A Methodology to Introduce Gesture-Based Interaction into Existing Consumer Product

Lorenzo Cavalieri(✉), Maura Mengoni, Silvia Ceccacci, and Michele Germani

Department of Industrial Engineering and Mathematical Sciences,
Università Politecnica delle Marche, Via Brecce Bianche, 12-60131 Ancona, Italy
{lorenzo.cavalieri,m.mengoni,s.ceccacci,m.germani}@univpm.it

Abstract. The continuous progress of interaction technologies reveals that we are witnessing a revolution that is leading to a redefinition of the concept of "user interface" and to the development of new ways to interact with the electronic devices of all sizes and capabilities. Current trends in research related to the Human-Machine Interaction (HMI) show a considerable interest toward gesture, motion-based and full-body based interactions. In this context, a User-Centered Design (UCD) methodology to implement these novel interaction paradigms into consumer products is proposed with the aim to improve its usability, intuitiveness and experience. A case study is used to validate the methodology and measure the achieved improvements in user performance.

Keywords: Gesture interaction · Design methods · User interfaces · User-Centered design

1 Introduction

The potentiality of Human-Machine Interfaces is progressively increasing to allow users to interact with electronic devices of any dimension and capability with a low cost impact and with an improved usability. Investigations on the effectiveness of natural, brain, gesture-based interfaces and so on and on the effects on the user experience appear to be really significant to drive the design and development of remote controls, electronic devices, game consoles, etc.

The paper focuses on a specific typology of electronic devices that are musical keyboards, where touchless interaction could play a key role to improve learning and use [1].

On the marketplace, there are some examples of musical applications exploiting gesture-based interfaces. One is the MiMu Gloves [2, 3] equipped with a motion tracker, an haptic motor, bend sensors, LEDs and x-OSC. Authors developed a robust posture vocabulary, an artificial neural network-based posture identification module and a state-based system to map the identified postures onto a set of performance processes. Another proposed technology is Titan Reality Pulse, that is a MIDI controller that allows the user to trigger and play software instruments using hands. It is pressure-sensitive and can recognize the number of fingers you're holding up to it. The device combines gesture interaction with the classic touch controller in a single object [4].

M. Kurosu (Ed.): HCI 2016, Part II, LNCS 9732, pp. 25–36, 2016.
DOI: 10.1007/978-3-319-39516-6_3

The above-mentioned applications exploit general-purpose gesture-based applications such as Kinect and Leap Motion that are useful tools for programming music controls. However the achieved results' performance is often left to the capability of developers and to the user's skill. Moreover, it is worth to notice that literature lacks of a design methodology to guarantee a robust and efficient interaction to control music parameters in a live performance or in a recording studio.

These are the main motivations that trigger out the present research work whose aim is both to define a User-Centered methodology to implement the gesture-based interaction paradigm into musical devices and to develop a usable, intuitive and friendly technology for non-expert performers. The paper describes the adopted approach, gives an overview of the implementation results and illustrates some preliminary experimental achievements.

2 Research Background

The main issues in gesture-based interaction in music are as follows: (1) interaction modalities, (2) recognition patterns of interaction, (3) enabling technologies and (4) the role of the music in the human conditions.

In the music world the main adopted interaction mode is touch. However, in the field of electronic music many touchless solutions have been implemented to allow the performer to play sounds.

Several research works are reported in literature. For instance some technologies exploit smartphones with an embedded digital compass [5, 6]. The touchless interaction is implemented by changing the magnetic field around the smartphone by moving a handle tiny magnet. Other works use infrared signals to enable the interaction [7]. Marrin [8] implemented the so-called *Digital Baton*, to allow a real-time gestural control. The hardware system consists of a baton, an external infrared sensor, a tracking unit, and a computer. The sensors on the baton include an infrared LED for positional tracking, five piezoresistive strips for finger and palm pressure, and three orthogonal accelerometers for beat tracking.

In the presented works, some opened issues still remain. They mainly regard the creation of an accurate pattern recognition model and a structured and repetitive method to map the recognized gestures with the functionalities to be enabled.

About the first issue, one aspect deserves the present research interest that is the design of proper recognition models to map the performed interaction with sound.

The major problems actually appear when the mode of interaction is based on gestures; in this case a precise detection system and an accurate process of mapping are required to obtain a usable system [9–11].

First of all, the most important aspect is the definition of gesture meanings in music. Godøy [12] explored the meanings of gestures in the different fields of the music (e.g. instrument-related gestures or director ones) and defined accurate models to recognize and map them with the proper musical event. Ng [13] presents a novel framework to map natural human languages (i.e. gestures) with multimedia event (i.e. audio output).

Gesture recognition and modeling are as accurate and repetitive as the movements are accurately captured. For that purpose, Vigliensoni [14] compared different three-dimensional (3D) position tracking systems (i.e. Vicon, Polhemus, Kinect, and Gametrak) in terms of static accuracy and precision, update rate, and shape of the space they sense. Vicon resulted to be the best one for gesture-based interaction in music. Other works are focus on the recognition of the gesture of a choir/orchestra director for an interactive music system [15–17]. A baseline study is the Morita's work [15]. The proposed gesture recognition system is based on a sensorized baton and a data glove to detect user intentions about the management of musical features (time pattern, velocity, instrument activation). Then an artificial intelligence supports the manipulation of MIDI data in order to properly reproduce the music. The work proposed by Morita represents a key reference for the present research, that is a step forward it as it overcomes the problems of data glove invasiveness and barriers of the baton by integrating an optical tracking system.

In another significant example, Je tries to understanding of four musical time patterns and three tempos that are generated by a human conductor of robot orchestra or an operator of computer-based music play system using the hand gesture recognition [16]. Je's work allows to have a valid reference point to the pattern recognition about tempo performance and its sound modulation due to the gesture decoding.

Finally, Marrin Nakra [17] presented the design and implementation of the Conductor's Jacket. The proposed wearable device is able to measure physiological and gestural signals to analyze, understand and synthesize expressive gesture in a musical context. The identified recognition patterns are exploited to develop a musical software system that analyzes and performs music in real-time based on the performer's gestures and breathing signals.

The second issue regards human conditions in relationship with music. Some works focused on the influence that musical parameters have on human perception [18], others on the monitoring of the brain activity while music plays [19, 20]. However, a more important line of research concerns the study of eye movement of musicians to understand their brain activity [21–23].

Therefore, inspired by works on interaction design in music, a methodology to support the user interface redesign once gesture-based interaction has been introduced is here developed.

3 A Methodology to Improve HMI Design

The integration of the above-described interaction paradigms into existing musical keyboards and electronic devices in general requires an adaptive design process.

It is divided into three stages as follows. In the first one, an *analysis of existing products* is conducted to identify the main usability issues, the usual user behavior, and the enabling technologies suitable to be integrated to the product itself. The second stage regards the *identification of possible design solutions* by starting from the identified critical issues and the technologies to be integrated in. Finally, the third stage consists in the *implementation and evaluation of the elaborated solutions* to fulfill user needs and expectations. Every stage generates functional outputs representing the input data to the next one.

Figure 1 shows the workflow of the Design Methodology to improve the HMI by adapting it to new interaction paradigms. It is called *M.I.I.D.*

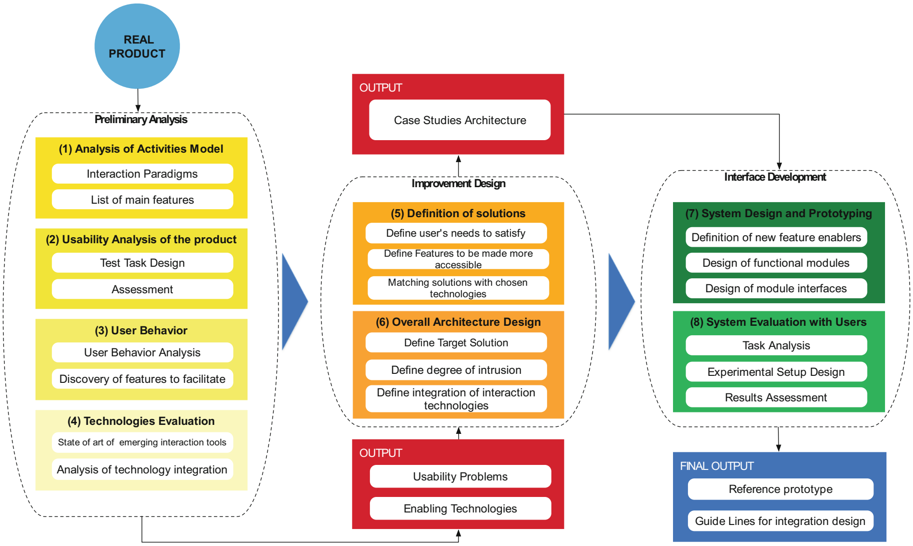


Fig. 1. *MIID Flowchart.*

3.1 Stage 1: Preliminary Analysis

The preliminary analysis represents the first stage of MIID; in this phase an analysis of the key-aspects of the product use scenarios is carried out. It takes care of four main aspects as follows: interaction modalities, performed tasks, user behavior and available technologies to support interaction.

The first aspect is studied by the construction of **an Activities Model** that tries to understand the nature of interactions and the functionalities (activities) activated by the user. It seeks to highlight the main interaction modalities and to define a list of the main features of the product. The *interaction modalities* can be found by directly analyzing the physical interface of the product that means the identification of buttons, touch screens, knobs, infrared sensors, joysticks. The *definition of main product features* is more complex because it also takes into consideration the elements influencing the affordance of the product in relation to the user needs.

The critical issues emerging in user-machine interaction are identified thanks to the **Usability Analysis of the Product.** A set of test tasks is defined to represent the overall range of choices the user has. A predictive evaluation according [24] is adopted to assess the usability performance, such as the number of steps to achieve the goal and the task execution time.

At the same time a **User Behavior Analysis** is carried out. It allows analysts to understand how consumers use the product interface focusing on the identification of primary and secondary behaviors. Direct user observations during usability tests and ethnographic techniques are used for that purpose.

In this case the analysis is aimed at identifying and decoupling a *primary behavior*, bound to actions the user makes to interact with the product and its functionalities, and a *secondary behavior*, made of all those gestures that are not directly connected with the product use but that have to be considered.

Two findings come from decoupling: (1) the strengths of the interface identified by possible uniform behavior of different users, but also (2) the weaknesses identified by an uneven behavior. This allows the identification of which modes of interaction to maintain, which require corrections, and which new types of interactions without interfering with the main ones improve the interface naturalness and intuitiveness.

Finally, **Technologies Evaluation** is used to investigate the tools improving interaction. It passes through the overview of the state-of-art of emerging HMI that can be applied on the target product category.

3.2 Stage 2: Improvement Design

Improvement Design aims to upgrade interaction and solve current interface failures and drawbacks. It follows the well-known adaptive design process.

The first task to be fulfilled is the **Definition of Solutions** at a conceptual and abstract level of detail. The complexity of this stage is directly proportional to the identified usability problems to be solved and the needs to be satisfied. The decomposition of the overall design problem into sub-problems can be useful to conceive targeted solutions and make the process more efficient. After problem framing, a plan of goals to be achieved is defined. This plan contains at first the functionalities not satisfied by the interface and a hypothesis of additional features and/or redesigned ones and the set of design requirements to start with the conceptual and detail design of each solution.

The **Overall Architecture Design** starts with the identification of each solving technology and the iterative evaluation of the changes propagation across the whole product structure in order to minimize the effects of modifications in terms of economic impact and technical feasibility (e.g. manufacturability, sustainability, etc.). Once conceived each solution, the evaluation of all technologies' integration is performed. The result of this stage is represented by the formalization of one or more case studies architecture, with the description of the functional parts of the solution and their own interconnections useful to reach the goal of the improvement design and to satisfy the user's needs.

3.3 Interface Development

The final stage of the MIID Methodology consists of (1) system development and (2) evaluation with end-users to test the reliability and performance of the improved interface solutions. It represents the implementation and validation phases of the overall redesign process: all collected data and taken design decisions are developed in this stage and tested in order to verify the achievement of initial objectives.

System development consists of **System Design and Prototyping**. The first aims at develop the conceived architecture and each constituting modules thanks to the adoption of a Systematic Engineering approach. Each module is characterized by its core functionality, all the attributes and the connections with other modules.

Enabling features and integrating technologies allow the designer to define the requirements for the design of each module. However, it is possible that the product under study does not directly support the selected interface features. In this case it is necessary to choose a platform compatible with the product. This represents a further requirement to add to the design requirements' list.

Each module is able to implement both high-level functionalities (exposed to the user) and low-level ones (exposed to the product). It also has to provide interfaces with other modules to enable data communication and interchange while keeping them independent each other.

Prototyping aims at implement the so-defined modular architecture by exploiting the proper platform and the conceived technologies to reach the fixed goals. It could be both low-fidelity prototype (e.g. paper-sketches) or high-fidelity one (e.g. virtual or physical mock-up). As previously mentioned, the choice of technology strongly influences the possibilities for integration. If the degree of intrusiveness is high, through suggesting a significant change within the current user interface, such choice is crucial thinking about a future integration.

The last and more important phase is the **Evaluation of the implemented solutions with end-users**. Being a methodology based on the user-centered approach, this step is a key point of the methodology. The prototype is tested to understand the degree of reliability, efficiency, effectiveness that the integrated solution proposes.

User tasks analysis is adopted for the evaluation. The assessment is made by separating the interaction of individual tasks in order to avoid interference among different tasks. In this way, it will be possible to evaluate the user's performance in terms of effectiveness (how many times the goal is achieved) and efficiency (in how much time the user reaches the target). In addition, to determine a full usability profile, an evaluation questionnaire is proposed to the user at the end of the performance in order to measure the degree of satisfaction in use.

Moreover the set-up of experimentation is defined. It takes into consideration aspects such as environment, used equipment (invasive or not), data flow management, experimental protocol to ensure repeatability and to define the boundaries of target user actions (what can and cannot he/she do?).

Finally, the results elaboration is carried out to validate improvements of the interaction with the proposed solution.

3.4 Prototype and Integration Guidelines

The final outputs of the overall methodology are a final reference prototype of the improved interface solution and a set of integration guidelines to drive designers into adaptive design processes in case of new interaction paradigms implementation.

The **Reference Prototype** can be either low-fidelity or high-fidelity according to the stage the development has stopped (i.e. conceptual, embodiment or detail design). If the design of the solution arrives at the concept stage, the final application will be developed by exploiting rapid prototyping techniques and paper-sketches without the need to implement the application logics (html wireframe, interactive interface builder). If it arrives at detail design, the prototyping will adopt a development platform compatible with the product and exploit a proper programming language to implement all features.

Design Guidelines contain useful instructions for the integration of the designed solution into a product at different level of detail according to the type of developed reference prototype. Results can be the definition of a format of the integration (library, scripting code, a bridge application, hardware design changes etc.) and the data interchange language to be used to allow the product to interact with the new solution.

4 A Case Study: Control Musical Parameter with Hands

4.1 Analysis of Korg Pa800 Musical Keyboard

The proposed approach is applied to re-design the user interface of an electronic musical keyboard (Fig. 2). Application aims to verify the methodology reliability. The analyzed target task is "*Create New Sound*". The usability test is composed by two main phases.

Fig. 2. Korg Pa800 musical keyboard

In the first phase, the study focuses on the *number of interactions* the user has to perform to reach the task and on the feedback the keyboard gives in reaction. Every interaction is classified into *three different typologies*, i.e. *wheel rotation, tap&list selection and tap interaction on display*. The goal of this assessment is the analysis of two aspects of interaction as follows: (1) the number of user interactions ($25 < \mathbf{N_i} < 40$) and (2) the paths done to reach the target task ($\mathbf{N_p = 3N_i}$), which depends from the three detected interaction modalities.

The second phase is focused on the occurrences of a given interaction typology. The analysis is carried out on the sound setting menu exploring the different sound modulation sections. In every tab, the interaction occurrences are monitored observing the different input mode. Results show the *wheel rotation* as the main interaction paradigm used in the performance of the task. The *tap&list selection* is the second one. However, this interaction mode represents a hybrid solution between the *tap interaction* and the *wheel rotation*: so it appears an alternative way to select menu items or combo box options instead the wheel movement. From these considerations, the role of wheel is clearly fundamental to the setting of keyboard functionalities.

The conducted study shows some critical drawbacks of the real product as follows: (1) *too many interaction steps are necessary to reach the task*; (2) *three different paths are vaguely chosen by the user* creating confusion; (3) some *buttons are not easily visible and finally (4)* some selected functions have *too many parameters* to be set. A workflow is used to illustrate the results of the analysis (Fig. 3).

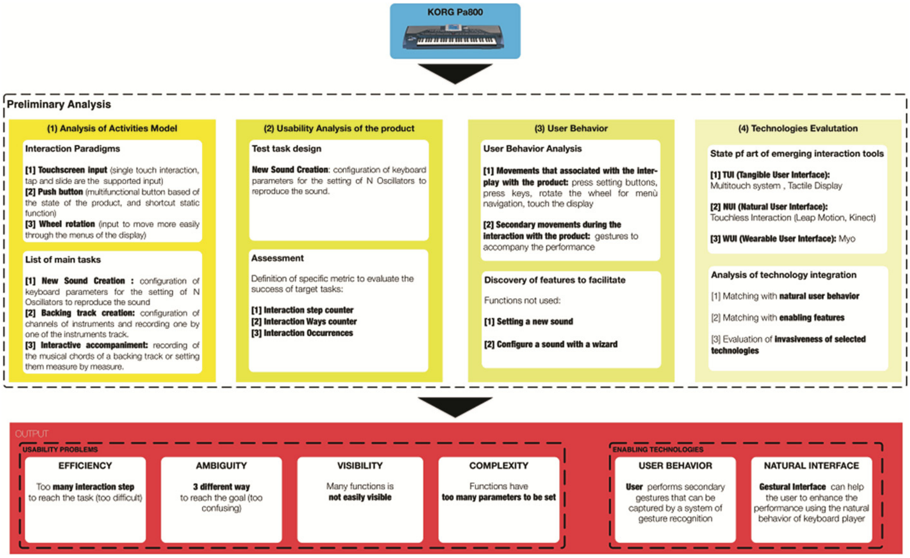


Fig. 3. Pre-analysis workflow

4.2 Design of the Case Study

The detected problems of *efficiency*, *ambiguity*, *complexity* and *visibility* drive the benchmark of natural interfaces and related enabling technologies to overcome them.

Three different solutions have been developed according to the structure of the second stage of the proposed design methodology. One has been evaluated the best and described below. It presents an improvement about the minimization of *visibility* and *complexity* as it avoids the use of a multi-level menu of the keyboard.

The solution comprehends an educational application to support orchestra conductors in performance direction that exploits hand gesture-based interaction modes. The application allows to manage the tempo, velocity and instrument activation in a target performance. The user can control the *volume level*, *tempo beats* and finally the *toggle of the orchestra instruments*.

An independent module is properly developed and prototyped to manage the new technology and implemented functionalities.

4.3 Interface Development

The development phase consists of the (1) *Definition of the architecture* and design requirement of the new module, the (2) *Design of the User Interface with a feedback system* and the (3) *Design of experimental setup*.

In this context, the architecture of the software application is shown in Fig. 4. It is composed by three main parts as follows: (1) a **gesture monitor module** that connects the Motion Controller with the core, (2) a **pattern recognition module** to identify the user gesture, and (3) an **actuator method module** that sends the MIDI commands to a sequencer for the modulation of the target MIDI file.

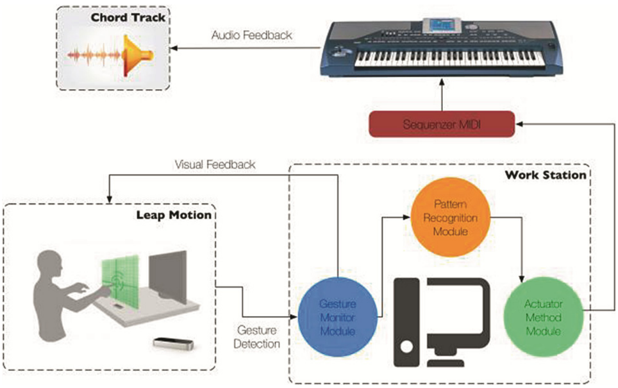
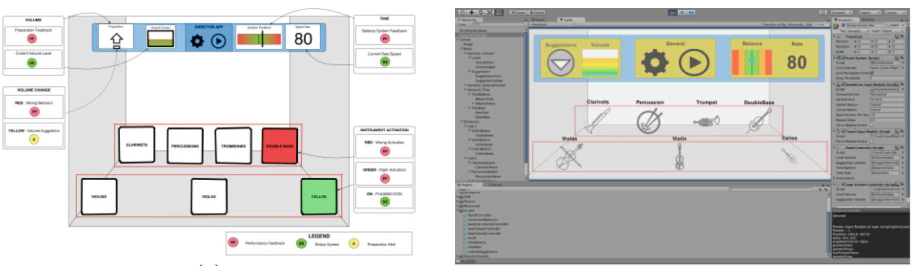


Fig. 4. Director App Architecture

The interface development phase consists in the design of the interaction modality according to the User Interface Design proposed by Norman [25]. In particular, the focus is on the definition of the necessary feedbacks to ensure a certain degree of user performance (see Fig. 5(a)).



(a) (b)

Fig. 5. Interface Design of application

A **prototype application** using Leap Motion SDK and Unity3D Environment connected with an external MIDI sequencer are developed. Furthermore, the core application is developed in C\# language and is able to be integrated with Korg Firmware through a

C-porting process and make a library package of the recognition and activation algorithms. The prototype is shown in Fig. 5(b).

4.4 Experimental Setup and Evaluation

A user-oriented experimental protocol based on task analysis is defined. The chosen tasks the user has to perform by the application are: (1) *tempo control*, (2) *velocity control* and (3) *instruments activation*. For each task usability metrics are measured. They are the following:

1. *Efficiency*, measuring the delay between the performance feedback trigger and the user action. The feedback system provides to suggest the user the action to be performed: once unleashed the feedback is measured how long it took the user to take the proper action.
2. *Effectiveness*, counting the number of failures of the task. If the user doesn't act the right action after the feedback visualization, a counter index increases.
3. *Satisfaction,* through compiling a Likert questionnaire about the sensations felt by the user during the performance. These data are useful to understand the user experience of the new interaction modalities.

Users perform the direction of a classic track in the following experimental environment: a large screen shows the user interface with an iconic representation of the instrument spatially arranged as in real orchestra; the musical keyboard are connected with a gesture controller in front of the user; a stereo audio system with two speakers are connected with the musical keyboard to enhance the audio feedback of the performance. The sample consists of a number of ten elements, with a mean age of 26 years. Users have experience in the music field, but they are not professional users.

Table 1 shows data related to the usability validators of the sample users in the three tasks: looking at the summarized results three considerations can be made:

Table 1. Results of user evaluation

USER	Efficiency (ms)			Effectiveness			Satisfaction
	T1	T2	T3	T1	T2	T3	
1	812	838	938	1	0	3	72,63 %
2	1050	1056	1056	3	0	2	73,68 %
3	857	1010	1046	0	0	2	85,26 %
4	991	833	917	4	2	3	63,16 %
5	999	908	1032	3	1	1	69,47 %
6	832	842	1091	2	1	3	70,53 %
7	1028	928	893	1	0	3	77,89 %
8	825	876	896	2	2	2	68,42 %
9	1070	824	823	1	1	3	74,74 %
10	866	966	846	2	1	3	65,26 %
TOTAL				19	8	25	
AVERAGE	933	908,1	953,8	1,9	0,8	2,5	72,11 %

- Efficiency, considering the update frame offset of 700 ms, turns out to be very low, showing the good responsiveness of the user to the system stimuli;
- In a typical performance session (9 min), the total average errors are 5, 2. The results indicates a good degree of effectiveness, even if the performance has room for improvement;
- Finally, the Likert Questionnaire base on IBM Template [26] show a satisfaction rate greater than 70 %.

5 Conclusion and Future Works

A methodology to introduce gesture-based interfaces into consumers' products as musical keyboards is presented and validated through a real case study. Experimental results demonstrate how the proposed methodology supports the interface re-design with the objective of simplifying and maximize the system usability. Moreover, it allows the definition of innovative applicative functions, which can be obtained through the combination of pre-existing functions and new interaction modalities. Future work will be focused on a deep investigation of user experience and achieved performance. This will lead to extend the technology to other products.

Acknowledgements. This work is carried out with the collaboration of Korg Italy SPA, which has provided the actual product and your skills and knowledge to support the analysis phase.

References

1. O'hara, K., et al.: On the naturalness of touchless: putting the "interaction" back into NUI. ACM Trans. Comput. Hum. Interact. (TOCHI) **20**(1), 5 (2013)
2. Hughes, D.: Technologized and autonomized vocals in contemporary popular musics. J. Music Technol. Educ. **8**(2), 163–182 (2015)
3. Mitchell, T.J.: Soundgrasp: a gestural interface for the performance of live music (2011)
4. Titanrealitycom: Titanrealitycom. Retrieved 11 February, 2016. https://titanreality.com
5. Ketabdar, H., et al.: MagiMusic: using embedded compass (magnetic) sensor for touch-less gesture based interaction with digital music instruments in mobile devices. In: Proceedings of the Fifth International Conference on Tangible, Embedded, and Embodied Interaction. ACM (2011)
6. Yuksel, K.A., Ketabdar, H., Roshandel, M.: Towards digital music performance for mobile devices based on magnetic interaction. In: 2010 IEEE International Symposium on Haptic Audio-Visual Environments and Games (HAVE). IEEE (2010)
7. van Dorp Skogstad, S.A., Jensenius, A.R., Nymoen, K.: Using IR optical marker based motion capture for exploring musical interaction, pp. 407–410 (2010)
8. Marrin, T.: Possibilities for the digital baton as a general-purpose gestural interface. In: CHI 1997 Extended Abstracts on Human Factors in Computing Systems. ACM (1997)
9. Cabral, M.C., Morimoto, C.H., Zuffo, M.K.: On the usability of gesture interfaces in virtual reality environments. In: Proceedings of the 2005 Latin American Conference on Human-Computer Interaction. ACM (2005)
10. Manresa-Yee, C., et al.: Hand tracking and gesture recognition for human-computer interaction. In: Progress in Computer Vision and Image Analysis, pp. 401–412 (2010)

11. Rico, J., Brewster, S.: Usable gestures for mobile interfaces: evaluating social acceptability. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM (2010)
12. Godøy, R.I., Leman, M. (eds.): Musical Gestures: Sound, Movement, and Meaning. Routledge, New york (2010)
13. Ng, K.C.: Music via motion: transdomain mapping of motion and sound for interactive performances. Proc. IEEE **92**(4), 645–655 (2004)
14. Vigliensoni, G., Wanderley, M.M.: A quantitative comparison of position trackers for the development of a touch-less musical interface. In: Proceedings of the 12th International Conference on New Interfaces for Musical Expression (NIME 2012), Vancouver, Canada (2012)
15. Morita, H., Hashimoto, S., Ohteru, S.: A computer music system that follows a human conductor. Computer **24**(7), 44–53 (1991)
16. Je, H., Kim, J., Kim, D.: Hand gesture recognition to understand musical conducting action. In: The 16th IEEE International Symposium on Robot and Human interactive Communication, 2007. RO-MAN 2007. IEEE (2007)
17. Nakra, T.M.: Inside the Conductor's Jacket: Analysis, interpretation and musical synthesis of expressive gesture. Ph.D. thesis Massachusetts Institute of Technology (1999)
18. Finney, S.A.: Auditory feedback and musical keyboard performance. Music Percept. **15**(2), 153–174 (1997)
19. Pfordresher, P.Q., et al.: Brain responses to altered auditory feedback during musical keyboard production: an fMRI study. Brain Res. **1556**, 28–37 (2014)
20. Zamm, A., Pfordresher, P.Q., Palmer, C.: Temporal coordination in joint music performance: effects of endogenous rhythms and auditory feedback. Exp. Brain Res. **233**(2), 607–615 (2015)
21. Bigand, E., et al.: Looking into the eyes of a conductor performing lerdahl's "Time after Time". Musicae Sci. **14**(2), 275–294 (2010)
22. Wurtz, P., Mueri, R.M., Wiesendanger, M.: Sight-reading of violinists: eye movements anticipate the musical flow. Exp. Brain Res. **194**(3), 445–450 (2009)
23. Gilman, E., Underwood, G.: Restricting the field of view to investigate the perceptual spans of pianists. Vis. Cogn. **10**(2), 201–232 (2003)
24. Nielsen, J.: Usability inspection methods. In: Conference Companion on Human Factors in Computing Systems. ACM (1994)
25. Norman, D.A.: The Design of Everyday Things: Revised and Expanded Edition. Basic Books, London (2013)
26. Lewis, J.R.: IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. Int. J. Hum. Comput. Interact. **7**(1), 57–78 (1995)