Integration.04 //

Dieter Vandoren

20th of March 2013

Short //

A realtime sculpture in light and sound.

Integration.04 is a live performance by Dieter Vandoren. His instrument projects light and sound structures in fog-filled space, immersing both him and the audience in it. He manipulates the ephemeral audiovisual shapes as if they were tangible materials, taking cues from the interaction between the human body and acoustic instruments. The abstract digital processes are thus projected into physical, experiential space and the emergent play becomes a strong embodied experience - for both performer and audience.

Documentation //

Media documentation can be found at:
http://vimeo.com/dietervandoren/integration04 (video)
http://dietervandoren.net/index.php?/project/integration04/ (video, stills)

Overview //

The goal of the Integration.04 project is to develop and perform an electronic instrument which is experienced at a comparable or higher level of embodiment than an acoustic instrument by both performer and audience.

An acoustic instrument is defined as the composite of the human body and the mechanical body intertwined in a complex interaction of vibration propagation, resonances and dampings. That intricate intertwining is the inspirational model for the instrument's development.

Volumetric and spatialized light and sound projections form the instrument's medium. Light rays projected into fog-filled space create dynamic spatial structures. Both performer and audience are fully immersed in it as to create not only visual and auditive connections but also bodily ones.

An abstracted 3D geometry lies at the base of the generative audiovisual model. The performer interacts directly with this geometry to sculpt and mould the spatial audiovisual projections. Interface inputs are provided by full-body motion tracking (Kinect 3D cameras), hand rotation and movement (accelerometers) and finger bending (flex sensors).

Background //

At the beginning of this project my horizon lies at the intersection of performative architecture (Iannis Xenakis, Kas Oosterhuis), immersive audiovisuals (Granular Synthesis, Ryoji Ikeda) and embodied electronic music (Michel Waisvisz). Previous projects in my *Integration* series, Integration.02 and 03, uncovered the necessity for tackling the instrument part of the concept. In hindsight those previous projects arrived somewhere in the middle of an interactivated dance piece and a truly instrumental performance. Integration.04 is to go beyond and reach the goal of creating a performative audiovisual environment that can be played as a full-fledged instrument.

Another angle is the spatiality factor. As Xenakis, Granular Synthesis and related have shown, electronic media can expand into space. It can define space. And in doing so connect with the whole body instead of just the eyes and ears. Yet the majority of digital media remains stuck in a 2-dimensional screen and speaker paradigm. Projection mapping aims at escaping from the projection surface but never truly achieves it, no matter how slick the production.

Spatiality holds another promise for digital media: to 'extract the interaction from the locked-in, impenetrable virtual space to the human-scale physical space' [Chris Salter, *Entangled: technology and the transformation of performance*, p. 329, 2010]. While digital tools allow for highly-dimensional generative processes for unheard and unseen media creation the majority of interfaces remain 1-dimensional (sliders, buttons) or at best 2-dimensional (XY-pads, MIDI notes). What if those processes are expanded into physical space and sculpted, moulded, grabbed and pushed around by the full range of human motoric capabilities?

This shift opens up countless interaction channels to the performer and at the same time lets the audience get a view at the connected mental and bodily processes involved in live performance. Electronic media performance could be read at the bodily level, just like acoustic instrument performance, which cannot be said from laptop performances and similar.

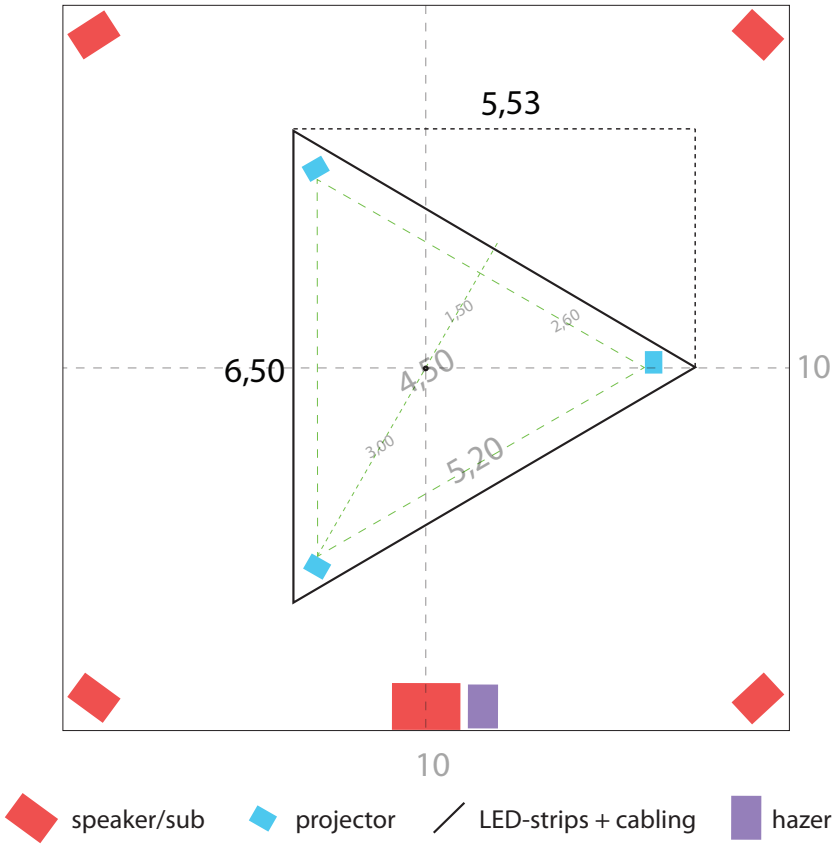
Spatial setup //

3 wide-angle projectors are set up on stands at eye-height in a circle around the performance area, aimed inwards to create a fully immersive dynamic light architecture. The performer stands in the middle of it while the light rays also extend to the outside where the audience is standing.

The 3 projectors define the triangular performance area. The triangle is delineated by RGB-led strips lying on the floor.

A quadrophonic sound system (+ subs) is positioned in the corners of the space to create an immersive sound field.

3D/depth-sensing camera's are mounted on 2 of the 3 projector stands.



(not in this drawing: audio and power cabling)

Interfacing //

Full-body tracking is provided by Microsoft/PrimeSense Kinect depth-sensing camera's. 2 are used because of the circular nature of the performance field, with no explicit front or back. They are positioned on top of 2 of the projectors, so with a horizontal offset of 120°. Merging the tracked data of both cams prevents from occluding the performer's body parts while turning around, giving 360° freedom of motion.

The relatively coarse and slow body tracking by the camera's is complemented by precise and fast hand-worn sensors. The 2 sensor gloves are built around SenseStage MiniBee microcontrollers with integrated Xbee socket and accelerometer. Each glove has additional bend sensors on thumb and middle finger.

The combination of cameras and sensor gloves makes for 2 distinct layers of motion tracking:

1. Macro: body skeleton, slow and low resolution
2. Micro: hands and fingers, fast and high resolution

The first macro layer comprises of the basic positions of the 15 tracked body joints and is expanded by a number of derivatives:

- speed and acceleration of joint movements
- inclination of limbs (azimuth-elevation coordinates)
- distance between selected joints, eg. the 2 hands

The second micro layer expands the range of control inputs with:

- 3-axis accelerations of the hands
- pitch and roll of the hands (coarse yaw can be derived from the skeleton tracking)
- bending of 4 fingers (interpretable as continuous or discrete data)
- speed and acceleration derivatives of each of the above

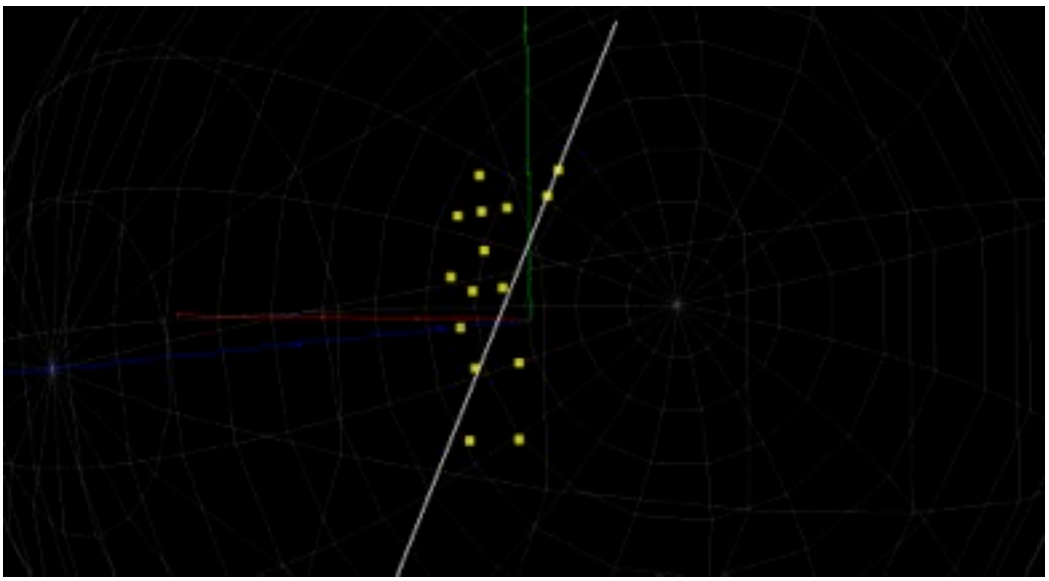
Augmented reality //

The 3D abstraction of the tracked body is reconstructed in a virtual OpenGL 3D context created in visual programming environment Max. This context takes into account the precise position of the 2 tracking camera's as well as the 3 beamers that will in turn project the virtual geometries into physical space. In other words, the virtual space can be overlaid 1-to-1 with the physical space. The movement of the body in physical space coincides with the movement of the tracked skeleton in virtual space.

A 3-channel graphic projection system is developed to actualize the virtually generated geometries in the physical performance space. 3 views of the virtual 3D model are taken with a horizontal offset of 120° (coinciding with the beamer positioning), rendered as orthographic projections and projected into the fog by the respective beamers. The result is a projection of the virtual model from 3 sides. A virtual point is projected as the intersection of 3 lines, 1 projected from each beamer. A virtual line is projected as the intersection of 3 planes, 1 from each beamer. This is not a 1-to-1 perspective-correct projection of the virtual model - a true holographic system would be needed for that - but it is a convincing actualization of the 3D generative system in the physical environment.

Thanks to the overlaying of virtual and physical space, precise bodily interactions with the projected elements are possible. Examples:

- a virtual point (visualized as the intersection of 3 light beams) can be 'grabbed' and dragged to a new position
- a virtual line (visualized as the intersection of 3 light planes) can be made to precisely follow the position of a limb, extending it into space



Generative model //

The software system consists of a suit of modular components:

- **Skeleton tracking:** interfacing to the camera's, 2 jit.openni processes on 2 separate machines
- **Skeleton merger:** combines the 2 tracked skeletons in 1 model
- **Motion abstracter:** extracts control data from the merged skeleton
- **Sensor gloves interface:** wireless interfacing with the gloves
- **Sensor gloves abstracter:** handles the accelerometer and bend sensor input streams
- **Main patch aka the 'skeleton':** data flow, system clock, settings manager, sequence control
- **Scene generators**: one or more for each scene, generate geometry and audio
- **Geometry renderer:** OpenGL output to TripleHead2Go and 3 or 4 projectors, jit.gl.node&camera make this a whole lot easier than before
- **Audio mix hub:** quadrophonic + subs output

The 'skeleton' component, scene generators and geometry renderer aggregate into the generative model which is at the core of the instrument. That model defines the multi-dimensional relations between motion tracking input and light and sound output. At the base lies a fairly simple geometry of vertices in 3D OpenGL space. Motion tracking input transforms this 3D geometry and the light and sound synthesis processes take it in as control data. In other words, it is the core model. It ties light and sound output together into one multi-dimensional shape.

The vertices of the core geometric model can be manipulated individually as points or all together as a polygon. Each sensor-driven manipulation maps to one or a combination of transform operations on this set of vertices. The basic transforms are translation, rotation and scaling.

The core geometry is the fundamental 'skeleton' of the projected light and sound shapes. The audiovisual generators are developed to display more 'organic' emergent behavior. Although inspired by acoustic instruments it doesn't try to emulate them. To the contrary, the digital nature of the system is embraced and cultivated.

The cues taken from acoustic instruments reflect in the interaction between the performer's body and the audiovisual generators. The performer gives impulses to and interferes with the generative processes. These

processes rely on a large amount of basic elements to create emergent auditive and visual shapes. Control does not happen on the individual element level but at a higher abstracted level, comparable to the plucking of a string that engenders a complex network of physical interactions which ultimately produce sounding air vibrations. These high-level inputs are derived from the positions, motion and internal geometric relations of the core model's vertices.

A broad range of different configurations can be built on top of this model. These are the 'scenes'. Each scene defines individual algorithms for:

- mapping motion tracking to core model transforms
- core model behavior (freeform polygon, particle system, etc)
- graphic rendering fed by core model (OpenGL primitives, texture feedback, etc)
- audio synthesis fed by core model (additive, granular, etc)

Results //

The development of Integration.04 resulted in an instrument with a satisfying and convincing depth. A depth in expressive dynamics and in plasticity of audiovisual shapes. The body is an integral part of this plastic medium. Shape is felt as an extension of the body.

In ideal situations gestures work directly in the shape instead of on a computer-human-interface control level. The technical user-interface is then transparent. This flow is sometimes interrupted when the 3D motion tracking loses sync. Camera-based, markerless tracking is still at an early stage of development and lots can be improved in accuracy and stability. Though in the beginning tracking glitches were disruptive for my performance flow, with practice I've learned to attenuate the hick-ups and get it back into shape fluently.

When the technology flows with the body, I am sculpting in space, light and sound. These merge in 1 multi-dimensional shape, bound by the underlying abstracted geometric model. With practice, manipulating this shape becomes a visceral, embodied experience. When that stage is reached spontaneous gestures start occurring. Not rationally planned actions but pre-thought, reflex-like moves triggered by a body that starts to integrate a new extension into its motoric channels. Much like learning to play an acoustic instrument.

At that point a language can start to develop. In Dick Raaijmakers' terms, a field is mapped in which routes can be explored [*Cahier "M"*, *A Brief Morphology of Electric Sound*, p.49, 2000]. The challenge in live performance then is to develop these in a meaningful way, for both audience and performer.