Interaction Design: The Mobile Percussionist

Tiago Reis, Luís Carriço, and Carlos Duarte

LaSIGE, Faculdade de Ciências, Universidade de Lisboa
treis@lasige.di.fc.ul.pt, {lmc,cad}@di.fc.ul.pt

Abstract. This paper presents the user centered iterative interaction design of a mobile music application. The application enables multiple users to use one or more accelerometers in order to simulate the interaction with real percussion instruments (drums, congas, and maracas). The ways through which the accelerometers are held, before and during interaction, define the instruments they represent, allowing the swapping of instruments during musical performances. The early evaluation sessions directed to the interaction modes created for each instrument enabled design iterations that were of utmost importance regarding the final application's ease of use and similarity to reality. The final evaluation of the application involved 4 percussionists that considered it well conceived, similar to the real instruments, natural and suitable for entertainment purposes, but not for professional musical purposes.

Keywords: Audio Interaction, Context Awareness, Mobile Interaction, Accelerometer-Based Gesture Recognition.

1 Introduction

Mobile music is a new field concerned with musical interaction in mobile settings through the utilization of portable technology [1], such as laptops, tablet PCs, mobile phones, and PDAs, amongst others. These devices have long been recognized as having potential for musical expression [2]. Such potential enables entertainment and artistic applications to leave the stationary living-room set, detaching themselves from the graphical and audio output of televisions (e.g. Wii), therefore, becoming available anywhere, allowing and encouraging multiple users to engage in spontaneous jam-sessions.

On another strand, the availability of an immense variety of sensors, their low prices, and their recent and rising inclusion on mobile devices enables the creation of mobile sensor-based musical applications. As we show in this paper, and as others have shown in previous experiments and products [4] (e.g. Wii), the use of sensor technology brings simulated musical interaction closer to reality [6]. It enables and facilitates the creation of input modalities that replicate musical interaction with different, well-known, real musical instruments (or families of musical instruments).

The research presented on this paper focuses the user centered iterative design of a multi-user mobile music application: The Mobile Percussionist. This application uses movement and grasp, both inferred from acceleration, as input, triggering percussion audio samples as output. It simulates three well-known percussion instruments, which

M.E. Altinsoy, U. Jekosch, and S. Brewster (Eds.): HAID 2009, LNCS 5763, pp. 109–118, 2009.

are played in different ways: drums (drumstick strikes), congas (hand strikes), and maracas (shaking movements); attempting to imitate, as much as possible, the interaction with these instruments' real counterparts. The application makes use of one or more accelerometers, wirelessly connected to a mobile device. These accelerometers are used in order to simulate the interaction with real percussion instruments, and the way through which they are held defines the instruments they represent, introducing context awareness principles in order to enable the intelligent selection and swapping of the different percussion instruments that are simulated (drums, congas, and maracas). The differences between how these instruments are played presented a motivating challenge for the interaction design of this application regarding both the selection of the instrument simulated by each accelerometer and the interaction modes created for each instrument. As it was being developed, the application was evaluated through strongly user-centered procedures, presenting very good results for entertainment purposes.

The following section presents the work developed in the mobile music field, emphasizing the differences with the research presented on this paper. Afterwards, we present and discuss the iterative user centered interaction design of the Mobile Percussionist. Subsequently, we describe and discuss the final evaluation of the different interaction modes available on the application, and, finally, we conclude and present future work directions for the Mobile Percussionist.

2 Related Work

The field of mobile music is developing fast over the past few years. A clear fact supporting this affirmation is the recent creation of the Mobile Phone Orchestra (MoPhO) [3] of the Center for Computer Research in Music and Acoustics of Stanford's University. MoPhO is the first repertoire and ensemble based mobile phone performance group of its kind and has already performed concerts. The ensemble demonstrates that mobile phone orchestras are interesting technological and artistic platforms for electronic music composition and performance.

The creation of the abovementioned orchestra was only possible because researchers have developed different mobile music applications, some of which are becoming very famous and broadly used. One good example is the Smule OCARINA [4], which is sensitive to blow, touch and movements. This application does not use pre-compiled riffs and allows users to select between diatonic, minor and harmonic scales. Moreover, it is a social application. Users can see and hear other Ocarina players throughout the world and rate their favorite performances, enabling other users to benefit from their judgments.

CaMus [5] - collaborative music performance with mobile camera phones - demonstrates that mobile phones can be used as an actively oriented handheld musical performance device. To achieve this, the visual tracking system of a camera phone is used. Motion in the plane, relative to movable targets, rotation and distance to the plane can be used to drive MIDI enabled sound generation software or hardware.

Closer to our goals, ShaMus [6] - a sensor-based integrated mobile phone instrument - presents a sensor-based approach to turn mobile devices into musical instruments. The idea is to have a mobile phone be an independent instrument. The sensors used are accelerometers and magnetometers. The sound generation is embedded on the phone itself and allows individual gestural performance through striking, shaking and sweeping gestures. The striking gestures presented enable users to play at most 4 drum components (base, snare, tom, and hi-hat). To prove this concept a Nokia 5500 was used, connected to a Shake unit [7], which is a small device that incorporates a range of high-fidelity sensors for rapid prototyping of mobile interactions. The Shake's core unit contains a 3-axis accelerometer, a 3-axis magnetometer, a vibration motor for vibro-tactile display, a navigation switch, and capacitive sensing abilities.

The Mobile Percussionist is significantly different from ShaMus. We do not limit the number of accelerometers used. This introduces the possibility of having multiple users playing music at the same time, while using only one mobile device with sound processing capabilities. Moreover, our approach enables users to perform natural actions that simulate the interaction with different percussion instruments, allowing them to swap instruments while playing. Finally, the interaction modes of the Mobile Percussionist were iteratively designed through strongly user centered procedures where accuracy rates for the different actions were registered. These procedures also considered users' opinions regarding the actions' ease of use and naturalness when compared to real percussion instruments.

3 Iterative User Centered Interaction Design

This section describes the user centered interaction design of the Mobile Percussionist. The process focused the creation and evaluation of different interaction modes for each one of the three instruments considered, as well as the creation and evaluation of a mechanism that enables the application to be continuously aware of which instruments are being simulated by each accelerometer in use.

Accordingly, the process comprised the development and evaluation of different prototypes. The prototypes created were developed using NetBeans IDE 6.1, Java 1.6.0_07, two SunSPOTS [8], one laptop, and a set of stereo speakers. SunSPOTS (Fig. 1) are small devices with processing capabilities. These devices include two buttons and a variety of sensors: one 3 axis accelerometer, one light sensor, and one temperature sensor. Furthermore, they can also work as a platform where other sensors can be connected. The mentioned accelerometer has a sensitivity of either 600 mV/G or 200 mV/G relative to the selected scale +/-2G or +/- 6G, respectively [8]. In our experiments we used the 6G scale.

The early evaluation sessions, which guided the iterative user centered design of this application, focused the different prototypes created and considered all the actions supported by each interaction mode available in the application. The main goal was to understand the accuracy rates for each action and users' opinions about the similarities between the interaction modes created and the interaction with real instruments. Four users were involved, all male, with ages comprehended between 20 and 30 years old. Two of these users had previous experience with percussion instruments while the other two had experience with simulated percussion instruments (e.g. Wii Drums).

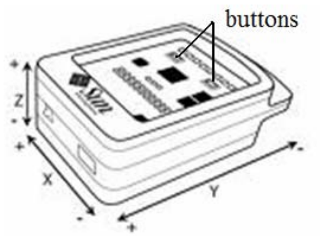

Fig. 1. SunSPOT and Accelerometer Axis

Following we describe the iterative user centered design of the interaction modes developed for each instrument considered, and afterwards, the design of the instrument selection and swapping algorithm.

3.1 Drums Interaction

Initially, in order to create an interaction mode that simulates real drums interaction, the design team focused only on accelerometer data. We defined the way the devices should be held, attempting to simulate drumsticks' grasp (notice that the leds are facing up in Fig. 2 and Fig 5). Following, it was necessary to accurately identify natural gestures on the accelerometer data. Three actions were defined, visually identified (inverted acceleration peaks), and labeled (Fig. 2): X-Strikes (fast movements performed sideways); Y-Strikes (fast movements performed forwards and backwards); and Z-Strikes (fast movements performed up and down).

In order to programmatically identify these actions, a simple algorithm was implemented and evaluated in terms of accuracy. The actions were evaluated in terms of ease of use and similarity to reality. Each user performed each available action 20 times, in a random order with the purpose of reducing bias. Average accuracy rates

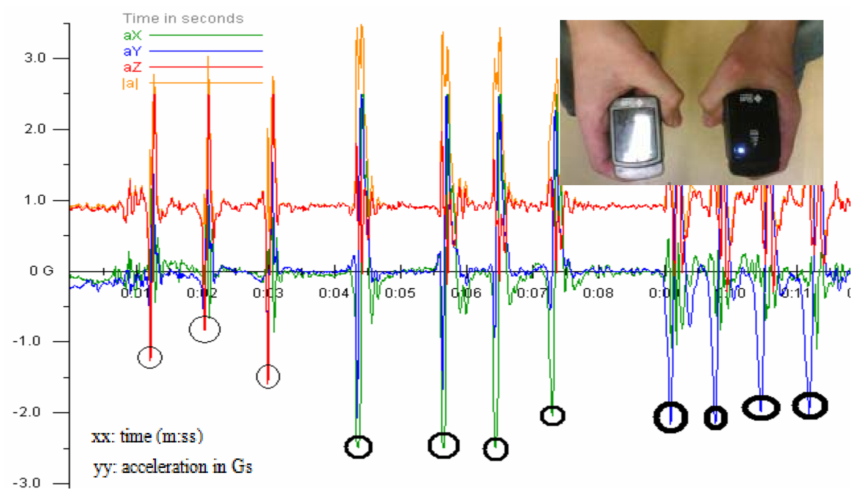

Fig. 2. Accelerometer Data: Strike Gestures (from the left to the right: Z, X, Y)

were of 90% for X-Strikes, 85% for Y-Strikes, and 95% for Z-Strikes. All users reported that performing Y-Strikes was very unnatural, and, therefore, more difficult than the remainder actions. Moreover, they all emphasized that Z-Strikes were the most natural ones regarding the interaction with drums. This early evaluation procedure was of utmost importance, since it revealed that only 2 of the 3 gestures were considered natural enough and that one of those was more natural than the other.

Afterwards, our team faced another design issue. Drums are composed by many components (e.g. snares, toms, crash, hi-hat, etc.), and when users are interacting with them they can "jump" directly from one component to another. Accordingly, the two gestures, which were previously considered natural enough for the interaction with drums, were not sufficient to simulate such a complex interaction. Therefore, in order to increase the number of samples available on each accelerometer, we considered the use of the 2 buttons available on each SunSPOT. The use of such buttons enables the definition of 4 different states. When these states are combined with the 2 previously elected strike actions, 8 different samples become available on each accelerometer. However, two questions still remained, whether to consider both strike actions or only one, and whether to consider combinations of buttons or not? In order to answer these questions, two prototypes were created and evaluated. Both prototypes considered the use of the 2 buttons available on the SunSPOTS. These buttons were used in order to select the audio samples, and the strikes to trigger those samples. The difference between the prototypes relied on the triggering of the samples. One considered only Z-Strikes and the other considered both Z and X-Strikes.

The same four users evaluated each prototype created. All of them reported that the use of Z-Strikes created an interaction mode that was more similar to reality. However, all the users agreed that the use of X-Strikes, despite being less natural, should also be considered, since it doubles the number of samples that can be triggered by each accelerometer. Finally, all users reported difficulties while having to press more than one button at a time, suggesting that the use of button combinations in order to increase the number of samples available is not a good approach.

Accordingly, the drums interaction mode available on the Mobile Percussionist considers the use of all the buttons (per se) available on the Sunspots, Z-Strikes, and X-Strikes, allowing the use of 6 audio samples per accelerometer.

3.2 Congas Interaction

The design of the congas interaction mode started by considering only accelerometer data. Similarly to the interaction design presented above, the main goal was to create an interaction mode that simulated as much as possible the interaction with real congas. We started by defining the way the devices should be held, in order for the application to be aware that the user is playing congas (notice that the leds are facing down in Fig. 3 and Fig. 5).

Following, we decided to consider 2 different actions that are used when playing real congas: center strikes (Fig. 3 on the left) and edge strikes (Fig. 3 on the right). Again, these actions were visually identified, and, afterwards, programmatically identified. Their distinction is inferred from the average acceleration on the Y-axis (○). When holding the device in the two different ways shown in Fig. 3 ways, the average acceleration on the Y-axis varies significantly enough to enable the application to be

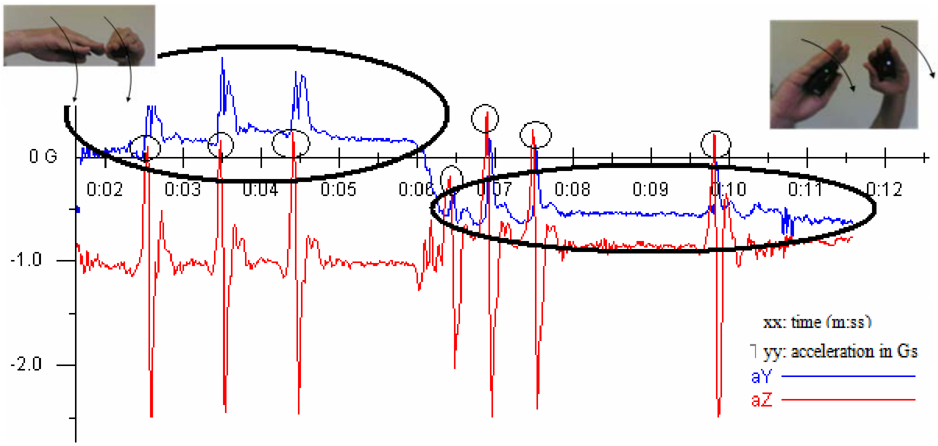

Fig. 3. Playing Congas

aware of users' intentions regarding the type of strike they are performing: center strikes consider a positive average acceleration, while edge strikes consider a negative average acceleration. Finally, Z-Strikes trigger the samples. Nonetheless, contrarily to the drums interaction mode, and since the way the accelerometers are being held is different, the Z-Strikes are represented by acceleration's peaks (○) instead of inverted peaks.

The users mentioned on the introduction of this section were all involved on the evaluation of this interaction mode. Again, each user performed each action 20 times, in a random order with the purpose of reducing bias. The procedure revealed average accuracy rates of 90% for both center strikes and edge strikes. Users reported that this interaction mode was very similar to the real interaction with congas. However, they disliked the limitation of using just one conga, since, usually, percussionists use more than one.

Similarly to what was done with the drums interaction, we decided to use the buttons available on the SunSPOTS, in order to increase the number of congas available. Two prototypes were created and evaluated by the same users. The difference between these prototypes relied on the way the buttons are used. One prototype considered holding the buttons pressed as state activators, while the other would activate and deactivate states when the users pushed the buttons.

The results were very similar for both prototypes. Users were more satisfied with the increased number of samples available, but all of them reported that interaction became a little less natural than before. Moreover, all the users agreed that holding the buttons pressed implied less cognitive effort than pushing the buttons in order to activate states, since the latter demands the memorization of the actual state in use.

Accordingly, the congas interaction mode available on the Mobile Percussionist considers the use of all the buttons (held pressed and per se) available on the SunSPOTS, Center-Strikes, and Edge-Strikes, allowing the use of 6 audio samples per accelerometer (3 different congas).

3.3 Maracas Interaction

The design of the maracas interaction mode considered only accelerometer data. As for the previous instruments, the main objective was to simulate reality as much as possible. We started by defining the way the devices should be held, in order for the application to be aware that the user is playing maracas (notice that the device is in a vertical position in Fig. 4 and Fig. 5).

Accordingly, two audio samples can be triggered by each accelerometer. One relates to shaking the maraca and the other transmits the sound of stopping the shaking. This interaction mode joins principles of the two abovementioned designs. Both peaks (○) and inverted peaks (**O**) are considered, the first triggering one audio sample and the latter the other.

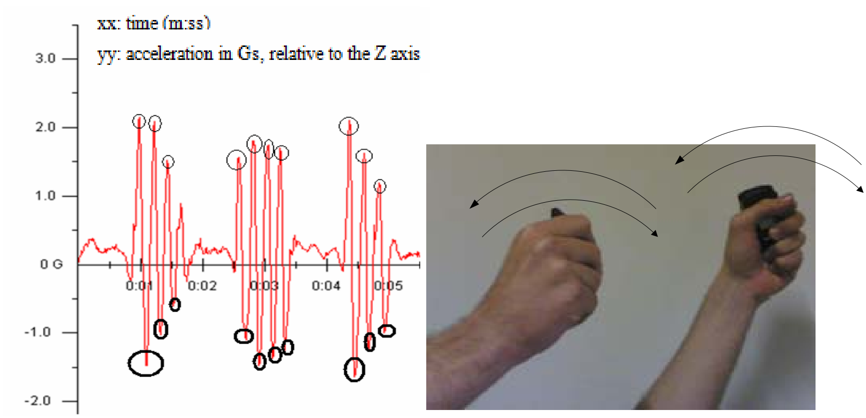

Fig. 4. Playing Maracas

The four users, mentioned on the introduction of this section, performed each available action 20 times. Average accuracy rates were of 95% for both actions. All users reported that this interaction mode was very similar to the real interaction with maracas and, therefore, very natural.

3.4 Instrument Selection and Swapping

Once the application is initiated, the users select the instrument they want to associate to each accelerometer in use. This selection (Fig. 5) is based on the way the devices are held, which is inferred from the acceleration on the Z-axis, and activates/deactivates the different interaction modes available. While playing an instrument, the acceleration on the Z-axis varies considerably. However, the average-based analysis behind the instrument selection and swapping mechanism enables the identification of the instrument being played also during interaction, allowing users to swap between instruments, depending on the way they hold the accelerometers.

The evaluation procedure considered the initial selection of the instruments and the swapping between instruments while playing. Again the order of the experiments was

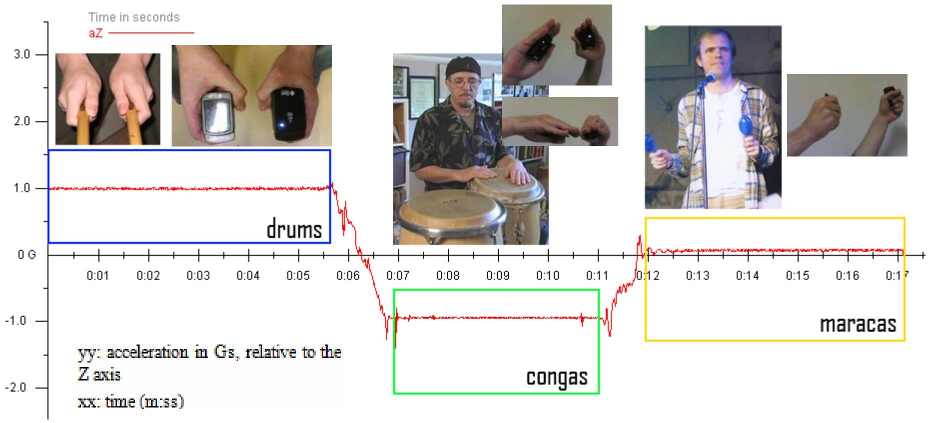

Fig. 5. Initial Step: Instrument Selection

defined in order to reduce bias. Firstly, each user involved performed 5 initial selections for each instrument, averaging 100% of accuracy. Afterwards, while playing the instruments, each user performed 5 swaps for each of the 6 possible transitions between instruments. Here, the average accuracy was also of 100%. However, the response time on the swapping between instruments caused the triggering of one unwanted audio sample on 30% of the swaps (e.g. playing a drum audio sample while the user was already holding the device as a maraca). Nevertheless, these unwanted samples were only triggered on the beginning of the experiment, suggesting a very short learning curve.

3.5 Discussion

The primary goal of this design process was to create natural interaction modes that simulate the interaction with real percussion instruments, allowing the swapping of instruments during musical performances. This goal could only be achieved through the iterations of the user centered design process conducted.

The first prototypes designed considered only accelerometer data, which was enough to enable the natural swapping of instruments. However, the experiments abovementioned, especially the ones regarding the interaction with drums, led us to use the buttons available on the devices, in order to increase the number of audio samples triggered by natural gestures. These gestures were elected by the users involved and all the possible combinations of buttons were experimented. The use of more than one button at a time decreased interaction's naturalness significantly, although the use of the each button per se did not. The use of the buttons was also included on the interaction with congas, but not on the interaction with maracas.

At the end of the process the users involved were satisfied with the different interaction modes created, their ease of use, and with the possibility of swapping instruments while playing.

4 Final Interaction Evaluation

This section is dedicated to the final qualitative evaluation of the Mobile Percussionist. The procedure involved four percussionists, none of them involved on the previous evaluation sessions, all male, with ages comprehended between 18 and 36 years old. There was a special concern in understanding if the users were capable of successfully playing rhythms with the application and make use of all the different instruments that are available. Accordingly, the users involved were taught to interact with the application through its different features, and were encouraged to use it freely to its full extent during approximately 20 minutes. Finally, the users expressed their difficulties and opinions during an interview.

All the users involved considered the concept of the application very interesting and useful for entertainment purposes, but not for professional musical purposes. They underlined the possibility of swapping between instruments while playing as a very innovative feature, emphasizing also the similarities between the interaction modes of the simulated instruments and their real counterparts as a very intelligent, usable, and natural feature, especially the ways the accelerometers are held.

Concerning the interaction with drums, all the users considered that six audio samples per accelerometer was a small number to simulate a drum set to its full extent. The same issue was reported regarding the interaction with congas, however, in this case, only two of the users considered three congas per accelerometer as a small number. Regarding the interaction with maracas, all the users considered that the number of samples was appropriate, however, all of them expressed that the buttons could be used in the same way they are for the congas and drums, enabling the use of three types of maracas instead of one.

Another issue, reported by all the users involved, consisted on the lack of expressiveness of all the interaction modes designed. This happened because our application makes use of accelerometer data only in order to trigger the samples used and not to define the strength with which these are performed (velocity). However, this data could also be used for this purpose.

Finally, all the users reported minor issues related to their own expressiveness while playing. All these users were percussionists, and, consequently, they had previous experience playing the real instruments our application tries to simulate. Accordingly, these users were experiencing difficulties during the first two/three minutes of their free play because their expressiveness caused an unwanted instrument swapping. However, after realizing this, these users voluntarily reduced their expressiveness being able to play fairly elaborated rhythms composed by the three available instruments.

5 Conclusion and Future Work

The research presented on this paper focused the iterative user centered interaction design of a mobile accelerometer-based percussion application. This application uses movement and buttons as input mechanisms, triggering percussion audio samples as output. It introduces context awareness principles that enable the swapping between the different percussion instruments available (drums, congas, and maracas) during musical performances. The different interaction modes available were designed and developed through iterative user centered procedures that significantly increased their similarities to reality. All the interaction aspects regarding the use and swapping of instruments were evaluated by four percussionists that considered that the application was innovative and simulated reality well enough to be used for entertainment purposes, but not for professional musical purposes.

Our future work plans, regarding the Mobile Percussionist, include porting the application to mobile phones, increasing the application's mobility factor and enabling the use of built-in accelerometers whenever those exist on the host device. However, we want to keep the possibility of connecting multiple accelerometers via wireless technology (e.g. SunSPOTS through Bluetooth), maintaining the multi-user facet of the application. Once this port is complete we intend to evaluate the application in order to understand discrepancies between its performance on a laptop and on a mobile phone. Moreover, we intend to test the application, both in the laptop and mobile phone versions, in order to understand how many accelerometers can be connected at the same time without generating or increasing latency significantly enough for it to be heard. We intend to use MIDI synthesizers in order to increase the expressiveness of the interaction modes available. Finally, we envision the inclusion of haptic feedback through the connection of vibrating devices to the SunSPOTS on the laptop version of the application, and through the use of mobile phones' vibrators on the ported version of the application.

Acknowledgments. This work was supported by LaSIGE and FCT through the Multiannual Funding Programme and individual scholarships SFRH/BD/44433/2008.

References

1. Gaye, L., Holmquist, L.E., Behrendt, F., Tanaka, A.: Mobile Music Technology: Report on an Emerging Field. In: Report - NIME 2006, Paris, France (2006)
2. Essl, G., Wang, G., Rohs, M.: Developments and Challenges Turning Mobile Phones into Generic Music Performance Platforms. In: Proceedings of Mobile Music Workshop, Vienna (2008)
3. Wang, G., Essl, G., Pentinnen, H.: MoPhO: Do Mobile Phones Dreams of Electric Orchestras? In: Proceedings of the International Computer Music Conference, Belfast (2008)
4. Smule OCARINA, `http://ocarina.smule.com/`
5. Rohs, M., Essl, G., Roth, M.: CaMus: Live Music Performance using Camera Phones and Visual Grid Tracking. In: Proceedings of the International Conference on New Interfaces for Musical Expression (2006)
6. Essl, G., Rohs, M.: ShaMus - A Sensor-Based Integrated Mobile Phone Instrument. In: Proceedings of the International Computer Music Conference (ICMC), Copenhagen, Denmark, August 27-31 (2007)
7. Hughes, S.: Shake – Sensing Hardware Accessory for Kinaesthetic Expression Model SK6. In: SAMH Engineering Services, Blackrock, Ireland (2006)
8. Sun Sunspots, `http://www.sunspotworld.com`