Traces – Body, Motion and Sound

Jan C. Schacher
Zurich University of the Arts
Institute for Computer Music and Sound Technology
Baslerstrasse 30 8048 Zurich, Switzerland
jan.schacher@zhdk.ch

Angela Stoecklin
The Fusion Projects
Neugasse 33
8005 Zürich, Switzerland
an.stoecklin@bluewin.ch

ABSTRACT

In this paper the relationship between body, motion and sound is addressed. The comparison with traditional instruments and dance is shown with regards to basic types of motion. The difference between gesture and movement is outlined and some of the models used in dance for structuring motion sequences are described. In order to identify expressive aspects of motion sequences a test scenario is devised. After the description of the methods and tools used in a series of measurements, two types of data-display are shown and the applied in the interpretation. One salient feature is recognized and put into perspective with regards to movement and gestalt perception. Finally the merits of the technical means that were applied are compared and a model-based approach to motion-sound mapping is proposed.

Keywords

Interactive Dance, Motion and Gesture, Sonification, Motion Perception, Mapping

1. INTRODUCTION

In this publication the question of motion analysis and mapping is regarded from a very specific angle. Starting from the experience of a contemporary improvising dancer, the issues of motion, gesture, flow and force are addressed. When dealing with elements that characterize a dance movement terms such as the motion description fundamentals start to appear: inertia, energy, space and temporal structure but also terms of expressive potential and of anticipation, perception and recognition of specific motion patterns. In an attempt to better understand these fundamentals a scenario for interactive dance that originates from a real-life artistic process is identified and defines a small exploratory study. A number of measurement techniques are brought to bear on a constrained set of movements, with a specific question in mind. The movements and the measured data are combined in an audification and sonification process, as well as in different technical visualisations. On a first level the differences between measurement techniques become apparent, since the underlying physical phenomena are directly informing the results. On a second, higher level of complexity and correlation it is less the direct relationships between the measured streams of data that are interesting, but – via the translation into a different sensory mode – the emerging salient features of movement or even gesture in dance.

NIME'11, 30 May–1 June 2011, Oslo, Norway.

2. BACKGROUND

Relating motion, actions and gestures to musical processes is an essential part of developing interaction systems involving technologically mediated instruments. One central aspect in this relationship is the mapping strategies applied. [14] A further important aspect is the type of interfacing / gesture acquisition technologies utilized, since this influences the type of information obtained. The NIME community's main focus lies in this area. A number of attempts have been made to classify these devices [11] and create a comprehensive overview of the affordances they offer. [9] Many of the projects presented in this context explore the capabilities of the usually most advanced technical solutions available. These works generate know-how about the application of these devices, their strengths and weaknesses for musical use – usually such a device is designed for a different context – and characterizes precisely the type of information generated.

The traditional music performance with instruments builds upon a relationship with sound directly through a physically sounding object (except for the voice, where the body is sounding directly). The schooling of instrumentalists involves a great deal of body conditioning and training, or imprinting of fine motor skills [6] related to and occurring in an adaptive loop with the production of sound. The movements and actions used for this task are almost completely informed by the physics of the instrument. Economy of motion is a guiding principle only to be transgressed when internal impulses demand expression. [15] In general four types of movement can be distinguished: reflex, locomotive, instrumental, and expressive movement. Musical actions are therefore essentially instrumental, and only a small percentage actually becomes expressive.

The term musical gesture is sometimes used in this context, without actually making a clear distinction between instrumental and expressive motion. In other fields, such as communications theory and linguistics, gestures denote a very specific type of motion. It is considered "an expressive movement that is not consciously thought out beforehand" [4] and serves to enhance thought and communication. Gestures also carry a signification: "Gestures are not just movements and can never be fully explained in purely kinetic terms. They are not just arms waving in the air but *symbols* that exhibit meaning in their own right". [10] This quality, which is present in the expressive part of music-related movement, has to be differentiated clearly. A term that originates from linguistics and which highlights a difficult challenge for motion analysis and mapping is that of co-articulation. [7]

In contemporary dance, however, gestures are considered higher-level expressive entities that convey more than just movement. It is important to understand the differing views between dance and music performance in this regard. Unlike in film, contemporary dance attempts to render movement into something abstract and detached from everyday connotations and situations. These abstract dance-movements represent traces of physical but also mental processes concerning the body in space. A trained dancer learns to circumvent everyday movements, to detach herself from them and to create new patterns and variations thereof. The motivation for movement might be musical or visual, although the intention does not always include the projected image of the body. Musical elements such as rhythm and pulse play a central role in structuring motion in dance. Most modern dance notations, after Laban [12], are foremost descriptive and not expressive. The main categories described in these dance-languages are energy, placements and motion paths of body parts, placement in space and shape of body motion with regards to physical properties such as momentum and inertia. An arm movement for example might be described as a circular movement going upwards to the tipping point, then swinging with its own momentum down and back in a pendulum arch, then swinging to the front with the remaining inertial energy.

Dance gestures on the one hand are always tied to their body, which is at the same time their medium and shows their result and are only quite recently being measured, stored and analysed with technical means. Music on the other hand can very well exist without a body, especially in technically stored forms. Furthermore musician execute their instrumental movements adding some expressive parts without ever consciously balancing the two, thus the level of abstraction lies in the music an not the motion. The question arises now about how to identify expressive elements of motion between dance and music performance, where to look and what categories to apply.

3. SCENARIO

An interesting case arises, when a dancer is put into an interactive situation, taking on the role of a musician, so to speak. The motivation for movement might stay the same but the rhythmic and dynamic execution changes, when sound is produced or controlled by movements. The scenario devised starts from the idea that a dancer will perform a dance sequence consisting of a chain of gestures that can be chunked into movement elements. In order to gain more precise information the situation is a reduction to a few core aspects and consists of short twenty-second phrases covering a limited space horizontally as well as vertically. The dancer choreographs the sequence and executes it numerous times while varying characteristics such as intensity, speed and effort. Unlike a more classic live-electronic approach [3], here the music is generated *after* the fact, there is no sound during the performance, the dancer is only following her inner rhythm not some exterior material. In order to avoid an excess of data and to be able to compare the different sensors used, the motion-capture is constrained to two marker-groups, one on each wrist, mirroring the accelerometers placed on the body.

Figure 1. The dancer in our lab wearing accelerator bracelets and motion-capture markers on her wrists. The insert shows one of the bracelets in combination with a rigid-body marker used for motion capture.

4. METHOD

The measurement technologies we use range from simple accelerometer bracelets, to more complex inertial measurement units, from frontal two-dimensional video tracking with classic image analysis to an eight-camera marker-based motion-capture system. Each of these techniques offers a specific perspective on the dancer's body. We chose to use them simultaneously because they represent on the one hand a rich palette of tools, and on the other hand we hope that the measurements can tell us something about the accuracy and performance of each system and might permit a qualitative comparison between the different measuring techniques. In the following section, the different sensors and their stage-worthiness are briefly described.

4.1 Sensors

The wireless sensor bracelets were described in detail in an earlier publication. [14] They consist of a three-dimensional accelerometer and also provide two dimensions of gyroscopes. The update rate is between 50 and 100 Hz. The wireless inertial measurement unit (IMU) provides three orthogonal data streams for each measurement type: acceleration, gyroscope rotation and magnetic orientation. These values can be combined to obtain an absolute reference heading value and the absolute attitude of the sensor. This information is interesting mainly with regards to the overall body attitude. The update rate is between 50 and 100 Hz. These two sensor are stage worthy and applicable to a variety of scenarios.

The frontal two-dimensional video camera is used for body silhouette and lateral spatial analysis. By using an industrial firewire camera sufficiently high frame rates are obtained to be useful in comparison with the other sensors. The update rate can go from 60 to almost 100 Hz. Since it uses traditional background subtraction techniques this system is very light dependant. It is only useful in stage situations where absolute control over the lighting can be exerted (see Figure 2.). [2]

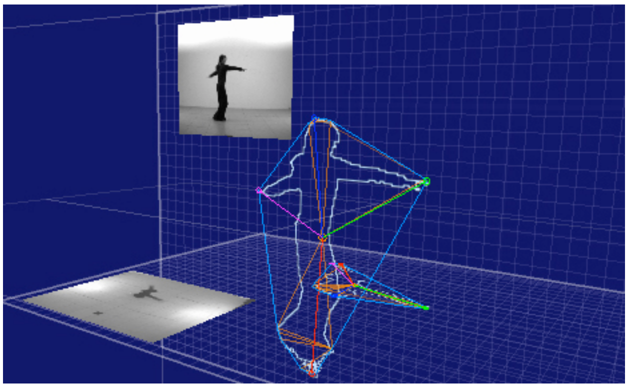

Figure 2. Dual camera silhouette analysis.

The motion-capture system we use is a smaller commercial eight-camera system. The three-marker solid bodies work very well in a space of the size of our lab (see Figure 1.). Although the system is one of the new generations of game-derived systems it delivers 100 frames per second and sub-millimetre precision. Since this is the biggest of the sensor-systems its application in a stage context poses a bigger challenge. [16]

4.2 Software

A variety of software applications are used to gather, store and analyse the sensor streams and their data. Apart from the commercial motion-capture system all the acquisition software was developed specifically for these sensors. The wireless sensors and their serial streams are parsed and transformed to OSC with dedicated proxy servers written in C++ using openFrameworks. [17] The frontal camera input and analysis is using Jitter and some custom code to calculate the convex hull and cardinal points of the silhouettes. [14] The motion-capture system runs its own software package on a dedicated windows machine and needs it's a proxy to translate from its native Nat NET protocol to OSC. This is a dedicated command line application written in C++. The data time tagging and storage is implemented with the Jamoma [13] module for GDIF [8] recording, which is based on FTM [1] and writes SDIF files. These modules run within MaxMSP, where all subsequent audification (straight data-to-sound mapping) , sonification (re-interpretation of data into sound), visualisations and calculations are executed.

4.3 Data Analysis

The resulting streams of values contain single integers and up to ten floating-point values. A central clock synchronizes the entire recording, and all data is recorded for each frame. In addition a video is recorded with the frame numbers inset for future synchronisation at playback. In order to reduce complexity further and to focus the analysis on a clearly perceivable element, only the accelerations from the motion-capture system and the wearable sensor bracelets are used.

The position data from the motion-capture system is analysed to its first two derivatives. These are purely spatial traces or kinematic calculations, no forces or masses are taken into account.

The accelerometer measurements are transformed to their summed absolute values, which only represents energy. In this form the values contain no more spatial or directional information. This data type differs from the former as it represents masses – actually a real micro-mass within the sensor – and the forces exerted onto them. Finally, in order to be able compare the two types of acceleration values we normalise them.

The visualisation tries to reintegrate as much information as possible into images and graphs. In an attempt to fuse spatial and temporal dimensions of the collected data, two graphical solutions – the timeline and the 3d representation with trails – are chosen. The actual video and imagery of the silhouette is added, since they enhance the perception, especially when viewed as moving images. The illustrations in this publication try to convey a sense of the temporally evolving values.

The audification uses filtered noise. The band pass filter is controlled on the frequency domain by the horizontal spatial x-axis, the output gain by the horizontal spatial y-axis and the steepness by the sum of acceleration on the bracelet. This algorithm presents a very basic one-to-one mapping but gives a clear sonic rendering of the motion and acceleration trajectories. (see Figure 5. upper half)

The sonification uses a granular cloud, where all spatial parameters obtained from motion capture are applied to the grain parameters and all accelerometer values influence the sample playback and windowing of the grains. (See Figure 5. lower half)

5. RESULTS

After combining the relevant streams and their cooked or analysed form into a range that puts them on the same levels, the comparative evaluation begins. Since the main question addresses qualitative rather than quantitative measures, no absolute numerical comparison was undertaken. The following illustrations show two frames from the motion sequence. This version of the sequence was executed in moderate dynamics with normal speed, so that the measured values vary moderately within their given ranges. The main feature visible is the circular motions of both wrists in space. The spikes in the accelerations corresponds to the changes in direction, which is clearly visible in the first frame's yellow line (Figure 3. upper half) and in the very last peak on the first time plot and the spike located at around 17.9 seconds immediately before the highest point in the second plot. (See Figure 3. lower half)

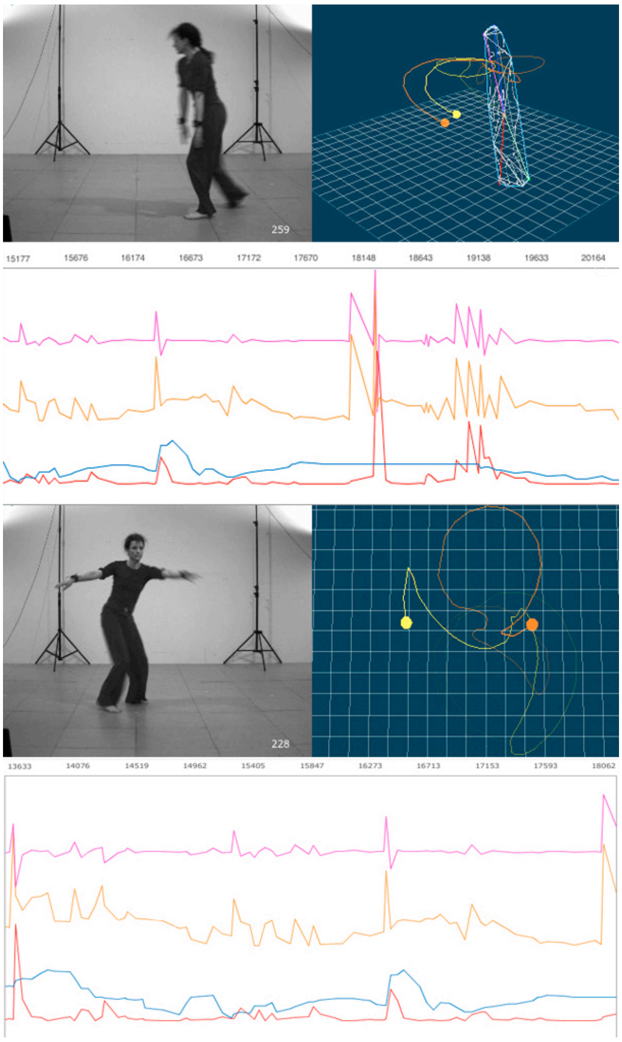

Figure 3. Visualisation of motion traces and wrist accelerations. The red trace shows the absolute kinematic acceleration, the blue trace depicts the dynamic acceleration.

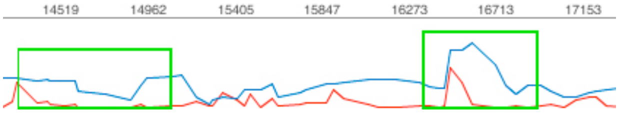

Figure 4. The red line depicts kinematic acceleration extracted from the motion capture; the blue line shows the accelerometer values. Note the overhang in the green boxes.

Considering the difference between the two measurement technologies, one being purely kinematic, the other a dynamic mass, there are obvious differences on how the values evolve. Where the kinetic movement stops, the dynamic mass in the accelerometer still has momentum to dissipate and goes into a negative acceleration with its direction vector inversed. As can be seen around 14.5 seconds in (Figure 4.) the inertial decrease of movement continues into the next onset and is carried over an even small absolute movement. At 16.3 seconds a very clear coupling of a jerky movement with a delayed but parallel decay in the dynamic mass is visible. Here again, the next small displacement is answered by a compensatory acceleration of the inertial sensor with the inverted direction vector.

The auralisation strategy demonstrates that low-level linking of parameters from the motion to the sound domains works well. The measured energy of the movement, particularly the inertial mass of the accelerometer translates well into sound energy as applied to this simple subtractive synthesis. (Figure 5. upper half)

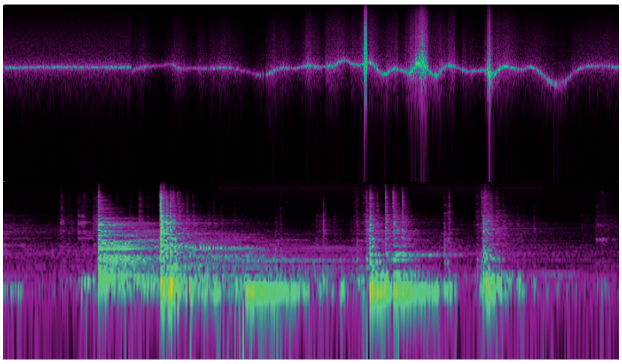

Figure 5. Sonogram of the audification of the motion fragment and of the sonification via granular synthesis.

This unsurprising fact is put into a different perspective, when a more complex sound generation algorithms are applied. Highly non-linear algorithms such as FM-synthesis or highly parallel processes such as granular synthesis offer a richer and more diverse sonic result. The granular synthesis algorithm used here produces a much richer sonic experience, but is less easily measured in terms of spectral content. (Figure 5. lower half)

Both hearing-based methods of data-display afford a perception of the motion sequences that emphasises the gestalt. The difference between the two methods, from the purely parametric mapping to a more subjective interpretation changes the richness of the sonic output.

6. DISCUSSION AND CONCLUSION

Comparing sensor data from purely kinematic measurement and an inertial mass sensor is obviously a strong reduction of information concerning a dancer's movement. Isolating and equalising two streams of values coming from the wrist of the dancer confirm a salient feature. The difference between kinematic and dynamic inertial measurement – but also the application of audification and sonification – shows that the physical properties of the moving body, such as its mass, momentum and inertia are more likely to be perceived as the carriers of expression. The energy or effort expended in the movement, which is the main category the dancer uses for creating and structuring a choreography, becomes more clearly visible in the data of the inertial sensor than the absolute spatial position acquired through the motion-capture system. The elements that comprise a gesture rather than a movement seem not to be clearly accessible in the data-streams, even though segmentation or chunking [5] is easily achieved at the rest-points of spatial, kinematic motion. This would indicate that in order to more naturally reflect the dynamic states of the body – which is what our perception principally anticipates and interprets – a physical model of the body should be introduced as a mapping category. This exploratory investigation also indicates that on-body sensors with their egocentric perspective remain a valid tool even when compared to the allocentric visual motion acquisition systems. When focusing on the perception of motion through other channels than the purely visual mode, the kinematic traces in space do not represent our motion as well is implied by the technology. The combination of the different types of sensor information with a model-based approach seems to promise the richest translation possibilities for body motion to sound.

7. REFERENCES

[1] Bevilacqua, F. Müller, R., Schnell, N. MnM: a Max/MSP mapping toolbox. In *Proceedings of the 2005 International Conference on New Interfaces for Musical Expression*, Vancouver, BC, Canada.

[2] Camurri, A., S. Hashimoto, M. Ricchetti, A. Ricci, K. Suzuki, R. Trocca, G. Volpe. Eyesweb: Toward gesture and affect recognition in interactive dance and music systems. In *Computer Music Journal.,* 24(1):57–69, 2000.

[3] Eckel, G., Pirro`, D., Sharma, G, K. Motion-enabled live Electronics. In *Proceedings of the SMC 2009 - 6th Sound and Music Computing Conference*, 23-25 July 2009, Porto.

[4] Gallagher, S. (2005) *How the Body Shapes the Mind,* Clarendon Press, Oxford.

[5] Godøy, R. I. *Systematic and Comparative Musicology: Concepts, Methods, Findings*, Chapter Reflections on Chunking in Music, pp. 117–132. Peter Lang, 2008.

[6] Godøy, R.I., Motor-Mimetic Music Cognition. In *Leonardo, Vol. 36, No. 4 (2003)*, pp. 317-319, MIT Press.

[7] Godøy, R.I., Gestural Imagery in the Service of Musical Imagery. In A. Camurri and G. Volpe (Eds.): *GW 2003, LNAI 2915*, 2004. Springer-Verlag Berlin Heidelberg.

[8] Jensenius, A.R. (2007) *Action – Sound, Developing Methods and Tools to Study Music-Related Body Movement*. Ph.D. Thesis, Department of Musicology University of Oslo.

[9] Magnusson, T. An Epistemic Dimension Space for Musical Devices. In *Proceedings of the 2010 Conference on New Interfaces for Musical Expression*, Sydney, Australia.

[10] McNeill, D. (1992) *Hand and Mind: What Gestures Reveal about Thought* Chicago, University of Chicago Press.

[11] Miranda, E.R, Wanderley, M.M. (2006) *New digital musical instruments: control and interaction beyond the keyboard.* A-R Editions, Inc.

[12] Laban, R. and F.C.Lawrence, (1974) *Effort: Economy of human movement,* second edition, MacDonald & Evans Ltd.

[13] Place, T. Lossius, T. Jamoma: A Modular Standard For Structuring Patches In Max. In *Proceedings of the International Conference on Computer Music* (ICMC'06) New Orleans, USA, 2006.

[14] Schacher, J.C. Motion To Gesture To Sound: Mapping For Interactive Dance. In *Proceedings of the 2010 Conference on New Interfaces for Musical Expression*, Sydney, Australia.

[15] Wanderley, M. M., B. W. Vines, N. Middleton, C. McKay, and W. Hatch. The Musical significance of clarinettists'' ancillary gestures: An exploration of the field. *Journal of New Music Research*, 34(1):97–113, 2005.

[16] http://www.naturalpoint.com/optitrack/

[17] http://www.openframeworks.cc/
URLs accessed in May 2011