

# THE COMBINATION OF VISUAL AND SONIC GENERATIVE PROCESSES IN A DIGITAL LANDSCAPE

*Felipe Otondo*  
Music Research Centre  
Department of Music  
University of York  
fo500@york.ac.uk

*Thomas Petersen*  
Artist at [www.crossover.dk](http://www.crossover.dk)  
Co-editor at [www.artificial.dk](http://www.artificial.dk)  
thomas@crossover.dk

## ABSTRACT

The text describes the procedure used as a part of a collaboration combining visual and sonic processes to build a generative electronic landscape for an installation. Initially the procedure used to develop the algorithm for the visual evolution of the landscape is described in terms of the visual process. Following this, the description of the sonic design of the mix to be controlled by the landscape is explained in terms of technical resources and approach as well as the implementation for a particular space. Finally some conclusions are outlined about the potential of the designed of shared generative processes for image and sound with perceptual and aesthetic considerations.

## 1. INTRODUCTION

The idea of an installation at the cultural event “Culture Night 2004” in Copenhagen lead to the planning of a collaboration combining image and sound focusing on the creative process [1]. The motivation was to investigate how far could one integrate sound and image using simple generative processes for their evolution in time in an installation. The goal of the project was to underline common features between evolving processes of sound and image in terms of relations and interdependences between - colour and timbre. Considering, as a basis, the evolution of the visual processes in an abstract digital landscape image, the idea was to design an evolving texture of sound streams that were controlled directly by colour variations in this landscape.

## 2. DEVELOPMENT OF A DIGITAL LANDSCAPE

The resulting project ‘SoundImage’ is a computer program (created in Actionscript for Macromedia Flash MX), which continuously constructs an abstract landscape image cycling through a range of potential colours determined by a system of random algorithms. Specific colour parameters of the landscape are considered as controllers for independent layers of sounds. The landscape can therefore be read as a visual expression of the current sound composition. Figure 1 shows an example of the digital landscape.

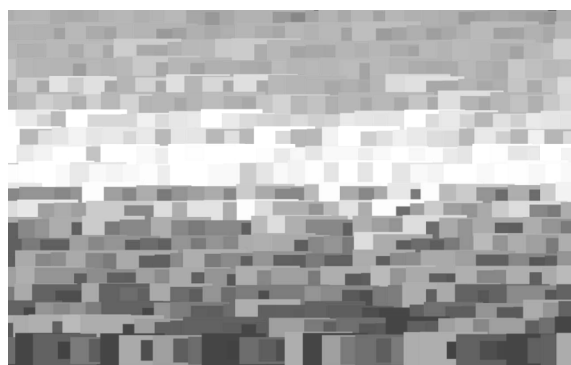


Figure 1. Example of the digital landscape. Colour examples can be seen at the reference [1].

### 2.1. The algorithm – step by step

The abstract landscape image consists of a grid of 35x20 coloured squares. After the program initially is started, the squares move step by step through a range of potential colours according to a system of random variables. After 15 minutes the program restarts in a slightly different position – but always in the lower end of the predetermined colour scales.

Every two horizontal lines have a separate predetermined colour database. Each square in a specific line pair can only choose colours from this colour database. For the 20 horizontal lines there are thus 10 separate colour databases – each with 20-50 potential colours. Each square’s colour is determined by its position in the colour database. This position is reflected as a number ranging from 1 to e.g. 50 (depending on the number of colours). If a square has the position of say 49 in its database then it acquires a specific colour stored at position 49. The database of colours for the top two lines in the landscape looks like this:

Position	Hexadecimal RGB value (red, green, blue)
0	023242
1	023a4a
[...]	[...]
49	7375a9

Table 1. Excerpt of colour database for the top two lines in the landscape. Left column: position in database. Right column: RGB colour values associated with the positions.

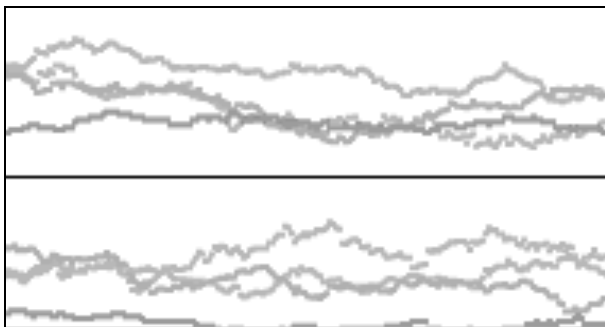
The hexadecimal colour values in the right column in the above table represent the specific mix of red, green and blue making up the final colour. If any square in the top two lines of the image has the position 0 in the colour database, its hexadecimal colour value will be 023242 (red=02, green=32, blue=42). This translates into a shade of dark blue. Typically a lower position in any of the colour databases is equal to a darker colour as a higher position is equal to a lighter colour. In the case at hand, position 0 is a dark blue and position 49 is a pale violet.

If the random algorithm chooses to change this particular square's colour, then its colour position is changed relatively to its former position. The position can change 0, 1 or 3 steps to a higher or lower position in the database. The square's colour position might therefore change to position 1 (RGB=023a4a), which translates into a lighter blue. For each instant cycle (fractions of a second between each cycle) this decision process is repeated for each square. The systems of random decisions are roughly equivalent to a series of dice throws.

The size of each square is determined according to its position in its colour database – a higher position always results in a larger square. The squares are stacked in an order beginning from the bottom rightmost square and ending in the top leftmost square. The combination of the varying sizes of the individual squares and the stacking creates many possibilities of an aesthetically interesting mosaic emerging.

## 2.2. Analysis of the colours in the image

At the end of each instant cycle, when every square's colour has been assessed and some of them changed, the colours of the landscape are analysed globally. The values obtained from these analyses are used later to control the volumes of specific sound layers.



**Figure 2.** Two separate evolutions over 3 minutes. Each graph reflects the fluctuating values derived from four parameters: the global colour values of red, blue and green, plus the purple event volume.

### 2.2.1. Total red, green and blue values

In this analysis, the fluctuations in the total values of red, green and blue in the whole image are found. First every square's colour value is broken down into three parts: the hexadecimal values of red, green and blue. All the red values in the total image are added up – the same

goes for the greens and blues. After this is done, the following questions are asked:

- How has the total red value of all the squares in the image changed compared to the initial starting point?
- How has the total blue value of all the squares in the image changed compared to the initial starting point?
- How has the total green value of all the squares in the image changed compared to the initial starting point?

In every instant cycle, these red-green-blue colour totals are compared to the initial total values when the program was started due to the fact that, as mentioned before, the starting positions are slightly different each time. The rise or fall in the red-green-blue totals are calculated as a percentage-wise rise or fall and then converted to a scale ranging from 1 to 100, which is used to set the volume of the corresponding sound layers, as it will be explained later. One loop is controlled by the changes in the global red colour value, and two other loops are controlled by the changes in green and blue.

### 2.2.2. Separate colour events

There have also been chosen certain 'event sounds', which are triggered by specific colour events in the image. On every instant cycle, the following questions are asked:

- How many purple squares ('purple tones in the sky') are there in the top half of the image and how bright are they?
- How many white squares ('clouds') are there in the area just above the horizon and how bright are they?
- How many yellow squares ('glints') are there in the area just below the horizon and how bright are they?

Every bright yellow square below the horizon within a certain colour range is given 'volume points' (1, 2 or 3 points) according to its yellowness. The total value of these volume points is applied to a specific 'yellow' sound loop. Many bright yellow squares therefore result in the possibility of a quite shrill sound. The same mechanism is applied to white and purple squares.

## 3. USING THE VISUAL PROCESS TO DESIGN SOUND LAYER INTERACTIONS

### 3.1. Sound mix controlled by the landscape

As a way to create a basis for the interaction between the landscape and sound, initially different types of sounds were related to particular colours in the landscape according to their characteristics in timbre and pitch, as well as their characteristics for being combined in a mix with other sounds. This was done considering as a basis long steady evolving sounds suitable for mixing and creating nuances in timbre variations. Table 2 shows the six sound layers used in the mix with their corresponding visual controllers in the electronic landscape.

As shown in table 2, the first three layers of sounds were those associated with overall total values of the colours red, green and blue. The sounds used in this case were mostly long-stretched tuba tones blended with (1) other tuba tones, (2) tones of a bass clarinet and (3) tones of a French horn. The procedure used to blend the tones was done using the convolution in similar way as done for previous works [2, 3]. The sound resulting from the three layers of different timbres was designed as a semi-continuous bass background where sound layers could evolve gradually with variations of intensity and contrasts of timbre according to the global changes in the overall colour components of the landscape.

Layer	Sound	Colour	Duration (min)
1	Tuba x Tuba	Total blue	12
2	Tuba x B. Clarinet	Total green	9
3	F. horn x Tuba (bass)	Total red	7
4	F. horn x Tuba (bright)	Purple events	5
5	Alto saxophone	Yellow events	5
6	Bass trombone (sharp)	White events	5

**Table 2.** Sound layers with associated colours in the landscape and particular duration.

The next sound layer consisted of a bright (middle frequency) sound obtained from the convolution of long tones of the French horn with tones of the tuba. This layer was controlled in intensity by the variations of the purple colour in the central region of the landscape and was used to create slowly evolving melodic contours that would blend with the bass sounds and also stand out with a more distinctive character.

Finally a third group of two layers of high pitched sounds of alto saxophone and bass trombone was used in connection with the variations in the yellow ‘glints’ below the horizon and the presence and intensity of clouds in the upper part of the landscape. These two layers of sound evolved punctually as sound events following those in the landscape.

### 3.2. Sound design and approach

Once the types of sounds were associated with specific graphic controllers in the picture the generative processes used to develop the picture were adapted to control each of the six layers in the global mix using two main sound parameters: intensity and spatialisation. The first parameter used to control the level of intensity of each layer in the mix was associated with a scale of numbers generated in the landscape for each of the colours of table 2. The numbers produced by the variations of each of the colours were associated using

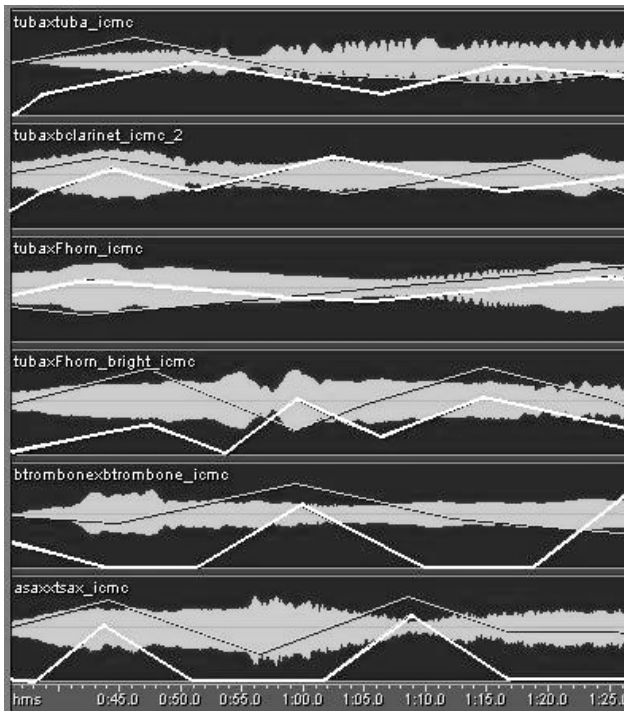
non-linear relations with the variations in intensity of each of the layers in the mix using a percentage scale. The particular non-linear relations were fixed according to the type of role that each of the layers could have in the final mix and in case more extreme or subtle changes were desired. Therefore, the variations in the intensities of the more bass sounds were less dramatic than those of the high-pitched sounds which appeared punctually with more extreme intensity fluctuations.

As an alternative to this lack of dramatic variations, a fixed starting volume for each of the bass sounds was set at the lower end of the scale (when the image restarts every 15 minutes). These fixed starting points were chosen to achieve the probability of the individual colour volume reaching 0 relatively often and thus giving the algorithm the possibility of turning off the particular sounds.

The second parameter manipulated by the visual controllers was the localisation of sound within the two loudspeakers. This was done using as a basis the intensity levels of the sounds as controllers for the localisations using a specific algorithm for each sound to determine the relationship between intensity and localisation. For some of the sounds an extreme intensity will result in an extreme right position as a low volume will result in an extreme left position. In other cases it is exactly opposite.

A model was designed in a sequencer to test aurally the variations of intensity and spatialisation of the six channel mix in a similar way as the 6-layer mix would react following the variations of the controllers of the picture in a normal evolution as the one shown in Figure 2. Figure 3 shows an example of the model used to test changes in intensity and spatialisation using information similar to the one obtained from the electronic landscape in a normal performance. As it can be seen in the picture the changes in intensity and spatialisation are different for each of the layers in the mix, being more gradual and constant for the bass sounds and more punctual for the high-pitched sounds.

Using this model the relations between the visual controllers and the parameters controlling the sound of the 6-layer mix were adjusted until obtaining a sonic framework that could work aesthetically with the evolution of changes suggested by the picture. Additionally to the use of the 6-layer mix a recurring set of clicking sounds controlled by the instant changes of the colours of the landscape was used. Each time a square changes colour there is the (random) possibility for a clicking noise to be played. The total volume for these clicks was connected to the volume derived from the total green value and the location of these punctual sounds within the two loudspeakers was also related to a random possibility. The use of these clicks was thought as a way to add instant liveness to the whole visual and sonic process and increase the awareness from the listeners to the micro changes of the landscape in contrast to the macro changes suggested by the gradual evolution of the mix of the continuous sounds.



**Figure 3.** Example of the model used to test the changes in intensity and spatialisation of the 6-channel mix. The white lines show the changes in intensity for each track and the grey lines the changes of spatialisation with left speaker on the bottom and right speaker on the top.

#### 4. INSTALLATION'S IMPLEMENTATION

SoundImage was created as an installation to be implemented in a particular space: Copenhagen Project Room, which is a small exhibition space in Copenhagen. The exhibition space consists of one room of approximately 25 square metres. The project was presented as a visual/sound installation in a space designed by artist Inez Mortensen as a way to supplement and enhance the visual and aural experience of the project. The exhibition space became a green carpet clad environment with soft red circles very suitable for the contemplative approach of the installation.

The technical implementation of the installation consisted basically in a computer that would show live and in cycles of one hour the evolving electronic landscape and would play the sounds stereo. These stereo sounds were then mixed and distributed separately in a four channel distribution through four loudspeakers with different equalisations. The two main loudspeakers were located in front of the audience at each side of the screen while the two other speakers were placed in the back of the room following a crossed stereo disposition to the main speakers of the screen. This was done as a way to enhance spatially the impulsive character of the clicks as perceived by the listeners in the centre of the room.

Furthermore, a website was created for the project where it was possible to freely download a version of

the SoundImage file [1]. The sample rate of this version was downsampled to shorten download speed, but the file is otherwise the same as the exhibited version in the installation.

#### 5. CONCLUSIONS

The collaboration presented in this work has shown that it is possible to develop and design shared generative processes for image and sound which can allow an integration of gradual evolutions of colour in connection to evolutions of timbre. It has also shown to be important in the design of these shared processes the adjustments related to perceptual and aesthetic characteristics of the evolutions in each media by the use of aural and visual evolution models. The use of these models showed clearly that interesting processes will not necessarily lead to interesting artistic results if things are not integrated and balanced in a coherent way.

Further developments of the ideas of this project could consider the use of more and diverse layers of sound related to different visual controllers using more direct and indirect variations. It would also be desirable to implement the project in a different programming environment to achieve a wider range of possibilities within the direct production of sound.

#### 6. REFERENCES

- [1] Link to the website of the project:  
<http://www.crossover.dk/lydbillede/indexeng.htm>  
This site includes photos of the installation and a downloadable version of the project.
- [2] Otondo, F. "Using the convolution to blend brass timbres", *Journal of Music and Meaning* (online journal), 1, 4, Fall, 2003.  
<http://www.musicandmeaning.net/issues/showArticle.php?artID=1.4>
- [3] Otondo, F. and Soto, J. "Three diagonal strategies for a sound installation", *Journal of Music and Meaning* (online journal), 2, 6, Fall, 2004.  
<http://www.musicandmeaning.net/issues/showArticle.php?artID=2.6>