ABSTRACT

This article documents a class that teaches gestural interaction and juxtaposes traditional instrumental skills with digital musical instrument concepts. In order to show the principles and reflections that informed the choices made in developing this syllabus, fundamental elements of an instrument-body relationship and the perceptual import of sensory-motor integration are investigated. The methods used to let participants learn in practical experimental settings are discussed, showing a way to conceptualise and experience the entire workflow from instrumental sound to electronic transformations by blending gestural interaction with digital musical instrument techniques and traditional instrumental playing skills. The technical interfaces and software that were deployed are explained, focusing of the interactive potential offered by each solution. In an attempt to summarise and evaluate the impact of this course, a number of insights relating to this specific pedagogical situation are put forward. Finally, concrete examples of interactive situations that were developed by the participants are shown in order to demonstrate the validity of this approach.

Keywords

gestural interaction, digital musical instruments, pedagogy, mapping, enactive approach

1. INTRODUCTION

Last year, the author had the opportunity to teach an elective class about gestural interaction to masters-level music students, majoring either as instrumentalists or composers. The course was offered twice, first at the prestigious summer-courses in Darmstadt and a second time during the fall term at the Music Department of the Zurich University of the Arts. In both editions of the class a maximum of twelve participants took part in four sessions, either on two consecutive days or on four afternoons spread out over the entire term. The student’s expertise ranged from concert diplomas as instrumentalists on piano, the flute, clarinet, baroque flute and the double bass to composers and in the arts school context media-artists and researchers in media-arts. The concept for a course on gestural interaction for performing musicians arose out of a series of artistic and research projects, which dealt with physical interaction and music, situated within the context of interactive dance and interactive audio-visual installations. The goal was to introduce musicians with no or very little prior exposure to music technology to an experience of embodied and gestural interaction with technological instruments, sensor-controlled live-electronics or so-called digital music instruments. Very little has been written until now in the NIME context about teaching and transmitting know-how about techniques, methods and strategies of dealing with a technology studies, the individual learning processes or are related to a specific interface types. This article tries to address the topic by providing some background reflections, methodical and conceptual categories and concrete solutions that were used in developing such a class.

2. BACKGROUND

The concept of gestural interaction with instruments forms the basis of most of our real-life experiences with musical performance. A foundational part of our education and cultural exposition to music is related to ‘playing’ musical instruments of the physical variety. Any musical training that surpasses the ‘mere’ act of singing has to deal with establishing a physical relationship with an instrument. The training on an instrument, in its initial stages at least, is mainly concerned with building this rapport between the body and the actions on the instrument that produce the sound. It serves to imprint the shape and sound of an instrument and its affordances and constraints into an adaptive, dynamic, extended, and perceptual body. Gallagher defines this as: “a body schema [that] is a system of sensorimotor capacities that function without awareness or the necessary perceptual monitoring” [5].

Highly integrated body-instrument relationships build on these schemata and permit to achieve fluent playing and reflex-like control. However, the more removed an instrument’s playing mode is from our normal object-handling-knowledge, the harder it will become to play it well and to achieve a good motor integration. On instruments based on simple acoustics and physics, basic modes of bodily actions are sufficient to produce an adequate sound, whereas on complex instruments a large number of fine sensorimotor adaptations have to be acquired in order to produce the desired output. A good case illustrating this point would be that of a toddler scratching on a violin compared to hitting on a drum. Of course, a percussionist ultimately develops the same level of differentiated control as any other instru-
mentalism, the differences show themselves in the type of adaptations and the different learning curves.

On an even more fundamental level, perception can be seen as necessarily involving “a relation to an object, or to some state of affairs that ultimately depends on an experience that is directed at one or more objects” [5]. In a music making situation, the objects of perception are on the one hand the instruments and on the other hand the music and sound processes that are produced and perceived. When a certain level of proficiency and fluency has been achieved in playing the instrument, the body becomes experientially ‘transparent’ [9] or will only be perceived peripherally. The difficulty of playing the instrument moves into the background and the intended output moves into focus.

2.1 The ‘Enactive’ Approach

These elements of perception relate to the traditional instruments, but also to new hybrid forms of musical interaction. When focusing on a musical practice with digital musical instruments, an approach that starts from a physical action or gesture rather than digital processes of sound transformation brings the basic principle of movement-informed perception to the foreground. This principle deals with the relationship or interdependence between the perceptual and the physical domains. Varela states that: “the enactive approach consists of two points: (1) perception consists in perceptually guided actions and (2) cognitive structures emerge from recurrent sensorimotor patterns that enable action to be perceptually guided” [20].

This means that the fine motor skills involved in playing an instrument form part of the perceptual system, or differently put, the body forms part of the adaptive loop that may start with the intentional sound-producing action, whose result rises to conscious perception in cognitive structures but is guided by or adapts via sensorimotor patterns embedded in our body.

And since not only perception but also conscious processes are guided by actions, the corporeal aspects of the loop become an inseparable part of the process. To emphasise this, Varela postulates that: “cognition depends on the kind of experience that come from having a body with various sensorimotor capacities and [...] these individual sensorimotor capacities are themselves embedded in a more encompassing biological, psychological and cultural context. [...] sensory-motor processes, perception and action are fundamentally inseparable in lived cognition” [20].

2.2 Digital Musical Instruments

It is clear that a digital musical instrument is constituted by more dimensions than just the physical and that these dimensions are also capable of eliciting perceptual experiences and even insights. The fact that the instrument is constituted by a physical as well as a symbolic layer compels us to view it under the dual perspectives of materiality and abstract structures. Contrary to traditional musical instruments that are physical contraptions built to be activated in a specific mode of physical sound production, the digital musical instrument combines metaphorical structures in the symbolic domain [8] and gestural affordances and constraints in its physical interface [6]. It embeds “ideas of music, musical culture, and musical work practices” [10] since its structure has to be designed if not distilled from an almost infinite field of potential and thus will invariably be informed by the conceptual capabilities and contextual choices of its creator.

In this sense the digital musical instrument serves as a vehicle or container for a type of knowledge about a specific music practice. This has probably always been the case to a lesser degree, especially because lutherie or instrument building in general have always been crafts situated at the forefront of available “technologies”. But with the inclusion of the ultimate symbolic machine, the computer, into a musical instrument, the notion of it carrying the full knowledge of music(s) and their structures with it, comes into its own. The question is, how this “embedding of music theory and other systems of knowledge in the instruments themselves” [10] alters, enhances or generates new insights.

2.3 Teaching NIMEs

In general, classes with NIME themes are given in the context of media and technology studies. An emphasis is usually put on the students learning the design concepts and becoming familiar with the technological aspects by using the various technologies in order to explore artistic concepts that combine interactions with interfaces. The knowledge and skills transmitted in these courses deal with basic technical understanding of sensors, interfaces, code, interaction design principles and the building of an individual, even idiosyncratic interface solution. In another trend, the arrival of laptop or mobile-phone orchestras proposes a different pedagogical practice [19, 4]. The students are guided in the exploration of electronic music performance with a device-and-software-combination that offers varying degrees of gestural affordances and sonic richness. The music is usually firmly anchored in a digital aesthetic and a large part of the learning process is concerned with experiencing the domain of purely electronic sounds. However, the group practice of these orchestras come closer to the musicianship of traditional ensembles, even though groups of identical digital musical instruments are used. The development of compositional strategies for ‘polyphonic’ or multi-voice electronic music is explored, sometimes in networked configurations, always in a reactive performance mode and usually ‘conducted’ in some way by a leader, or interconnected via a network [21]. Apart from these two paradigms, no other courses exist, to our knowledge, that try to bridge the gap between an instrumentalist practice and a gestural musical interaction through NIMEs.

3. THE COURSE

The course is designed against this backdrop of reflections about physical agency and perception on the one hand and embedded musical knowledge of instrumentalists on the other hand. The target audience for this class is situated in a context of higher level (academic) music training in either classical music or contemporary art-music, be it as an instrumentalist or composer. The participants have experience in playing contemporary music or new music (some also specialise in old music), while some of the instrumentalists have acquired the experience of interpreting repertoire pieces from live-electronic music. Most of them, however, have not been exposed previously to a culture of electronic music performance in more popular styles and are not necessarily the technology-savvy types.

Taking these background considerations into account, a basic interaction model was designed for this class, that can exemplify a musical situation likely to be encountered by the participants in their professional careers and that demonstrates one fundamental way of working with electronic sound-processing. By mixing the practise of live-electronic sound-transformation of real instrumental sound with a gestural control paradigm, the hybrid musicianship that is posited in the title is created: this forms a contrast to the predominant styles in interactive electronic music, which are characterised by abstract synthesised or sampled
sounds that are coupled with gestural interaction through an interface. The juxtaposition between an existing instrumental skill of playing and producing sound and the new activity of sound-transformations through gestures in real-time produces a blending of skills and provides the experiences that is aimed at in this pedagogical context. For a traditionally trained musician the addition of a new layer of sonic materials to her instrumental skills is appealing. That way it is not a new musical idiom that has to be learned, but merely an extension of the existing practice. Maintaining the sense of agency both in the electronic as well as the acoustic domain reinforces the musical experience, something that in conventional live-electronics is not necessarily the case. The choice of using live-audio input serves the purpose of anchoring the musical action in the bodily domain as well. From a perceptual and ‘enactive’ point of view, the connection between the physical act of playing a sound combined with a corporeal gesture that controls transformation, is more strongly reinforced than if the sound were produced in a abstract non-corporeal manner.

3.1 Methods
In order to be able to provide an experience of the interaction with sound through body and gesture, most of the complexities of a digital musical instrument have to be hidden away. Therefore it is important that the participants will not have to deal with the questions of building sensor hardware or configuring software themselves and are thus able to concentrate more fully on applying their interaction ideas with a set of given tools. An emphasis is put on imagining short ‘scenes’ or scenarios made of a movement and gesture that will alter the sound of the instrument. Simple exercises are set up with the different sensing techniques, where a type of gesture, a mode of playing the instrument and a mapping are combined into a short sketch.

In order to experience the entire workflow and instead of spending much effort on trying to solve single technical or conceptual issues with the technology – something which is a common pitfall of working with sensors – a lot of importance is given to these immediate hands-on experiments. The ‘simple scenes’ assignment also forces the participant to think in terms of musical effect or the possible expression of the final outcome of an idea. This emphasis permits to thematise the experience of performing and at the same time provides a means of reflecting the connection between artistic intention and concrete musical application.

Two fundamental types of sensing interfaces are provided in this class. The first one is directly related to the body-space or instrument (on-body inertial sensing); the second one provides information about the position in space of the body or a view of the body-shape (camera based silhouette tracking). These two concepts have been described in the context of interactive dance [18, pp.30-31], coining the terminology of inside-in and outside-in sensing, but they apply equally well to an instrumentalist’s action space.

The sequence of introduction of the elements in the class is arranged in such a way as to permit the participants to completely explore one sensing mode and gain a clear point of reference before being exposed to the second one. A central pedagogical goal of this class is to convey the importance of the concept of cross-domain translation and more specifically that of mapping. The translation of essential aspects of movement from the physical domain into the discrete and abstract world of numbers and parameters needs to be understood, in order to connect real instrument sounds and digital sound processing. To that end the basic concepts of mapping, such as scaling, routing, one-to-one vs. many-to-one and one-to-many, are first presented in a theoretical way. However these types of mappings are then immediately put to action in an experimental trial phase, in order to experience directly how they can be applied.

3.2 Interfaces
A set of soft- and hardware interfaces was assembled and is put at the disposal of the participants. The first type is a sensor interface with inertial sensing. Making use of open-source and commercially available sensor and microcontroller packages, the small inertial sensing node offers a minimal but essential set of sensing capabilities. Built with the wireless, rechargeable battery-driven Arduino FIO platform or with an Arduino mini pro, the motion sensor for acceleration and rotation measurement and two buttons are mounted in such a way as form produce a very compact little package (see Figure 1).

Figure 1: The inertial sensor node attached to a flute.

The device resembles in size and scope the one designed at IRCAM that became the MO-objects [2, 14] or the Sense/Stage modules originally developed at Concordia and now supported by STEIM [1].

The inertial node is ideal for hand-held actions but also small enough to be strapped to a wrist or an instrument. As an alternative the same wireless motion-sensing device is mounted on a small, flat bracelet intended to be worn while playing (see Figure 2).

This technology was deliberate chosen in order to present a technical solution that stays within the reach of the participants, should they decide to want to build their own, even if this exceeds the scope of this class.

In addition to the sensing node, a dedicated proxy-server software is provided. Its sole function is to translate the serially transmitted sensor values to an OSC stream and relay it on the internal network bus. It may be compared to what STEIM’s junXion software\(^3\) or the OSCulator\(^4\) provide, but is explicitly tailored to the device in a ‘fire-and-forget’ way. It also helps to isolate the user from the sometimes irritating problems of running virtual serial ports within the same software as the sound processes.

The second type of interface that is used is the – now almost ubiquitous – Microsoft-Xbox Kinect camera. Again, apart from providing the hardware, a software solution in the form of a small, single function standalone software is

\(^3\)http://steim.org/product/junxion/
\(^4\)http://www.osculator.net/
all URLs last accessed in February 2013
put at disposal, that handles all the camera-specific controls. This software leverages open-source libraries giving access to all the data from the depth-camera and like the previously mentioned sensor proxy server was coded in openFrameworks. The depth-image is used not to extract the skeleton – which was found to be too slow for musical use – but to compute a simple silhouette analysis on the person’s body, and provide only the five essential points of the silhouette in coordinates corresponding to the actual space (see Figure 3). More details about this type of analysis using 2D camera images were presented in an earlier NIME-publication about interactive dance [16].

Figure 3: Tracking a moving instrumentalist. Inset a view of the camera analysis tool. Note the failure of the kinect to capture the shiny surface of the flute.

3.3 Software

Finally, and in order to complete the digital musical instrument, a software instrument was designed according to the paradigm of the workshop. This tool consists of two principal parts: a chain of audio-processing modules and a mapping layer providing access to all the parameters of the audio layer. The audio processing part is deliberately kept simple. Apart from a direct input and a file-playback module, the audio transformation modules comprise a delay, a reverb (freeverb), a filter (resonant), a module for ring-modulation and a granulation unit (munger). Each of the modules has a dry-wet mixing and bypassing controls, and all the parameters of the effects are exposed to the mapping layer via a dedicated namespace. The audio chain is geared towards processing live-audio and the omission of a more powerful, loop-based sample player is a deliberate choice. The configuration of the audio-system is stored in persistent settings and can be interpolated where it makes sense. For the participants, learning to use the sound transformation doesn’t present a big challenge, because the possibilities offered by these sound-transformation modules is limited.

Learning to use this mapping layer, however, presents the big challenge for the participants. The mapping layer is structured in a modular fashion, connecting OSC input streams to output nodes. The connections happen via dynamic address-based routing. Each data-stream is conditioned using pre-scaling, smoothing and if desired a transfer function, and finally is forwarded to the target parameter through an address-based output routing. In addition, a system of thresholds using simple boolean operators is implemented, that permits the conversion from continuous streams to discrete events i.e. triggers. The entire mapping is stored in persistent settings and can be interpolated where it makes sense. In order to make it easier to understand the behaviour of the sensors and the numerical ranges of the values, a display for incoming values and the possibility to automatically capture the minima/maxima ranges is provided.

This mapping layer is structured in a highly methodical fashion. It tries to convey all the basic mechanisms needed to achieve a meaningful connection without incurring the cognitive load of a high flexibility and complexity. For more advanced users, the application of a more mature system, such as the ‘libmapper’ tools, would be a better choice [11]. Again, in order to give more advanced users the chance to leverage this software for their own use, it is provided in both the finished form and as source in a collection of MaxMSP patches and externals, to be tinkered with by the participants themselves (see Figure 4).

Figure 4: The audio and mapping parts of the software (made in MaxMSP).

To round off the description of the materials used in the class, a word about the audio equipment. Each participant is provided with a microphone, an audio-interface, their own laptop or a computer provided by us, and one loudspeaker located in their immediate vicinity. This simple and
straightforward setup helps to emphasise the integrity of the body-instrument relationship together with the interactive situation and facilitates the acoustic blending of instrumental and electronic sound. For the class it also helps to have a large enough space so that the different experiments don’t get in each other’s way. And of course, feedback loops and volume have to be constantly kept in check.

4. DISCUSSION

After two editions of this class within the span of half a year, a few insights were gathered that should help to shape future teaching of this subject. One of the essential tasks of the teacher is to balance the complexity of the technical elements with the simplicity of traditional instrument playing. Fortunately, the participants in both editions of the course were motivated enough to willingly tackle that hurdle and keep going through more complex topics and managed to maintain a grip on their musical ideas throughout the course. Apart from fundamental pedagogical principles valid at any level of expertise, in this specific mixture of themes and material configuration the following facts became apparent:

- Understanding the concepts of mapping and sound transformation is only the first step to meaningful gestural interaction. Having a clear scenario or idea for a performance situation based on a model from the physical world helps to focus the entire workflow from instrument to augmented sound.
- Giving examples of gestural interaction, in forms of videos or a live demo, help to motivate the participants. However, they also bias them and skew the ideas into the direction of the examples. Let the participants bring their own ideas first, possibly formulated in natural language, before demonstrating too much prior art.
- For people with no mathematical background, to understand the notion of or to mentally represent a physical phenomenon such as acceleration or rotation with a stream of numbers poses a big challenge.
- Spatial movement is easier to understand and control than energy measured with inertial sensors. Let the participants work on the harder method first; the easier method will free them to think more musically.
- A good way of balancing the cognitive load can be achieved by always leading back to the bodily domain and the traditional instrument playing expertise. This helps to anchor the experiments in real-world experiences (and emphasises the ‘enactive’ approach).
- It is extremely important to let the participants experiment on their own.
- Small steps covering each element of this multi-part workflow are essential.
- Not every instrument is easy to move with, but all instruments have one aspect of movement as their fundamental interaction form. All instrumentalists move in idiosyncratic ways, this is enough to start the exploration.
- Movements are not necessarily gestures. They only become that, if they manage to convey some sort of meaning. Building interaction scenarios on that principle is a good starting point.
- Using sound generated by the machine, for example sample playback, is always the easy way out. Resisting that temptation forces the musician to re-evaluate her musical and interaction ideas.

This listing is by no means exhaustive and it is certainly quite subjective. However, it shows the scope of the issues presented by the concept of this class. It is interesting to note the different levels that have to be taken into account for this concept to work. All technical requirements and problems have to be solved before starting the class. In order for non-technically trained participants to understand the concepts of the course, a simple and clear model has to be presented. A run-through of the complete workflow needs to be completed by the participants at least once as quickly as possible before musical issues can be approached. The experience of the interactive situation is only complete with a direct musical feedback.

By placing this course in a context that is outside the usual media- and technology track, people get exposed to the NIME ideas that wouldn’t normally explore these issues. Incidentally, most of the participants, once they get comfortable with the basic premises, also start to explore other situations and ideas, some of them musical, some of them not. This is a welcome development, since it shows that the imagination can be triggered in other directions than what is presented at the outset. Thanks to this openness, many of the participants use the course as an opportunity to explore the possibilities of the gestural interaction through these technologies in pointing directions that are not just in the sense that we proposed. This represents an interesting collateral benefit. Many participants bring in their own ideas, be it for a specific performance type, an interactive installation or a different type of interaction they want to explore.

Some of the scenarios or short scenes that were shown by the participants at the end of each course were quite expressive:

- Two double-bass players for example developed a good gestural correspondence between the gesture of quickly pulling back of the bow in a ‘sforzato’ style and immediately freezing the sound through a reverb into a noisy sound surface.
- A composer/accordionist established a image-sound relationship by opening and closing the bellows imitating breathing and changing the filter and reverb quality of the amplified air sound.
- A composer working with a flautist used the difference in body surface as seen by the camera to map the sequence of moving from a crouch to a fully raised flute to the size of the reverb space, going from dry to very reverberant (see Figure 3).
- One of the more technically advanced participants decided to try to control a small servo-based moving head with a torch to have it follow the top of the tracked silhouette like a search light.
- A performer working gesturally with contact microphones on cardboard boxes managed to map the position of her hands on the box and her raised arm to specific transformations: lowering the arm at the front of the box opens the reverb, and scratching the surface from front to back changes the amount of ring-modulation of the cardboard sound.
- A musician coming from a hip-hop background discovered a way of using the typical hand-pumping gesture to change the delay and reverb amounts applied directly to the hip-hop track being played.

5. CONCLUSION

This class is structured with a specific paradigm in mind, that of an ‘enactive’ experience and connecting the motor skills of trained instrumentalists with technologically...

---

\(^{5}\)for example the piano, but we even had somebody using that as gesture!

\(^{6}\)The control of lights or projection was the second biggest theme that cropped up in the class.
mediated sound processes. A big part of the effort demanded of the participants is to understand the connection between the traditional instrument, the gesture and the sound-transformations. The goal is to demonstrate in an exemplary fashion the complete workflow of working with interactive gestures, and by that means provide the experience of the transformation of instrumental live-sound on a musical level in a manner closely related to a musician’s practice. By proposing a situation where a musical action is modified by gesture and provides a direct feedback, the same perceptual adaptations that are given through traditional instrumental skills can be experienced despite the filter by a symbolic machine. David Wessel justly remarks the following: “both traditional instrumental and vocal musical practices require rich sensory-motor engagement. [...] modern computer-based musical instrumentation [however] remains far from involving the body” [22].

It will prove interesting in future editions of the class to let the participants explore more complex mapping-schemes. Ideally the possibilities offered by machine-learning or generative algorithms could be integrated to provide yet another level of translation between the bodily, instrumental and the abstract symbolic domains. In the future, the class could also be structured to include the possibility of working with small teams of dancers and musicians. Finally, with more time for the class a slower, more in-depth learning process will be established, where more careful exploration and experimentation become possible.

By bridging the two worlds of traditional instrument practice and gestural interaction based on digital music instruments, we believe that the participants get to experience a truly hybrid musicianship.

6. ACKNOWLEDGMENTS
I’d like to thank my colleague Sébastien Schiesser and the sound-engineer Simon Kônz for their active participation in this pedagogical endeavour.

7. REFERENCES