COMPOSING GRAPHIC SCORES AND SONIFYING
VISUAL MUSIC WITH THE SUM TOOL

Sara Adhitya
IUAV/ EHESS/ STMS Lab
(IRCAM/CNRS/UPMS)
sara.adhitya@ehess.fr

Mika Kuuskankare
Sibelius Academy,
CCRMA, Stanford University
mkuuskan@sibi.fi

ABSTRACT
This paper will explore the potential for the SUM tool, intended initially for the sonification of images, as a tool for graphical computer-aided composition. As a user library with a graphical user interface within the computer-aided composition environment of PWGL, SUM has the potential to be used as a graphical approach towards computer-aided composition. Through the re-composition of the graphic score of Ligeti’s Artikulation, we demonstrate how SUM can be used in the generation of a graphic score. Supporting spatio-temporal timepaths, we explore alternative ways of reading this score. Furthermore, we investigate the claim of certain visual artworks to be ‘visual music’, by sonifying them as graphic scores in SUM.

1. INTRODUCTION
The SUM tool [1] is a user library with a graphical user interface within the computer-aided composition environment of PWGL [2]. Initially designed for the sonification of images, through a user-defined mapping process, it also allows a graphical approach to computer-aided composition.

The field of graphic-based computer-aided composition spans from Xenakis’ pioneering UPIC, to CAC software such as OpenMusic and PWGL, and image-sonification toolkits such as SonART [3]. However, the former was limited to piano-roll lecture, the second allows the description of object-based notation [4], and the third allows importation of images which can be read by raster-scanning or real-time probing.

SUM can be seen to span all three approaches, allowing both the pixel-by-pixel exploration of a score as an image, as well its temporal reading from any number of angles, and the ability to generate images internally. It also supports the use of multiple time paths, allowing the lecture of the image in numerous ways – as the open graphic score was intended.

Thus this paper will explore the use of the SUM tool in the computer-aided composition of graphic scores. We will first present an overview of the structure of the SUM tool and its relationship to the traditional compositional process. Then we demonstrate how a graphic score can be composed in SUM through the recomposition of Rainer Wehinger’s representation of Ligeti’s Artikulation. We then demonstrate the ability of SUM to modify and play this piece in multiple ways, before attempting to play selected artworks as graphical scores through the process of image sonification. Thus we hope to demonstrate the potential of this tool to be used as a flexible graphical approach to composition.

2. SPATIO-TEMPORAL ORGANIZATION
The following section outlines the definition of space and time in SUM through the definition of its component elements of image layers and time paths.

2.1 Spatial structure
In SUM, images are used as musical ‘data-sources’. Raster images can be imported, and the spatial scale depends on the resolution of the image in dpi. Vector objects can also be created within the tool. SUM supports the superimposition of multiple images, which can be seen to form a 3D matrix. This layered structure of image data-sources can be seen as the equivalent of simultaneous parts. Images can either be read independently or in combination with each other.

2.2 Temporal structure
The temporal structure of SUM is path-based, user-defined by the drawing of a vector path and its application to one or more images. The path samples the image(s) according to Bresenham’s line algorithm, as shown in figure 1 [5].

![Figure 1. Diagram of Bresenham’s line algorithm, showing sampling order](image-url)
The user defines the start-time and speed of a path, which then determines the sampling rate and thus the timing of the score.

Multiple paths are supported, allowing the score to be read from different spatial directions, like an open graphic score. Since they each have their own independent start-times and speeds, it is possible to read an image at the same or different speeds, allowing the user to flexibly manage the time of a composition.

Thus through the combination of multiple image layers and spatio-temporal time paths, SUM can be seen to support the creation of a multi-dimensional graphic score.

3. GRAPHIC MAPPING PROCESS

In a graphic score, sound is often represented through non-traditional symbols or images. With its ability to import raster images as well as create vector graphics, SUM becomes a flexible tool for composers to develop their own graphical notation. This is implemented through the definition of one or more mappers.

A mapper translates the color attributes retrieved from the image into audio events, defining the sound attributes of pitch, volume, articulation and timbre. A group of mappers is termed a ‘mapper-group’.

The parameter-mapping process is defined through the mapping menu of the graphic user interface (figure 3). The composer selects the data-source to be used and assigns each legend value with a sound value (in MIDI units). This can be implemented by directly inputting the values for each legend value in the GUI, or by using Lisp for more complicated mappings. There is also the option to output the data through OSC for further mapping in an external program such as MaxMSP or PureData.

As the definition of each sound attribute (pitch, duration, volume, timbre) is independent of another, any combination of images can be used to define one sound. Thus one mapper can refer to multiple data-sources and different mappings can be generated from the same dataset of data-sources. This gives a composer much freedom in defining the relationship between image and sound, and developing his or her own graphic notation.

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

Mapper-groups can be activated and deactivated so that any combination can be used to create a SUM score, to be explained in the following section.

4. SUM COMPOSITIONAL STRUCTURE

As a part of the compositional process, a ‘SUM score’ is generated consisting of any number of parts. Here a ‘SUM part’ is the result of the application of a path to an image data-source via a mapper. The resulting compositional framework can thus be represented as a network of paths and mappers, as shown in figure 4.

Figure 3. Mapping menu of the graphic user interface

Figure 4. An example of a SUM score - one possible network of paths and mappers
5. CREATION OF A GRAPHIC SCORE

This section will demonstrate how a graphic score can be created in SUM through the use of an image data-source. As an example we will ‘recompose’ Rainer Wehinger’s graphic representation of Gyorgy Ligeti’s Artikulation, developed 12 years after the work’s initial composition in 1958. Wehinger represented Ligeti’s various sound objects graphically by using different forms and colors according to the legend shown in figure 5.

Since sound in SUM is defined through the use of color, we can easily import this graphic score and ‘play’ it through the definition of a color-key. In order to retain the identity of overlapping objects, we have separated the score into color-coded layers. This can be represented as a matrix of superimposed 2D layers as shown in figure 6.

6. PLAYING GRAPHIC SCORES IN SUM

As explained earlier, SUM supports the co-existence of multiple paths at differing speeds, and thus the playing of an image as an open graphic score. Here we give examples of the different ways in which our ‘recomposed’ graphic score of Ligeti’s Artikulation can now be played, as shown in figure 7.

The top of the score shows how the score can be played in the intended traditional piano-roll style by drawing the paths from left to right over the objects of interest.

However, as the temporal structure of a SUM score is not restricted to one axis, we can also approach the reading of the score from any direction, for example with a curved path in reverse, as shown at the bottom of figure 7. Since the timing of the score is determined by the start-time and speed of each path, as defined by the user, these paths can be either synchronized or not. The composer can thus play a composition at multiple speeds.

7. GRAPHICAL TRANSFORMATIONS

Here we explore the possibility to create new image layers in SUM to apply graphic transformations to an existing graphic score. Through image masking and graphic filtering we can add sound objects or alter audio parameters.

7.1 Image masking

SUM allows the option of masking one image with another. With its internal vector drawing ability, new graphic objects can thus be drawn over an existing score in SUM with the effect of modifying its sonic output.

This masking ability can be used to subtract areas of the score by returning them to their background color. In the top of figure 8, we have added an extra layer of white rectangles, in order to visually ‘erase’ existing sound objects. New objects can also be added to conceal existing objects, such as the black triangles at the bottom of the score.
7.2 Graphic filters

Graphical filters can also be created to globally control a sonic parameter over an existing score, for example its overall volume or timbre. This is implemented by parameter-mapping the filter layer to the audio parameter concerned, as seen in figure 9 in which the superposition of a gray-scale gradient globally reduces the volume from left to right across the score.

8.1 The musical lines of Kandinsky

Many artists have explored the line as the primary element of movement, most notably Wassily Kandinsky in his book [8]. In SUM we can explore the potential for such line drawings to be used for the structuring of time in a graphic score. We have used the line study by Kandinsky, seen in figure 10, as a basis for our path-based approach to image-reading.

8.2 The color compositions of Mondrian

Composition of color in space is another basic technique of generating visual rhythm. Piet Mondrian produced a series of paintings entitled ‘Composition’, in which he explored the spatial composition of the primary colors of red, blue and yellow. In his painting *Broadway Boogie Woogie* (1942-43), he intended to express the rhythm of the ‘boogie woogie’ through the organization of these colors along a gridded structure resembling the streets of New York [10]. By mapping each of these colors to different percussion sounds in SUM, and drawing paths of movement over the areas of interest, we were able to listen to the rhythmic results (figures 12 and 13).
Playing the colored rhythms of Mondrian’s *Broadway Boogie Woogie* [10] in SUM

The mapping of colors to tuned percussion instruments and the color-coded piano-roll generated

9. CONCLUSIONS

In this paper we have explored the potential of the image sonification application of the SUM tool as a more graphical approach to computer-aided composition. With its flexible structure of multiple data-sources, paths and mappers, the generation of numerous combinations of parts, and thus scores, is possible. The structuring of time along user-drawn spatio-temporal paths allows an image to be read from a multitude of directions like an open graphic score, and alternative readings easily proposed and tested.

As seen in the examples presented above, the musical results are as varied as the images and mappings themselves. In the future we will explore different mappings, particularly in terms of timbre through instrumentation or sound synthesis, in order to generate a more effective image-sound vocabulary for the interpretation of each image.

What can be seen so far is the potential of the SUM tool as a means of exploring the relationships between compositional techniques in the visual arts and those in music. Drawing such audio-visual connections can only inspire new compositional techniques in music and generate a more multi-sensorial approach to music composition.

Acknowledgments

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).

We would also like to thank IRCAM, and in particular the Perception and Sound Design Team, as well as CCRMA, Stanford University, for hosting the respective authors.

10. REFERENCES


