# Movement-Music Relationships and Sound Design in MotionComposer, an Interactive Environment for Persons with (and...



Andreas Bergsland
Norwegian University of Science and Technology

# Movement-Music Relationships and Sound Design in MotionComposer, an Interactive Environment for Persons with (and without) Disabilities

[1]Andreas Bergsland, [2]Robert Wechsler

[1] Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway.
andreas.bergsland@ntnu.no

[2] Der MotionComposer, Karl-Haussknechtstr. 11, 99423 Weimar, Germany.
robert@motioncomposer.com

**Abstract.** MotionComposer (MC) is a device that turns human movement into music using state-of-the-art video based and 3D sensing motion-tracking technology combined with tailor made sound generating software. The device is developed particularly for disabled people, both for therapeutic reasons (encourage activity, stimulate motor control and sensing, coordination etc.), emotional expression, entertainment, and aesthetic pleasure. In this article, we will give a brief outline of how the device works. We then discuss movement-music relationships in general and how these can be mapped in video based systems like MC. In particular, we go into details about these questions regarding the *Particles* environment, one of several interactive environments included in the MC. For the latter, we also discuss issues of sound design, including narrative versus abstract strategies and different ordering approaches.

## 1 Introduction

Motion tracking technologies have gradually found their way into a number of areas in society the last decades, including in research, arts and performance disciplines [1]. These technologies detect human movements and turn them into analogue and/or digital signals, which in turn can be analyzed, converted and/or manipulated. Camera based motion tracking is an important part of this technological landscape, and the advantages of being non-intrusive and being able to detect holistic body movements has made it popular in dance, music and art installations [2]. These advantages also make this technology suitable for applications that aim at letting non-musicians express themselves musically through bodily movements and most especially persons with disabilities. MotionComposer is one such application. In this paper we will discuss some of its salient features.

## 2 MotionComposer – the Device

The MotionComposer (MC) is a device that turns human movement into music using state-of-the-art video based and 3D sensing motion-tracking technology combined

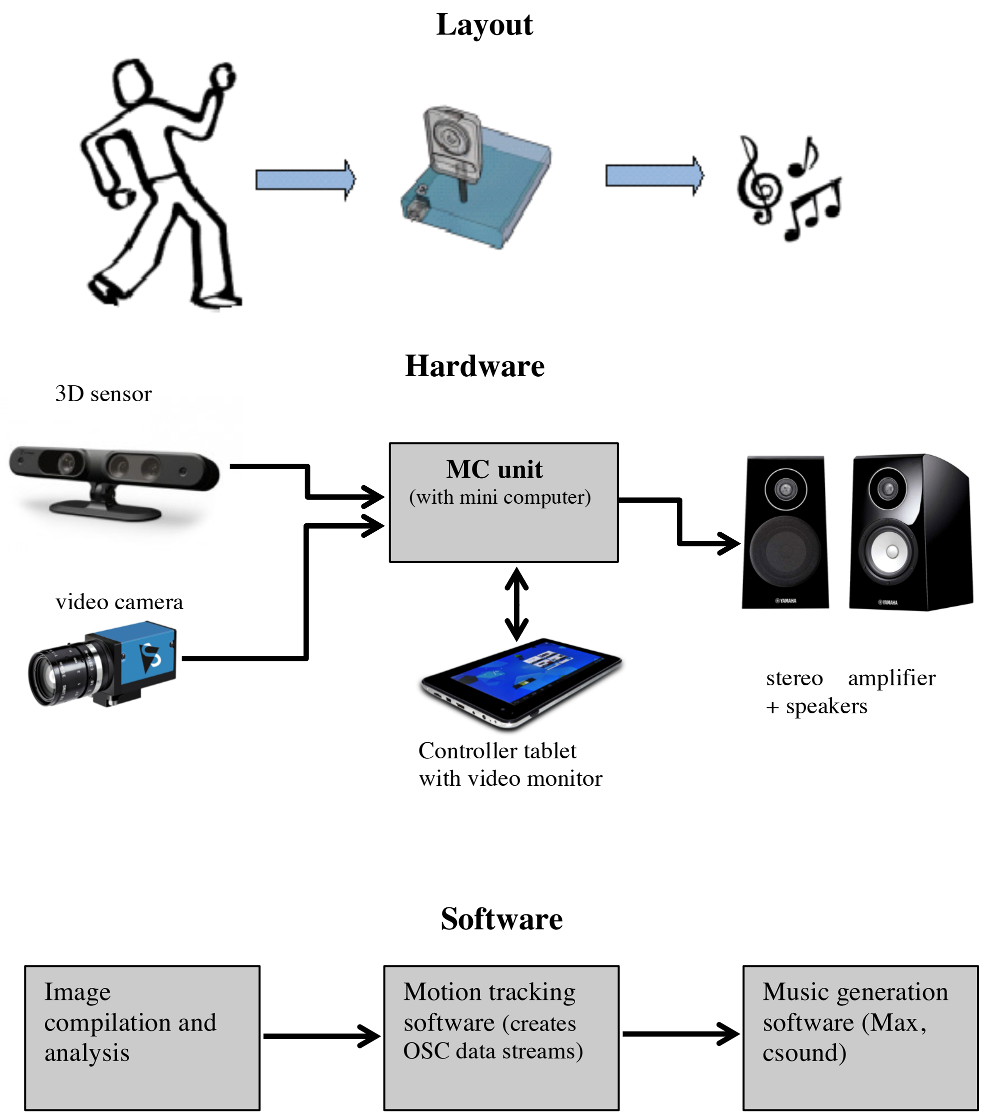



**Fig. 1.** The MotionComposer device. Layout (top), hardware (middle), and software (bottom)

with tailor made sound generating software (Figure 1).[1] It contains a video camera providing low-latency and high-resolution images, crucial for achieving "immediate" response from the system and with high accuracy. In addition, the device will be equipped with a custom 3D sensor, providing depth information, necessary for producing data about the user's localization and posture in front of the sensor, among other things.

[1] It is currently under development, but most of the crucial hardware and software elements are already fully functional. While the technological elements are not yet integrated in a physical unit, the development plan includes the eventual integration of all technology for motion tracking and sound production in one single chassis.

The high-resolution video images and 3D data are first fed into a mini computer running custom motion tracking software. This software then produces many streams of high-rate control data (using the OSC protocol), which is subsequently passed to sound generation software where sound output is created based on the data streams. A tablet, wirelessly connected to the box, provides control of the functionality of the box along with a video feed from the camera. The user can thus localize him/herself in front of the camera and choose between available settings while the box remains stationary on a stable surface facing the user.

While the MC device is thought to be a commercial product directed towards disabled users, we also want to apply the device in artistic projects, both because many of the developers have a background in artistic and performance disciplines, but also because we feel that the dynamics of artistic processes can open up new possibilities for the device. Lastly, the state-of-the-art technology implemented in the MC device lends itself very well to a range of artistic applications, including interactive dance performance and art installation.

## 3 Movement-Music Relationships

### 3.1 General Considerations

Dance and music are two expressions of a single urge. Many languages have a single word for both dance and music (English does not). In Portuguese, when I say "Capoeira", I am referring to something that is in equal parts a music and movement tradition. Small children begin to dance and make music without any understanding of what they are doing or clear distinction between the two. Dance and music have been practiced by every people who have ever walked the earth.

There are many ways to turn movement into movement; musical instruments for example. Still, the MC differs from musical instruments in a number of respects. For one thing, musical instruments are generally played with the extremities: fingers, hands, mouth, sometimes the feet. Dance, it is said, comes from the center of the body, the solar plexus. Dance is based on full-body movements, shifts of weight, swings, extension and contraction, angular momentum (turning movements). The MC project concerns encouraging creative movement of all kinds, not just the "controlling" kind we use when we want a repeatable result.

Indeed, the MC limits ones control over the sound. For example, while it might give the user a high degree of temporal control, the choice of *which* sound is heard might be elusive and, to an extent, pre-selected by the software of the device with the intent that the sounds that come out of it are attractive. And there is another difference: When we make musical sounds directly through our gestures in space, it alters our proprioception. Our self-image is extended beyond the boundary of the physical body and into the (aural) environment around us. When these sounds are beautiful, and when we are convinced that we made them, an innate sense of joy and confidence results and it means that one movement leads to the next and the next. Explorative, creative movements is the result 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 [2].

To convince the user that he or she has made music merely by making gestures in front of a camera, it is of crucial importance that the *mapping* of the system, i.e. how the movement analysis data is used to control musical parameters, is carefully set up. Through extensive experimentation and user testing, the MC team has over the years established a set of convincing mappings. In Figure 2, we have included a few of the motion tracking parameters we use along with some suggestions of how they can be applied (for a more comprehensive list, see [2]).

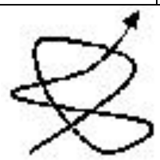
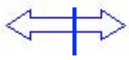
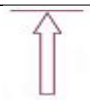
| # | pictogram | name | Description |
|---|---|---|---|
| 1 |  | Small Movements | When the body is still, our focus naturally goes to the small and precise movements of our fingers, hands, eyes, mouth, etc. Because these gestures are discrete -- with clear beginnings and endings -- they lend themselves to strong association with discrete sounds, for example single notes in a scale or sound particles. |
| 2 |  | Gestural Movements | These are the typical movements we do with arms, hands and head (though any body part may be involved). They are fairly discrete, but in contrast to Small Movements, we tend to experience them as continuous modulations, curving up and down along with our movements. |
| 3 |  | Large Movements | Adults generally make large movements only during sports, running for a bus or in a genuine emergency. In any case, Large Movements are associated with high volume, high velocity (in musical instruments) and perhaps distortions of the same sounds applied to Gestural Movements. They can be mapped as individual "bangs", or bursts of high energy and, like Small Movements, are Boolean in nature. |
| 5 |  | Stillness (not moving) | While often overlooked, stillness is actually a special activity. It is not merely the absence of motion, since it generally requires concentration with specific intent. Nevertheless, it is something most people can do without much practice. |
| 6 |  | Center-X | Assuming we have a sense of communication in our movements, we usually assume a direction in the room we call "front". Movement perpendicular to this direction we refer to as Center-X. Center-X offers a one-dimensional location-orientation and maps well to content-bearing sound elements. |
| 7 |  | Width | Width is like expansion. When we stretch out our arms or legs we grow in size. Increasing loudness is a simple option, but there are other sound transformations that may more closely resemble this action. For example, think of the formant changes that occur as the mouth is opened wider and wider while speaking or singing. |
| 8 |  | Height | Height maps intuitively to pitch (up = higher pitch), but as with width, there are other implications as well. e.g. low = rumbling, tumbling, grumbling, growling vs. high = thin, flighty, suspended, stretched. Indeed, when a voice goes from low to high it is not simply the pitch that changes! |

**Fig. 2.** Some of the motion tracking parameters applied in the *Particles* environment, along with suggestions of mappings.

### 3.2 Movement-Music Relationships in the *Particles* Environment

One approach we are using is to base the aural environment on sound particles, short sound objects with duration of less than one second (0.15-1s.).[2] Naming the

[2] This is what Curtis Roads would define as belonging to the *sound object time scale*: "A basic unit of musical structure, generalizing the traditional concept of note to include complex and

environment *Particles* was therefore quite evident. These sound particles are either played individually, or in most cases, chained together in sequences. Thus, the sound synthesis part of the system for this environment could be characterized as a form of corpus based concatenative synthesis, with some similarities to systems like CataRT [4].[3]

The sound palette of the *Particles* environment is based a high number of unique pre-recorded sound files organized in sets, each defining the sound "flavor" of one single environment. The user (or the therapist) can then switch between these sets using the controller tablet. One set should preferably have more than 50 sound particles to ensure enough sonic variability in the environment, and in many cases, one set can have several hundreds. In addition to providing variability, a high number of sounds in a set will warrant that even small changes in location will cause a change in sound particle, since the more particles one has in a set, the less one will have to change in localization to change the sound. This is important, because repeated triggerings of the same sound can quickly feel tiring to the ear.

**Basic mapping.** The mapping of this environment is based on the idea that that the localization of the user in the room (*Center-X* in Figure 2) determines the *choice* of sound particle (illustrated in Figure 3), and that the quantity of bodily activity (*Small Movements*, *Gestural Movements* and *Large Movements* in Figure 2) determines the *rate* in which the chosen particle(s) are played back (illustrated in Figure 4). Together, these two basic features make up an environment that is both highly intuitive, in that the user can feel and hear every movement, and very interesting sonically, at least if one is careful in the choice of sounds and their spatial layout. The idea is that the *Particles* environment should arouse not only an urge to move around, but at the same time feel these movements in the sound, but also that it should evoke interest and curiosity in the user for exploring the sonic richness and possibilities that the environment offers.

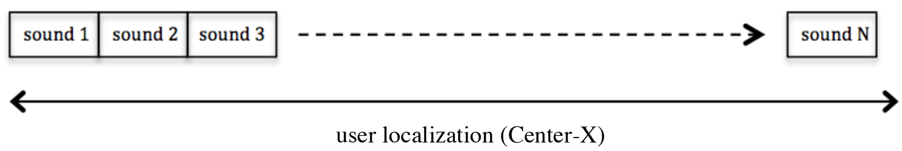



**Fig. 3.** User localization orthogonal to camera direction is mapped to choice of sound particle.

**Sensitives.** From a mapping perspective, the *Particles* environment can offer the possibility of marking out discrete and precise points in time.[4] This creates an opportunity for the direct correlation of single sound particles and *Small Movement*

mutating sound events on a time scale ranging from a fraction of a second to several seconds" [3].

[3] Although this system has implemented off-line feature analysis and sorting based on these features, the system does not currently use it.

[4] This depends on, however, that the sound particles have an abrupt and marked attack that offer maximal precision to the temporal articulation.

actions with a high degree of temporal precision. Short and discrete actions, like opening or closing an eyelid or moving a finger, can thereby result in a similarly short and discrete sound (see Figure 4c). We label this mapping element *sensitives*.

For more complex movements of longer duration we apply two different strategies: The first is based on what we call the *Creaking Door Principle* and the second we call *Clouds*:

**The Creaking Door Principle.** This principle says that typical human gestures (mapped as *Gestural Movements*) can be interpreted as a series of discrete particles much the way a creaking door sound is comprised of many small "creaks". When the door moves faster, the increased density of the creaks will eventually create a pitched sensation. Thus, as movements progress from tiny to large, the transition is continuous and the creaks will take on a pitched quality, where pitch fluctuations correspond to variations in the movement dynamic (see Figure 4 a and b).

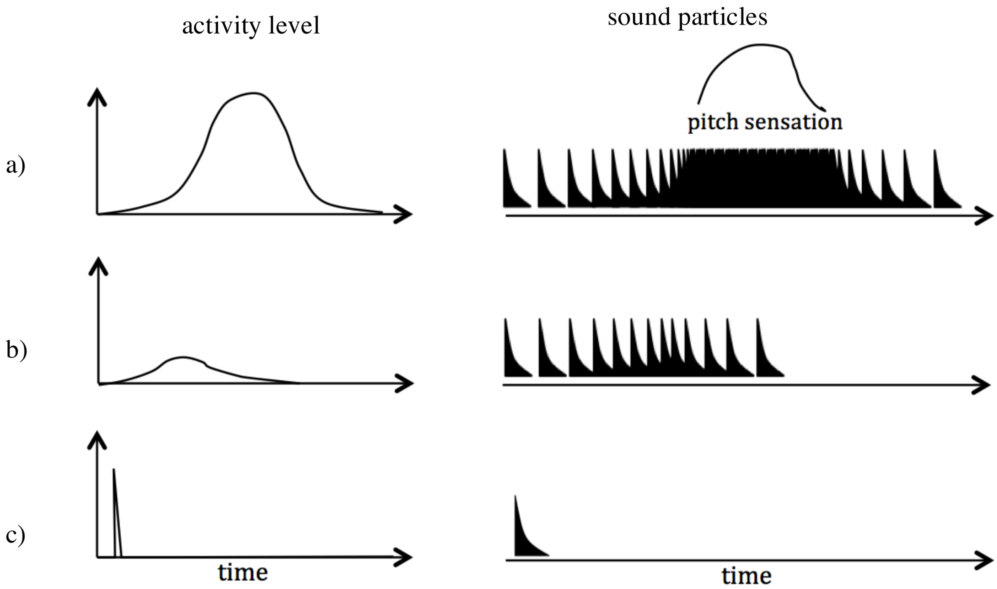



**Fig. 4 a, b and c.** Activity level is mapped to sound particle triggering rate in the *Particles* environment. In a) and b), the *creaking door* principle is illustrated. Low activity, as in b), will give a chain of sound particles with a triggering rate corresponding to activity level. Activity above a certain level, as in a), will cause the rate of triggered sound particles to cross the threshold of pitch, so that a tone with a pitch corresponding to the activity level can be heard. In c), a discrete movement causes the triggering of a single sound particle.

All this requires, however, that all the sound particles one triggers have a marked attack with a relatively homogeneous sound quality. Hence, it means that one needs to be careful in the choice of sounds to achieve the effect. When this is successfully achieved, one will experience a gliding transition between discrete temporal articulations and continuous pitched sequences, similar to those explored in compositions like Stockhausen's *Kontakte* (1960) and Henry's *Variations pour une* *porte et un soupir* (1963).[5] An additional effect caused by triggering a lot of sound particles in a short time, is that the sound intensity will increase markedly, thereby constituting a simple activity-to-intensity mapping, as suggested by Winkler [5].

**Clouds.** This second movement-music relationship strategy for larger movements of longer duration depends less on precise temporal articulation and more on flow. By using a set of more heterogeneous sounds with gradual attacks one hears dense masses of sounds as the triggering rate increases above a certain level. Thus *high activity* will in such cases create a cloud with *high density*, whereas *lower activity* will make the sonic clouds *sparse*, until the point where clouds will dissolve and individual sound particles can be identified.

Through these two strategies, one can see how the characteristics of the sound particles and their assignment to correlates in human movement are integral to the design of intuitive movement-music environments. The choice of sounds and their ordering within such a system represents an equally important aspect of the process and will therefore now be discussed in more detail.

## 4 Sound design in the *Particles* Environment

In addition to an intuitive movement-music relationship, creating environments that spur explorative behavior and heighten aural curiosity and sensitivity for the users has been a second area of concern for the authors. Here, we have particularly been interested in exploring different weightings of variability versus continuity/coherence, locally as well as globally. Furthermore, we have explored different techniques for aiding the user's cognitive structuring of the sonic environment, among them using sounds with a high degree of referential potential and using over-all ordering based on acoustic properties. We will have a look on each of these issues in turn.

It became clear very early in the development that one of the attractive features of the *Particles* environment lay in having a high degree of variability in the sound particles within each set. However, a lot of experimentation has taught us that variability in most cases needs to be balanced against sonic continuity, or the effect can quickly become one of chaos.[6] E.g. if one localizes many sounds with highly different characteristics in the immediate vicinity of each other easily, the relatively disordered result can in some cases be interesting, but tends to be tiring over time. At the other end, having too smooth transitions between sounds, e.g. using interpolated morphs in between recorded sound particles, tends to sound uninteresting and mechanic.

A nice compromise between variability and continuity can be achieved by creating local zones for which the intra-variability is relatively small (using e.g. recordings of different excitations of the same physical objects), and where there is greater inter-variability between zones. This can potentially make the user want to explore the finer

[5] Whereas the former piece of music uses synthetic sounds, the latter uses the actual sounds of creaking doors.

[6] In some cases a chaotic feeling can still be interesting. E.g. we have a set designed by Goran Vejvoda using wholly synthesized sounds, which has a lot of interesting variety often verging on the chaotic.

details within the zones, and subsequently seek greater changes by switching zones. This approach also has the advantage of creating cognitive markers for the users, thereby providing a mental structure for the sonic spatial layout.

Another strategy for creating a good balance between variability and continuity has been to order the sound particles based on certain acoustic properties, in particular spectral brightness and/or pitch. In that way, one can make the user go from duller to brighter, or low pitched to high pitched sounds when she or he moves across the space. This way of organizing the sounds is easily apprehended and understood by the users, and also allows for larger contrasts between neighboring particles.

**Set coherence.** We have also been concerned with giving each *Particles* set a particular identity or character, thus giving it an inner coherence. To pursuit this goal, we have explored the use of sounds made with a particular material or object, like glass or metal containers, sounds belonging to a common category, like orchestral instruments or synthetic noise sounds, and vocal sounds made by a single person. All these strategies have seemed to gain some success in creating an environment with coherence. Probably, it is not just the acoustic similarity, but also the sounds' potential for evoking associations that plays an important part here. Sound's associative qualities can in our experience play an important part in creating clear cognitive markers for the users, both for a whole set, and for a restricted zone within a set. By using zones with different associations, but which nevertheless are bound together by a more overreaching association, one can even build up complex environments. For instance, we are now working with putting different sounds together in a set associated with a forest: birds, water, animals and trees. Hence, we are approaching a highly narrative sound design, similar to what one can experience in video games and movies.

## 5. Conclusion

The collaborations of these authors, one a choreographer and one a composer, have been in the area of mapping and the search for interactive sound environments that give users a natural synaesthetic response, i.e. the sense of "hearing movement". This investigation, which seems at first to be straightforward, quickly grows in complexity into a deeply psychological quest involving kinesthetic awareness, proprioception and psychoacoustics. The interplay of these (in Western cultures) separate art forms – dance and music – correlated interactively, represents a rich and heretofore little-explored area of research, one with important implications for motion tracking-based performances and therapy for persons with and without disabilities.

The approach of the current work, involving sound particles – which are at once discrete, yet at the same time, when clustered, malleable and manifold in expressive quality – offers a promising general approach.

## References


1. Ng, Kia C. Music via Motion: Transdomain Mapping of Motion and Sound for Interactive Performances. In Proceedings of the IEEE, 92(4) (2004) 645-55.
2. Kontogeorgakopoulos, Alexandros, Robert Wechsler and Wendy Keay-Bright. Camera-Based Motion Tracking and Performing Arts for Persons with Motor Disabilities and Autism. In Disability Informatics and Computer Access for Motor Limitations, ed. Georgios Kouroupetroglou, IGI Global (2014) 294-322.
3. Roads, Curtis. Microsound. MIT Press, Cambridge (2004) 3.
4. Schwarz, Diemo, et al. Real-time corpus-based concatenative synthesis with CataRT. In Proceedings of Digital Audio Effects (DAFx), Montreal, Canada (2006)
5. Winkler, Todd. Making motion musical: Gesture mapping strategies for interactive computer music. In ICMC Proceedings. (1995).


## Additional reading


Acitores, A.P. and Wechsler, R., Danza interactiva con niños parálisis cerebral, In XXVII Congreso de la Asociación Española de Logopedia, Foniatría y Audiología, 7, 8 y 9 de julio de 2010, Valladolid, Spain.

Dietz, Josepha; Der MotionComposer: Bewegung in Musik verwandeln, Interview. URL: http://www.checkpoint-elearning.de/article/10614.html, accessed 26.8.2013.

Billerbeck, Uwe. Bewegung und Klang - Eine Woche mit dem MotionComposer, In Unterstützte Kommunikation 2(2012), 27-29.

Godøy, Rolf Inge and Leman, Marc. Musical Gestures: Sound, Movement, and Meaning. Routledge, London (2010).

Wechsler, Robert. Strategies for coherent mapping in movement-to-music interactive systems. In Proceedings of the International Forum on Cultures – Arts – Technologies – Creations – Disabilities, April 2013, National Technical Library, Prague.

Wechsler, Robert. Applications of Motion Tracking in Making Music for Persons with Disabilities. In Proceedings of the 4. Workshop of Innovative Computerbasierte Musikinterfaces, Mensch&Computer, Konstanz, Oldenbourg Wissenschaftsverlag, (2012).