Mapping Out Instruments, Affordances, and Mobiles

Atau Tanaka
Culture Lab
Newcastle University
Newcastle upon Tyne, NE1 7RU UK
+44-191-246-4640
atau.tanaka@ncl.ac.uk

ABSTRACT
This paper reviews and extends questions of the scope of an interactive musical instrument and mapping strategies for expressive performance. We apply notions of embodiment and affordance to characterize gestural instruments. We note that the democratization of sensor technology in consumer devices has extended the cultural contexts for interaction. We revisit questions of mapping drawing upon the theory of affordances to consider mapping and instrument together. This is applied to recent work by the author and his collaborators in the development of instruments based on mobile devices designed for specific performance situations.

Keywords
Musical affordance, NIME, mapping, instrument definition, mobile, multimodal interaction.

1. INTRODUCTION
As the NIME community enters its second decade, we take stock of its centers of interest, the knowledge that has accumulated, and the new directions that are taking shape. We review the literature on two of the federating themes in NIME research: that of *defining the instrument*, and that of *mapping*. We first look at work establishing the field leading up to the first NIME conference in 2001, followed by developments in technology that have broadened the field. We then reconsider the notion of instrument and approaches to mapping together, and see how they can inform the development of a series of new instruments built on commonly available platforms.

The late Michel Waisvisz placed an importance on the viscerality of a electro-instrumental system [12]. Ryan reported on the STEIM approach [32], and also reflected on the importance of expressivity in his own work [33]. Cadoz had been interested in ways in which types of gestures in musical performance [20] were relevant to live performance of computer music [6].

Winkler presented early techniques of best practice for programming interaction in live music performance in [43]. Wanderley and Battier created the first compendium of texts in this area, where many of the early work in the field [40], including that of the present author, are presented [35].

With the establishment of NIME in 2001 as a workshop at the Computer Human Interaction (CHI) conference [31], its establishment as a full conference in 2002, and the archiving of its proceedings in the Association for Computing Machinery (ACM) Digital Library in 2005, an existing musical community coalesced into and established it as a research field. Common themes began to emerge, and forms of best practice could be transmitted. Two of the predominant themes were the definition of a electronic musical instrument, and approaches to mapping gesture to sound. Throughout this, there is the underlying goal to make advancement in the creation of expressive musical systems.

2. DEFINING THE INSTRUMENT
Miranda and Wanderley [26] describe an interactive musical instrument as being a system comprised of three basic subsystems:

- Sensor input
- Mapping
- Sound synthesis

More recently with the increasing use of mobile devices in NIME, Essl, Wang, and Roh describe generic qualities as desirable characteristics of electronic instrument building, in their case, on mobile instruments [13, 14]. The characteristics defining generic for them are flexible without specific prefiguring, designed without a specific musical work as a driver. Meanwhile, the present author has written about idiomaticity of instruments and the distinction of an instrument from a tool [36]. Camurri introduces the notion of entrainment in expressive musical gesture [7]. Here the notion of a generic system meets the situation-specific instrumental circumstances that lead to forms of corporeal motor instincts such as muscle memory that constitutes entrainment.

It is interesting noting that early in the history of NIME, considerable discussion that took place on whether the “I” of NIME should be the word, *Instrument*, or the word, *Interface*. The choice of interface reflects a desire to open up the field, but also perhaps a recognition of the challenge in arriving at a single vision of what a NIME instrument should or could be.

In an attempt to leave room for new developments, yet identify federating characteristics, one proposed working definition of the instrument is that of a system that is autonomous and open-ended [36]. By autonomous we mean self-contained and self-sufficient. This could mean that audio content, sonic modification, and control interaction co-exist on the same device. By open-ended, we mean extensible systems. This could mean modifications adding to or enhancing an instrument’s sonic characteristics.

This drew upon evolution of musical instruments from acoustical instruments to multi-component systems like the electric guitar and associated effects pedals and amplifier, to turntable as musical instrument where choice of content (in this case the vinyl record used) made it open-ended.

Notions from instrumental practice have been used to inform interaction design. Beaudoin-Lafon defines Instrumental Interaction as a set of techniques to draw upon tool usage in the real world to inform graphic user interface (GUI) design [1]. Interaction instruments mediate the space between GUI users and domain objects. Instrumental properties include degrees of indirection (spatial and temporal offsets), integration (mapping of degrees of freedom), and compatibility (consistency of action and response modes). Verillon [39] calls on the act of instrumentalisation to distinguish between tools and instruments, a theme that is directly relevant to music, and an issue of differentiation that has been discussed by this author in distinguishing musical instruments from tools [36].

We can draw upon taxonomies of embodiment developed in the field of Tangible User Interfaces to categorize the self-contained nature or level of distributedness of a musical instrument. Fishkin defines embodiment as the extent to which input focus of a system is bound to the output focus [15]. He goes on to discern four levels of embodiment:

- Distant – where system output is remote with respect to the input device
- Environmental – where output surrounds the input but is not graspable
- Nearby – where output is proximal to input
- Full – where the output device and input device coincide

This range of embodiment can be applied to musical instruments in different ways. A pipe organ in a church can be considered the clearest case of distant embodiment where keyboard actioning is far from actual sound output, as compared to the violin in which the coincidence of the gestural articulation of bowing and actual acoustical output represents full embodiment. Electric guitars, an example of open ended in [36], can from the viewpoint of embodiment, be considered nearby embodiment, coupling the guitar, via pickup to amplifier. Finally, acousmatic music in this sense could be thought of as environmental embodiment where a composer performs an orchestra of speakers in a space from a central console.

In the case of NIME instrument building, systems that capture performer gesture and produce sonic musical output mean that full embodiment is an implicit goal and criteria of success. The notion of expressivity is a core principle, and indeed, part of the NIME acronym. It is assumed, but is elusive as a concept. From Cadoz to Camurri, researchers have tried to characterize expressivity. Meanwhile for performing artists like Ryan, expressiveness in musical performance is taken to be intuitive and empirical. For purposes of the musical interaction concepts discussed in this paper, we will take expressivity of a musical instrument to mean specific musical affordances of an instrument that allow the musician performing on the instrument to artfully and reliably articulate sound output of varying nature that communicates musical intent, energy, and emotion to the listener. The success of expressivity resides not just in the effectiveness of communication, but in the sense of agency that the system gives back to the performer.

3. MUSICAL AFFORDANCE

Affordance is a concept fundamental to interaction design practice. Arising from Gibson's seminal work in perceptual psychology, it maps potential action relationships between subject and object based on qualities of the object and capabilities of the subject [18]. Norman introduced affordance in design practice, providing guidance for ways to create designs for objects that project perceived possibilities of usage [27]. Perception is crucial to both Gibson and Norman, but treated more explicitly by Norman. While Gibson discusses emergent properties, perceived or not, of an object giving to its affordance for action, Norman very practically encourages a designer to create perceptions to suggest usage. Gaver reconciles and extends these notions by proposing hidden and perceptible affordances, as well as types of complex affordances [17].

Affordance has been called upon in describing computer music systems. Magnusson applies the concepts of affordance to inform the design of screen based interactive music systems. Gurevich [21], Braasch [5], and Dillon and Brown [11] draw upon affordance theory to inform the design of network music systems. Cook and Pullin apply affordance in design practice to facilitate creating musical interfaces from everyday objects [9]. This becomes in some ways a musical manifestation of Chung and Ishii's Mega-Affordance objects [8].

Affordance has been used in the sociology of music by DeNora to look at music-as-practice, and the sociality of music, to begin to characterize what might be afforded by certain pieces of music to listeners of that music [10]. This has an interesting parallel in sensory motor studies by Godøy where subjects are asked to trace gestures as they listen to music to infer an affordance innate in the musical sound [19].

4. MAPPING

Mapping is commonly defined as the translation layer correlating gesture to sound. Hunt and Wanderley conducted an early literature review and established basic principles of mapping in gestural music [22].

They identify 4 types of mappings

- 1 to 1
- 1 to many (divergent)
- Many to 1 (convergent)
- Many to many

They distinguish perceptual parameters from abstract parameters and conclude by proposing a two-layer model.

Wessel introduces early use of tablets interfaces and multi-touch surfaces and speaks of expressiveness in forms of intimate interaction – levels of subtlety that are satisfying for the performer [42]. He proposes several metaphors for meta-mapping interaction:

- Scrubbing
- Catch/Throw
- Dipping

Paine applied user centered design method in conducting an interview-based design of mapping for a specific interface device (the Thummer) [29]. Tarabella considers the mapping problem by using gestural metaphors from established instruments (i.e. piano) as a point of departure. He applies this to investigate the potential of spontaneity and improvisation of computer music [38].

4.1 From Mapping to Musical Phrases

While the existing literature investigates mapping techniques to yield expressive musical situations, they either focus on specific instrument examples, or on isolated gestural events. Conceiving of effective strategies for creating higher level musical output requires deploying layers of different mappings that work in tandem to allow the articulation of musical phrases. A phrase here is used in a more general sense than in music theory, and includes high-level musical events encapsulating phrases, meta-events, and complex musical structures.

Articulation of musical phrases is not typically executed by a single mapping. Here we propose a model of minimum three types of mappings that work on conjunction to formulate the articulation of a musical unit. The proposed mapping types are:

1. Binary mapping
2. Basic parametric mapping(s)
3. Expressive mapping(s)

If we consider a traditional instrument such as the violin, the elements necessary for articulation of sound can be seen as: 1) contact of the bow with the string (binary), 2) selection of the length of string (basic parameter), and 3) vibrato (expressive). The definition of these three mapping types, then, should not be seen as limiting, but as a point of departure.

We can apply this tri-partite mapping structure to the metaphors proposed by Wessel and Wright:

Scrubbing:

- Sound on/off (Binary)
- Sound selection (Basic)
- Scrubbing position (Expressive)

Catch/Throw:

- Capture live sound (Binary)
- Choice of effect (Basic)
- Effects parameters (Expressive)

Dipping:

- Dip on/off (Binary)
- Dip volume (Basic)
- Additional transformation parameters (Expressive)

This model can also be applied to higher order musical meta-events, where the "binary" can articulate a group of note or sound events, with "basic" representing a parameter such as playback speed of a sequence, and "expressive" modulating synthesis parameters that affect the timbre of the sequence and clusters that are sounding.

4.2 Complex Mappings

There are several ways to arrive at the three levels of mapping needed to generate a musical event as described. The Binary mapping can be seen as an activation of the sound, the Basic mapping as a fixed parameter in the articulation of the sound, and the Expressive mapping being a continuously varying parameter that follows the gesture and is mapped to modulation of the sound. In order to derive these three levels of mapping from one sensor, a single input must be processed in different ways.

From a given input, we can derive the Binary mapping by setting a threshold level to trigger the event. The Basic mapping can be derived by looking at the slope of the input curve, it's rate of change over a small time window, at the time of the trigger. The Expressive mapping can then follow the sensor input to modulate a sound synthesis parameter. The use of a single sensor input to generate three types of mappings (albeit to control a single musical event) can be considered using Hunt and Wanderley's classifications as a "One to Many", or divergent type of mapping.

Techniques for deriving rich interaction in this configuration include splitting the sensor input and applying different levels of filtering for each type of mapping. The Binary trigger might be driven off of the sensor input with slight filtering with a programmed Schmitt trigger to avoid accidental multiple triggering [34]. This puts in place two comparators, one for the main trigger, and a second, downward comparator in a hysteresis loop allowing new event triggers only after the system is "re-armed" by crossing the downward threshold. Meanwhile the Basic mapping in this case can be derived from a branch that applies smoothing in the form of a low pass filter on the sensor input. The Expressive mapping could be derived from the raw sensor data with no filtering for maximum responsiveness, or with a slight level of smoothing to distinguish it from the basic mapping, depending on the musical situation at hand. Responsiveness of each branch can be set independently with different latencies inherent in the types of filtering used. We can call this Complex Mapping, where one gesture is subject to several post-processing branches to yield multiple mappings.

4.3 Compound Mappings

Another approach to derive the three levels of mapping described is by the application of Hunt and Wanderley's "Many-to-One", or convergent technique. Here "many" refers to multiple sensors on a single gesture, or the capturing of multiple gestures, and "one" refers to a single musical event being shaped by multiple parametric control. This approach can be addressed by applying techniques of multimodal interaction.

In the classical multimodal interaction, a single gesture can be captured by multiple sensing modes. Applied to the present context of deriving three mappings in order to articulate a single musical event, each mode can be optimized to yield the kind of signal best suited for a binary, basic, or expressive mapping. On a Nintendo Wii remote, this can be thought of as the button, infrared camera, and accelerometer. We have applied the notion of multimodal sensing to arm gesture, captured by biosensors and accelerometers [37].

The Many-to-One scenario can be realized also by the use of multiple gestures to articulate a single musical event. This could mean articulating a sound with one gesture while modulating it with another. Conditional situations can be set in he forms of double triggers, logical AND where a Binary mapping is set only if multiple thresholds are all crossed, or a Schmitt trigger is set while an auxiliary sensor is above a threshold, or in a logical OR where any number of multiple sensor inputs could trigger a sound, later to be modulated. More sophisticated interrelations can be created, such as Activity Windows where a range of sensor values in one input open a window in which a second sensor might influence a certain parameter in the sound being articulated. A series of windows can be established creating a situation where Sensor 2 modulates different sound synthesis parameters depending on the window activated by Sensor 1. In this way, these mappings can be considered Compound Mappings.

5. DEMOCRATISATION AND MOBILITY

Alongside evolution in the field, with the establishment of communities of practice and the NIME conference have been the increased accessibility of sensor technologies and industrial and consumer applications of them. The early systems described by Waisvisz required development of custom hardware and software systems. While a focus of NIME research remains the invention of new systems, this now customarily takes place on standardized software environments such as Max, PD, SuperCollider, and accessible hardware platforms such as the iCube or Arduino. Meanwhile the explosion of consumer devices with built-in sensors such as mobile phones and game controllers such as the Nintendo Wii-remote have opened up and democratized NIME practice outside the realm of academic research.

5.1 Participation

This potential of broadening participation has a profound impact on the contexts in which NIME research finds itself, and new opportunities for the application of knowledge created in our field over the years. These new contexts include

- Enhanced musical interaction in consumer products
- New forms of amateur music making
- Increased dissemination of NIME performances

The integration of sensors in and the increasing sophistication of game controllers has led to the increasing richness of musical interactions in computer games. Blaine surveys the use of novel input devices in musical video games [3]. Bott has looked at 3D spatial interaction with standard game controllers for expressive amateur music activities [4]. We have coined the notion of “music one participates in” to capture participatory musical activity [24].

Meanwhile mobile devices such as smartphones have accelerometers, and touchscreens for input, and processors capable of realtime DSP. This has led to consumer applications for amateur music making such as the Ocarina [30].

The widespread availability of NIME hardware and software in convenient portable form factors has led to the creation of groups and ensembles. In the field of academic and experimental music, we have seen a proliferation of mobile phone orchestras [41].

5.2 Mobile Instruments

We apply the above principles in guiding the design of musical instrument performance systems built on the iPhone mobile phone. We draw upon Wang and Essl’s notion that the mobile phone has robust potential to be a musical instrument by virtue of it having sensor input, signal processing capability, and sound output on a single device. While Essl seeks to create generic instruments, we were interested in looking at specific configurations of the hardware and associated software to create distinctive instruments with specific modes of musical articulation.

An advanced mobile phone such as the iPhone is a compelling musical instrument for its various sensing modes:

- Multitouch screen
- 3D accelerometer
- Audio input
- GPS
- Compass
- Camera

It has network capability and signal processing capability. The Audio input can be used in several modes: 1) to treat audio, as in Wessel/Wright’s Catch and Throw scenario, 2) as a breath sensor as in Wang’s Ocarina, and 3) as a data acquisition channel for auxiliary sensors, as in Jo’s Inaudible Computing [23].

With Marek Bereza [2] we have experimented with use of the camera and the multi-touch screen. By placing one’s hand over the camera, we could get a very rough approximation of the distance of the hand through image luminosity. The auto-exposure feature of the camera is hardwired and could not be overridden, and ultimately defeated the use of camera as luminosity sensor.

Bereza’s SampleToy implements granular synthesis controlled by the multi-touch screen. Each point detected on the multi-touch screen played a voice of the granular synthesizer, with X-axis modulating the time stretching, and the Y-axis the pitch. We can apply the tri-partite mapping model from above here, to consider that the Binary mapping was the presence of a point on the multi-touch screen activating a synthesizer voice, with sliding on the Y-axis setting the Basic mapping of pitch, and the X-axis being the expressive mapping of time stretching.

The multiple sensing modes of the iPhone make it apt for multimodal interaction. The different modalities we have experimented with include

- Use of GPS and compass
- Simultaneous use of touchscreen and accelerometer
- Simultaneous use of multiple iPhones by one performer

With Bernhard Garnicnig we have been experimenting with use of the GPS for geographic location alongside the compass to detect orientation. We have been implementing a prior locative audio work, *Craving*, initially realized on computer based systems [16] spatializing a theatre text according to latitude/longitude and radial orientation. Extensions to this include the addition of accelerometer data to couple local motion with the cruder location data of the GPS. In this way, high frequency accelerometer data can serve to trigger the Binary mapping to activate a line in the script, with GPS location giving the Basic mapping of determining which line, and orientation giving the Expressive mapping of binaural spatialization.

With Adam Parkinson we have been using the Pure Data implementation on iPhone, RJDJ to create concert performance instruments. These have focused on performance modes allowing one-handed playing. This came out of the observation that most modes of touchscreen interaction with the iPhone involve two hands. In the case of the multi-touch interaction with SampleToy, for example, in order to take advantage of the multi-touch interaction, one hand needs to hold the device while multiple fingers on the other hand articulate synthesis voices. In the instruments created by Parkinson, the accelerometer is used in conjunction with the touchscreen where buttons onscreen activate individual sounds (Binary mapping), sliders set Basic mapping parameters such as granular synthesis grain size, leaving the X and Y tilt of the accelerometer to modulate Expressive mappings of pitch and time stretching. These instruments can be played by one hand grasping the iPhone, with the thumb manipulating the touchscreen one parameter at a time, and with hand and wrist motion performing the accelerometer. This creates a compelling performance gesture which becomes all the more dynamic with the use of two such instruments, one in each hand.

6. DISCUSSION

The iPhone/RJDJ instrument developed with Adam Parkinson provides a working system to apply the concepts of mapping and affordance. The system has been used in live performance contexts where each performer in a duo has a device in each hand, creating a four-hand iPhone performance.

A granular synthesis patch is controlled by the device's accelerometer and touch screen using the following mappings:

- X tilt – time stretching
- Y tilt – pitch
- On-screen buttons – sample and preset select
- On-screen sliders – pitch quantization, grain duration
- X-Y on-screen control – filter cutoff frequency and resonance

Using the mapping types from Section 4.2, the buttons can be thought of as activation mappings, the sliders as articulation mappings, and the accelerometer as modulation mappings. These are controlled by a single hand grasping the object with the thumb manipulating the on screen interface.

The use of screen interaction and object manipulation together in one hand lends to integrated, compound gestures to articulate sound. These gestural components are captured through two independent input modes, creating a form of multi-modal interaction.

The system has distinguishing instrumental qualities in that it is a graspable object that is at once sensor input device and sound synthesis device. In this, input and output modes coincide and represent a case of Fishkin's full embodiment.

In order to broach issues of affordance, we must consider both the physical object as well as its software based sound synthesis capabilities. Norman's use of affordance is useful to understand the instrument's origins as a consumer electronics device. Its form factor make a graspable device that can be operated with one hand. The touch screen affords tactile interaction, of which our software affords a certain subset of all possible interaction. The use of sliders and an X-Y on screen control afford single touch, and not multi-touch interaction.

Culturally, the iPhone is an iconic device that affords communications (as a telephone) and music listening (as a personal music player). Rather than play upon these cultural associations and constraints, we imbue the device with sonic affordances that afford forms of gestural, musical expression not typically associated with the device as product.

Gaver's distinction of perceptible and hidden affordances can be applied here – perceptible affordances being the properties generally associated with the iPhone as consumer device, and hidden affordances being the sound synthesis capabilities programmed in using the RJDJ software. The performer on the instrument draws upon both types of affordances. The spectator at a concert, on the other hand, begins with expectations conditioned by the perceptible affordances and discovers the musical expression afforded by the hidden affordances, contributing to the concert experience. This ultimately leverages the compound invariants of Gibson in a dynamic of exteroception channeled through proprioception.

With this we arrive at a characteritic of musical performance that take affordance beyond the single-user orientation of Norman, Gaver, and Gibson. Musical affordance is situated in music as a social, public activity. Affordance the instrument provides in this case concerns not just the primary user, the performer, but also the audience as observer and perceiver. The instrument in the first instance must afford expressive musical action to the performer at the same time that it affords communicable musical articulation in a concert setting. These two levels of affordance are conditioned by form factor, embodiment, and cultural association. This opens up a rich area for further investigation.

7. CONCLUSIONS

The self-contained nature of a mobile device for Essl and Wang bring the computer music instrument closer to a traditional instrument. Meanwhile the position of an acoustical instrument as gold standard is put in question by Overholt, Roads, and Thompson [28] who observe that electronic instruments are distinct from traditional instruments in the multiple timescales on which they operate, and the breadth of parametric control that is possible and often required.

While not confronting the relative merits of acoustical instruments as compared to computer-based instruments, we can take qualities of embodiment as defined by Fishkin as a way to look at the expressive potential of an instrument, and define types of musical gesture and the mappings that can support those gestures to define the musical capabilities of a digital, computer music performance instrument.

The autonomous, standalone quality – the embodied nature – of a mobile instrument allows specific, possibly indiosyncratic, mappings to be implemented. In this, the hardware itself may remain a general purpose device, and the underlying software framework may be generic, but the instantiation of these resources in the form of an instrument are specific for any given musical situation. We can draw upon Gibson's theory of affordance [18], and the notion of constraint [25] to conclude that this distinctness creates a form of *musical affordance* that is propitious to entrainment, and therefore intuitive expressiveness.

8. ACKNOWLEDGMENTS

The work described in Section 6. are the result of collaborations with Adam Parkinson, Marek Bereza. Bernhard Garnicnig.

9. REFERENCES

[1] Beaudouin-Lafon , M. "Instrumental interaction: an interaction model for designing post-WIMP user interfaces." In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '00)*. 2000.

[2] Bereza, Marek. http://www.mrkbrz.com/ Retrieved April 10, 2010.

[3] Blaine, T. "The Convergence of Alternate Controllers and Musical Interfaces in Interactive Entertainment." In *Proceedings of the 2005 International Conference on New Interfaces for Musical Expression (NIME05)*. 27-34. 2005.

[4] Bott, J., Crowley, J., LaViola, J. "Exploring 3D gestural interfaces for music creation in video games." In *Proceedings of the 4th International Conference on Foundations of Digital Games, FDG* 18-25. Orlando, Florida, 2009.

[5] Braasch, J. "The Telematic Music System: Affordances for a New Instrument to Shape the Music of Tomorrow." In *Contemporary Music Review*. 28, 4/5 (August 2009), 421 – 432.

[6] Cadoz, C. "Instrumental Gesture and Musical Composition." In *Proceedings of the 1988 International Computer Music Conference*. San Francisco, International Computer Music Association. 1-12. 1988

[7] Camurri, A., Mazzarino, B., Volpe, G. "Analysis of Expressive Gesture: The EyesWeb Expressive Gesture Processing Library." In Camurri, A., Volpe, G. (Eds.), *Gesture-based Communication in Human-Computer Interaction*, LNAI 2915, 460-467, Springer Verlag, 2004.

[8] Chung, K., Ishii, H., "Fusing computation into mega-affordance objects." In *CHI 2009 Workshop on Transitive Materials*. 2009.

[9] Cook, A., Pullin, G. "Tactophonics: your favourite thing wants to sing." In *Proceedings of International Conference on New Interfaces for Musical Expression (NIME07, New York)*, 2007.

[10] DeNora, T. *After Adorno: Rethinking Music Sociology*. Cambridge University Press. 2003.

[11] Dillon, S. C. and Brown, A. R. "The educational affordances of generative media in arts education." In *Proceedings of International Technology, Education and Development Conference (INTED2010)*. Valencia, Spain. 2010.

[12] Dykstra-Erickson, E., Arnowitz, J. "Michel Waisvisz: The Man and the Hands." In *ACM Interactions* 12, 5 (September + October 2005) 63 – 67. 2005.

[13] Essl, G. Wang, G. and Rohs. M. "Developments and Challenges turning Mobile Phones into Generic Music Perfor- mance Platforms." In *Proceedings of the Mobile Music Workshop (MMW-08)*, Vienna. 2008.

[14] Essl, G. "SpeedDial: Rapid and On-The-Fly Mapping of Mobile Phone Instruments." In *Proceedings of International Conference on New Interfaces for Musical Expression (NIME09, Pittsburgh)*, 2009.

[15] Fishkin, K.P. "A taxonomy for and analysis of tangible interfaces." In *Personal and Ubiquitous Computing*. 8, 5 (September 2004), 347-358.

[16] Garnicnig, B. and Haider, G. "Craving, a Spatial Audio Narrative." In *Proceedings of the International Workshop on Mobile Music Technologies (MMW)*. Amsterdam. 2007.

[17] Gaver, W. "Technology Affordances" In *Proceedings of the CHI'91 Conference*. 79-84, 1991.

[18] Gibson, J. J. *The Ecological Approach to Visual Perception*. Houghton Mifflin, 1979.

[19] Godøy, R. "Geometry and Effort in Gestural Renderings of Musical Sound." In Dias, M., Gibet, S., Wanderley, M., Bastos R. (Eds.) *Gesture-Based Human-Computer Interaction and Simulation LNAI 5085*. Springer. 2009.

[20] Gritten, A. and King, E. (Eds.) *Music and Gesture*. Ashgate Publishing, Surrey. 2006.

[21] Gurevich, M. "JamSpace: a networked real-time collaborative music environment." In *CHI '06 extended abstracts on Human factors in computing systems*. 2006.

[22] Hunt, A. and Wanderley. M. 2002, "Mapping Performance Parameters to Synthesis Engines." In *Organised Sound*, 7, 2 97–108, 2002.

[23] Jo, K. "Inaudible Computing: An Extension of Physical Computing using Audio Signals." In *Extended Abstracts CHI09*. 2009.

[24] Jo, K. and Tanaka, A. "The Music Participates In." In Schroeder, F. (Ed.) *Performing Technology: User Content and the New Digital Media*. Cambridge Scholars Publishing. 2009.

[25] Magnusson, T. "Affordances and constraints in screen-based musical instruments." In *Proc. 4th Nordic Conference on Human-Computer Interaction (NORDICCHI)*. 2006.

[26] Miranda, E. and Wanderley, M. "New Digital Musical Instruments: Control and Interaction Beyond the Keyboard." *Computer Music and Digital Audio Series* 21. A-R Editions, Middleton, WI, 2006.

[27] Norman, D. A. *The psychology of everyday things*. Basic Books, New York. 1988.

[28] D. Overholt, C. Roads, J. Thomspon. "On Musical Gestures and New Performance Interfaces for Electronic Music." In *Proc. 5th International Gesture Workshop*, Genova, Italy, 2003.

[29] Paine, G. "Gesture and Morphology in Laptop Music Performance." In Dean, R. (Ed.) *The Oxford Handbook of Computer Music*, Oxford University Press, (2009) 214-232.

[30] Perry, T. "Ge Wang: The iPhone's Music Man." In *IEEE Spectrum*. September 2009.

[31] Poupyrev, I., Lyons, M., Fels, S., Blaine, T. "New Interfaces for Musical Expression: CHI 2001 Workshop Report." In *SIGCHI Bulletin* 34(2): 13-14,21-26 (March/April 2002).

[32] Ryan, J. "Some remarks on musical instrument design at STEIM." In *Contemporary Music Review* 6, 1 (1991) 3-17.

[33] Ryan, J. "Effort and Expression." In *Proc International Computer Music Conference*. 1992.

[34] Schmitt, O. "A Thermionic Trigger." In *Journal of Scientific Instruments* 15 (January 1938) 24–26.

[35] Tanaka, A. "Musical Performance Practice on Sensor-based Instruments." In Wanderley, M., Battier, M. (Eds.) *Trends in Gestural Control of Music*, IRCAM, Paris (2000) 389–405.

[36] Tanaka, A. "Sensor-Based Musical Instruments and Interactive Music." In Dean, R. (Ed.) *The Oxford Handbook of Computer Music*, Oxford University Press, (2009) 233-257.

[37] Tanaka, A. and Knapp, R.B. "Multimodal Interaction in Music Using the Electromyogram and Relative Position Sensing." In *Proceedings of International Conference on New Interfaces for Musical Expression (NIME02)*, May 24-26, Dublin), University of Limerick, (2002) 43-48.

[38] Tarabella, L. "Improvising Computer Music: An Approach." In *Proc. Sound and Music Computing '04 (SMC04)*. Paris. 2004.

[39] Verillon, P. and Rabardel, P. "Cognition and artifacts: A contribu9tion to the study of though in relation to instrumented activity." *In European Journal of Psychology of Education.*10, 1 (March 1995) 77-101.

[40] Wanderley, M., Battier, M. (eds.) Trends in Gestural Control of Music. IRCAM, Paris (2000).

[41] Wang, G., Essl, G. and Pentinnen, H. "MoPhO: Do mobile phones dreams of electric orchestras?" In *Proceedings of the International Computer Music Conference*. Belfast. 2008.

[42] Wessel, D. and Wright, M. "Problems and Prospects for Intimate Musical Control of Computers." In *Computer Music Journal*. 26, 3. 2002.

[43] Winkler, T. *Composing Interactive Music Techniques and Ideas Using Max*. Cambridge, MA, MIT Press. 1998.