Rhythm'n'Shoes: a wearable foot tapping interface with audio-tactile feedback

Stefano Papetti
Dept. of Computer Science,
University of Verona
Strada Le Grazie, 15
37134 Verona, Italy
stefano.papetti@univr.it

Marco Civolani and Federico Fontana
Dept. of Mathematics and Computer Science,
University of Udine
Via delle Scienze, 206
33100 Udine, Italy
name.surname@uniud.it

ABSTRACT

A shoe-based interface is presented, which enables users to play percussive virtual instruments by tapping their feet. The wearable interface consists of a pair of sandals equipped with four force sensors and four actuators affording audio-tactile feedback. The sensors provide data via wireless transmission to a host computer, where they are processed and mapped to a physics-based sound synthesis engine. Since the system provides OSC and MIDI compatibility, alternative electronic instruments can be used as well. The audio signals are then sent back wirelessly to audio-tactile exciters embedded in the sandals' sole, and optionally to headphones and external loudspeakers. The round-trip wireless communication only introduces very small latency, thus guaranteeing coherence and unity in the multimodal percept and allowing tight timing while playing.

Keywords

interface, audio, tactile, foot tapping, embodiment, footwear, wireless, wearable, mobile

1. INTRODUCTION

In many cultures, music and dance performers make use of foot tapping, from folk fiddlers and street buskers to flamenco and tap dancers. For instance, a fiddler stomping on a pub's wooden floor can cheer on the audience meanwhile supporting his or her own playing by adding a simple percussion part; buskers often include foot drums in their setup to add even complex percussion parts to their guitar playing. Moreover, traditional musical genres exist where players make extensive use of foot percussions (*podo-rhythm*) as main accompaniment. As for dance, foot tapping can have both an expressive and rhythmic function, to the extent that some dance genres are centered on the musical and gestural performance produced by the dancer's feet.

On the other hand, in everyday life many musicians and music enthusiasts alike find themselves "tapping songs" with their fingers, hands and feet. Such tapping may represent the song's main melody, rhythm, or even accurately simulate its percussions part.

The gesture of playing rhythms with the feet offers spontaneity and expressivity, at the same time enabling an embodied experience.

Taking inspiration from these observations, and starting from a prototype shoe-based interface we had previously realized for interactive walking purposes [18], we implemented a wearable controller for foot tapping that we have named "Rhythm'n'Shoes". A peculiarity worth noticing is that the interface provides the user with foot-level audio-tactile feedback through exciters embedded in the shoes' sole.

A similarly immediate approach to playing rhythms that avoid the use of virtual drum interfaces, but instead takes inspiration from the common experience of hitting the chest or thighs with the hands, is depicted in [2]: the interface consists of a pair of gloves embedding piezo microphones that are used as sensing devices. Several recent studies take into account novel percussion instruments [1, 6] and interfaces for percussion tasks [9, 25].

As for foot-based interfaces, various works exist that describe instrumented shoes and floors for interactive dance [21, 19] or other musical purposes [16, 11]. Such examples present higher latencies and lower sampling rates compared to our prototype (see Section 2.1). Moreover, those interfaces only act as controllers tracking the user's gestures, while they do not directly provide any feedback. A few notable exceptions out of a musical context are [23, 24], where foot-level haptic feedback is provided.

Various researches consider the use of haptic feedback in digital musical interfaces and instruments [5, 17, 15]. With regard to interfaces for percussion tasks, haptic feedback is exploited in e.g. [12] and [4].

2. INTERFACE DESIGN

This section describes the design of the interface from the hardware implementation to the software level.

2.1 Hardware

Starting from the top left of Fig. 1, a pair of sandals are equipped with four force sensing resistors (Interlink 402 FSR) fixed under the insole, one at the toe and one at the heel. The FSR sensors are connected to the analog inputs of an Arduino Duemilanove board (*force data transmitter*). Here the force signals are sampled and encapsulated using a custom protocol [7] and sent to a 2.4 GHz wireless transceiver module based on the nRF2401A chip by Nordic Semiconductor. A one-directional wireless line is realized by connecting a specular system: another nRF2401A module receives the data stream and routes it to an Arduino board (*force data receiver*). The latter is interfaced via a USB connection with a personal computer running Pure Data (Pd). Here the received data are processed to generate audio-tactile signals[1] to be sent to the sandals (see Section 2.2).

[1]As described in Section 2.2, while the system can be directly interfaced with MIDI and OSC compatible instru-

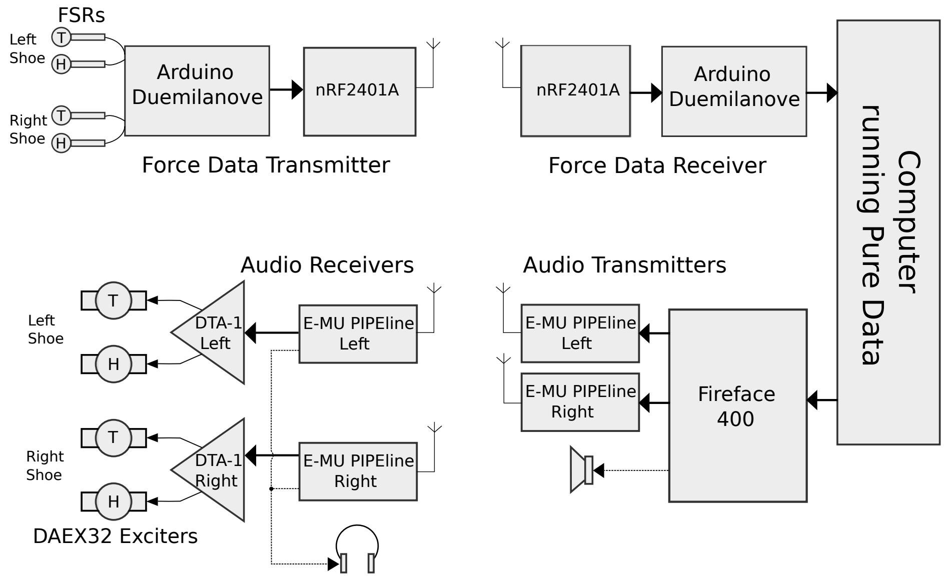

Figure 1: Block diagram representing the low-level hardware setup. The upper and lower flows illustrate respectively the sub-systems implementing force data acquisition and audio-tactile feedback.

Each sandal embeds two exciters that are driven by the outputs of a RME Fireface 400 multichannel audio interface. In detail, starting from the bottom right of Fig. 1, four output channels of the RME are grouped into two stereo pairs and each pair is routed to an E-MU PIPEline wireless audio transceiver. Each E-MU on the computer side (*audio transmitters*) is paired with another one on the user side (*audio receivers*), thus obtaining a four-audio-channel simplex wireless connection. Finally, the outputs of each E-MU receiver are injected into a Dayton Audio DTA-1 stereo amplifier which drives two Dayton Audio DAEX32 exciters fixed under each sandal – one at the toe and one at the heel – thus closing the interaction loop.

Since the setup is conceived for live performance, freedom of movement must be ensured, and therefore wireless communication represents an ideal choice. On the other hand the overall latency must be kept as low as possible, and the interface needs to be wearable (i.e. lightweight and small). The whole hardware system has been designed to satisfy such requirements, moreover using readily available components. At the user side, one Arduino with the attached nRF2401A module (*force data transmitter*), two E-MU units (*audio receivers*) and the two DTA-1 amplifiers are carried into a small backpack worn by the performer, together with standard batteries. Each sandal is then connected to the backpack via a single multi-conductor cable, this way minimizing encumbrance. The round-trip latency exhibited by the system – measured as the delay between the onset of an impulse at the FSR sensors and the arrival of the corresponding feedback signals to the exciters [7] – amounts to about 20 ms.

2.1.1 Details on data acquisition

Several solutions for musicians and performers have already been proposed which offer wireless acquisition of control signals[10, 22, 8], however most of them are based on custom hardware and/or are quite expensive. In our prototype, on the contrary, the wireless transmission of force data is managed by two readily available and low-cost transceivers based on the nRF2401A chip. Moreover our choice of developing a custom data protocol and send it over a dedicated wireless connection was necessary to avoid the latency and sampling rate drawbacks that other standard solutions (e.g. WiFi, Bluetooth or ZigBee) would have imposed.

Despite its low cost, the Arduino Duemilanove offers high-performance signal acquisition functionalities [7]. We have configured its microcontroller's ADC to uniformly sample up to six analog channels with a fairly high-rate and 10 bit resolution: the sampling frequency per single analog channel depends both on the serial data rate and number of channels [18]. With four channels, as in our case, the resulting frequency is 1050 Hz per single channel.

The latency introduced by the data acquisition system amounts to about 1.2 ms.

2.1.2 Details on feedback

The interface provides the performer with four-channel audio-tactile feedback:two E-MU PIPEline are used to send four audio signals introducing a delay of 5.5 ms (from official specifications).

The used exciters are meant to generate audio-rate vibrations, therefore abundantly covering the bandwidth required for haptic display [15].

2.2 Software

At the software level, three modules realized in Pd are organized in a bottom-up hierarchy: 1) at the first layer, the data stream generated by the FSR sensors is conditioned and analyzed in order to detect tapping events. As soon as one of such events is detected, this module outputs a measure of its energy; 2) the second layer maps the detected events alternatively to MIDI or OSC messages, or directly to the parameters of a sound synthesis engine running in Pd; 3) the third and last layer implements a physics-based impact sound model, which is driven by the detected tapping events. These three layers are described in detail below.

2.2.1 Data conditioning and analysis

The force data are received and unpacked, this way obtaining four separate streams respectively corresponding to the four FSR sensors (left/right heel and toe).

These streams are then conditioned and optimized in view

ments, the prototype already provides a synthesis engine implemented in Pd, this way offering a self-contained setup. For the sake of simplicity, in what follows we refer to the included synthesis engine.

of the following processing stage: the data is passed through threshold gates to filter out signal noise and avoid unwanted rebounds in the impact detection process.

The pre-conditioned force data are then analyzed in order to detect the onset of tapping events and measure their energy. To this end we made use of a Pd object called `bonk~` [20], which decomposes the incoming signal into frequency bands and computes the power in each of them, then it looks for sharp edges in the spectral envelope of the signal, enabling a very accurate detection of percussive events. As `bonk~` is meant to analyze audio signals, before being sent to it the pre-conditioned data streams are converted accordingly: they are oversampled from the original sample rate of 1470.5 Hz to Pd's internal audio rate by using the Pd object `sig~`. The resulting signals are then processed by a simple anti-aliasing filter. The output provided by `bonk~` consists in a measure of the energy of the detected event, calculated as the sum of the square roots of the amplitudes in each frequency band.

2.2.2 Mapping

The energy values of the detected tapping events are used to drive the control parameters of different instruments. In particular, a threefold path has been implemented, complying with three distinct protocols:

MIDI: the energy values are converted into integer values in the range 0-127 to comply with MIDI velocity values. As soon as a tapping event is detected, a "note on" message is generated and associated with such velocity value. A "note off" message is produced following each "note on" message, after a settable delay time. The "channel" and "note number" for each of the four data streams can be assigned to interact with any MIDI instrument, however the default configuration already offers a common drum setup according to the General MIDI standard.

OSC: since OSC-compatible instruments require custom messages, the user can modify the generated OSC messages to taste. For example using Pd's object `maxlib/scale` the original range of energy values (0-100) can be converted to any range of choice. The default configuration already offers predefined messages for communicating the onset of tapping events and their energy, while energy values are expressed as floating-point numbers from 0 to 100.

SDT: energy values are converted into physically-consistent velocity values expressed in m/s, that are sent to a physics-based impact sound model.

The three mappings described above can be selected alternatively via a switch implemented in Pd.

2.2.3 Sound synthesis

In order to provide a self-sufficient system, a sound synthesis engine was included in our prototype, this way transforming the interface into a complete instrument.

The sound synthesis engine makes use of a physics-based impact model [3] which is part of a library for Max and Pd called Sound Design Toolkit (SDT).[2] The model simulates a mass (object 1) colliding with a resonator (object 2), and the model's output represents the vibrations of the latter. Therefore the synthesized signals are particularly suitable to drive both audio and vibrotactile feedback.

In more detail, the contact between the two objects is accounted for by a nonlinear spring with dissipation, while the resonator is modeled according to the modal synthesis paradigm. The available control parameters give access to: the mass m (in Kg) of object 1; the resonating modes of object 2, namely their frequencies $f_{0..n}$ in Hz (where n is the number of modes), their decay times $t_{0..n}$ in s, and their gains $g_{0..n}$; the nonlinear spring, namely its nonlinearity exponent α and its stiffness k in Kg/N$^{\alpha}$. Such parameters enable the user to design sounds that simulate a wide variety of object's sizes and materials, like wood, plastic, metal and glass.

Each force data stream is mapped to a separate instance of the impact model, resulting in a different sound for each tapping position.

[2] Freely available at `http://www.soundobject.org/SDT`.

Figure 2: A performer wearing the interface, tapping the feet while sitting.

3. THE INTERFACE IN USE

The system has currently been calibrated for playing in a sitting position (see Fig. 2), which minimizes the detection of spurious tapping events. On the contrary, the calibration required for playing while standing up is obviously trickier, as the performer inevitably has to adjust his/her posture, e.g. to balance.

Thanks to Velcro straps, the sandals can easily fit a wide range of foot sizes, both bigger and smaller than their native European size 44 (corresponding to U.S. male size 10 1/2).

As shown in Fig. 1 the user can connect headphones and/or external loudspeakers to the interface, e.g for rehearsing purposes or for performing on stage.

Although the system is especially suited to play percussion instruments, it is not just limited to them. Indeed the availability of MIDI and OSC controls on the one hand allows to connect the interface to potentially any electronic or computer-based instrument, on the other hand it enables the implementation of complex mappings for supporting the experimentation of further musical styles and aesthetics.

Digital musical instruments usually lack the tactile feedback that is inherently conveyed by most traditional instruments. Such vibrations stimulate the mechanoreceptors in the skin [15]: in particular, the fingers are sensitive to vibrations up to 1000 Hz with a peak at about 250 Hz, and while it is generally acknowledged that the foot is less responsive than the hand, similar sensitivity figures are found for the foot sole [13]. Sensitivity thresholds also depend on the area of contact and the nature of the stimuli.

As explained in Section 2.1, the exciters embedded in the sole are driven by audio signals, therefore the resulting vibrotactile feedback ensures a tight coupling with the action of tapping. Informal evaluation done while testing the interface showed that such energetic consistency gives rise to a fairly convincing experience: in particular, by using the included physically-consistent impact model both the audio and tactile feedback improve on dynamics and realism.

Despite the fact that a maximum of 10 ms latency is generally suggested for music controllers [14], from informal evaluations we have found that our system is very responsive, and guarantees coherence and unity in the multimodal percept (see Section 2.1 for the measured latency figure). This is possibly partly due to the fact that the feet are not as sensitive as the hands, thus resulting in a higher tolerance to foot-level delays. As a result, the user is able to play with remarkable accuracy even fast paced and complex rhythms.

Also, tests showed that the implemented wireless communication is solid and reliable, independently of the performers' movements and within a range of about 15 meters.

Since the interface does not require any visual skill and provides vibrotactile feedback, it is perfectly suitable for both blind and hearing impaired persons. For example, as an alternative usage the hearing impaired could exploit the interface's feedback as a personal monitoring system, especially effective for feeling rhythmic parts or just the tempo.

4. CONCLUSIONS

The act of tapping the feet to play rhythms guarantees spontaneity and expressivity, while allowing the skilled performer to use more than one instrument (or other devices) at a time.

The "Rhythm'n'Shoes" interface is suitable for traditional musical genres where players make use of foot drums, as well as other types of performance where enhanced control over electronic (e.g. MIDI or OSC compatible) devices is required.

Even if proper testing is still needed, preliminary informal evaluation shows that the system exceeds by far the expressivity offered by simple trigger-based interfaces. Furthermore, thanks to the provided audio-tactile feedback, the interface offers a truly embodied interaction, even while playing an electronic instrument. Additionally, the use of physics-based sound models for generating both the audio and tactile feedback provides a consistent and realistic experience.

5. ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Community's Seventh Framework Programme under FET-Open grant agreement 222107 NIW - Natural Interactive Walking.

6. REFERENCES

[1] R. M. Aimi. New expressive percussion instruments. *M.S. Thesis. MIT*, 2002.

[2] M. S. Andresen, M. Bach, and K. R. Kristensen. The lapslapper: feel the beat. In *Proc. Int. Conf. on Haptic and Audio Interaction Design (HAID)*, pages 160–168, Berlin, Heidelberg, 2010. Springer-Verlag.

[3] F. Avanzini, M. Rath, and D. Rocchesso. Physically-based audio rendering of contact. In *Proc. IEEE Int. Conf. on Multimedia and Expo (ICME)*, volume 2, pages 445–448, 2002.

[4] E. Berdahl, W. Verplank, J. O. Smith III, and G. Niemeyer. A physically intuitive haptic drumstick. In *Int. Computer Music Conf. (ICMC)*, page 363–366, Copenhagen, Denmark, 2007.

[5] C. Chafe. Tactile Audio Feedback. In *Proc. Int. Computer Music Conf. (ICMC)*, Tokyo, Japan, 1993.

[6] K. Chuchacz, S. O'Modhrain, and R. Woods. Physical models and musical controllers: designing a novel electronic percussion instrument. In *Proc. Int. Conf. on New Interfaces for Musical Expression (NIME)*, pages 37–40. ACM, 2007.

[7] M. Civolani, F. Fontana, and S. Papetti. Efficient acquisition of force data in interactive shoe designs. In *Proc. Int. Workshop on Haptic and Audio Interaction Design (HAID)*, Copenhagen, Denmark, 2010. Springer.

[8] T. Coduys, C. Henry, and A. Cont. TOASTER and KROONDE: high-resolution and high-speed real-time sensor interfaces. In *Proc. Int. Conf. on New Interfaces for Musical Expression (NIME)*, 2004.

[9] M. Collicutt, C. Casciato, and M. M. Wanderley. From real to virtual: A comparison of input devices for percussion tasks. In *Proc. Int. Conf. on New Interfaces for Musical Expression (NIME)*, 2009.

[10] E. Fléty. The WiSe Box: a multi-performer wireless sensor interface using WiFi and OSC. In *Proc. Int. Conf. on New Interfaces for Musical Expression (NIME)*, 2005.

[11] J. A. Hockman, M. M. Wanderley, and I. Fujinaga. Real-time phase vocoder manipulation by runner's pace. In *Proc. Int. Conf. on New Interfaces for Musical Expression (NIME)*, 2009.

[12] S. Holland, A. J. Bouwer, M. Dalgelish, and T. M. Hurtig. Feeling the beat where it counts: fostering multi-limb rhythm skills with the haptic drum kit. In *Proc. Int. Conf. on Tangible, Embedded, and Embodied Interaction (TEI)*, pages 21–28, New York, NY, USA, 2010. ACM.

[13] J. Kekoni, H. Hämäläinen, J. Rautio, and T. Tukeva. Mechanical sensibility of the sole of the foot determined with vibratory stimuli of varying frequency. *Experimental Brain Research*, 78:419–424, 1989.

[14] T. Mäki-Patola and P. Hämäläinen. Effect of latency on playing accuracy of two continuous sound instruments without tactile feedback. In *Proc. Int. Conf. on Digital Audio Effects (DAFx)*, pages 11–16, Naples, Italy, October 2004.

[15] M. T. Marshall and M. M. Wanderley. Vibrotactile feedback in digital musical instruments. In *Proc. Int. Conf. on New Interfaces for Musical Expression (NIME)*, pages 226–229, Paris, France, 2006. IRCAM.

[16] B. Moens, L. van Noorden, and M. Leman. D-Jogger: Syncing Music with Walking. In *Proc. Sound and Music Computing conf. (SMC)*, volume online, pages 451–456, Barcelona, 2010. Universidad Pompeu Fabra.

[17] M. S. O'Modhrain. *Playing by feel: incorporating haptic feedback into computer-based musical instruments*. PhD thesis, Stanford, CA, USA, 2001.

[18] S. Papetti, F. Fontana, M. Civolani, A. Berrezag, and V. Hayward. Audio-tactile display of ground properties using interactive shoes. In *Proc. Int. Workshop on Haptic and Audio Interaction Design (HAID)*, Copenhagen, Denmark, 2010. Springer.

[19] J. Paradiso, K. yuh Hsiao, and E. Hu. Interactive music for instrumented dancing shoes. In *Proc. Int. Computer Music Conf. (ICMC)*, pages 453–456, 1999.

[20] M. Puckette, T. Apel, and D. Zicarelli. Real-time audio analysis tools for Pd and MSP. In *Proc. Int. Computer Music Conf. (ICMC)*, pages 109–112, Ann Arbor, Michigan, USA, 1998.

[21] P. Srinivasan, D. Birchfield, G. Qian, and A. Kidané. A pressure sensing floor for interactive media applications. In *Proc. 2005 ACM SIGCHI Int. Conf. on Advances in Computer Entertainment Technology*, ACE, pages 278–281, New York, NY, USA, 2005. ACM.

[22] D. Topper and P. V. Swendsen. Wireless dance control: PAIR and WISEAR. In *Proc. Int. Conf. on New Interfaces for Musical Expression (NIME)*, 2005.

[23] Y. Visell, A. Law, and J. R. Cooperstock. Touch is everywhere: floor surfaces as ambient haptic displays. *IEEE Transactions on Haptics*, 2009.

[24] J. Watanabe, H. Ando, and T. Maeda. Shoe-shaped interface for inducing a walking cycle. In *Proc. Int. Conf. on Augmented Tele-existence*, ICAT, pages 30–34, New York, NY, USA, 2005. ACM.

[25] D. Young and I. Fujinaga. AoBachi: A new interface for Japanese drumming. In *Proc. Int. Conf. on New Interfaces for Musical Expression (NIME)*, 2004.