SUM: from Image-based Sonification to Computer-aided Composition

Sara Adhitya[1] and Mika Kuuskankare[2]

[1] EHESS/ IUAV/ STMS: IRCAM/CNRS/UPMS
sara.adhitya@ehess.fr
[2] Sibelius Academy / CCRMA Standford University
mkuuskan@siba.fi

Abstract. This paper will discuss the development of the SUM tool, a user library with a graphical user interface within the computer-aided composition environment of PWGL, aimed at the integration of image and sound. We will discuss its internal structure, consisting of image layers, mappers, and paths. We will explain the mapping process, from the retrieval of graphic data to its translation into audio parameters. Finally, we will discuss the possible applications of SUM in both image sonification and computer-aided composition, resulting from this structure.

Keywords: image sonification, graphical computer-aided composition, open graphic score, structure

1 Introducing the SUM Tool

The SUM tool allows the integration of image and sound through a graphic user interface. It was originally developed as an audio-visual representation tool in urban planning [1], an applied discipline involving the spatial composition of temporal systems. The traditional use of multiple 2-dimensional graphic maps makes it difficult to represent dynamic flows, as well as synthesise multiple layers due to legibility constraints. Thus SUM provides a more temporal approach to spatial composition through sonification – the representation of data through auditory means [2].

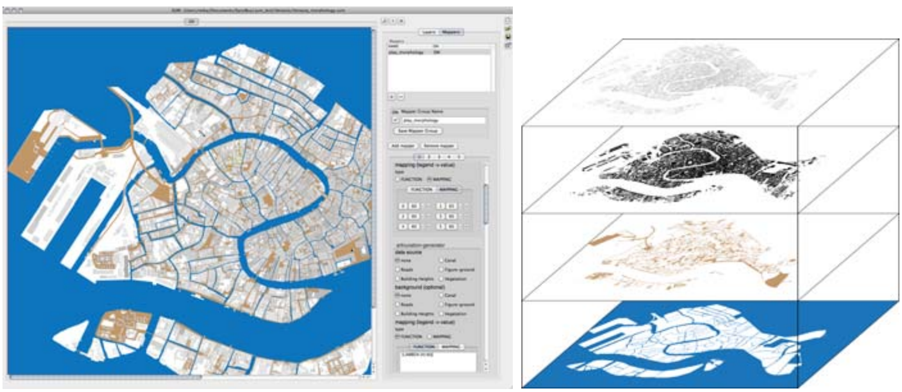

Fig. 1. The sonification of multiple urban maps in the SUM tool

9th International Symposium on Computer Music Modelling and Retrieval (CMMR 2012)
19-22 June 2012, Queen Mary University of London

Due to its design application, SUM supports both the importation and creation of multiple image layers (raster and vector) as data input. This data is then retrieved through the drawing of one or more vector paths over the areas of interest, and their graphic attributes mapped to sound attributes results in the generation of audio parts. Thus SUM supports a multi-dimensional spatio-temporal approach to image sonification, which sets it apart from other image sonification toolkits such as SonART [3].

As a user library within PWGL [4], a widely-used Lisp-based visual computer-aided composition environment, SUM can also be used as a graphical composition tool. PWGL's internal music notation editor strictly allows the description of object-based graphical scores [5], rather than the pixel-by-pixel exploration of a score as an image. Other graphical computer-aided composition environments, such as HighC [6] and Iannix [7], inspired by Xenakis' UPIC system [8], support the drawing of graphic objects but are limited to a single horizontal time axis. However SUM, with its ability to create and read objects along multiple spatio-temporal paths, allows an image to be composed and played as an open graphic score from multiple perspectives.

This paper will discuss the structure of SUM, which supports a multi-dimensional approach to both image sonification and graphical computer-aided composition.

2 The Structure of SUM

The SUM tool consists of three main components: images; paths; and mappers. The following section will explain each of these components and their inter-relationships.

2.1 Images

SUM uses images as data-sources. Each image is described by a 'color-key', in which each color of interest is allocated an arbitrary numerical value, to be referenced in the sonification mapping process. SUM supports the superimposition of multiple images, which allows the synthesis of overlapping graphic information, visualisable as a '3D' matrix of data as shown in figure 2. A group of data-sources is called a 'dataset', from which any number of image layers may be drawn upon as data-sources in the mapping process.

SUM allows the co-existence of raster and vector images. The flexibility of raster importation permits any visualization, including that produced by other software, to be sonified. The tool's vector drawing ability allows it to be used as a computer-aided design tool, such as Adobe Illustrator or AutoCAD, with graphic changes able to be made internally.

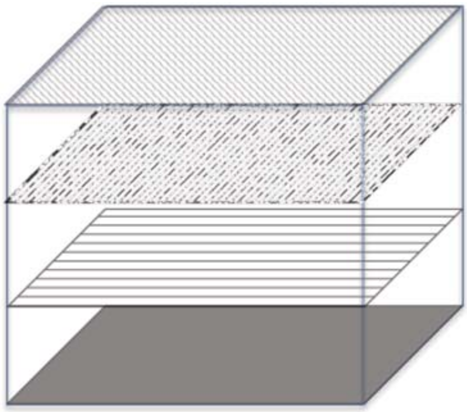

Fig. 2. Visualisation of a 'dataset' of 2D images as a 3D matrix

2.3 Paths

A path is responsible for defining the connection between the graphic space and musical time. It is a spatio-temporal object consisting of the following qualities: location; direction; delay; duration; and speed. The path is drawn as a vector polyline by the user over the area of interest, and then assigned a speed and delay. SUM supports the co-existence of multiple paths of various speeds and delays.

2.2 Mappers

A mapper is responsible for defining the sound output of the mapping process. It translates the graphic attributes retrieved from the image into discrete audio events, defining the sound attributes of pitch, volume, articulation and timbre. The definition of each sound attribute is independent of another. Thus one mapper can refer to multiple data-sources. A group of mappers is termed a 'mapper-group'.

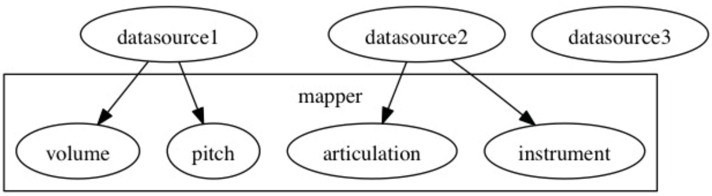

Fig. 2. The SUM mapper: one possible definition of sound attributes by data-source

3 The SUM Mapping Process

The SUM mapping process from image to sound is a two-fold process: graphic data is retrieved from a data-source by a path; it is then applied to a mapper for transformation into audio attributes.

3.1 Data Retrieval

The SUM mapping process is path-driven. Data is retrieved through the drawing of a vector path on an image, and the sampling of the image along this path. The vector path is rasterized according to Bresenham's line algorithm [9] in order to break it down into discrete sampling points, while retaining the order of the points to determine the direction of the path along which the time progresses. Thus for a line extending upwards and to the left, the pixels would be sampled in the order shown in figure 3.

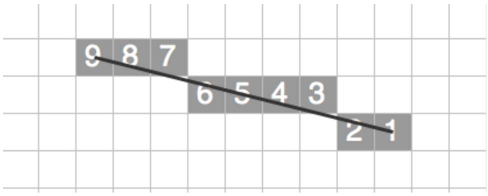

Fig. 3. Diagram of Bresenham's line algorithm, showing sampling order

Each raster map image is then sampled pixel-by-pixel to retrieve the data of interest per each sample-point along the path. The user-defined start-time and playback speed determines the temporal structure of the mapping process.

3.2 Parameter-Mapping

After retrieval of the graphic information along a path, these values can be applied to a mapper in order to generate the desired sound attributes of an acoustic signal (pitch, volume, articulation, and timbre). The parameter-mapping process is defined by assigning a legend, from a given data-source, with a sound value. This can be implemented either directly through the graphic user interface or by using Lisp for more complicated mappings.

Application of a path to a mapper produces a set of sound parameters, which can then be used to drive a wide-variety of internal or external instruments. PWGL has its own internal synthesizer as well as MIDI and OSC output. This allows connection to external sound synthesis engines such as Max/MSP and flexible possibilities for sound output.

It should be noted that a path and a mapper are independent of each other in terms of data-source/s. Thus different mappings can be generated from the same dataset of data-sources.

4 The SUM Compositional Process

This section will relate the SUM process to the compositional process. Here we introduce the concept of the SUM score, consisting of multiple SUM parts.

A SUM part is a sequence of audio events, the qualities of which are defined by the retrieval of data from an image with a path, and applying this path to a mapper. Thus the generation of a SUM part is a path-driven process. Application of multiple paths to one mapper will produce multiple SUM parts of the same timbral quality, but of variable temporal structure. Application of the same path to multiple mappers will produce multiple SUM parts of the same spatio-temporal quality, but of variable timbral qualities. Different combinations of paths and mappers allow the generation of numerous SUM parts from the same dataset. Figure 4 shows one possible network of paths and mappers producing a SUM score.

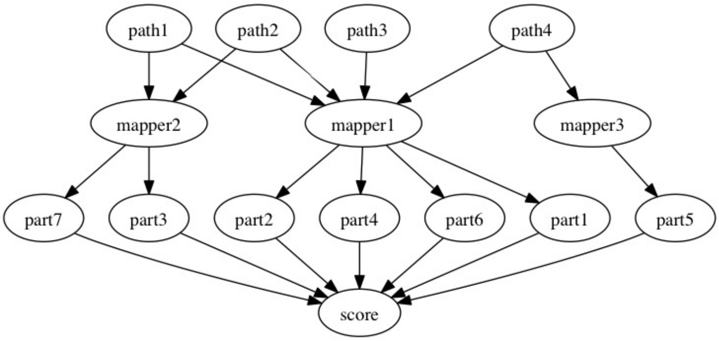

Fig. 4. An example of a SUM score - one possible network of paths and mappers

5. Applications

The flexibility of the mapping process established between image and sound has the potential for application in both image-based sonification and computer-aided composition.

5.1 Image Sonification – Playing of 'Visual Music'

The SUM tool, with its image-based input and user-defined mapping process, supports the sonification of any color-coded image. This means that any bitmap image can be sonified according to its own color-key.

One artistic application is in the playing of 'visual music' – the generation of musical concepts such as rhythm through graphic means. One visual composition technique is through the spatial arrangement of colour, as explored by Piet Mondrian in his series of paintings entitled 'Composition' utilizing the primary colours of red, yellow and blue. Here we demonstrate the sonification of his work *Woogie Broadway*

Boogie (1942-43), in which he attempted to express the musical rhythm of the 'boogie woogie' through colour, and in addition along a gridded structured resembling the streets of New York[10]. By separating the painting into each of its colours, and mapping each colour to a different sound parameter, such as pitch, volume or timbre, we can not only see but listen to this rhythm along each of the paths.

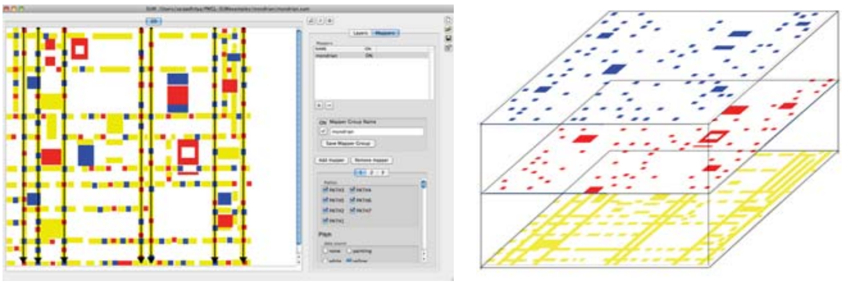

Fig. 5. Sonifying the colour rhythms of Mondrian's *Broadway Boogie Woogie* [10]

Through the sonification of such visual artworks in SUM, we can explore the application of visual composition techniques to musical composition. We can also see the potential for SUM to play any image as an open graphic score. In the following section, we will demonstrate the use of SUM as a tool for computer-aided composition, leading to the generation of a graphic score.

5.2 Computer-aided Composition – Generation of a Graphic Score

The SUM tool, with its vector drawing capability, also supports the creation of graphic scores. The user-defined mapping process means that a composer is free to create his own graphic-sound vocabulary. It supports the creation of a multi-layered graphic score (ie. multiple spatial dimensions), and its playback from any direction, time and speed (ie. multiple temporal dimensions).

As an example, we will show how the graphic score created by Rainer Wehinger for Gyorgy Ligeti's Artikulation, can be generated in SUM and used to explore its playback.

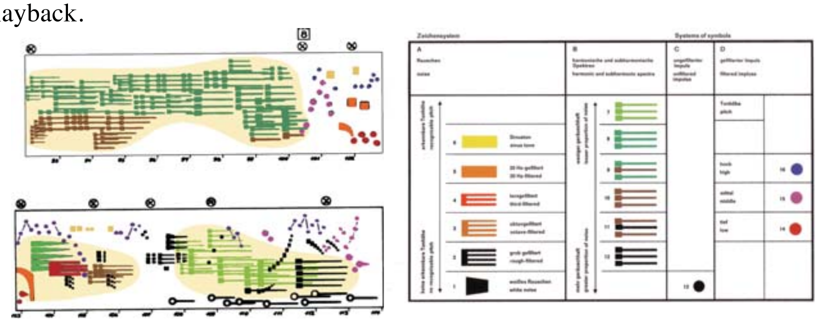

Fig. 6. A section of Wehinger's graphic score for Artikulation (Ligeti) with accompanying colour-coded legend [11]

Wehinger represented each of Ligeti's sound objects graphically, in terms of different forms and colors (see figure 9). As different colors are read as different sound objects in SUM, we can structure our SUM score similarly.

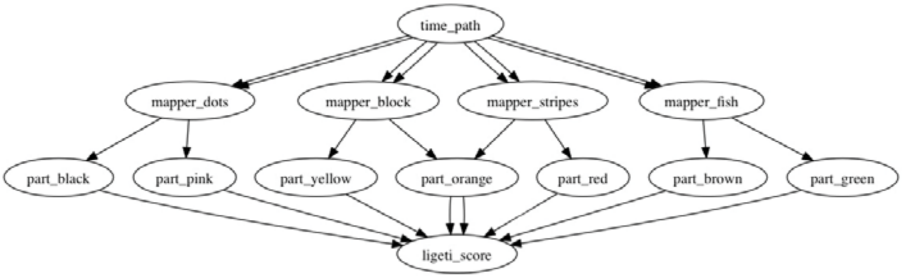

Fig. 7. A possible SUM score structure of Artikulation

The subsequent reading of our SUM score, by any number of user-defined spatio-temporal paths, frees it from its intended linear reading from left-to-right. As seen in figure 8, the same segment of Wehinger's score can be played from different directions and at different speeds.

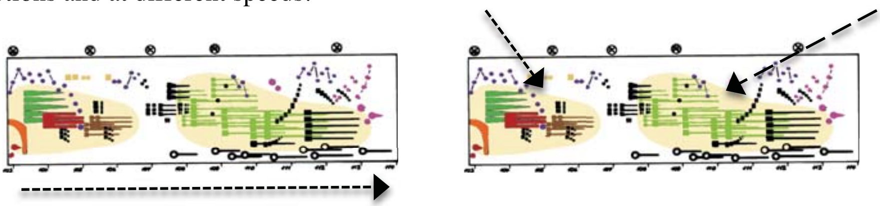

Fig. 8. Different ways of reading Artikulation – linearly as a pianoroll or as an open score

This opens up new possibilities for existing graphic scores to be played in alternative ways and to generate new musical results.

6. Conclusions

As seen above, the structure of the SUM tool supports the integration of image and sound in multiple spatial and temporal dimensions. Growing from the objective to sonify urban maps for a more temporal representation of urban systems, as seen in this paper, we can also use it to compose a multi-dimensional graphical musical score and play it back from numerous perspectives. The flexible structure of SUM allows the audio-visual representation of multiple spatio-temporal relationships in general, from an urban system to a musical score.

Future improvements include the automatisation of the retrieval of the image color palette, and thus the generation of the color-key. We also aim to improve our path-sampling approach in order to more accurately determine the duration of a path.

Acknowledgments.

The authors would like to thank IRCAM and CCRMA for their hospitality during this collaboration. The work of Sara Adhitya has been supported by the John Crampton Scholarship Trustees of Australia. The work of Mika Kuuskankare has been supported by the Academy of Finland (SA137619).

References

1. Adhitya S., Kuuskankare M.: The Sonified Urban Masterplan (SUM) Tool: Sonification For Urban Planning And Design. In: 17th ICAD International Conference on Auditory Display (2011)
2. Kramer, G.: Some organizing principles for representing data with sound. In: Auditory display. Sonification, Audification, and Auditory Interfaces, Addison-Wesley, pp. 185– 221 (1994)
3. Yeo, W. S., Berger, J., Lee, Z.: SonART: A framework for data sonification, visualization and networked multimedia applications. In: 30th ICMC International Computer Music Conference (2004)
4. Laurson M., Kuuskankare M., Norilo V.: An Overview of PWGL, a Visual Programming Environment for Music, In: Computer Music Journal, vol. 33, no.1, pp.19–31 (2009)
5. Kuuskankare M., Laurson M., Expressive Notation Package. , In: Computer Music Journal, 30(4), pp. 67–79 (2006)
6. Baudel, T.: High C draw your music, a graphical composition system inspired by Xenakis' UPIC, http://highc.org/history.html
7. IanniX, a graphical real-time open-source sequencer for digital art, based on Xenakis' works, http://www.iannix.org/en/index.php
8. Unité Polyagogique Informatique du CEMAMu (UPIC), a computerized graphical musical composition tool developed by Iannis Xenakis, Paris, 1977
9. Flanagan, C.: The Bresenham Line-Drawing Algorithm, http://www.cs.helsinki.fi/group/goa/mallinnus/lines/bresenh.html
10. Based on the painting by Piet Mondrian, Broadway Boogie Woogie, MOMA, NY (1942-43)
11. Rainer Wehinger's graphic score (1969) for Gyorgy Ligeti's Artikulation (1958), http://www.tumblr.com/tagged/artikulation