Composing Interactive Dance Pieces for the MotionComposer, a device for Persons with Disabilities

Andreas Bergsland
Norwegian University of Science and Technology (NTNU)
7491 Trondheim, Norway
andreas.bergsland@ntnu.no

Robert Wechsler
Motion Composer
Karl-Haussknechtstr. 11
99423 Weimar, Germany
robert@motioncomposer.org

ABSTRACT

The authors have developed a new hardware/software device for persons with disabilities (the MotionComposer), and in the process created a number of interactive dance pieces for non-disabled professional dancers. The paper briefly describes the hardware and motion tracking software of the device before going into more detail concerning the mapping strategies and sound design applied to three interactive dance pieces. The paper concludes by discussing a particular philosophy championing transparency and intuitiveness (clear causality) in the interactive relationship, which the authors apply to both the device and to the pieces that came from it.

Author Keywords

Interactive dance, interactive music, mapping, persons with disabilities, abilities, dance therapy, music therapy

ACM Classification

I.4 [IMAGE PROCESSING AND COMPUTER VISION] Scene Analysis, J.5.8 [ARTS AND HUMANITIES] Performing arts --- Motion

1. INTRODUCTION

The authors have worked together since 2011 on the parallel tasks of motion tracking and designing interactive music environments in search of stimulating and aesthetically pleasing movement-music experiences. This has been done as a part of the development of the MotionComposer (MC), a device that uses video and 3D sensor-based motion tracking to turn movement into music. It was designed with special consideration for users with disabilities and has during the last three years been developed and tested in over a dozen workshops with users in this category. The design of the interactive music environment *Particles* [1], which is one of six different environments implemented in the MotionComposer device, has been the main focus of the authors' collaboration. However, during this period, the MC has also proven to be highly suited not only for therapeutic use, but also for purely artistic purposes, involving performers both with and without disabilities. More specifically, we have developed three performance pieces, *Jeu de modes*, *La dance II* and *Songshan Mountain*, all within the genre of *interactive dance* [12]. These pieces have all been presented at international festivals and conferences during the last few years with different performers.[1] After a brief presentation of the technical aspects of the devicewe will proceed to discuss mapping strategies and sound design issues in these pieces.

NIME'15, May 31-June 3, 2015, Louisiana State University, Baton Rouge, LA.

2. MOTION TRACKING WITH MC

The MC is a device combining motion-tracking sensors with sound generating software running on a mini computer (Mini-ITX motherboard with an Intel i7 processor). It uses a Kinect along with a 1.3 mega pixel Ethernet camera. The Ethernet camera provides the low-latency and high-resolution images that are crucial for achieving a synaesthetic (or sense-confusing) response, i.e. this would not be possible with the Kinect alone do to its limitations.[2] In addition, the device is equipped with a Kinect1 3D sensor, used for tracking the user's location and posture, among other things.[3] It has similarities with different video-based tracking systems that have been developed over the years like the Very Nervous System [17], Big Eye from STEIM [13], and EyeCon by Weiss and Wechsler [16], all falling under the category of *outside-in systems*, according to Alex Mulder's categories, as presented by Siegel [12].

The video images and 3D data are interpolated and analyzed by software developed by Simone Ghisio and Paolo Coletta in the EyesWeb environment [2]. The software then produces up to 20 streams of high-rate control data passed via OSC on to the music software. These parameters are sent to the real time sound generation software where sound output is created. The parameters relevant to the discussion in this paper are (see [8] for more details):

- Quantity of activity (QoA)
- Centre-X (centre of the user's horizontal position)
- Height

---

[1] This includes CYNETART (Dresden, Germany), New York City Electroacoustic Music Festival (NYCEMF), re-new (Copenhagen), and ICMC/SMC (Athens, Greece). Videos can be seen at http://www.palindrome.de -> videos -> recent.

[2] Latency in such interactive systems is often misunderstood. It is rarely related to frame rate, but rather to much more significant delays resulting from the busing and processing of the signal. More details on the technical aspects of MC can be found at http://internal.motioncomposer.com/technical (accessed January 26th 2015).

[3] The developer team is currently doing tests with Kinect 2 with the aim of implementing it in the system.

3. MAPPING STRATEGIES

The issue of *mapping* is often taken to be crucial importance in creating expressive performances with digital music instruments (DMIs) [15] [11]. During the last few decades, the issue of mapping has been explored by a number of researchers, both theoretically, experimentally and related to specific applications [11] [5-7; 9] [14]. Of the two main approaches to mapping delineated by Hunt and Wanderley, *generative* and *explicit*, the latter approach appears to have had more attention by researchers [7]. This approach is usually classified into four (or sometimes three) *mapping strategies* depending on whether one or several performance/control parameters are mapped to one or several synthesis parameters [7] [11; 14]:

- one-to-one
- one-to-many or *divergent*
- many-to-one or *convergent*
- many-to-many

In several of the referenced studies it is implied that the *many-to-one/convergent* and *many-to-many* strategies provide richer, more interesting or more expressive interaction than the one-to-one strategy [4; 6; 11]. While this might sometimes be the case, we would argue that all of these strategies, from simple one-to-one to more complex many-to-many can provide expressive possibilities. The critical issue in terms of user engagement is how the environment evolves over time, i.e. how the user is guided back and forth between "causal-ordered-predictable" and "intuitive-improvised" processes. This is to say that either, at exclusion of the other, can quickly lose interest.

3.1 Simple mappings

In working with users with and without disabilities over the years we have learned that users as well as those watching them seem to enjoy experiencing a relatively direct and intuitive relationship between the users' movement and the resulting sound.

One type of mapping we have experienced as effective in that respect is letting QoA, calculated by frame-by-frame subtraction [3], be mapped to sound intensity via a gain factor. This simple one-to-one mapping is highly dependent on low-latency response of the Ethernet camera component of the system to achieve the sensation of immediate response for the user. With a static synthetic sound, this simple mapping strategy might be experienced as dull or uninteresting. However, if the gain factor modulates the output of a sound sample with a more complex quality and some temporal variation, the result can be quite interesting for the user and/or audience.[4] Admittedly, the consistency in the movement-sound relationship and the sense of *control* won't be absolute, but the close relationship between the experienced amount, size or energy of movement and the amount or intensity of sound is nevertheless maintained. With careful design of the sound sample, balancing identity/recognizability with variation and complexity, the user experience might be rewarding.

We have applied this mapping and sound design strategy to several of the works mentioned above, for instance in of the intro solos of *Jeu de modes* [1:31-1:55]. In this case, the interactive sound is combined with a fixed sound file functioning as a sonic backdrop that gives the overall composition a direction at the same time as it binds the different sections together. The same mapping strategy is applied in section 3 of *La dance II* [2:00-4:20] using live-sampled sounds, and the two B-sections of *Songshan Mountain* [0:46-1:30, 1:58-2:38] using pre-recorded sounds.[5]

3.2 The *6s*: Layered mappings

Another, and somewhat more complex mapping strategy is applied in the *6s* section of *Jeu de modes* [03:20-06:03]. In this trio, it is the quantity of activity for all dancers together that is tracked. While this might seem to imply the individual dancer's sounds would be more difficult to discern, in fact the opposite can be true. "Unison" or synchronicity is a bit of a myth, since artists are not perfect copies of one another. Dance "breathes" both literally and figuratively, implying adjustment and change. Thus this section becomes interesting (and interactive) precisely because the system makes the small differences (to wit, inaccuracies) between the dancers' movements discernible though sound.

In this section, the QoA parameter was used to control the output gain of seven synchronized rhythmical streams. These streams contained chains of sound samples with an impulse-resonance morphology, all with different subdivisions of a 6/4 pulse (Figure 1). By letting each of the streams have different mapping functions/curves where the higher subdivisions would generally "kick in" at a higher activity level than the lower ones, we got the effect that more movement would produce both a higher overall sound intensity and a more busy and multi-layered groove.

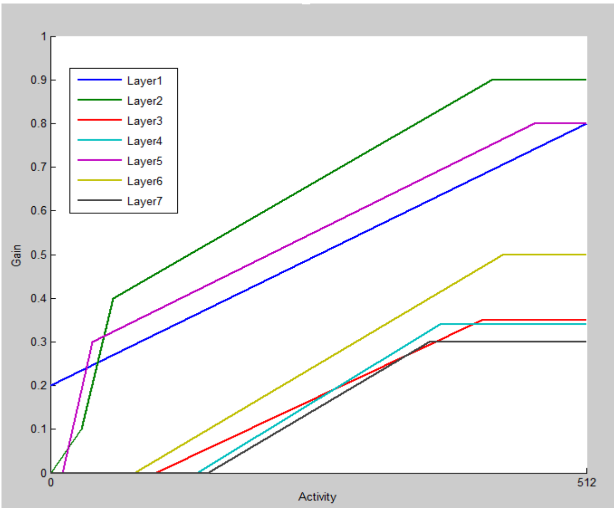

Figure 1. Quantity of activity mapped to layer gain.

While this mapping creates relatively interesting rhythmic textures in itself, we wanted to add further sonic interest and development matching the structure of the choreography of the *6s* section with its five different "variations", applying different but loosely related movements. To achieve this we used a relatively large bank of samples containing a lot of variation and organized it according to sonic similarity. The different variations in the section would then cycle through the sample bank, ensuring variation and interest, but also a continuous development through the section.

[4] This naturally requires a sound with a relatively constant gain level.

[5] In the latter case, however, we realize that the fixed background sounds take on too much of a foreground function, thereby making it more difficult for the audience to experience the interactive part (a noisy texture).

3.3 *Particles* environment: complex mapping

In two of the pieces, *Jeu de modes* and *Songshan Mountain*, we have applied a mapping strategy that we have developed over some years through the work with the *Particles* environment of the MC [1]. As implied by its name, the environment is based short sound objects with duration of less than one second (0.15-1s.). These short sound objects are organized in sets, each defining the sound “flavor” of one single environment. The particles can then be played individually, but in most cases, they will form sequences or more complex sonic textures. Thus, the system could be characterized as a form of corpus based concatenative synthesis, with some similarities to systems like CataRT [4], but where the combination of samples is based on pre-defined musical criteria.

The basic idea behind this environment is that the localization of the user in the room orthogonal to the camera determines the *choice* of sound particle (Figure 2, reprinted from [1]), and that the QoA determines the *frequency* with which the chosen particle(s) are triggered. Together, these two basic features make up an environment that is highly intuitive, in that the user can feel and hear every movement, and when provided with a well-designed set of sounds, it can be very interesting sonically. However, with the goal of making the environment more organic and minimizing the feeling of static and repetitive sample playback, we have over the years made several modifications to the basic idea.

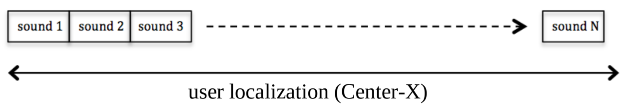

Figure 2. User localization orthogonal to camera direction is mapped to choice of sound particle.

In the current version, the mapping strategy is a combination of one-to-one, one-to-many, and many-to-one (Figure 3). Of the three parameters controlling the main mode of the environment, QoA clearly has the most complex mapping, affecting both the triggering frequency of the sound particles, overall gain and the sound particle envelope, in addition to affecting the random deviation of the choice of sound files.

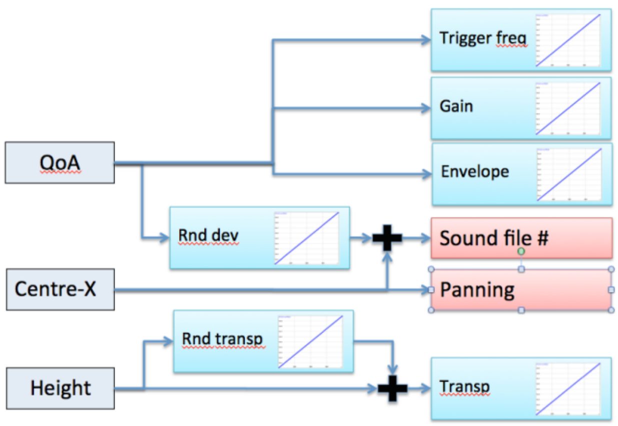

Figure 3. Parameter mapping in *Particles*.

Another factor adding to the complexity is that all of these parameters index a set of transfer functions consisting of break-point tables, thus changing the behavior for different ranges of the parameters. For instance, we have taken care to let smaller gestural movements differ qualitatively from larger ones, so that for the former, one can often discern single particles, whereas for the latter the result resembles more a sound mass or cloud. Moreover, when standing completely still users will have access playback of single sound particles, what we like to call *sensitives* [1].

The mapping of height to sound file transposition adheres to the conventions linking physical height and pitch. This may seem naïve, but together with the sonic complexity created by the high number of sound files, it nevertheless becomes a much-appreciated feature that users relate to strongly while avoiding a banal sonic result.[7]

3.4 Interaction affordances

The interactive environments that we have composed in the three discussed pieces all have their particular *affordances*; certain potential action relationships between the qualities of the environment and capabilities of the user [14]. At the core of all three, however, is that the environments invite the users to move, and to feel that there is an overall correspondence between the size, amount or energy of movement and the intensity of the sounds.[8] Hence, the environments afford exploration of a wide dynamic range of movements, from the tiniest eye blinks, to the most energetic leaps off the ground.

But, not all kinds of movements will sound equally good. For instance, keeping a relatively constant level of movement by walking around will tend to produce a continuous flow of sound that can be tiring to listen to in the long run. When the users try to shape *intended gestures* having defined beginning and ending with their body movements, however, the results tend to be sonically more interesting. Therefore if a user is listening to what she is playing, the environment will *afford* these intended gesturally shaped movements.

Moreover, the high sensitivity of the motion tracking hardware has enabled us to explore the active use of *stillness* as a parameter in the interaction. The *sensitives* in the particles environment suddenly makes stillness much more interesting than what is common in interactive environments, precisely because it is only by being completely still, and then move a little, that one can produce these single sound particles.

Lastly, for the two pieces applying *particles*, the environments also afforded aural exploration, since by moving in the room one would hear different groups of sounds. For some of our dancers who were used to moving to fixed music, learning to actually listen to the musical result of their movements took a bit of practice, but after a time they developed a sense of what sounded good in their ears and started to implement this in their structured improvisations. In other words, they developed a musical expressivity along with their bodily expressivity.

4. THE MC – A MUSICAL INSTRUMENT?

The MC differs from traditional musical instruments as well as DMIs based on the musical instrument paradigm in a number of respects. For one thing, musical instruments are generally played with the extremities: fingers, hands, mouth, and sometimes the feet. Dance, it is said, comes from the center of the body, the solar plexus. Dance is based more on full-body movements, shifts of weight, swings, extension and contraction, angular momentum (turning or spinning movements). The MC project concerns encouraging creative movement of all kinds, not just the accurate and task-oriented movements that we tend to think of when we think of musical instrument playing.

---

[7] See http://palindrome.de/motioncomposer for user statements.

[8] This is partly related to, but not the same as, *effort*, since it can actually take a lot of effort to stand completely still.

Another way that this system differs from traditional music instrument playing is that the degree of control that one has over the music varies over time. As the user interacts with the environment, our concerns are, 1) that it engages their interest from the outset, and, 2) that it remains interesting.

This simple-sounding demand belies deeper psychological principles. On the one side, the experience needs to be convincing – the user must "get" the causal relationship. This demands gestural and aural discreetness, repeatability and relative simplicity (since there is not a haptic or tactile experience involved, it is easy to lose the sense of causality).

On the other side, there is the need for an artistically satisfying experience – *even for users who are not musicians*. Hence one needs to separate the experience that one is making music, from *how the music sounds*. A delicate balance must be achieved between discreet, repeatable (causal) events, and more subtle, complex and non sequitur elements. Note: it is not simply a matter of blending the two, for they are dichotomous. Rather an alternation is required if we want the experience to remain engaging over any length of time.

5. CONCLUSION

In one study [10] using this device, it was shown that the level of engagement depends little on a person's ability. That is, persons with and without disabilities were evaluated for their level of engagement and this was found to be roughly the same in both groups. This early finding would seem to support the assertions made above.

With careful attention to issues such as those described in this paper, such environments can precipitate creative movement in most users, and this has wide ranging health benefits both physical and psychological, particularly for persons who, for reasons for mental or physical disability are limited in expression [8].

While designing NIMEs in which the user can develop skills and virtuosity through practice has been advocated as a design goal by many (e.g.[4]), we have instead focused on making a device that will respond to any movement with aesthetically pleasing sounds. In the future, one could imagine systems capable of meeting both goals – which are both immediately satisfying and also allow for sophisticated control and musicianship in the traditional sense of user-achievement.

6. ACKNOWLEDGMENTS

Our thanks to Bauhaus University, Franz Liszt College of Music (facilities and equipment); Marcello Lussana (technical support); Madeleine Shapiro (NYCEMF); Teoma Naccarato, Annika Dörr, Jossia Clement and Tomas Mader (dancers); Fond for utøvende kunstnere (FFUK); Music Norway, the Norwegian Embassy in Athens, Norwegian University for Science and Tehcnology (financial support); the IMM-group and Infomus (development of Motion Composer device).

7. REFERENCES

[1] Bergsland, A. and Wechsler, R., 2013. Movement-Music Relationships and Sound Design in MotionComposer, an Interactive Environment for Persons with (and without) Disabilities. In *Proceedings of the re-new* (Copenhagen, 2013), 56-62.

[2] Camurri, A., Coletta, P., Varni, G., and Ghisio, S., 2007. Developing multimodal interactive systems with EyesWeb XMI. In *Proceedings of the 7th international conference on New interfaces for musical expression* (New York, 2007), ACM.

[3] Camurri, A. and Moeslund, T.B., 2010. Visual Gesture Recognition. In *Musical Gestures. Sound, Movement, and Meaning*, R.I. Godøy and M. Leman Eds. Routledge, London, 238-263.

[4] Dobrian, C. and Koppelman, D., 2006. The 'E' in NIME: musical expression with new computer interfaces. In *Proceedings of the Proceedings of the 2006 conference on New interfaces for musical expression* (Paris, France2006), IRCAM; Centre Pompidou, 1142283, 277-282.

[5] Hunt, A. and Kirk, R., 2000. Mapping strategies for musical performance. In *Trends in Gestural Control of Music*, M. Wanderley and M. Battier Eds. Ircam, Paris, 231-258.

[6] Hunt, A., Wanderley, M., and Paradis, M., 2003. The importance of parameter mapping in electronic instrument design. *Journal of New Music Research 32*, 2, 429-440.

[7] Hunt, A. and Wanderley, M.M., 2002. Mapping performer parameters to synthesis engines. *Organised Sound 7*, 02, 97-108.

[8] Kontogeorgakopoulos, A., Wechsler, R., and Keay-Bright, W., 2014. Camera-Based Motion Tracking and Performing Arts for Persons with Motor Disabilities and Autism. *Disability Informatics and Computer Access for Motor Limitations, ed. Georgios Kouroupetroglou, IGI Global*, 294-322.

[9] Murray-Browne, T., Mainstone, D., Bryan-Kinns, N., and Plumbley, M., 2011. The medium is the message: Composing instruments and performing mappings. In *Proceedings of the Proceedings of the International Conference on New Inerfaces for Musical Expression 2011* (Oslo, 2011), 56-59.

[10] Peñalba, A., Valles, M.J., Partesotti, E., Castañon, R., Sevillano, M.A., and Wechsler, R., 2015. Types of interaction in the use of Metabody Box (MotionComposer), a device that turns movement into sound. In *Proceedings of the International Conference on Multimodal Experience of Music* (Sheffield, UK, 2015).

[11] Rovan, J.B., Wanderley, M., Dubnov, S., and Depalle, P., 1997. Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. In *Proceedings of the Kansei, The Technology of Emotion. Proceedings of the AIMI International Workshop* (Genova, October 3-4 1997), Associazione di Informatica Musicale Italiana, 68-73.

[12] Siegel, W., 2009. *Dancing the Music: Interactive Dance and Music*. Oxford University Press, Oxford.

[13] STEIM, n.d. Steim - legacy products. http://steim.org/product/discontinued-products/. 28.01.2015.

[14] Tanaka, A., 2010. Mapping Out Instruments, Affordances, and Mobiles. In *Proceedings of the International Conference on New Interfaces for Musical Expression* (Sydney, Australia, 2010), 88-93.

[15] Wanderley, M., 2001. Gestural control of music. In *International Workshop Human Supervision and Control in Engineering and Music*, Kassel, Germany, 101-130.

[16] Wechsler, R., Weiß, F., and Dowling, P., 2004. EyeCon: A Motion Sensing Tool for Creating Interactive Dance, Music, and Video Projections. In *Proceedings of the Proceedings of the AISB 2004 COST287-ConGAS Symposium on Gesture Interfaces for Multimedia Systems* (Leeds, UK2004), Leeds, UK, 74-79.

[17] Winkler, T., 1997. Creating interactive dance with the very nervous system. In *Proceedings of the The 1997 Connecticut college symposium on art and technology* (New London, Connecticut, 1997), 212-217.