The Mitt: Case Study in the Design of a Self-contained Digital Music Instrument

Ivan Franco* and Marcelo M. Wanderley

CIRMMT, McGill University, Montreal, Canada
ivan.franco@mail.mcgill.ca

Abstract. *Many Digital Music Instruments (DMI) are composed of an input controller connected to a general-purpose computer. But as computers evolve we witness a proliferation of new form/factors, bringing us closer to the possibility of embedding computing in everyday tangible objects. This fact may have a considerable impact on future DMI design, through the convergence between gestural interface and processing unit, materialized into Self-contained DMIs.*

By bypassing general-purpose computers and their imposed interaction modalities, these instruments could be designed to better promote embodied knowledge through enactive interfaces, while still maintaining many of the capabilities of computer-based systems. This context suggests the research of novel interaction models and design frameworks.

"The Mitt" is a Self-contained Digital Music Instrument which explores the capture of high-resolution finger gestures through its tangible interface. It is a first implementation using an ARM embedded system, customized for sensor data acquisition and sound synthesis, and capable of dynamic re-configuration.

Keywords: Digital Music Instrument, Embedded Systems, Granular Synthesis, BeagleBone Black, SuperCollider.

1 Introduction

In this paper we start by discussing the motivation behind the study of Self-contained Digital Music Instruments, after which we describe a supporting technical system and the particular implementation of The Mitt. Lastly we draw conclusions from the development of this first proof-of-concept and discuss future research directions.

Many Digital Music Instruments (DMI) have traditionally followed a morphology that relies on several distinct technical apparatuses, grouped together to form what is considered to be the instrument. The most common case is the use of input controllers connected to general-purpose computers. The controller acquires information about the performance gesture and sends this data to the computer using a standard protocol. The computer is then responsible for sound

* Gratefully acknowledges the financial support from the Portuguese Foundation for Science and Technology.

synthesis and mapping, functions that often require a considerable amount of processing power and that are intricately related to the musical properties of the instrument[1].

Undoubtedly this is a convenient architecture, since it allows the musician to flexibly reconfigure the functionality of the instrument, by programming new states into the machine and radically changing its behavior. Yet, instruments that depend on general-purpose computers often fail to provide the sense of intimacy, appropriation and determinacy that traditional instruments offer, since they enforce non-musical activities and rely on fragile technological systems. These problems are less prominent in a category that we will define as dedicated devices - electronic music instruments that use digital processing but are focused on the musical activity, like traditional hardware keyboard synthesizers or drum-machines.

By observing market trends at music trade shows, like NAMM or Musikmesse, we conclude that the public continues to show a large preference for dedicated hardware devices when it comes to musical instruments. While many musicians have promptly embraced the use of computers as virtual substitutes for production studios, they have largely unexplored their potential as expressive performance instruments. This could be due to many different factors, so it is relevant to analyze the strengths and weaknesses of both DMIs and dedicated devices to better inform new design proposals.

2 Dedicated Devices versus General-purpose Computers

It is difficult to compare dedicated devices to DMIs, considering the diversity of existing instruments, ranging from traditional keyboard synthesizers to the idiosyncratic practices of skillful makers and artists, that create their own instruments and practices. Still it is useful to try to analyze some of the differences at both theoretical ends of the spectrum.

Dedicated electronic music devices are readily available to be played and do not require complex connections, operating systems or the launch of specific computing processes to reach a ready-state. The bypass of several non-musical interaction layers brings dedicated devices closer to the immediacy of acoustic instruments, which "just work". One could use the expression "pick & play" for the definition of this quality, which is not present in most computer-based systems.

Another important distinction is that dedicated devices have relatively static functionality, while the computer can virtually recreate any type of instrument through software. It can also radically shift its behavior by dynamically changing synthesis and mapping on-the-fly. This capability for reconfiguration is possibly one of the clearest distinctions of DMIs. Additionally some DMIs can also incorporate other parallel interaction tasks, like co-play through algorithmic processes or networking with other devices to share musical data.

A common complaint from computer musicians is related to the effort in dealing with non-musical technicalities. Although system complexity can be easily accommodated by tech-savvy individuals or years of practice with specific tools, it is undeniable that having to involve a computer is discouraging for many users. Due to their fixed architectures, dedicated devices have a significantly increased resilience and reliability, making them less sensitive to problems of longevity. Contrarily, today's DMIs will most certainly not work in future computer architectures and operating systems. Manufacturers are currently moving to yearly operating system updates, often breaking functionality of relatively recent software and hardware. Personal computers also tend to be dedicated to multiple other activities, which often results degraded performance or compatibility.

Computers are also associated to particular musical sub-genres and practices, such as live coding, which deliberately embraces computer aesthetics and programming as part of an artistic manifesto. While this is certainly a valid musical approach, it relies more on the construction and steering of rules/models and less on direct motor skill[2], impacting the nature of the musical performance. Many musicians value instruments that are oriented to the maximization of dexterity and embodied knowledge, possibly easier to achieve through dedicated devices and their music-focused interfaces.

Finally, another important characteristic of the computer is that it incites continuous tweaking, which could be detrimental to the progress of skill-based knowledge. This problem was often referred by Michel Waisvisz[3], who was capable of delivering intricate and highly expressive performances with his instrument, "The Hands". Waisvisz often referred that his acquired skill was due to a voluntary decision to stop development and spend another ten years learning to play. The achievement of competence could be strongly molded by static affordances and constraints[4], often clearer in dedicated devices than computer systems.

Each of these aspects requires in-depth study. To support our experiments we have decided to first concentrate on surpassing the model of the controller-computer combo, by developing an embedded computing platform useful in the fast prototyping of Self-contained DMIs.

3 Previous Work

The use of embedded computing in digital music instruments is not new. Most digital synthesizers since the 80's use some sort of dedicated DSP chip, with compact form/factors, high-performance and low cost[5]. Field Programmable Gate Arrays (FPGA) are also gaining market in audio applications, due to their extreme efficiency in hard realtime DSP processing[6].

The recent push in low-powered computing for smartphones and the internet-of-things has also greatly contributed for the advance of ARM processors, which are being integrated into circuits called System-On-Chip (SOC), coupling processor, clocks, memory, interfaces and mixed signals into a single package. Closer to computer architectures, SOCs run simple operating systems and take advantage of mature software libraries, with a ease-of-use and flexibility more difficult to achieve with DSP or FPGA solutions.

There are already instruments and effects processors developed on top of ARM architectures. Notable examples are the digitally-reconfigurable stompboxes Owl and Mod Duo. These use a companion software to design and upload programs to the standalone instrument, actively promoting exchange of user setups on the Internet. Satellite CCRMA[7] is a Linux distribution that ships with pre-compiled binaries of popular computer music languages and runs on the Raspberry Pi, an ARM single-board computer(SBC). The Cube[8] is another instrument that uses a Beaglebone Black (another ARM SBC), housed inside a wood box that can be opened to expose a breadboard, modifiable in an exploratory fashion akin to circuit bending techniques. Although these are relatively different instruments, they all incite a model of flexible reconfiguration of sound synthesis or interaction.

4 System Architecture

To support the quick development of standalone instruments we built a system based on the Beaglebone Black, a modest ARM SBC equipped with a 1 GHz Cortex-A8 processor. A custom stackable board, equipped with another Cortex-M3 microprocessor, is responsible for signal acquisition with 12-bit resolution, and can in turn be expanded through up to ten additional 5 x 2.5 cm boards with 8 channel multiplexers. With this setup the Beaglebone Black can receive up to 80 simultaneous analog sensor signals, via an internal serial UART connection, offloading these processes from the main Cortex-A8 processor.

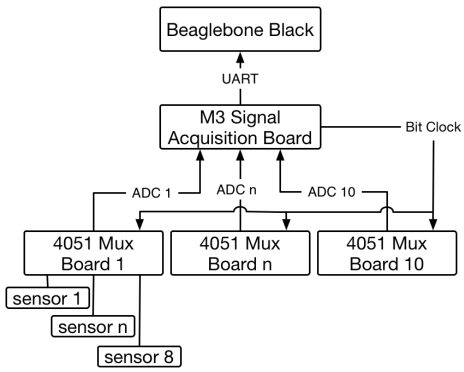

Fig. 1. Signal Acquisition Architecture

Sound input and output is done through the addition of an external codec chip, that connects to the Beaglebone Black via USB or through the Multichannel Audio Serial Port (MCASP), exposed by the Beaglebone's General-Purpose Input Output (GPIO) pins.

Audio processing is done using SuperCollider on top of Linux, running at a sampling rate of 48 KHz and 16 bit depth. SuperCollider programs can be either edited directly on the device (via network connection) or uploaded in a micro SD Card. We chose SuperCollider due to the convenient features of an interpreted language, like on-the-fly reconfiguration without compilation stages or audio interruption. Although this choice implies a performance trade-off, our previous tests have revealed the ability to run relatively complex audio synthesis[9].

5 "The Mitt"

"The Mitt" is a first instrument built using the previously described system architecture. It aims to explore the fine motor skills of the human hand and the performative aspects of micro gestures. In the next section we describe some of the instrument's features.

5.1 Tangible Interface Morphology

The Mitt's tangible interface is constituted by an array of five vertically disposed channels, each composed of a highly sensitive thumb joystick, three potentiometers and a button.

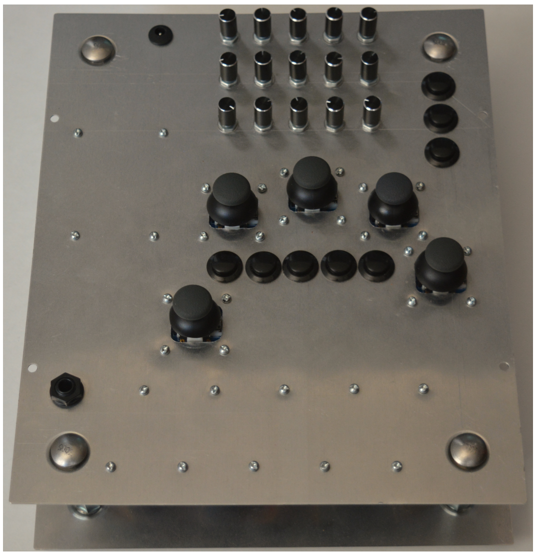

Fig. 2. The Mitt

Instead of using a perfect alignment, the five joysticks are distributed to naturally accommodate the morphology of the first author's dominant hand, so that each finger can easily rest on top of its respective joystick. They are equipped with inner spring, facilitating the unguided return to a neutral central position, and their vertical shaft measures 25 mm, with a 45-degree bend end-to-end. This implies a travel projection on the horizontal plane of about 2.5 cm, resulting in a rough estimation of 1 mm accuracy, considering a single byte reserved for the representation of each axis (256 possible values).

The case is made of two aluminum sheets sustained by vertical rivets and the front panel also accommodates a power connector and a 1/4” stereo audio output jack. Three additional general-purpose buttons are used for system interactions.

5.2 Sound Synthesis

In this particular implementation we arbitrarily decided to explore granular synthesis. Each of the 5 channels produces grains with maximum duration of 1 second, driven at a frequency of up to 15 KHz and with a maximum of 20 overlapping grains. Each channel has an independent sound buffer, to which new sounds can be freely loaded during performance. By continuously changing the sample read position it is possible to induce a sense timbre morphing. Harmonic relations are possible by using pitched samples and converting notes to the equivalent playback rates. Finally each channel is chained to a master effects bus with panning delays and long-tailed reverberation.

5.3 Mapping

The main instrumental articulation is performed through the joysticks, while the potentiometers are used for parameter fine-tuning. The bi-dimensional position of the joystick is converted to polar coordinates and the resulting vector size is applied to the granular voice amplitude, while angle affects playback rate (pitch). A switch allows for pitch quantization, defined by a scale and distributed across an interval of octaves. In turn the three potentiometers control trigger rate, grain size and position in the buffer.

Although this a case of explicit one-to-one mapping[10], with each control mapped to a single synthesis parameter, interesting behaviors emerge from the mechanical and physiological constraints implicit by the design. Since it is difficult for a human to move fingers independently, any hand gesture will potentially influence the five points of control simultaneously. It is possible to use gestures such as opening, closing, translating or rotating, adding dynamism to the five voices simultaneously. Future mappings could further explore this implicit correlation by applying gestural classification and interpretation.

By mapping the joystick’s vector size directly to amplitude, the Mitt requires continuous energy input from the user[10], due to joystick’s natural return to a central position. The result is an instrument that has little interruption tolerance but that in return highly promotes skill and nuanced playing[2]. An alternative mapping that also encourages energy input is the association between sample position and angle, controlling timbre morphing with rotational movements.

6 Discussion

6.1 Musical Expressiveness with The Mitt

The fast response and fine accuracy of the joysticks result in a control with a high range of expressiveness. The extreme gestural amplification derived from the chosen mappings causes small movements to have a significant impact in the resulting sound, facilitating dramatic variations from loud, frantic and noisy to quiet, contemplative and delicate. This quality may be deterrent to the audience's understanding of cause-effect, due to the imperceptibility of such small movements, even if the performer is actually involved in an exceptional physical effort required in the sustaining of a musical event. On the other hand it empowers the user with a very direct connection between motor function and sonic result, exploring human sensitivity and skill with detailed finger gestures. Although initially idealized for slow-evolving sounds, the Mitt is also very appropriate for playing short events that can be swiftly released, due to the strong pull-to-center exerted by the joysticks.

6.2 Connectivity and Integration into Existing Setups

The distribution of tasks across devices offers the possibility for the musician's cockpit[11] to be composed of several independent instruments and sound processors, connected to each other to achieve complexity. This notion also extends to control data since the instruments can be networked. Considering modularity and a shift to more task-focused devices, the limited processing of the ARM processor seems less significant. In turn the capability for quick re-routing of audio and control functions could promote an increased quality of immediacy, further stimulating quick experimentation and tacit knowledge.

6.3 Reconfigurability

Although we have described a specific synthesis application, the capability for DMIs to change their behavior is one of their defining traits. The Mitt can smoothly cycle between stored patches, that will gracefully fade out to be substituted by new DSP chains without audio interruptions or the need to compile and relaunch applications. By separating main data handling programs from instrument definitions, new behaviors can be created with succinct computer code.

6.4 Towards Better Metaphors

The Mitt's goal is to explore human motor functions through an abstract interface, but we admit that there might be advantages in designing interaction models that are more tightly coupled with particular sonic creation metaphors. This is the case with the Pebblebox[12], where granular synthesis is represented by a box full of pebbles that can be freely manipulated. These types of interfaces could prove to be beneficial to self-contained instruments, by imposing a strong identity in detriment of adaptability.

7 Conclusions

In this paper we have presented the motivation for the development of digital music instruments that promote immediacy and ease-of-use through self-contained designs, while maintaining attributes that are singular to computer music. We have also presented a first test with the Mitt, an instrument that explores a particular control method, but that can be easily reconfigurable to perform any type of synthesis and mapping functions.

Future steps will concentrate on further understanding the influence of affordances and constraints imposed by tangible interfaces and how they might influence usage patterns and levels of customization.

References

1. Miranda, E. R., Wanderley M. M.: New Digital Musical Instruments: Control and Interaction Beyond the Keyboard. A-R Editions, Inc. (2006)
2. Malloch, J., Birnbaum, D., Sinyor, E., Wanderley, M. M.: Towards a new conceptual framework for digital musical instruments. In: Proceedings of the 9th International Conference on Digital Audio Effects, pp. 49-52. Montreal, Quebec, Canada (2006)
3. Krefeld, V., Waisvisz, M.: The Hand in the Web: An Interview with Michel Waisvisz. Computer Music Journal. 14 (2), 28-33 (1990)
4. Norman, D.: The Design of Everyday Things Revised and Expanded Edition. Basic Books, New York (2013)
5. Wawrzynek, J., Mead, C., Tzu-Mu, L., Hsui-Lin, L., Dyer, L.: VLSI Approach to Sound Synthesis. In: Proceedings of the 1984 International Computer Music Conference. San Francisco, USA, (1984)
6. Saito, T., Maruyama, T., Hoshino, T., Hirano, S.: A Music Synthesizer on FPGA. In: Field-Programmable Logic and Applications: 11th International Conference, pp. 377–387. Springer, Belfast, Northern Ireland, UK (2001)
7. Berdahl, E., Wendy J.: Satellite CCRMA: A Musical Interaction and Sound Synthesis Platform. In: Proceedings of the 2011 International Conference on New Interfaces for Musical Expression, pp. 173-178. Oslo, Norway (2011)
8. Zappi, V., McPherson, A.:Design And Use Of A Hackable Digital Instrument. In: Proceedings of the International Conference on Live Interfaces, pp. 208-219. Lisbon, Portugal (2014)
9. Franco, I., Wanderley, M. M.: Practical Evaluation of Synthesis Performance on the BeagleBone Black. In: Proceedings of the 2015 International Conference on New Interfaces for Musical Expression, pp. 223-226. Baton Rouge, USA (2015)
10. Hunt, A., Wanderley, M. M.: Mapping Performer Parameters to Synthesis Engines. Organised Sound. 7 (02), 97-108 (2002)
11. Vertegaal, R, Ungvary, T., Kieslinger, M.: Towards a musicians cockpit: Transducers, feedback and musical function. In: Proceedings of the International Computer Music Conference, pp. 308-311. Hong Kong, China (1996)
12. O'Modhrain, S., Essl, G.: PebbleBox and CrumbleBag: Tactile Interfaces for Granular Synthesis. In: Proceedings of the 2004 Conference on New Interfaces for Musical Expression, pp. 74-79. Hamamatsu, Japan (2004)