Interactive Spaces. Models and Algorithms for Reality-based Music Applications

Publisher:
Association for Computing Machinery (ACM)

Published version:
DOI: dx.doi.org/10.1145/2817721.2820986

Interactive Spaces: Models and Algorithms for Reality-Based Music Applications

Marcella Mandanici
CSC - Sound and Music Computing Group
Department of Information Engeneering
University of Padova, ITALY
mandanici@dei.unipd.it

ITS '15, November 15-18, 2015, Funchal, Portugal
ACM 978-1-4503-3899-8/15/11. http://dx.doi.org/10.1145/2817721.2820986

Abstract
Reality-based interfaces have the property of linking the user's physical space with the computer digital content, bringing in intuition, plasticity and expressiveness. Moreover, applications designed upon motion and gesture tracking technologies involve a lot of psychological features, like space cognition and implicit knowledge. All these elements are the background of three presented music applications, employing the characteristics of three different interactive spaces: a user centered three dimensional space, a floor bi-dimensional camera space, and a small sensor centered three dimensional space. The basic idea is to deploy the application's spatial properties in order to convey some musical knowledge, allowing the users to act inside the designed space and to learn through it in an enactive way.

Author Keywords
Blended interaction; Sound augmented reality; Learning environments.

ACM Classification Keywords
H.5.1 [Artificial, augmented, and virtual realities]; H.5.2 [User Interfaces]; H.5.5 [Sound and Music Computing]

Figure 1: The floor interface of the "Harmonic Walk" application while being tested by a high school student

Motivation and Research Approach

The widespread utilization of cameras, camera software systems and motion tracking devices like Kinect (https://www.microsoft.com/en-us/kinectforwindows/) and the Leap Motion Sensor (https://www.leapmotion.com/), has brought to the attention of a continously growing audience of researchers, designers and practitioners the great potentialities of newborn interaction styles. Physical environments lying in the range of these sensors, allow the users to move freely in everyday spaces, without anything to wear or to hold in their hands, and to produce some audio or/and visual feedback from a computer system, connecting in this way the digital contents to their movements or actions. My research focuses on these kinds of spatial interaction, and on how interfaces belonging to the real world can be arranged to deploy the uniqueness of these interaction styles. In particular, my design experiences concern music expressive interaction, music learning environments (see Fig.1) and music production (see Fig.2). Many musical features, like harmony or melodic movements, have been historically depicted by meaningful spatial representations, e.g. the *Euler's Tonnetz* (see Fig.3, literally "web of tones", a spatial schema showing the triadic relationships upon which tonal harmony is based), or the Gregorian Chant *Neumatic Notation* and *Chironomy*, a gestural system expressing melodic contour variations. The very notion of musical instrument involves precise spatial rules and element displacement, whose study has produced a large body of research about new music production interfaces in the field of the Sound and Music Computing. Another interesting example of music spatial interaction is the conductor model, where expressive performance informations are transmitted to performers through free-hands open-space gestures. This led me to consider the possibility of employing physical two dimensional surfaces or 3D spaces to project geometrical representations of musical concepts, thus enabling users to enter and to navigate conceptual maps representing some musical knowledge. The experimental hypothesis upon which my work is based, is that in this situation implicit and tacit knowledge may emerge and drive the users to also accomplish very complex tasks, like melody harmonization, without delivering them any previous information. Implicit knowledge refers to our brain's ability to acquire an abstract representation of any environmental stimulus (like for instance an unknown language), learning its rules in an unconscious way. Implicit learning is linked to the idea of enactivism, that is the ability of learning by doing. These important cognitive and psychological concepts are both involved in reality-based learning environments, which can deploy such powerful tools to convey information in a faster and more direct way. The aim of my work is also to study the cognitive content of musical knowledge and to try to expand the same learning methodology towards other domains, like the STEM disciplines (Science, Technology, Engineering and Mathematics), which very often benefit from being introduced through creativity and arts-based learning.

Research Outlook

In 2010 while attending the 7^{th} Sound and Music Computing Conference and Summer School in Barcelona I was introduced to the *Stanza Logomotoria* [18], a motion tracking, camera-based application where children are asked to match a story tale with sounds positioned on the mapped surface (see Fig.5). Further, the application was presented in some experimental sessions with music didactics students in Conservatories and elementary schools. Nevertheless, my actual research begun in 2012, when I graduated in Electronic Music with a master thesis on *Disembodied voices*, the first of the three case studies

Figure 2: A child playing a Disklavier through gestural input detected by a Leap Motion Sensor

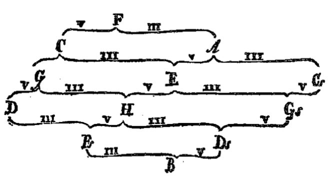

Figure 3: Euler's Tonnetz, 1739.

described below. This experience disclosed to me the great potentialities of reality based, full body and free handed interaction, especially when coupled with a well established human-to-human communication model, like the one of the music conductor. During my first year of PhD at the Department of Information Engineering of the University of Padova with the Sound and Music Computing Group (2013), I examined in depth various theories about the representation of music harmony spaces and music image schemas. In 2014 I realized my first camera space music application, the *Harmonic Walk*, the second of the three case studies reported below. By the end of 2015 I will conclude my PhD studies, after having developed several applications for entertainment, learning and creative and expressive music interaction. Presently, together with other members of the CSC research group, I am working to a spatial collaborative game to enhance music listening (*Good or Bad*) and to a spatial display for an interactive image sonification installation.

Related Work

My work is based on two different kinds of background. Firstly the analysis of the origins ([4] and [3]) and of the state of art of interactive spaces like

- *Google's Interactive Spaces and Liquid Galaxy* (http://www.interactive-spaces.org/),
- *Aarhus University Research Centre* (http://www.interactivespaces.net/,
- *Blended Interactions Design Studio, Rochester (NY)* (http://blendedinteractions.com/)
- *UCLA REMAP University of California* (http://openptrack.org/)

and blended interaction theory [5], providing also a short survey of interactive spaces definitions, typologies, available platforms and themes. I pay also particular attention to existing learning environments like

- *WizeFloor* (https://www.wizefloor.com/)
- *Smallab* (http://smallablearning.com/)
- *STEP* (http://remap.ucla.edu/research/cultural-civic-computing/791-science-through-technology-enhanced-play-step)

research centres, arts and culture installations and music production environments. Secondly, all the theoretical studies concerning space cognition ([11] and [17]), implicit knowledge ([14] and [16]), enactivism [2] and image schemas ([6] and [1]), which form an interconnected interdisciplinary field, very useful to understand how reality-based applications work. In Fig.4 a tentative conceptual map of the relationships existing among reality-based applications, implicit knowledge and space cognition is proposed, outlining reciprocal interplay and dependencies.

Three Case Studies

In this Section three case studies of music applications are presented. The three cases employ different motion tracking devices and convey musical concepts through geometrical interpretation and spatial representation. Also the projections refer to different spatial models, as the first employs the 3D spherical-polar coordinate system, the second 2D Cartesian coordinates on a flat floor surface and the third 3D space with x, y Cartesian coordinates plus z plane for depth data.

Disembodied Voices

[9] is an interactive environment designed for an expressive, gesture-based musical performance. The motion sensor Kinect, placed in front of the user, provides the computer with the 3D polar coordinates of the two

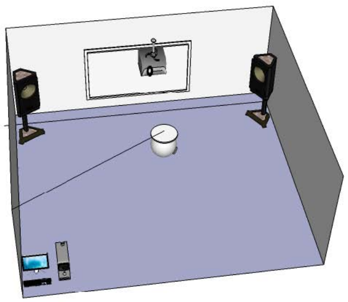

Figure 5: The "Stanza Logomotoria" basic configuration, with computer, audio monitors and ceiling mounted camera.

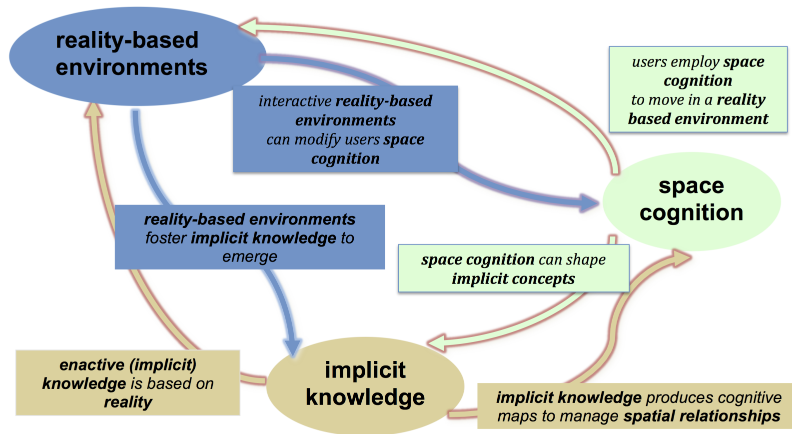

Figure 4: A conceptual map of the main relationships existing among reality-based environments (blue connections), space cognition (green connections) and implicit knowledge (brown connections).

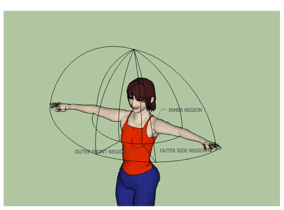

Figure 6: The hemispherical regions for user hands interaction in "Disembodied Voices". For the left hand there is an inner region and two outer regions, one in front and the other at the side of the user.

hands. The application is designed according to the interaction model of the music conductor: the user, through gestures, is able to run a score and to produce a real-time expressive interpretation. The conductor moves her/his arms and hands in the space around her/his torso and in the direction of the performers. Movement analysis ([10] and [12]) as well as the teaching practice[15] subdivide the role of the two hands; in general, the right executes musical cues while the left is devoted to iconics, metaphorics and dynamics. As can be seen from Fig.6, the geometrical interpretation of the conductor's interaction space is a hemisphere with the center at the level of the breastbone of the conductor and the diameter corresponding approximately to the two stretched arms length. Following the conductor's interaction model, the hemisphere is subdivided in two parts, one for the right and the other one for the left hand. For the left hand three different regions are shaped, triggering various digital sound processing effects. (video available at `https://www.youtube.com/watch?v=oyf7GrMMrL8`)

The Harmonic Walk

[8] is an interactive physical environment designed for experiencing a novel spatial approach to musical creation. In particular, the system allows the user to get in touch with some fundamental tonal music features in a very simple and readily available way. The application's interface consists of a camera placed on the ceiling which can trace the presence of a user who walks on a flat surface within the camera's view. The *Harmonic Walk*, through the body movement in space, can provide a live experience of tonal melody structure, chord progressions, melody accompaniment and improvisation. Enactive knowledge and embodied cognition allow the user to build

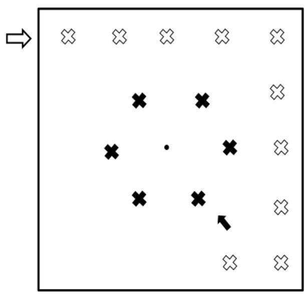

Figure 7: Visual tags of the straight and circular path of the "Harmonic Walk's" interface.

an inner map of these musical features, which can be acted by moving on the mapped surface with a simple step. Listeners interpret a tonal melody grouping the perceived sequence of events after a metrical and harmonic frame [13]. This produces a segmentation of the composition into different harmonic regions which, with the underlying harmonic structure, are the leading features of a tonal composition. The time proceeding of the various musical units is led by the melody, whose metaphoric scheme is expressed by the so-called "source-path-goal" schema [6]. Following this metaphor and employing the simplest motion in space a human can do - the walk - it is possible to represent a tonal composition as a sequence of spatial blocks, where each step corresponds to the next musical unit (white crosses in Fig.7), while the harmonic space is represented by a six parts sliced circular ring containing the six roots of the tonality harmonic space (black crosses). (video available at https://www.youtube.com/watch?v=0jwXfzq_CkU&index=1&list=UU1E9xCq8TWqlzessRIzUGxw)

Hand Composer
[7] is a gesture-driven composition system, based on the analysis of the existing relationships between music generative models and musical composition in the context of the XX century music history background. The system framework is based on a number of interactive machines performing various patterns of music composition and producing a stream of MIDI data to be compatible with a Disklavier performance. Hand gestural input, captured by the Leap Motion sensor (see Fig.8), can control some parameters of the music composing machines, changing interactively their musical output. (video available at https://www.youtube.com/watch?v=mdsn9_5Ig_A&list=UU1E9xCq8TWqlzessRIzUGxw)

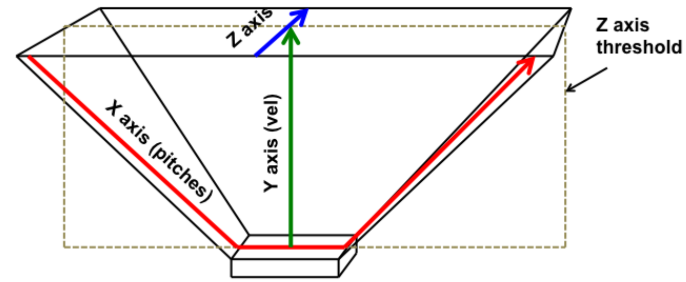

Figure 8: The 3D interaction space of the *Hand Composer* application. x, y Cartesian coordinates are employed to map the two-dimensional vertical plane, while z coordinates map the depth data.

Thesis Scope and Expected Contributions

The experimental hypothesis upon which all my research is based, is that abstract knowledge like musical harmony or tonal music composition structure can be conveyed to the users through the enactive learning induced by interactive space environments. Thus, my thesis will focus on the relationship between space and cognition, and on the ways knowledge can be spatially represented. As I have worked with various kinds of interactive spaces (bi-dimensional floors as well as three-dimensional volumes), I will try to work out a comprehensive, unifying framework for all of them and to highlight their respective properties and differences. If my experimental hypothesis is verified, reality-based learning environments design could benefit of very powerful cognitive tools, which could be extended beyond music also to other knowledge fields, like science, mathematics and technology.

Acknowledgements

I am grateful to my supervisor Sergio Canazza for his precious support and help. I also want to express my gratitude to Antonio Rodà who followed step by step the development of my work and to Davide Rocchesso for his advices and presence.

References

[1] Brower, C. A cognitive theory of musical meaning. *Journal of Music Theory* (2000), 323–379.

[2] Bruner, J. S. The act of discovery. *Harvard educational review* (1961).

[3] Engelbart, D. C. Augmenting human intellect: a conceptual framework (1962). *PACKER, Randall and JORDAN, Ken. Multimedia. From Wagner to Virtual Reality. New York: WW Norton & Company* (2001), 64–90.

[4] Harrison, S., and Dourish, P. Re-place-ing space: the roles of place and space in collaborative systems. In *Proc. ACM conference on Computer supported cooperative work*, ACM (1996), 67–76.

[5] Jetter, H.-C., Reiterer, H., and Geyer, F. Blended interaction: understanding natural human–computer interaction in post-wimp interactive spaces. *Personal and Ubiquitous Computing 18*, 5 (2014), 1139–1158.

[6] Lakoff, G., and Johnson, M. *Metaphors we live by*. University of Chicago press, 2008.

[7] Mandanici, M., and Canazza, S. The "hand composer": gesture-driven music composition machines. In *Proc. 13th Conf. on Intelligent Autonomous Systems* (2014), 553–560.

[8] Mandanici, M., Rodà, A., and Canazza, S. The harmonic walk: an interactive educational environment to discover musical chords. In *Proc. ICMC-SMC Conference* (2014).

[9] Mandanici, M., and Sapir, S. Disembodied voices: a kinect virtual choir conductor. In *Proc. 9th Sound and Music Computing Conference* (2012), 271–276.

[10] Marrin, T., and Picard, R. The 'conductor's jacket': A device for recording expressive musical gestures. In *Proc. International Computer Music Conference* (1998), 215–219.

[11] Montello, D. R. *International Encyclopedia of the Social and Behavioral Sciences*. Oxford, Pergamon Press, 2001, ch. Spatial Cognition, 14771 – 14775.

[12] Murphy, D., Andersen, T. H., and Jensen, K. Conducting audio files via computer vision. In *Gesture-based communication in human-computer interaction*. Springer, 2004, 529–540.

[13] Povel, D.-J., and Jansen, E. Harmonic factors in the perception of tonal melodies. *Music Perception 20*, 1 (2002), 51–85.

[14] Reber, A. S. Implicit learning and tacit knowledge. *Journal of experimental psychology: General 118*, 3 (1989), 219.

[15] Rudolf, M., and Stern, M. *The grammar of conducting: A comprehensive guide to baton technique and interpretation*. Schirmer Books, New York, 1994.

[16] Tillmann, B., Bharucha, J. J., and Bigand, E. Implicit learning of tonality: a self-organizing approach. *Psychological review 107*, 4 (2000), 885.

[17] Tversky, B. Functional Significance of Visuospatial Representations. *Handbook of higher-level visuospatial thinking Handbook of* (2005), 1–34.

[18] Zanolla, S., Canazza, S., Rodà, A., Camurri, A., and Volpe, G. Entertaining listening by means of the Stanza Logo-Motoria: an Interactive Multimodal Environment. *Entertainment Computing 2013*, 4 (2013), 213–220.