Versum: audiovisual composing in 3d

Tarik Barri
me@tarikbarri.nl

Abstract

This paper introduces the new audiovisual sequencing system "Versum" that allows users to compose in three dimensions. In the present paper the conceptual soil from which this system has sprung is discussed first. Secondly, the basic concepts with which Versum operates are explained, providing a general idea of what is meant by sequencing in three dimensions and explaining what compositions made in Versum can look and sound like. Thirdly, the practical ways in which a composer can use Versum to make his own audiovisual compositions are presented by means of a more detailed description of the different graphical user interface elements. Fourthly, a short description is given of the modular structure of the software underlying Versum. Finally, several foresights regarding the directions in which Versum will continue to develop in the near future are presented.

Keywords: audiovisual, sequencing, collaboration.

1. Introduction

Versum is an audiovisual sequencer. Whereas most sequencers such as Ableton Live and Logic use two-dimensional images to represent sequences of notes and sounds, Versum uses three-dimensional representations and allows the user to edit and run the sequence in three dimensions. For example, the sounds represented in Versum can not only be heard from left to right, they can also be heard from right to left, from up to down, form back to forth, in a spiraling motion, etc.

2. Philosophy

2.1. Breaking down barriers

Versum was born out of a need to break down the barriers that usually exist between an electronical musical composer and his audience. I felt that by letting the audience literally see how the composer works and how the composition is constructed, great value could be added to the listening experience.

Wanting to have the audience understand what's going on in a composition however does not imply that the goal is to make all musical and visual events predictable. On the contrary, having the audience create expectations based upon their understanding of the perceived audiovisual structures can provide the basis for surprises, which are caused by the composer's decision to deny these expectations. Both these expectations themselves and how they are resolved (or not) influence the way people perceive tension in music [1].

2.2. Audiovisual tension

Perceived melodies and rhythmical patterns can instigate expectations in the listener with regard to future musical events. Musical tension can be created when the composer plays with these expectancies, by denying them at some times and realizing them at others. To paraphrase music theorist Leonard Meyer: when expected events do not occur, or when they are delayed, these expectancies produce a state of suspense in the listener. It is this suspense that music cannot do without [2].

In Versum the listener has not only his ears to base predictions upon, but also his eyes. Therefore Versum enables the composer to create expectancies and build up tensions that are specifically audiovisual.

2.3. Spatial relationships

In clinical experimental settings spatial relationships between sounds have frequently been shown to interact with perceived musical attributes in systematic ways. Also the composer Hector Berlioz (1806-1869) has argued that the disposition of instruments in space should be considered an essential part of the composition [3]. Within Versum, the composer can easily make use of - and experiment with - the musical implications of spatial distributions of sounds, with clear visual feedback that enables the composer to keep track of the different positions where the sounds are coming from.

3. Seeing and hearing Versum

3.1. *Entities*

Within Versum, compositions consist of routes being traveled through a virtual universe. When viewing and listening to a composition, the software shows the viewpoint of the *actor* - a virtual moving camera with virtual microphones attached - in an infinite space. This can be seen in the Actor Window (see Figure 1). As the camera travels through this space, it travels past sound-emitting three-dimensional shapes called *entities*.

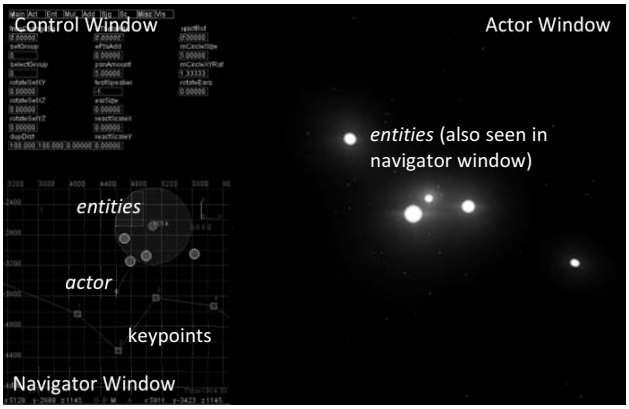

Figure 1. The three windows of Versum

The closer an *entity* is, the bigger the amplitude of it's sound, as recorded by the *actor'*s microphones. The sounds recorded by these microphones are then sent to their respective speakers, for the listeners to hear. Also, the perceived direction from which the sound comes will be the same as the direction in which the *entity* is visually located. Every *entity* has parameters that influence both the sonic and the visual output, making sound and image analog to each other. For instance, smaller *entities* produce a sound with a smaller amplitude than bigger *entities* do and an *entity* depicted with a less smooth surface will produce a more noisy sound.

The user of Versum can place these *entities* anywhere in the virtual universe and can define relationships between them in terms of motion. For instance, *entity A* can have *entity B* orbiting around it, at a speed and radius defined by the user. *Entity B* may also have another two *entities* orbiting around it, and so forth.

In summary it is the combination of the properties and positions of *entities* and the position of the *actor* that determines the whole auditory and visual experience.

3.2. Sound speed

Versum also has a built in sound speed, meaning that each *entity*'s sound is delayed by a time proportional to the distance it has from each of the *actor*'s microphones. When the *actor* travels past *entities*, the distance of an *entity* to the microphones will decrease or increase with time, and so will the delay time of the sound. When the distance will decrease or increase at high speeds, this changing of delay times will cause a clearly audible Doppler effect, increasing a sense of movement in the listener. At lower speeds the increasing and decreasing of sound delays will cause subtle phasing effects, thus bringing slight variations and liveliness to otherwise more static and monotonous sounds.

4. Versum's windows

4.1. The Control Window

Versum generates three output windows in total (see Figure 1). The Actor Window, through which the viewpoint of the *actor* can be seen, has already been discussed above. Then there are also two graphical user interface windows: the Control Window and the Navigator Window. The Control Window contains interface elements which enable the user to set specific parameters of the *entities* such as pitch, noisiness, low pass cutoff frequency, etc.

The Navigator Window contains a two dimensional map, similar to the kind of map a radar would provide, depicting a portion of the virtual universe. The Navigator Window allows the user to zoom in and out, allowing for both very detailed and very broad views of the positions of *entities*. Within this map *entities* can be easily added, deleted, copied, dragged and selected by using the mouse and several keyboard shortcuts. The Navigator Window also allows the user to quickly and easily move the position and camera angle of the *actor*, thus changing what can be heard and seen in the Actor Window.

5. Future development

An objective for the near future is to create the possibility of having multiple composers work within the same instance of Versum simultaneously. This would mean that each composer has his own interface and his own *actor* to hear and see the virtual universe with. Thanks to the fact that the internal software of Versum actually consists of several modules working together, and the fact that these modules communicate via network protocols, there is no necessity for all of them to run on the same computer. One could therefore have multiple interfaces running on separate computers connected to the same instance of Versum via network cables or a wireless network. Composers could then work together on populating a virtual universe with *entities*, creating their own constellations or edit those of each other.

When this feature is implemented, Versum's ability to have several *actors* roaming within the same space would mean that audience members could see and hear the same instance of Versum from different auditory and visual perspectives simultaneously, with each *actor* travelling a separate path and generating a unique composition in the process.

6. References

[1] Farbood, M. "A Quantitative, Parametric Model of Musical Tension", PhD thesis, MIT, 2006, p. 24.

[2] Meyer, L. B. *Emotion and Meaning in Music*. Chicago: University of Chicago Press, 1956, pp. 26-27.

[3] Bornstein, M. H. *Psychology and its Allied Disciplines, vol 1.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1984, p. 170