to designing a gestural controller with a sound engine providing tactile feedback was to separate each one of its five sections and investigate them separately, assembling all of them afterwards. Below is a description of each one of the sections.

3.1. Sensors and gesture acquisition

The Parallax's *PING)))* ultrasonic sensors were selected to capture hand's position and movements due to their linear response in the required distance detection range, power requirements, narrow acceptance angle, shape, and low cost. Each of them measure the distance from the microphone stand to the performer's hand by emitting a short 40 kHz burst. This signal travels through the air hitting the performer's hand and then bounces back to the sensor. It provides an output pulse that will terminate when the echo is detected, hence the width of this pulse corresponds to the distance to the target. Although each sensor can measure from two centimeters up to three meters, in our configuration the sensed space for each hand is limited to sixty centimeters, allowing the performer the open both arms to a maximum of 1.2 meters, thus avoiding measuring wall reflections and other possible objects close to the perfomer. In addition, a pedal footswitch was implemented to switch the buffer recording in the software side on or off.

[1] Egolocation refers to the awareness of one's overall position within a defined space, or with respect to objects in that space [10]

The data acquired from the sensors is sent to the open-source Arduino platform and handled by its Wiring based language.

3.2. Vibrotactile and visual feedback

To provide the performer with cues about the sensed space without requiring a computer screen, two actuators are used to provide vibrotactile feedback. In addition, it is expected that this feedback will provide most of the information needed for expert performance [12]. Given the frequency response requirements for the fingers and the system, two motors were selected to transduce the measured distance to vibrating movement [7]. The voltage signal driving the motors is scaled from the hand to the microphone stand distance. The lower the value, the faster the motors vibrate. If the hands go beyond the sensed space, the motors do not vibrate at all. Two digital pins with pulse wave modulation capability in the Arduino board are used to drive the actuators. Furthermore, the motors are attached to the box device through cables, giving the performer the possibility to use or let them loose, depending on his/her needs. With this solution, the performer can obtain haptic feedback without losing their mobility.

For performance contexts, in order to provide the audience and the performer additional visual cues, we installed two large LEDs on each side of the box device. With this little theatrical trick, we are giving extra information about what is being sensed, making it easier to relate the performer's gestures with what is being heard. Thus, the system is helping the performer to communicate with the audience [2]. Figure 1 shows the ultrasonic sensors and actuators setup in the device.

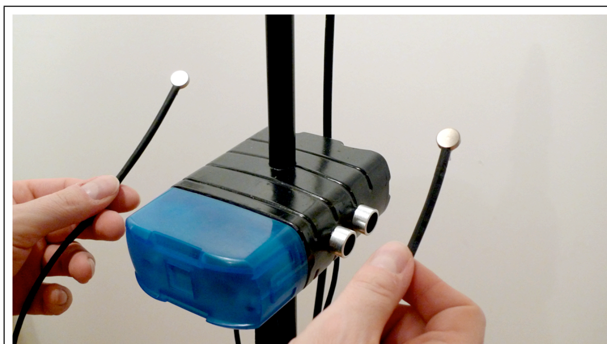

Figure 1. Detail of the actuators providing vibrotactile feedback.

3.3. Sound Processing Engine

The sound processing engine for SoundCatcher is based in the Max/MSP environment using two concurrent approaches to provide a more complex experience for the performer. On one hand, through a time domain approach, a buffer is recorded and its playback start and end points are controlled by the ultrasound sensors. It is always running so if the performer wants to silence it, she must fill it with silence. On the other hand, a frequency domain approach using previous research by J. F. Charles is used to freeze a sound in real-time. His implementation resynthesizes several frames continuosly with a stochastic blurring spectral technique allowing smoother transitions and avoiding frame effect artifacts [3]. The patch developed also allows the performer to crossfade between two different portions of audio in a desired time. Using both techniques at the same time, the musician can compose and perform using a looping technique in a long time scale, and freezing and crossfading audio frames in a short time scale, both in real-time.

To provide an easier experience working with other musicians, sequencers, and digital audio workstations, the buffer loop start and end points can be synchronized to MIDI clocks, thus ensuring a synchronized performance if wanted.

3.4. Mapping

Although SoundCatcher was designed to provide the performer straightforward control of both loop points, thus making its use apparently simple, we decided to map the same variables to more parameters in order to give more expressive possibilities, such as the frame blurring level described above, and signal dry/wet ratio for the time-domain looper section. In addition, the skilled use of the footswitch allows the performer to create complex rhythms and melodic structures. Thus, the performer can use the controller in both analytic and holistic ways, being conscious of the control parameters if needed, but being more explorative and integral way if she wants to try more unexpected sounds. Also, it is important to consider that the sound output depends on the audio signal addressed to the system, so *performance mode*, as defined by Hunt and Kirk [5], assumes that the singer needs to explore his/her vocal expressive possibilities as well as the gestural controller to obtain more artistic results.

3.5. Setup

The device is placed in the microphone stand, allowing height changes depending on the singer's expressive needs and performance convenience. The vibrating motors and footswitch are connected directly to the device, so the entire controller can be considered a one out-of-the-box product. The computer running the software side of SoundCatcher does not need to be on stage. Both LEDs can be adjusted in order to light up the performer's hands or his/her face, achieving different expressive and theatrical effects. Figure 2 shows the clean SoundCatcher setup in the microphone stand.

Figure 2. The SoundCatcher stage setup.

4. MUSICAL APPLICATIONS

With the current parameter mapping, SoundCatcher can be used to augment vocal performances in several different ways, such as: creating new melodic and rhythm structures, repeating desired audio portions with or without synchronization, freezing and crossfading different audio frames, and creating delay lines. It can be used for composing and/or performance. The device visual feedback through very brilliant LED's can be used in live contexts to enhance the audience's understanding of the gestural controller and sound engine behavior. Although the setup was designed for a vocalist, it can be used by other musicians.

A first approach to the use of SoundCatcher in a musical performance context can be seen in `http://vigliensoni.com/blog/soundcatcher`

5. CONCLUSIONS

We have developed a device capable of controlling a looper sound engine through the use of open-air gestures. The mapping strategies selected make SoundCatcher an integral or holistic gestural controller, giving high exploration possibilities to the musician. As an addition to sonic feedback, visual and vibrotactile feedback was provided in order to know when the performer is inside the sensed space and how close or far is of its limit boundaries. The design also allows the performer to escape the interface at any moment because she is not hard-connected to the interface. The sound engine was developed to give the musician two different approaches to the instrument, one through time-domain, *looper*, and other through a time-domain manipulation *freezer*. As a controller designed for modern musical performance, in which it is possible to find other synchronized instruments, MIDI clock synchronization for the looping part is possible.

All in all, a gestural controller was developed that allows high control levels and explorative possibilities for a looper and freezer sound engine.

6. FUTURE WORK

Even with little practice time, we have realized SoundCatcher's high level of artistic possibilities. As usual in the development stage of new musical digital instruments, there are several possible improvements for SoundCatcher. In the gestural controller side, we will investigate on new mappings to allow the control of other or more sound parameters in one-to-many configuration. Furthermore, more sensors in the device box or attached to it can be used to control other parameters (e.g., the space in between the singer and the microphone stand is not presently being sensed, and more footswitches can be used). As a following step, a gesture vocabulary can be developed in order to control the behavior of the sound engine (e.g., interrupting the sensed space very fast several times can be used to activate a new algorithm patch and disable another, or changing the sensed distance very fast can make the patch to go to other mapping strategy). On the sound engine side, once the sound is already in the internal patch buffer, it is possible to think of new audio processing techniques to achieve new expressive possibilities, such as pitch-shifting, time-compress expansion, etc.

Although all the work mentioned above is possible, we think that in order to master a new digital instrument or controller, practice and performance time is needed to play and discover it, so each one of the mapping changes mentioned above should be carefully designed to avoid continuously re-learn what has been learned.

7. REFERENCES

[1] D. Birnbaum, "Musical vibrotactile feedback," Master's thesis, McGill University, Montreal, Canada, October 2007.

[2] B. Bongers, "Physical interfaces in the electronic arts," in *Trends in Gestural Control of Music*, M. Wanderley and M. Battier, Eds. IRCAM - Centre Pompidou, 2000, pp. 41–70.

[3] J. F. Charles, "A tutorial on spectral sound processing using Max/MSP and Jitter," *Computer Music Journal*, vol. 32, no. 3, pp. 87–102, 2008.

[4] D. Hewitt and I. Stevenson, "E-mic: Extended mic-stand interface controller," in *NIME*, 2003, pp. 122–128.

[5] A. Hunt and R. Kirk, "Mapping strategies for musical performance," in *Trends in Gestural Control of Music*, M. Wanderley and M. Battier, Eds. IRCAM - Centre Pompidou, 2000, pp. 231–258.

[6] E. Jessop, "The vocal augmentation and manipulation prothesis (VAMP): A conducting-based gestural controller for vocal performance," in *NIME*, 2009, pp. 256–259.

[7] M. T. Marshall, "Physical interface design for digital musical instruments," Ph.D. dissertation, McGill University, Montreal, Canada, 2009.

[8] A. Mulder, "Towards a choice of gestural constraints for instrumental performers," in *Trends in Gestural Control of Music*, M. Wanderley and M. Battier, Eds. IRCAM - Centre Pompidou, 2000, pp. 315–335.

[9] A. Nowitz. (2009, December) Voice and live-electronics using remotes as gestural controllers. [Online]. Available: http://cec.concordia.ca/econtact/10_4/nowitz_voicelive.html

[10] J. Rovan and V. Hayward, "Typology of tactile sounds and their synthesis in gesture-driven computer music performance," in *Trends in Gestural Control of Music*, M. Wanderley and M. Battier, Eds. IRCAM - Centre Pompidou, 2000, pp. 297–320.

[11] D. Toop, *Ocean of sound : aether talk, ambient sound and imaginary worlds*, 5th ed. Serpent's Tail, 2001.

[12] R. Vertegaal, T. Ungvary, and M. Kieslinger, "Towards a musician's cockpit: Transducers, feedback and musical function," in *Proceedings of the International Computer Music Conference*. International Computer Music Conference Association, 1996, pp. 308–311.

[13] M. Waisvisz, "The hands, a set of remote midi controllers." in *Proceedings of the International Computer Music Conference*. Burnaby, BC: International Computer Music Conference Association, 1985, pp. 86–89.