CRAFTING SOUNDS WITH THE POTTER

DIBRIS-University of Genova
Viale Causa 13, 16145 Genova, Italy

Giacomo Lepri
giacomo.lepri@unige.com

Corrado Canepa
corrado@infomus.org

Gualtiero Volpe
gualtiero@infomus.org

Simone Ghisio
ghisio@infomus.org

Antonio Camurri
toni@infomus.org

ABSTRACT

This paper introduces *The Potter* a cross-modal reflexive interactive application for learning-by-exploring sound and its qualities. It is conceived mainly for children. The child is guided to perceive different basic qualities of sound, such as timber, pitch, loudness, and density, i.e., sound morphology. The adopted interaction paradigm is the manipulation of a sound object. Interaction with sound is simultaneous, i.e., sounds are produced as the child moves. The metaphor it builds upon can be summarized as follows: As the potter takes the clay and moulds it yielding a pot, in the same way the sound artisan takes a sound and moulds it yielding a crafted sound.

1. INTRODUCTION

If on the one hand, sound and music are nowadays more and more ubiquitous and accessible even to non-professionals, on the other hand such a trend risks to put music in the background, as something that more or less constantly accompanies human activities, but which is not really the focus of our attention. To give music its cultural and social role back, it is important to start music education since early childhood. The first fundamental step in this process is to stimulate attention to listening to sound and its qualities.

This paper introduces *The Potter*, an application having a twofold objective: (i) stimulating the attention of a child with respect to sound, (ii) learning-by-exploring sound morphology and its major qualities. It is conceived mainly for children 5-9 years old. The child is guided to perceive and scout different basic qualities of sound, such as timber, pitch, loudness, and spectral density, i.e., sound morphology.

The interaction paradigm The Potter adopts requires that the user images to manipulate with the hands a virtual object. Since the pioneering concept of Theremin [5], the design of interfaces for real-time moulding of sound and music by means of gesture, with a particular focus on hand gesture, is a very active research topic in the Sound and Music Computing research community. This includes, for example, design of novel digital musical instruments (see [9] for a review), control of digital audio (e.g., [13]) and gestural control of sound synthesis (see [14] for a review).

Applications exploit either computer vision techniques (e.g., [3] and [4]) or dedicated hardware (e.g., [10] and [7]). However, most of such work is intended to support music players in live performances, whereas considerably little attention is devoted to its possible application to music education. Among the examples in this area, for the science center Città dei Bambini, Camurri and colleagues, developed several interactive hands-on applications to manipulate and experience sound by means of hand and body gesture [2]; Bevilacqua and colleagues developed a complete gestural interface built to support music pedagogy [1]; *SoundBlocks* is a tangible environment where young users connect blocks for digital sound manipulation [6]; Machover carried out *The Toy Symphony* project, providing an integrated series of activities as an alternative entry for children into music. In particular Music Toys is an instruments that do not require any special skill but which do reward curiosity, imagination, and expression [8].

The Potter grounds on the concept of reflexive interaction [11]. This is based on the idea of letting users manipulate virtual copies of themselves. Reflexive interactive systems are designed as intelligent mirrors of the user's behavior. The Potter addresses reflexive interaction since the crafted sound, which is produced and stored, is the direct result of the manipulation of the child, i.e., it is a kind of cross-modal mirror reflecting her movement.

Section 2 discusses the metaphor The Potter is grounded on, Section 3 describes the set-up of the application, Section 4 presents how the application works, i.e., how sounds are crafted, Section 5 provides a technical description.

2. METAPHOR

The metaphor The Potter builds upon can be summarized as follows: As the potter takes the clay and moulds it yielding a pot, in the same way the sound artisan takes a sound and manipulates it yielding a crafted sound object. Indeed, the work in a traditional potter workshop implies an intuitive discipline of gestures and basic conventions.

Moreover, the metaphor creates a context enabling affordances to gestures to manipulate sound: continuous and detailed iterative manipulation of sound is central to the process. By gestural interaction the user has the ability to listen, shape, and “play” a particular sound. Interaction with sound is simultaneous, i.e., sounds are produced as the child moves. Rules apply to movement only, whereas sound manipulation is open to the free and creative interaction of the child. For each session of crafting the focus is on a specific sound object. All the different sounds can be easily related to a well-known physical object/element (e.g., bells, human voice, animals, natural elements, and so on). Each sound is designed to be constantly manipulated through this association. The focal point is the concept of sound object exploration inspired by the theories introduced in Musique Concrète by Pierre Schaeffer (see [12] and [13]). The adopted pedagogical paradigm hence draws on researches on electroacoustic music experiences. The user manipulates the sound which, as a concrete sound object, is characterized by an imaginary physical autonomy. As a piece of clay detains an original shape that can be manipulated, also the sound object has an original morphology, which can be modified through the manual user intervention. It is therefore required from the user a process of objectification that leads to identify the sound in something external and independent. A key issue is also the role of memory: actions and gestures made previously will guide the future experience of the user.

From a more general point of view, the metaphor of the Potter is extremely useful to give children the possibility to experience, through the game, a simple and effective working methodology – such as the artisan’s work discipline. In addition, this working process, involves important preliminary abilities related to the learning of a musical instrument. Amongst the others: (i) rational usage of space (proper use of the working areas), (ii) capability of following methodologies and tasks over time (execution of simple procedural steps aimed to carry out the sound manipulation), (iii) development of cross-modal creativity (association gesture/sound, internalization of the relations that occurs between the two media and experience a free creative control over both).

3. SET-UP

In her experience with The Potter, the child operates in front of a piece of furniture that, according to the adopted metaphor, recalls what happens in a potter’s workshop. It consists of the following areas, corresponding to the major areas a potter uses during her work (see Figure 1).

3.1 The draft area

The draft area is a repository where the sound artisan keeps the sound objects (the left table in Figure 1). Sound objects are contained in sound pots. Sound objects are sounds of limited duration with specific features. The draft area corresponds to the basket holding the clay in the potter workshop.

3.2 The final area

The final area is a repository where the sound artisan puts the crafted sounds (the right table in Figure 1). The final area corresponds to the table or basket that will hold the manipulated crafted pots.

3.3 The working area

The working area, i.e., the area where the sound artisan carries out her work. This area consists of the personal space of the sound artisan herself. Moreover, the sound artisan requires an area where she can listen to a sound object. This is a table (the small table in the middle of Figure 1) in between draft and final areas.

Figure 1. The Potter’s set-up.

4. SOUND CRAFTING

Sound crafting is carried out with the supervision of a teacher. Through a graphical user interface (the teacher interface, see Section 5), the teacher chooses a set of sound objects. Then, she metaphorically fills the sound pots (physical objects) with the selected sound objects. Finally, she chooses the movement mapping, i.e., which morphological qualities of the sound objects are subject to manipulation. This may range from one single quality to a collection of qualities, depending also on the pedagogical objectives of the teacher. For example, the teacher may decide to focus on a single quality at a first stage in order to make the child aware of the variation of such a quality, and move at a second stage to more complex mappings. This allows for different pedagogical paradigms. For instance a linear approach: the different sound qualities are introduced one by one so that the child can train in perceiving them gradually. A reticular approach can also be adopted: all or most of the qualities are introduced at the same time and the child can experience the interaction between them simultaneously. The child takes a sound pot from the draft area. She puts the pot in the listening area, to listen to the sound object the pot contains. The whole sound object is reproduced once. Then, she grasps the pot and takes in her hands the contained sound. The transition between the state in which the user can listen to the original sound and the state in which she manipulates it, is related to the action of holding the sound objects with the hands. This transition is carried out by the user through the symbolic gesture of inserting one of the two hands inside the pot in order to pick the sound. After "wearing" the pot (see Figure 2) she can start manipulating the carried sound object with her movement. The sound object is a virtual object, so manipulation is performed by moving the hands in the air, rather than acting on a physical object (e.g., the pot). The processed sound is reproduced in loop as long as the child keeps moving. Finally, if the child freezes for a few seconds, the resulting sound is stored. The child can then put again the sound in its pot and place it in the final area. This will be the crafted sound. At a first stage, the interaction of the child with The Potter can even be limited to grasping a sound pot, putting it in the listening area, listening to the sound object, and putting again the pot in the draft area. Especially for young children, this is already an important step, since it stimulates the attention to sound and makes the child appreciate the difference between different sounds. Figure 2 shows a child crafting a sound. The child can explore and manipulate the following sound qualities.

4.1 Pitch

Once the child has taken the sound object in her hands, she can move it in whatever direction. The pitch of the sound object changes with the height of the hands with respect to the floor. The child can even throw the object toward the ceiling or the floor: the object bounces and pitch changes accordingly.

4.2 Lateralization

Keeping the sound object in her hand, the child can move it to her left or her right, listening to the sound moving coherently in that direction. She can also throw the sound object from her left and collect it on her right, listening to the sound moving.

4.3 Dynamic (volume accent)

The child can make the sound object bounce in the working area (e.g., on the floor or on the table between the draft area and the final area). This is associated to a percussive attack. The attack is proportional to the impulse given to the sound object, a percussive sound being associated to a sharp gesture and a slow attack being associated to a smooth gesture.

4.4 Density (variation of the spectral energy)

Compression and stretch of the sound object with both hands determines a variation in the density of the sound timbre. For example, the sound of water can change from the sound of a single drop, when hands are very close to each other, to the sound of rain when hands are in an intermediate position, to the sound of a river or of a waterfall when hands are far from each other. In the same way, multiple voices (e.g., from two to a choir) can be added to a sound (e.g., a bell, a human voice).

4.5 Distortion

The strength applied in the above-defined compression/stretch is associated to a temporary distortion of sound. If the gesture is smooth and effortless no distortion is applied. To experiment a strong distortion without varying density the child can press the sound object with both hands on the table in the working area. The amount of distortion changes with the strength applied to the pressing gesture, making the child perceive a dirtier and noisier sound when the sound object is pressed onto the table. Since the child does not wear any sensor on her hands, the distance from the table provides a first approximation of such strength. A quick sequence of compressions and expansions makes the child clearly perceive the difference between original and distorted sound.

Figure 2. A child crafting a sound with The Potter.

5. TECHNICAL DESCRIPTION

The Potter runs on a personal computer equipped with a Microsoft Kinect sensor. Figure 3 shows its architecture. Thick arrows represent data flows and thin arrows represent specific settings the teacher can configure by means of the teacher interface. The kernel of The Potter is the crafting module. It includes a movement detection and feature computation module operating on the video and Motion Capture (MoCap) data Kinect captures, and an audio processing module, operating on the audio stream an audio reader provides. Kinect is placed in front of the piece of furniture described in Section 3. Indeed, besides conceptual and interaction design issues, such a set-up also enables to accommodate technical needs. For example, by putting the Kinect sensor in front of the tables, the child is naturally induced to move in front of the sensor, yielding a more robust and reliable analysis of her movement. Kinect was chosen, instead of common video cameras, in order to simplify the set-up so that The Potter can be prospectively installed at schools or kindergartens. Indeed, Kinect can perform background subtraction and motion tracking in almost non-controlled environments, thus removing constraints such as the need of constant background and lights. Moreover, Kinect does not need any further hardware (e.g., frame grabber cards), it can track multiple users, and it provides 3D MoCap data reliable enough for the aims of the application. The movement detection and feature computation module is responsible for (i) tracking the movement of the sound pots, and (ii) extracting features from the movement of the child's hands. In order to track sound pots, these are endowed with a band of tape reflecting infrared light. Each pot has a band of different height. The infrared light emitted by Kinect is thus reflected and captured by the video camera. The different areas are used to detect which pot is currently moving and to track it. This approach has also the advantage that it does not need any special set-up (e.g., specific lighting for colour tracking), being the Kinect sensor enough to exploit it. Kinect Mo-Cap data are processed in order to obtain (i) 3D position of the right and left hand, used to control pitch and lateralisation, (ii) impulse in the hands movement, used to control dynamic accent, (iii) distance between hands, used to control density, (iv) distance between hands and table, used to control distortion. The audio processing module applies real-time signal processing techniques to the audio stream in order to get a strict connection between the digital sound processes and the child's gestures. In order to modify pitch without changing the time dimension (reading speed of the sample), a pitch-shifting algorithm based on the use of variable audio delays is applied. As for lateralisation, sound is spatialised on the right and left channels using a square root pan algorithm to avoid reduction of volume when the sound is in a central position. Dynamic accent is obtained with an impulsive exponential increment of the volume. Variation of density depends on two different audio processes: (i) when the distance between hands is short - usage of filtering through a chain of band-pass filters, and (ii) when the distance between hands is large - activation of multiple delays with different delay times and/or utilization of harmonization algorithms with different simultaneous tuned voices. Distortion is achieved by applying granulation and wave-shaping to the sound material. Motion features and crafted sounds are stored in an archive containing what a child produced along a session. Through the teacher interface, the teacher can choose the draft sound to be used (selected sound set) and which sound qualities the child can explore (selected mapping). A particular possible setting consists of enabling the child to work with her own voice. In this case, one sound pot remains empty. When it is put in the listening area, the child can fill it with her own voice by means of a microphone. Once the pot is filled, the new sound object, i.e., the voice of the child, can be manipulated as usual.

The teacher interface is displayed in Figure 4. The radio buttons on the left enable the teacher to select the sound qualities to be used in the experience. The teacher can enable just one of them, or a subset, or all of them. The box in the middle allows the selection of the sound-set. Several sound sets are currently available (e.g., bells, animals, voices, and so on). The panel on the right enables settings the parameters for recording a new sound to be included in the sound set. In this way, the teacher can for example allow a child to record her own voice and play with it. Finally, the widgets in the bottom allow controlling the volume of the audio output and let the teacher know when the application is ready for recording the sound the child is actually making. Indeed, both for pedagogical and for technical reasons (time needed for filling audio buffers), the child has to play with a sound object for some time (in the order of 10s) before being enabled to record it and to move it to the final area. Movement detection and feature extraction are developed in EyesWeb XMI[1]. Audio processing is implemented in EyesWeb XMI and in Pure Data[2]. These are connected through Open Sound Control[3] (OSC). The different phases of The Potter and the selection of different sounds and mappings are implemented with the support of the MetaEyesWeb platform. This platform enables developing Finite State Machines (FSMs) and to control EyesWeb XMI depending on the current state of an FSM. Python scripts are used to specify FSMs and to control the MetaEyesWeb. The teacher interface is developed with EyesWeb Mobile.

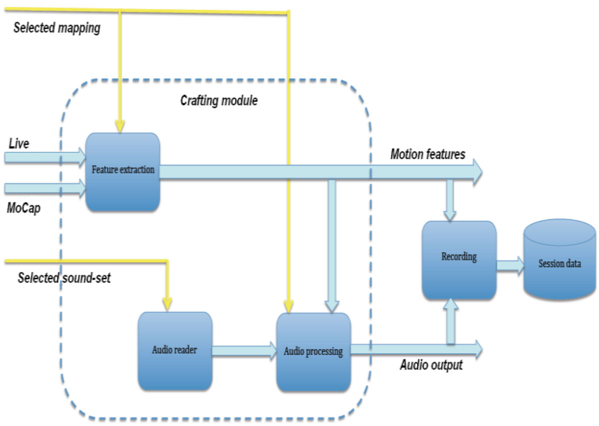

Figure 3. The Potter's architecture. Blue thick arrows represent flows of data and yellow thin arrows represent

[1] http://www.infomus.org/eyesweb_ita.php
[2] http://puredata.info
[3] http://opensoundcontrol.org/introduction-osc

specific settings the teacher can select through the teacher interface.

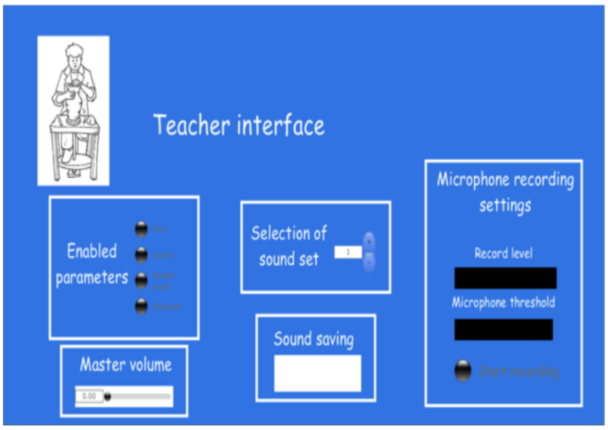

Figure 4. The Potter's teacher interface.

6. CONCLUSION

This paper presented The Potter a cross-modal reflexive interactive application for learning-by-exploring sound and its qualities. Since The Potter is an interactive musical system that requires a simple set-up (a laptop, two speakers and a Microsoft Kinect device) it can be easily installed at schools or kindergartens. The application was tested with children in several occasions. An initial version was presented at Festival della Scienza 2011, a science festival yearly hold in Genova, Italy. Further, a preliminary qualitative evaluation was carried out by psycho-pedagogical partners in specifically organized sessions at University of Gothenburg, Sweden, in April and October 2012. From a musical-pedagogical point of view, the application offers to the users the opportunity to experience an important process of internalization of different sound qualities. Consequently such process allows the acquisition of fundamentals knowledge useful for any kind of musical practice (e.g., playing a musical instrument or listening to and understanding a piece of music). From the interviews with children, it emerged that children show different degrees of awareness of the different sound qualities. The most evident quality was the difference in pitch. Children generally enjoyed interacting with the sound and the technology.

7. ACKNOWLEDGMENTS

This work was partially supported by the EU-ICT Project MIROR (Musical Interaction Relying On Reflexion, www.mirorproject.eu). The authors thank the teams of University of Gothenburg (Åsa Bergman, Bengt Olsson, Pernilla Lagerlöf, Sarah Mercieca) and of University of Bologna (Anna Rita Addessi, Marina Maffioli) for the useful feedback. Special thanks to Giovanna Varni, Paolo Coletta, Michele Pizzi, and Luca Serra for their precious contributions to this work..For a demo of The Potter application visit *https://www.youtube.com/watch?v=_PB29Z_Nq9I&feature=youtube&app=desktop*.

8. REFERENCES

[1] F. Bevilacqua, F. Guedy, N. Schnell, E. Flety, and N. Leroy: "Wireless sensor interface and gesture-follower for music pedagogy", *Proceedings of the 7th international conference on New interfaces for musical expression* (*NIME '07)*, pp. 124-129, 2007.

[2] A. Camurri and P. Ferrentino: "Interactive environments for music and multimedia*". ACM Multimedia Systems*, 7(1) pp. 32-47, 1999.

[3] J. M. Comajuncosas, A. Barrachina, J. O'Connell, and E. Guaus. Nuvolet: "3D Gesture-driven Collaborative Audio Mosaicing", *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 252- 255, 2011.

[4] S. Şentürk, S. W. Lee, A. Sastry, A. Daruwalla, and G. Weinberg: "Crossole: A gestural interface for composition, improvisation and performance using kinect.", *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME '12), pp.* 21-23, 2012.

[5] A. Glinsky: "Theremin: Ether Music and Espionage", *University of Illinois Press*, Illinois, 2000.

[6] J. Harrison: "Soundblocks and soundscratch: Tangible and virtual digital sound programming and manipulation for children", *Massachusetts Institute of Technology, School of Architecture and Planning, Program in Media Arts and Sciences*, Boston 2005.

[7] J. Homan: "Q&A: The sound catcher", *Nature International weekly journal of science*, pp. 486:472, 2012.

[8] T. Machover: "Shaping minds musically" *BT Technology Journal*, No. 22(4) pp. 171-179, 2004.

[9] E. R. Miranda and M. Wanderley: "New Digital Musical Instruments: Control And Interaction Beyond the Keyboard", *Computer Music and Digital Audio Series* A-R Editions, Inc., Madison, WI, USA, 2006.

[10] A. Mulder, S. Fels, and K. Mase: "Mapping virtual object manipulation to sound variation", *IPSJ Sig Notes*, No. 97(122) pp. 63-68, 1997.

[11] F. Pachet: "Enhancing individual creativity with interactive musical reflective systems", *Musical Creativity: Multidisciplinary Research in Theory And Practice*, I. Deliège and G. Wiggins, editors, Psychology Press, 2006.

[12] P. Schaeffer: "À la recherche d'une musique concrète", *édition du Seuil*, Paris. 1952.

[13] P. Schaeffer: "Traité des objets musicaux", *édition du Seuil*, Paris 1966.

[14] V. Verfaille, M. M. Wanderley, and P. Depalle: "Mapping strategies for gestural and adaptive control of digital audio effects" *Journal of New Music Research*, No. 35(1) pp. 71-93, 2006.

[15] M. Wanderley and P. Depalle: "Gestural control of sound synthesis" *Proceedings of the IEEE*, No. 92(4) pp. 632-644, April 2004.