ORGANIZED DATA FOR ORGANIZED SOUND
Space filling curves in sonification

Florian Grond

ZKM - Center for Art and Media
Institute for Music and Acoustics
Lorenzstraße 19, D-76135 Karlsruhe, Germany
grond@zkm.de

ABSTRACT
In this paper, we introduce space filling curves (SFC) as a useful possibility to organize data for sonification. First, we give a brief overview about the history of SFCs and their graphical construction. Then we focus on the mapping properties of SFCs from 2D to one dimension. We present the acoustic results of an implementation of the described method, in which we took the Hilbert curve as one particular example of an SFC. The actual sonification program features different methods for real-time interaction. These methods take advantage of the particular properties of SFCs. We further discuss their restrictions, how they can be circumvented, and give an outlook to future applications, where we also make suggestions as to how the properties of SFCs can be combined with methods of data reduction.

[Keywords: Space filling curves, data handling, sonification of high dimensional data, interactive sonification]

1. INTRODUCTION

One particular advantage of sonification is the possibility to perceive complex data in a holistic way. If we consider the meaning of complexity, which is that complex systems cannot be reduced to subsystems - or at least not in a trivial way -, we realize that a possibly direct translation of data into sound is an important prerequisite for sonification to unfold its full potential.

Unfortunately there are only few data types that allow for a direct translation into sound, which is known as audification.

Since sound develops in time, these are mostly data that are by their nature organized in a linear way, such as records from earthquakes [1] [2] or ECG [3] data. Many different sonification methods for such data have been developed. Sometimes apparently high dimensional data like multi-channel measurements from EEGs are primarily still linearly organized, since each channel corresponds to a scalar developing in time. The multi-channel EEG measurements can for instance be played back by a multi-channel speaker setup, such as the audio dome of the “Denkgeräusche” project [4][5], or the multichannel seismographic art installation “Circum Pacific” [6].

In general, however, the sonification of multidimensional data requires methods of data reduction or transformation in one way or another. For multi-dimensional data for instance, principal curve analysis was successfully used [7]. Other possibilities include employing data sonograms or particle trajectories calculations in order to prepare the data for sonification [9]. In some specific cases, when the data allow for certain assumptions about their inherent structure, like in face recognition tasks for instance, methods of feature extraction can be exploited [8]. But particularly if we want to use sonification for exploratory data analysis [9], we should try to start without any preconceived assumption about the underlying data structure. In this case we try not to reduce their complexity, because we want to listen to the data rather than to the algorithms that are applied onto them.

In this paper, we specifically want to address the sonification of data types where each time frame consists of more than one dimension. We focus on data such as digital images often encountered in many fields of sciences, which deliver information in a parallel manner. We will extend the method we suggest to image sequences, as in movies. Sonification strategies for stacks of images were already developed in [9]. In our paper we want to particularly focus on questions of data handling and mapping that allows for the transmission of the complexity of the underlying data and further allows to navigate through acoustical clues in these data.

Let us assume for the beginning that the amount of 2D data developing in time - the bit rate of a movie - would be as small as a bit rate typically encountered in audio. In this hypothetical case, we only face the problem to organize these bits of each frame in a principal way for sonification. We could start out with a simple horizontal scan, just as image data is organized on a digital storage medium. This scan method places the data in a linear order, as encountered in any uncompressed pixel image file format and does not reduce their amount; yet the inherent data structure, i.e., the relations of shapes and color in the image, is altered. In this paper we discuss different possibilities for reorganizing 2D data sets in one dimension for sonification. We will focus on space filling curves (SFCs) as promising candidates for scanning 2D data sets. We will discuss the results of sonifying movies with various scan methods as a proof of concept implementation in Pure Data (PD) and the computer graphic library GEM. We want to emphasize that movie data primarily serves as an example for high dimensional data that develops in time. The method can easily be applied for data types with 2 or more dimensions. Our main aim is to introduce an existing method of data handling to sonification.
Thus, our emphasis is not on the data material itself as a particularly promising candidate for sonification.

2. SPACE FILLING CURVES AS A MEANS BY WHICH TO ORGANIZE DATA:

SFCs were invented in 1890 by Giuseppe Peano [10]. It is often mentioned that his approach was entirely analytic, and that there were no accompanying drawings. Today SFCs are popular because of their graphical construction and representation. Although mostly of academic interest at the beginning, SFCs stimulated interdisciplinary approaches early on [11]. In 1914 the Russian mathematician and philosopher Pawel Florensky [12] linked SFCs with problems of representing reality in the fine arts and issues of the central perspective. In his best known essay, “The Reversed Perspective,” all his arguments about the central perspective are derived from the particular mapping properties of SFCs. As an example he uses a particular SFC constructed by his contemporary David Hilbert [13], which is often referred to as the Hilbert curve. Today SFCs reach far beyond purely mathematical interests. They are applied in different fields, such as load balancing for parallel computing, indexing databases for efficient range queries, and in methods for image processing. [For a good overview about fields of applications and further literature, please refer to Ref. 14 and 15]. In whatever field of applications they are found, researchers always exploit the SFCs’ basic properties in order to translate complex high dimensional data into linear sequences, onto which algorithms can work very efficiently. In the sequel we will briefly recapitulate how an SFC is graphically constructed. A good introduction to this topic can be found in [16]. For mathematical details and implementation methods we refer to numerous sources on the internet [17].

![Figure 1. graphical construction of a space filling curve (Hilbert curve).](image1)

Any SFC starts out with a simple building block. In the case of the Hilbert curve, it is an upside-down “U”. Each step in the iterative construction process involves copying the building block, rotating the copies, connecting them and resizing them to the square unit. The result, after infinite repetitions, is a line that covers the square unit in a self-similar meandering manner, as demonstrated graphically in figure 1. A finite number of iterations suffices to cover a finite amount of digital 2D data. The fact that the line never intersects itself ensures a one to one mapping from 2D to 1D. We mention in passing that similar linear structures can be constructed to fill 3D and higher spaces. These self-similar structures have interesting mapping properties. Not only do they represent a unique mapping, they are self-avoiding for all finite approximations, they also tend to preserve neighboring relations of a 2D image on the 1D line. This roughly means that shapes within an image, such as areas of one color, can be found as connected regions on the line. Therefore SFCs are able to roughly keep and transfer the essence of the original image onto its one-dimensional rendering.

It is very instructive to illustrate this property by comparing it with a spiral. On an image covered by a spiral, random pairs of neighboring points are also neighbors on the line of the spiral, as long as they happen to fall in the close vicinity of the center of the spiral. If they are found away from the center, their probabilities to be neighbors on the line of the spiral tend to measure zero. Having this picture in mind we can think of the self-similar meander like a spiral, in a spiral, in a spiral, which almost keeps neighboring relations equally well everywhere. This particular clustering property of SFC can be seen in figure 2 and compared with a line scan in figure 3.

![Figure 2. distances of pair points of the Hilbert curve versus distances on the unit square in 200 * 200 bins. In the lower left corner in the foreground the clustering behavior can be recognized.](image2)

![Figure 3. distances of pair points of the line scan versus distances on the unit square in 200 * 200 bins. High Peaks and gaps are distributed close to the diagonal.](image3)
3. FROM THE SCAN LINES TO SOUND PROCESSING:

The data, reorganized into lines through SFCs (we take the Hilbert curve as our SFC), are now ideally prepared to use for sound synthesis. It would be possible to produce a typical audification by sending them, as they are, to the digital analog converter. But this approach does not produce any sounds that reveal the particular properties of the underlying data set, because information about a single point in an image would be reflected through a single sound sample. Considering the fact that an image transmits information in a synchronous way it would be desirable to find a qualitatively similar sonification. An overview of existing strategies can be found in [9]. Our approach was mostly inspired by Joachim Goßmann’s audio fractal [18], this is why we took the linearized data as frequency information. Through the use of the inverse FFT we can turn this information into sounds that reflect the relations within the whole image through the relations of frequency ranges in the sound.

Let us compare the effects of different data scans on their structures, and likewise, their effects on the resulting sounds (figure 4). If we take a horizontal scan line, we observe repetitive patterns of shapes in the image. These patterns repeat approximately with the numbers of samples necessary to scan from one side of the image to the other. If we look at these patterns from a sonic perspective we recognize something like an equally distanced overtone series. These overtones have a very pronounced effect onto our acoustic perception. Of course their intensity varies with the content of the image, but their structure remains independent of it. This fact leads us to conclude that some parts of what we hear in this case are structural artifacts of the scan process. If we take a spiral instead, the overtone patterns are more complex, but are essentially still there. By using the Hilbert curve, we avoid such patterns, thanks to the locally progressing nature of this scan process.

4. IMPLEMENTATION DETAILS AND SOUND DESIGN:

In order to test and listen to the different scanning possibilities of frames in movies, as described above, we implemented a prototype program in PD. The Hilbert curve was implemented as an external GEM Object programmed in C++, since iterative structures are difficult to implement in the graphical programming environment of PD. The graphical output of this object was a Hilbert curve or other line scans onto which an image could be textured. Further, this object had access to the image data and could read out the RGB(A) information. These RGB(A) data were reordered depending on which scan method was switched on, and sequentially piped through an outlet. The object could be fed with data from existing GEM objects from the pix class, which could load either images/image stacks or films.

As a creation argument the iteration depth of the Hilbert curve scan could be set. The Hilbert curve object had several methods that could be called by messages. These methods allowed switching between 4 different scan modes:
1. a horizontal line scan
2. a square spiral from the center of the image outwards
3. random sampling on a regular square grid
4. the Hilbert curve

If the random message was sent, the sample points were reshuffled on the regular square grid. All these four scan methods had as many scan points as the Hilbert curve defined through the given iteration depth of the creation argument.

Another set of messages allowed to transform the orientation of the scan line either by flipping it horizontally or vertically, or by rotating it around the diagonal. The generic PD message “bang” was bound to a method which released the output of the reordered image data. The sequential RGB output of the image data was scaled between 0 and 1 and further read into three different audio buffers of 1024 samples. These buffers were used in subtractive audio synthesis as a filter bank for white noise. The three different resulting audio streams were sent to a pair of stereo channels, with one stream equally distributed to both of them.

Figure 4. This image shows different scan methods applied to the Homer Simpson movie of GEM tutorial about the pix class. The vertical axes are the scan points and the horizontal axes are the 85 movie frames. From left to right: spiral scan, line scan and SFC scan (Hilbert curve) The horizontal axes are the Frame numbers and the vertical axes are the 1024 scan points.
There were some parameters to tune the filter bank. One was a simple gain. The others are in the following function (1) of the amplitudes of the frequency bins freq(i) (0,1) in the filter bank:

\[
\text{freq}(i) = (a \cdot b \cdot \text{amp}(i))^2
\]  

(1)

Through this function the contrast in the filter between the different frequency bins could be tuned by changing the parameter \( n \) to support the acoustic differentiation of the structures and shapes in the image. The parameters \( a \) and \( b \) were either set to \( a = 1 \) \( b = -1 \) or \( a = 0 \) \( b = 1 \). This was necessary to enable the sonification of generally white frames with little dark structures in it, which would otherwise let most of the noise pass through the filter. It is similar to taking the negative of the image data.

5. INTERACTION POSSIBILITIES:

In order to use the program to explore movies, we implemented methods to interact with the program in real-time during sound synthesis. We thereby focused on the complementary potentials of the visual and acoustic senses. We wanted to design a tool that would, on the one hand, visually highlight structures that we can hear, and on the other, sonify regions in the film where we expect something of interest.

One general method was to rescale the size of the area covered by the Hilbert curve or other line scans. In this way, we could acoustically zoom into or out of the image. Apart from that, there were two other interesting interaction features in the program prototype. We designed the following two use cases to suggest how people might possibly interact with the data:

1. Exploring the data from the visual perspective allowed restricting the scan to a 2D region in the image. RGB information was then sent to the sound buffers from these regions only. Through this feature, the sonification of these restricted parts of the movie image could immediately help to evaluate any hypothesis about structures in the time development of this image region.

2. Acoustic structures that the user would hear could be repeating oscillations in a certain frequency range. In this case, the program allowed to restrict the scanning to this range, and mute all the other frequencies. The segment of the scan line that represented this particular range was displayed as a line textured with the scanned image in the OpenGL Window. In this way, a correspondence between acoustically interesting frequency regions and the relating parts of the image could easily be established.

The self-similar Hilbert curve is useful for expressing the relationship between sound and visual representation. This is particularly true for the second approach, where the driving sensual information is acoustic. Any other line scan would have distributed the visual patterns over a much wider frequency range with a repeating overtone series. This would make it impossible to narrow down the corresponding frequency range, which would correspond to the area in the image that caused the sound. In this use case, the application of SFCs basically corresponds to range queries in databases as described in [14].

As mentioned above, it was possible to change the orientation of the different scan types. The sonification of horizontal line scans through frames of a movie often sound very different to vertical ones. This is particularly the case when the movements of shapes in the movie have pronounced directions. Scans with the Hilbert curve are more invariant to transformations like rotations, although their sound character changes too. In the paragraph about future possibilities we will suggest other SFCs with a higher symmetry.

Figure 5. the 4 possible rotations of the Hilbert curve (iteration depth 2) on a square. For each scan the index order can be reversed giving rise to 8 possible scans.

6. DISCUSSION OF SONIFIED SAMPLE MOVIES:

Please note that you can find this paragraph and all sound samples at: http://sonification.kommerz.at.

We sonified two different movies with the program described above. The first sample movie of our choice was taken from the PD documentation pix class tutorial about movies. You can see different scans of each frame of this movie in figure 4. The movie features Homer Simpson in 85 frames, running down a street from the right upper corner to the bottom center of the image. In the background we see trunks of trees from a forest. In the beginning, the background moves as there is a pan from a long-shot to a close-up. When Homer Simpson arrives in the middle of the frame, he falls onto his knees. We sonified this movie with all different implemented scan methods. For the Hilbert curve and the scan line, we sonified all 8 possible orientations. (sound files 1-8 and 9-16). For the line scans, it is interesting to note the aforementioned acoustic difference between the vertical and horizontal orientation of the lines. For the spiral we sonified 4 orientations with the lower frequencies in the middle (sound files 16-20). We also sonified 4 randomly sampled versions (sound files 20-24).

The second movie is an infrared satellite view of the earth over 4 days from 02.10.2006 to 06.10 2006 from the following source [19]. It shows the weather dynamics on the globe. We chose this movie because cloud formation is a more abstract phenomenon of scientific interest, and it potentially exhibits oscillating patterns. Auditory weather forecast has also been used for sonification [20]. In this sample sonification, we particularly want to demonstrate the possibilities for real-time interaction with this program. In the first sound sample of the second movie we hear the whole frames covered by a SFC scan. We can hear only weak oscillations, then we zoom in and restrict the scanning to interesting frequency ranges (soundfile 25). We manage to identify two different oscillations, and we can separate them by restricting the scanning to the respective frequency regions.

In the second sample, we again hear both regions on the globe where we can find a pronounced day/night oscillation in anti-phase. Looking back to the movie we find that the Sahara heats up over the day, and cools down at night. This is
accompanied with nightly cloud formation around the equator. (soundfile 26).

Figure 6: This image shows the weather dynamics as sonified in the implemented program. The sound of sample 26 corresponds approx. to the scan of the 1st image in the second row.

7. PROBLEMS/RESTRICTIONS:

As mentioned above, scanning the 2D data set of an image does not reduce the data, but simply rearranges them for further processing. Therefore, the above method only allows for a limited bit rate to be transmitted directly into sound.

Taking the linear organized data as frequency information automatically introduces the limits of the audio blocksize for inverse FFT. The Hilbert curve, with an iteration depth of 4, leads to 1024 scan points. Every further iteration step multiplies this number by 4, which gives 4096 scan points for the next iteration. A big but still reasonable blocksize for the inverse FFT is 8192 samples. Due to Nyquist’s theorem, we can use only half of the blocksize to distribute the information of the scan. At a first glance this would suggest to use iteration depth 5, which gives exactly 4096 samples. There are mainly two reasons to reject this choice and to go for 1024.

1. The lowest possible frequency that is represented by an audio blocksize of 8129 equals 5.425 Hz. This frequency cannot be perceived as well as higher ones, if at all. Thus, the lower part of the theoretically possible frequency range has to be discarded in order to transmit information.

2. Ignoring point 1, if we distribute the information of the scan linearly onto the range between 5.425 to 22050 Hz (corresponding to Nr. 0 – 4096 in the frequency bins of the spectrum with an audio blocksize of 8192 samples) we have squeezed most of the information into frequencies of an indistinguishable high pitch. This is because of the exponential frequency pitch relation.

We therefore need to discard the frequencies that are too low. Then, we must try to densely distribute the information of our scan in the remaining low frequencies, and more sparsely in the higher ones, preferably following an exponential function. In our implementation we used an exponential function to map the scan points to frequency bins. Whenever two scanpoints would fall into one bin, the next higher bin was chosen. Therefore we had in the lower frequencies a linear mapping function between the scan points and the frequency bins.

Particularly for rapid data scanning the audio blocksize for the inverse FFT should be small in order to acoustically render as many data frames per second as possible. The restriction through the defined audio blocksize for inverse FFT brings methods of data reduction into play. This method can be easily combined with the SFCs, as outlined in the next section.

8. FUTURE POSSIBILITIES:

Subtractive synthesis is of course not the only possibility to deal with spectral information and there are many other ways of possible sound design. But instead of pointing at issues of sound synthesis, we want to focus in this paragraph more on the structural possibilities of SFCs for sonification, which are much wider than illustrated so far. Particularly if the restrictions of distributing data in a frequency range, as described above, are cleverly combined with data reduction, the sonification approach can exhibit a huge potential.

Let us illustrate such a combination of scanning an image through an SFC together with data reduction. As a first step the iteration depth of the SFC could be set as deep as necessary to cover every pixel in a given image. The resulting linear data can then be reduced by averaging windows of a certain size leading to an amount of data that is usable as spectral information and sound synthesis. Not only would the averages of the windows be of use, but also other statistical properties, such as standard deviation, skewness, quantiles, etc. All these or other measures could serve as interesting parameters for sound synthesis. In a first approach, the playback of these parallel streams of information could be spatialized similarly to the RGB approach of color movie sound synthesis, as described in the section about sound synthesis. Ideally, the size of the averaging windows could be rescaled in real-time just as the iteration depth of the SFC. Then scanning into and out of the image would be a continuous movement, spanning from an acoustic overview to a more detailed data analysis.

Other possibilities are to break down the 2D dataset into a matrix of 2^2 subsets that are scanned by SFCs and converted into sounds. Depending on the available loudspeaker setup and the additional use of virtual sound sources it might even allow for 3^3 subsets. These approaches could be combined with HRTF techniques and headphones together with a head-tracking device. Like this, an image could be transferred from a usually vertically and frontally encountered phenomenon into something that would be horizontally perceivable between both ears.

In this paper we focused on the qualities of SFCs mapping 2D images and movie frames to one dimension, but space filling curves can also map higher dimensions to one. It must be investigated which SFCs keep neighboring relations in those cases. Once they are identified, SFCs in combination with data reduction methods can play an interesting role to acoustically navigate in 3D data that develop in time. Particularly in combination with data reduction the method of context-based SFCs [21] offers interesting possibilities.

It must also be mentioned that there are other types of SFCs apart from the example of the Hilbert curve we used. It has
still to be tested how these other SFCs would affect the acoustic output of this method. For instance, there exists an SFC very similar to the Hilbert curve designed by Eliakim Moore, where the building block “U” can be a closed “O”, see figure 7. Links which show examples of such SFCs can be found in the www, under ref. [22]. Because of the higher symmetry in this SFC, and since there is no obvious beginning for the frequency spectrum to start with, slightly different interaction possibilities would arise: Instead of changing the orientation of the SFC on the image by rotation, it would be interesting to continuously shift the spectrum along the structure of the SFC. This would allow for moving any interesting acoustic pattern into a frequency range were they are best perceived.

9. CONCLUSION:

As we have shown in this paper, SFCs have useful properties for data handling in the field of scientific sonification. A prototype implementation of the suggested methods was shown to transmit interesting qualities of the underlying data structure. The suggested method and the implementation of SFC included different fertile possibilities for real-time interaction during the sonification process. These possibilities took advantage of the close relation between the image data and the produced sounds, so that any perceived sound pattern could be well identified to its visual cause. The biggest strength of SFC is to open up the possibilities to point from perceived acoustic patterns from the sonification directly into complex high dimensional data. Clever combinations with methods of data reduction still have to be designed to open up even more possible applications.

10. ACKNOWLEDGMENTS:

We would like to thank Mathieu Bouchard for his help with GEM programming. Hannes Breitschädel contributed valuable Pure Data code. Hans Dibner and his idea of performative science provided a stimulating framework for working on interdisciplinary questions. We like to thank Ludger Brümmer of the Institute for Music and Acoustics at ZKM and Peter Weibel to support scientific sonification research. For stimulating discussions we want to thank Chandrasekhar Ramakrishnan and Joachim Gossman. Special thanks to Tamar Tembeck for text revision.

11. REFERENCES

[5] Denkgeräusche (sounding thoughts)
[17] Links to space filling curves in the www:
http://www.cut-the-knot.org/do_you_know/hilbert.shtml
http://en.wikipedia.org/wiki/Hilbert_curve
http://www.geom.uiuc.edu/docs/reference/CRC-formulas/node36.html


http://www.cut-the-knot.org/do_you_know/hilbert.shtml
http://www.cs.utexas.edu/users/vbb/misc/sfc/Oindex.html