EGGS in Action

Pietro Polotti
Department of New Musical Technologies and Languages
Conservatory of Music G. Tartini
Via Ghega, 12, Trieste, Italy
Interaction Group
IUAV, University of Venice, Italy
pietro.polotti@conts.it

Maurizio Goina
Department of New Musical Technologies and Languages
Conservatory of Music G. Tartini
Via Ghega, 12, Trieste, Italy
maurizio@goina.it

ABSTRACT
In this paper, we discuss the results obtained by means of the EGGS (Elementary Gestalts for Gesture Sonification) system in terms of artistic realizations. EGGS was introduced in a previous edition of this conference. The works presented include interactive installations in the form of public art and interactive onstage performances. In all of the works, the EGGS principles of simplicity based on the correspondence between elementary sonic and movement units, and of organicity between sound and gesture are applied. Indeed, we study both sound as a means for gesture representation and gesture as embodiment of sound. These principles constitute our guidelines for the investigation of the bidirectional relationship between sound and body expression with various strategies involving both educated and non-educated executors.

Keywords
Gesture sonification, Interactive performance, Public art.

1. INTRODUCTION
In this paper, we report three years of explorative work about gesture sonification on the basis of the EGGS system introduced in NIME08 [5]. We present various realizations both in the form of public installations and interactive performances. The main principle stated in 2008 is respected in every work: elementary sounds are defined and employed for the sonification of a small number of gesture categories. Roughly speaking, these categories can be subdivided into the two main classes of straight and circular movements. In all of the cases, we defined various and alternative instances of elementary gestalts in terms of sonification sound sets, intended as unitary cognitive structures activated by gestures according to the particular goal and context of the artistic realization. We also realized extensions of the principles to the visual domain in a multimodal sense. This is also meant as a possible way for defining effective indirect mapping between visual forms and sounds with gesture as a sort of interpreter.

As already illustrated in [5], we think of gesture as generated by sequences of points forming trajectories in space that can be combined in order to generate complex cognitive objects at different possible levels. Our approach to gesture analysis is, thus, essentially abstract, different from an expressive/emotional feature extraction as, for example, in the work by Camurri and al. [2]. Expressiveness is a consequence of the combination of sonic/cinematic relationship in articulated structures jointly to the effect of further parameters such as sonic/gesture dynamics relationships. In this sense, one of the sources of inspiration of this work are the theories of Paul Klee. In particular, we inherit from Klee the idea that a dot is the ur-element that is the atomic element, whose movement generates lines and planes. This concept is well illustrated in Klee's "Pedagogical Sketchbook" [8], a book intended as the basis for the course in Design Theory at Bauhaus.

The paper structure is the following. The next section provides a brief review of the EGGS principles and system. Section 3 describes the public art realizations. Section 4 illustrates the interactive performance works. In Section 5, we discuss possible future developments and applications, and draw our conclusions.

2. THE EGGS SYSTEM
EGGS provides a basic system for gesture sonification. Arbitrary articulation and combination of elementary mappings can be defined. The system allows applications for performing arts and interactive dance as well as for public interactive installations. Differently from what pursued in other works on gesture-sound mapping [1], [10], we opt for a simple cinematic and dynamic gesture analysis valid for general categories characterized by some abstract and elementary properties. In order to do this, we consider gestures as spatial trajectories of moving points, the point being a hand, an elbow or a knee, and we define segmentation criteria based on simple geometric considerations. In EGGS, visual data of a point-wise source are processed by a trajectory tracking routine that returns different indexes corresponding to five categories: straight movements, circular clock wise and circular counter clock wise movements, direction inversion and stillness (see [5] for more details). A variety of second order parameters are derived from the primary data related

NIME'11, May, 30- June 1, 2011, Oslo, Norway

to the purely cinematic trajectory that is the movement scalar and vector velocities and accelerations. Up to now only a 2D tracking was taken into consideration.

Using the EGGS system, as in any musical practice, involves learnability issues (see, for example, [7] for a discussion about apprenticeship in new musical interfaces). Exercise is important in order to understand the possibilities of the instrument and obtain relevant results. This is the case, when working with professional performers. However, the system is immediate and easily usable by anybody as any simple gesture produces a meaningful sonification. The latter is the approach taken into consideration in case of public installations, where the visitors should be immediately able to get a satisfying, enjoyable and stimulating feedback from the system. Indeed, differently from the perspective introduced in a recent work on interactive dance [9], our work is intended to spur the performer (or user) to control, adapt and explore her/his gesture according to the received continuous sonic feedback in an enactive way.

Given these founding criteria, the fashion we carried on our investigation and development of the system was in spirit of experimentation of multiple alternative realizations of the same principles in different artistic applications. We consider this way of proceeding to be in analogy to a design practice, in which the following realization depends on the previous results in a cyclic way and a comparative evaluation of the different realizations is a matter of enrichment and deepening of knowledge about some subject; in our case the effectiveness of sound as continuous feedback for controlling gesture and the expressivity of the sound-gesture couple in a multimodal sense. The latter point is considered both from the spectator point of view (audio and visual multimodal aspects) and from the actor point of view (audio and proprioceptive multimodal aspects).

An extension of the system to the case of a multimodal feedback was also considered. We investigated the possibility offered by using gesture as a control of both sound and graphic generation, where the correspondence between sounds and images was in general arbitrary. In this sense, we refer to Chion's definition of audiovision [3], according to which any association of images in movement and audio produces a composite object that lives in a third dimension, which is multimodal: a sort of vector product. On the other side, a concurrent generation of sound and graphics by means of the same movement analysis allows to search for novel relations between sound and image through the juxtaposition of abstract and elementary categories "unified" by the same gesture.

3. INTERACTIVE INSTALLATIONS IN A PUBLIC ART FASHION

Visual Sonic Enaction (VSE) is a multimodal and interactive installation that allows to generate an audiovisual representation of one's gestural expressivity. VSE is usually presented in a ludic fashion, by introducing it to the public through the metaphor of graffiti painting: the visitors are encouraged to paint on a large wall by means of an "electric torch/spray can" controlling different graphic and sound processing algorithms. The sound elicits and guides the movements of the user and immerges she/him in a bodily-visual-auditive experience, by producing a multimodal and continuous feedback to the gesture. Indeed, sound plays the role of connective element of the three components of VSE.

In the current version [14], three sets of elementary sounds and three groups of simple graphic signs were defined and employed for the sonification and visualization of two basic main categories of gestures/movements: straight and circular. The visitor could experiment only one sonic set at a time for each visual-sonic self-portrait. On the contrary, in the context of a single portrait, the user could change graphic set in any moment and any number of times, by shaking one of three coloured bottles put aside the interactive area. The bottles were equipped with wireless accelerometer sensors. The three graphic typologies are illustrated schematically in the bottom-right corner of Figure 1. The three kinds of sounds were experimented one after the other in three different visual-sonic portraits. The three sound sets, represented iconically in the top-right corner of Figure 1, included swishing and metallic sounds, low pitch FM synthesis sounds and glass and crystal tinning sounds generated by the physics-based sound synthesis package Sound Design Toolkit (SDT) [4]. The sound typology changed only when the user decided to save the current portrait and to start another one from blank. Different kinds of mappings were implemented, more or less variable in timbre and other characteristics in relations to gesture categorization. Some mappings implemented a discrete separation between straight and circular movements and other were modulated continuously with the curvature of gesture, in order to obtain a certain variety in each one of the nine possible sonic-graphic combinations. An example of a portrait is reproduced in Figure 1.

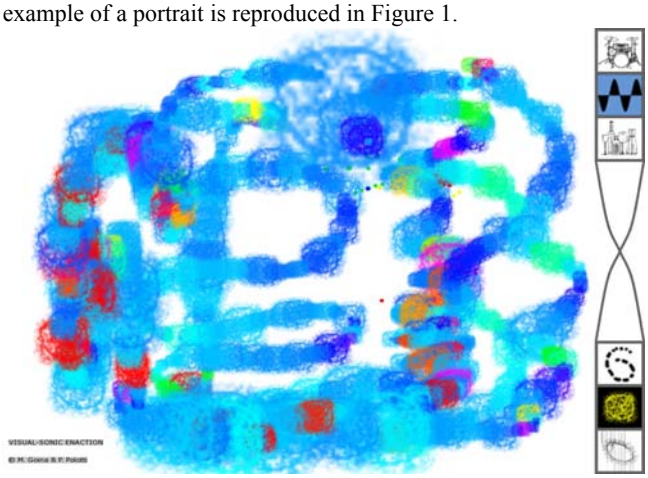

Figure 1. The VSE "canvas" with a visual-sonic self-portrait of one of the participants.

As a final product, the visitors received an audiovisual file as a record of the abstract visual-sonic representation of their gestural expressivity. Besides, anyone can watch and listen to the bodily expressivity of any visitor on the VSE website, where the "visual-sonic self-portraits" are uploaded. Someone among the visitors was able to interpret very quickly the spirit of the installation and adapted her/his gesture to the different sonic/graphic combination in order to reach a coherence of all of the three aspects involved in the installation (see for example Serena's portraits [14]).

In VSE, the EGGS principles are applied to the visual domain as well. The aim is not to paint. Rather, what appears on the wall or on the computer screen is a visualization of the expressivity of gesture. At the same time, in an enactive way, the visual feedback spurs the user to modify and control her/his own gesture also

according to different type of visualized graphic. The use of different graphic types is fundamental in order to uncouple the gesture from the idea of painting as well as from a particular sonic set. In the future, another aspect we want to investigate is if the definition of abstract (gestural) categories and the definition of effective, however independent, mappings for sound and graphics generation could reveal unexpected relations between images and sounds.

In another recent public installation, Sonic Walking (SW), we concentrated on gait expressiveness [13], therefore, shifting the focus from the upper part of the body to the lower part and from a creative to an everyday context, where a visitor has just to walk freely along a straight path in an ordinary indoor space. The gait of the visitors is sonified by means of ecological sounds related to nature and, more specifically, to the four basic elements fire, earth, air and water. In particular, we employed the sound of a big fire in a forest, the sound of a rain stick recalling that of sand or grains and the sound of a strong wind. The water had two versions: quite waves on a beach and a sound giving the impression of being underwater. The visitors experienced the five sounds in a fixed order: water, earth, fire, air, underwater.

Before starting to use the system, we told the visitor that they would listen to her/his footstep and that their footsteps would first dabble, then rustle, then crackle, then blow and, finally, go underwater. The visitors could walk along a path of approximately 8 meters and wore two lights tied to the external side of their knees, so that every light was detected by one of the two cameras located on the two side of the path. Also, the users wore wireless headphones in order to experience a more immersive and internal feeling of her/his body movements given by the continuous sonic feedback. The audio was as well reproduced by four loudspeakers located at the extremities of the path, so that the audience passing by could hear the gait sonifications. This case of EGGS application was the least articulated, since no cinematic analysis was taken into consideration, and only the dynamic aspects of movement drove the sonic feedback. A further development of SW, integrating visual elements and a 3D detection aiming at including circular trajectories, is previewed.

4. INTERACTIVE PERFORMANCES

In case of a stage context, working with a professional performer/dancer, sound is meant as an effect of the choreographic gesture and a representation of her/his gestural expressiveness. EGGS becomes what we could denote as a "choreophone": the performer/dancer does neither follow a musical piece, nor controls the execution of a musical piece, and not even generates any music with her/his movement [5]. Rather, (s)he listens to her/his gesture, enactively, modifying and controlling her/his performative action according to the produced sound. The sounds, thus, is a representation of the movement, a sonic consequence and a continuous feedback, in no way external to the gesture itself. In this fashion, sound is intended as augmenting the proprioception of the performer.

In the context of the latest SMC conference, we presented a performance entitled “Swish ‘n’ Break” [6]. The performance is conceived as a controlled improvisation on a predefined score of sounds and gestures, in the style of the previous performances realized by means of the EGGS system [12]. The sounds used in this performance are all derived from the Freesound project [11] and retrieved by means of a number of keywords defined in advance, in the spirit of a programmatic compositional approach. The keywords are: 1) Swish, 2) Nature (Air – Water – Fire – Earth), 3) Break. The choice of the keywords determines the overall structure of the performance, which is fixed and divided into three sections. The Nature section, the richest in sounds, is conceived as a gradual passage from a natural open-air soundscape to an indoor soundscape. Within each section, the sound-gesture mapping is configured according to the general principles of simplicity underlying the EGGS project, and based on a decomposition of the gesture cinematic into segments belonging to the already mentioned five categories: straight, clockwise, counterclockwise, direction inversion and stillness. A camera detects the bi-dimensional coordinates of two electric bulbs handled by the dance performer (see Figure 2). A correspondence between the dynamic of gesture and the dynamic and other parameters of sounds adds a further expressive layer. Within certain constraints, the live electronics players can change the mapping of the sound as well as the quality of the dynamic response of the system, engaging a dialogue with the performer/dancer. The final result is a performance based on a mostly predetermined sonic-choreographic score defined along a large number of rehearsals, in which the performer experiments how to create and adequate her/his gesture according to the system sonic feedback and, reciprocally, how to condition and control the sonic response by means of her/his gesture. All the choices in terms of evolution and refinement of the performance and of the EGGS system as a whole, are taken during the rehearsals as an agreement among all of the members of our group through discussions, trials, selections and optimizations of the gesture-sound mappings and their combination and concatenation. This corresponds to a creative methodology, where working in group and going through brainstorming and debate phases is a firm point.

In NIME 2011, we are going to present a new performance, entitled “Body Jockey”. The idea is to introduce embodiment in club culture and musical styles. The technical setup is the same as in Swish ‘n’ Break. Part of the sounds employed has been retrieved from the Freesound project, another part has been composed by the authors. The trio of performers acts as if being in a DJ and VJ set. The dancer triggers and modulates sounds by mean of her body, while the laptop performers change sounds and mappings, as well as the quality of the dynamic response of the system. The result is a dialogue of the laptop performers with the dancer, who follow a predetermined score, however leaving space to a controlled improvisation. A graphical representation of sounds and mapping is projected on the screen in order to add a video layer to the multimodal experience of the performers and the audience. Through gesture sonification, music becomes embodied in the dancer herself, and this feeling is transmitted to the audience attending the performance as well – the purpose is to create an enhanced disco-club environment, where body, music and video are jointly engaged in the audience experience. This is also an attempt to provide in the future a version of EGGS not limited to trained dancers but available to everybody.

Figure 2 The performer in action.

The overall structure of the performance is fixed and divided in three sections. Mapping and sounds change in every section. In the first part metronome is fast and fixed. Sounds are percussive and rhythmically constrained. The most used elementary gesture is the trajectory inversion, employed to trigger sounds. The main rhythmic patterns are the usual dance even meters. The sound volume is fixed and dynamic changes are obtained by modulating the density of sound events. The second part is more free-style and based on long sound events that are not rhythmically constrained. Here, the usual types of EGGS elementary gestures that is the straight and circular trajectory elements are important and used to modulate the sound parameters. The third conclusive part is similar to the first one in a disco dance style.

5. CONCLUSIONS

In this paper, we presented a number of artistic implementations of the EGGS system. In all of the works, we applied the EGGS principles consisting in i) treating sound as a representation of gesture and ii) working with elementary cognitive and abstract units in terms of gesture analysis, segmentation and sonification. Such an approach has the advantage of enhancing sound embodiment in interactive installations and performances: the sound-gesture binomial is cognitively fully integrated in a multimodal sense, auditory/proprioceptive from the actor side, and auditory/visual from the spectator side. Also, the basic strategy adopted for the gesture analysis and sonification provides a clarity of interpretation of the system response both in case of a professional performer and of an ordinary user. Reproducibility and learnability, in fact, are part of the main issues EGGS aims at.

These aspects allow to envisage many other potential applications for the EGGS system in other fields where body movements and its control in time are crucial, ranging from sensory-motion recovery to musical pedagogy and others. Concerning the latter aspect, a research project is in progress involving experimentation in primary schools employing a version of the system adapted to pedagogical purposes.

6. ACKNOWLEDGMENTS

This project is funded by the *Servizio università, ricerca e innovazione of the Regione Friuli Venezia Giulia* and co-funded by the *Fondazione Cassa di Risparmio di Trieste*.

7. REFERENCES

[1] Bevilacqua, F., Müller, R. and Schnell, N., MnM: a Max/MSP mapping toolbox, *Proceedings of the 2005 International Conference on New Interfaces for Musical Expression,* (*NIME05*), Vancouver, Canada, 2005.

[2] Camurri, A., Lagerloef, I. and Volpe, G. Recognizing emotion from dance movement: Comparison of spectator recognition and automated techniques. *International Journal of Human-Computer Studies*, 59(1):213–225, July 2003.

[3] Chion, M. *Audio-Vision -- Sound on Screen*, New York, Columbia University Press, 1994 [Original edition: *L'audio-vision. Son et image au cinéma*. Editions Nathan, Paris, 1990].

[4] Delle Monache, S., Polotti, P. and Rocchesso, D. A Toolkit for Explorations in Sonic Interaction Design, *Proceedings of Audiomostly '10*, Pitea, Sweden. Sept., 15 – 17, 2010

[5] Goina, M. and Polotti, P. Elementary Gestalts for Gesture Sonification, *Proceedings of the 2008 International Conference on New Interfaces for Musical Expression* (*NIME–08*). Genova, Italy, pp. 150–153, 2008.

[6] Goina, M., Polotti, P. and Taylor, S. Swish & Break - Geschlagene-Natur, *Concert around Freesound, SMC 2010, 7th Sound and Music Computing Conference*, Universitat Pompeu Fabra, Sala Polivalent, Barcelona, Spain, 22 July 2010.

[7] Jordà S. Digital Instruments and Players: Part I – Efficiency and Apprenticeship, *Proceedings of the 2004 International Conference on New Interfaces for Musical Expression* (*NIME04*), Hamamatsu, Japan, 2004.

[8] Klee, P. *Pedagogical Sketchbook,* trans. Sibyl Moholy-Nagy. Frederick A. Praeger, New York, 1965.

[9] Schacher, J. C. Motion to Gesture to Sound: Mapping for Interactive Dance, *Proceedings of the 2010 International Conference on New Interfaces for Musical Expression* (*NIME 2010*), Sydney, Australia, 2010.

[10] Van Nort, D. and Wanderley, M. The LoM Mapping Toolbox for Max/MSP/Jitter, *Proceedings of the 2006 International Computer Music Conference ICMC 06*, New Orleans, USA, 2006.

[11] www.freesound.org (Mar. 30, 2011).

[12] www.visualsonic.eu/eggs_in_action.html (Mar. 30, 2011).

[13] www.visualsonic.eu/sw.html (Mar. 30, 2011).

[14] www.visualsonic.eu/vse.html (Mar. 30, 2011).