Relating Timbre and Shape in the Audiovisual Composition S Phase

Lance Putnam

Received 31st March 2017
Revision 16th September 2017

1. Introduction

S Phase is a computer-generated audiovisual composition that was led by a desire to create an abstract form that had a sense of living and breathing and to associate sound and graphics on a deep level. The title refers to the S-phase (synthesis phase) of the cell cycle where DNA is replicated. The DNA must be precisely replicated during this phase to avoid cell abnormalities or death, therefore it is an especially critical moment in the propagation of life. The work incorporates a continuously evolving audio waveform that is presented visually as a three-dimensional space curve (Figure 1). The resulting audiovisual form lapses into moments of replication and symmetry to convey the similar essences characteristic of many types of organic forms.

2. Background and Motivation

The use of technology to visualize abstract curves dates back to at least the 18th century with the introduction of the geometric pen [5]. An early example of sound visualization is Thomas Young's experiment of attaching a reflective wire to a piano string [15]. In 1815, Nathaniel Bowditch traced the motion of a pendulum suspended from two points [12]. Later, Jules Antoine Lissajous visualized similar curves using an apparatus of light reflecting off two mirrors vibrating in perpendicular directions [10]. The geometric chuck is an attachment to a lathe that allows the precise generation and tracing of compound harmonic motion [8]. Later advances in the geometric chuck show an impressive array of results [14]. As these mechanical devices evolved, one sees a clear pattern of increased control, precision, and rapidity with respect to the generation of abstract curves.

S Phase finds its main aesthetic inspiration in Ben Laposky's *Oscillon* series of images [2, 3]. The *Oscillon* images were generated on an oscilloscope in "x-y" mode that was driven by waveforms output by a custom-built electronic synthesizer. The oscilloscope images cover a wide spectrum of results from precise, geometric curves to

Musica/Tecnologia, 11-12 (2017-2018), ISSN online 1974-0050, ISSN print 1974-0042, pp. 111-119
DOI: 10.13128/Music_Tec-23803 – http://www.fupress.com/mt

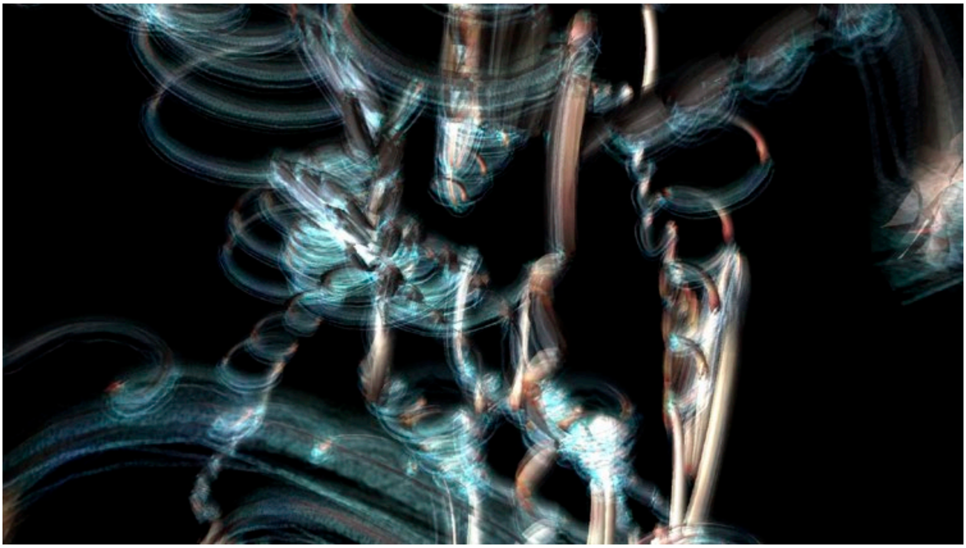

Figure 1. Screenshot of *S Phase*.

organic forms. As such, it seemed well worth exploring this technique further using a computer and advancing similar earlier computer-generated imagery [7, 9].

A key motivation for the work is to explore mathematical relationships between sound and visuals. Abbado's notion of timbre-shape is one way to link aural and visual events [1]. The starting point is identifying timbre and shape as fundamental elements that provide identity to aural and visual objects, respectively. Abbado associates timbre to shape mostly through feeling and intuition. For example, harmonic and inharmonic sounds correspond to smooth and jagged shapes, respectively, and loud sounds associate with bright shapes. We can only speculate that the space of timbre-shape is largely unexplored as it was only with the advent of the computer that abstract forms could be generated both quickly and precisely and mapped into sensory events. The goal of *S Phase* was to explore more automated types of timbre-shape links through various types of signal processing. An especially interesting link is the "bouba/kiki effect" where most people associate rounded words like "bouba" with a rounded shapes and non-rounded words like "kiki" to angular shapes [11, 13, 4].

3. Signal Generation

The signal generator used for the piece is primarily a sound synthesizer as it generates samples at audio rate. The synthesis flow diagram is shown in Figure 2. On the left are synthesis parameters and on the right are output signals. The synthesizer has different paths for visual and aural output. The x, y, and z output signals are used to draw a space curve and the L and R signals are the left and right channels of audio output.

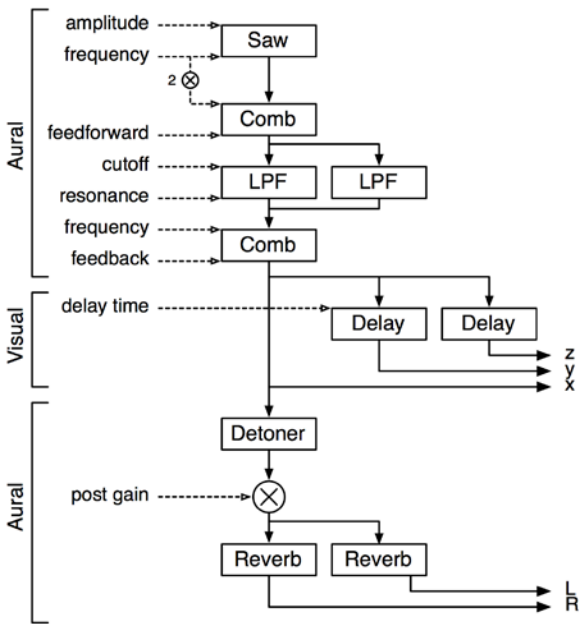

Figure 2. Synthesis flow diagram of *S Phase*.

The first aural stage of the synthesizer employs subtractive synthesis to shape the spectrum of a harmonically rich sound, in this case a saw wave, with several filters. The first filter is a pitch-tracking comb filter that sums the waveform with a delayed version of itself. The frequency of the comb filter notches are set to twice the frequency of the oscillator. Changing the feedforward mix of the delayed version adjusts the strength of the even harmonics. When the feedforward amount is – 1, the output contains only odd harmonics. This is followed by two resonant low-pass filters in parallel to allow more dynamic shaping of the waveform. The last filter is a feedback delay line that effectively allows "copies" of the signal to be generated along with more complex types of long-term motion.

In the final aural stage, the signal is detoned by subtracting a copy of itself delayed by its fundamental period. This has the effect of only passing through changes in the waveform over time and eliminates an otherwise monotonous buzz. Thus, when there is a change in the image, one hears a corresponding sound event. This creates a familiar perceptual scenario where a sounding object typically also exhibits some kind of motion. Reverberation is added at the end to add spaciousness to the sound.

In the visual stage, the audio waveform is expanded in dimension to generate a three-dimensional space curve. To do this, we use a technique known as embedding or the method of delays [6]. The general idea is to construct additional dimensions (in this case two) from time-delayed versions of the original signal. For example, we can construct the x,y coordinates of a circle by using a sinusoid and a version of itself delayed by one quarter period. In the visual processing stage of this system, the amplitude of the source waveform is mapped to the x position, and two delayed copies of it are mapped to y and z positions to obtain a point along a curve in three-dimensional Euclidean space.

To obtain a sufficient amount of visual detail in the space curve while maintaining tightly correlated sound and visual events, we found it necessary to use a relatively high audio sampling rate of 192 kHz. For a typical graphics frame rate of 40 Hz, this equates to a maximum of 4,800 new space curve points per frame. This can be understood as the synthesizer needs to map fixed intervals of time into space and there is a limit to how many audio samples can be generated per frame of graphics.

4. Graphics Rendering

Given the three-dimensional space curve as a sequence of points, additional geometry is generated to make the curve smoother and more solid appearing. The smoothing involves inserting three extra points between each original point using a cubic spline. Without this step, the curve will often appear jagged and overly geometric when there is high-frequency content present in the waveform. Next, two ribbons are generated along the curve at right angles to one another. One curve is opaque to give the curve a certain amount of solidity and weight and the other is transparent to achieve interesting blending effects and texture.

The colors of the curve are derived from synthesis parameters as well as features of the curve. In general, the colors must be subject to as much variation as the geometry and the sounds. That is, they must vary in complex ways both along the curve and as a function of time. We begin by selecting a hue based on the pitch of the waveform. This is done by assigning the frequency to an octave class according the approximate frequency range of visible light from 400 THz - 800 THz. The hue, in [0, 1], is calculated from $x - [x]$ where $\log_2(f/(400\cdot 10^{12}\cdot 2^{-40}))$\bmod1 and f is the frequency in Hz. Frequencies of 90, 180, 360, ... Hz produce a hue of 0 (red) while frequencies a perfect fifth up, 135, 270, 540, ... Hz, produce a hue of 0.5 (cyan). To expand the color palette, the hue is varied along the curve by adding an offset proportional to the curvature. To obtain a proper amount of contrast, a single directional light is applied to the ribbon surfaces from above.

5. Timbre-Shape Relationships

The work displays several connections between the spectrum of the waveform and the shape of the corresponding space curve. The shape of an object is that which remains invariant under uniform scaling, rotation, or translation of the object. If only the amplitude of a sound waveform visualized using embedding is changed, then only the size of the corresponding curve changes. This would not be considered a timbre-shape association as the operation on the sound does not change the shape of the curve. In the following sections, we discuss three basic timbre-shape relations exploited in the work.

5.1 Harmonic Spacing and Rotational Symmetry

Perhaps the most interesting timbre-shape relationship displayed in the work is between harmonic spacing and spatial symmetry. A space curve constructed from odd-harmonic waveforms will exhibit 2-fold improper rotation (S_2) symmetry. That is, rotating the curve on a plane by 180° and then reflecting it with respect to that plane will produce the same figure. When there are three odd-harmonic waveforms driving the *x*, *y*, *z* axes, then there will be three mutually perpendicular planes of S_2 symmetry. Figure 3 shows a transition in the time frame 1:12-1:20 (min:sec) from an exclusively odd-harmonic state to an all-harmonic state.

Figure 3. Transition of the space curve going from a waveform with odd to all harmonics (left to right). The odd-harmonic space curve has three perpendicular planes of 2-fold improper rotation (S_2) symmetry.

5.2 Feedback and Translational Symmetry

Another type of symmetry exhibited in the piece is translational symmetry. This symmetry becomes apparent whenever a waveform is passed through a delay-line with a relatively high feedback amount. Figure 4 is a 1-second section of the piece starting at 4:11 (min:sec) when the feedback of the delay line is set near 1, progressively creating a translational symmetry.

The period of the delay line relative to the pitch of the source sound determines the nature of the "copies" that are produced. Shorter delays create less copies that are more spatially distant, while longer delays create more copies that are closer together. With a more spatially compact source, such as an impulse, the translational symmetry becomes more apparent.

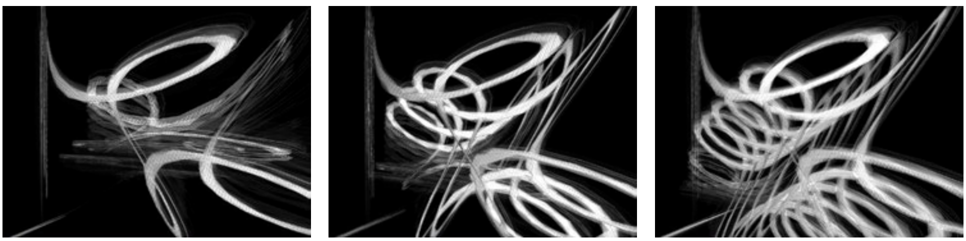

Figure 4. Temporal evolution of a saw wave passed through a delay line with feedback close to 1.

5.3 Spectral Brightness and Roundedness

The distribution of harmonics determines to a large extent how spread out in space the curve is. For dark sounds with more low frequency content, the resulting space curve is generally more spread out and simplistic. In contrast, for brighter sounds with more high frequency content, the resulting space curve tends to wind around and create a more compact shape. As the low-pass filter blocks out more of the high frequencies, the curve becomes more rounded and circular. Emphasizing high frequencies will add more bending energy to the curve creating sharp corners. If isolated frequencies are resonated, loops will be created along the curve. The number of loops is determined by the ratio of the resonant frequency to the fundamental frequency of the waveform.

6. Composition System

The piece was composed using a custom-built interactive program that allowed the synthesis parameters of the system to be modified, stored, and sequenced over time. The overall goal was to have a way of exploring state spaces interactively with the ability to store and subsequently morph between sequences over time.

The interface is divided into three separate components: the observed state, state set, and state sequence (Figure 5). The observed state interface provides control and

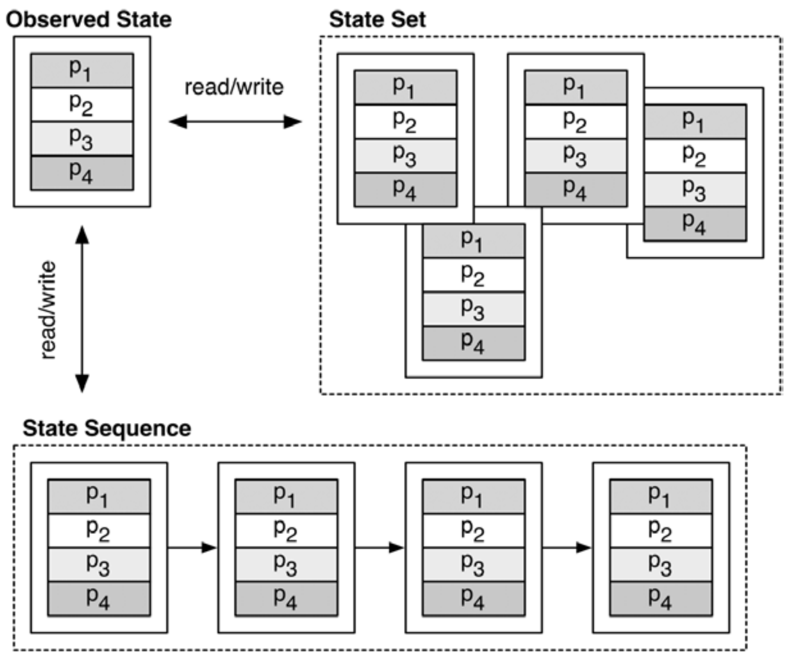

Figure 5. State space interface model.

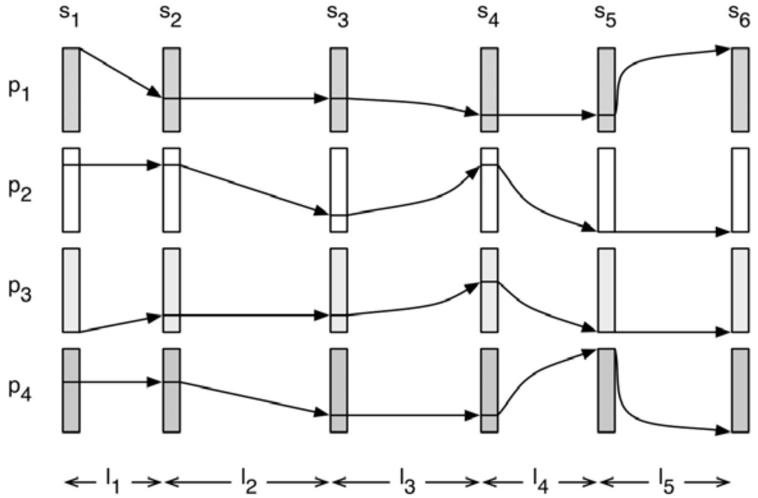

Figure 6. State sequence model where p_n is the nth state parameter, s_n is the n^{th} state in the sequence, and l_n is the transition length between the states n and $n + 1$. The arrow shapes indicate the curvature of transitions between elements.

reflection of the current system state and the primary means for interactive control of the synthesis parameters. It is rendered as a column of numerical sliders that are mapped to each parameter of the state space with appropriate warping functions and ranges. The state set interface stores an unordered collection of states and its main purpose is for saving interesting states that can be used later for sequencing a composition. The state set can store the current system state as a new element or set the current system state from one of its existing elements. The state sequence interface operates identically to the state set, but has a specific ordering of the elements. It provides a means of smoothly interpolating between different parameter states. The edges between state elements, called transitions, have two unique properties. First of all, each transition has a length or duration. Secondly, the transitions can be curved rather than linear. In a normal linear case, the relative time within the transition is mapped directly to the fraction of linear interpolation between the two connected states. With a curved transition, the mapping from relative time to interpolation fraction can follow a power or s-shaped curve. Figure 6 illustrates interpolation across several states in a sequence.

7. Conclusion

Several insights into the mapping of sound into space were made during the production of this piece. First, sound synthesis and transformation methods provide an intuitive and nuanced language for controlling global characteristics of forms. For instance, the overall curvature can be controlled through low-pass filters and the rotational symmetries can be derived from simple harmonic relationships. Another key insight is that there is an inherent memory in the animated forms as a result of the feedback used to transform the sound wave. The more feedback that was used, the more spatiotemporal history of the form was present. This gave more natural and coherent transitions of the form over time to complement the basic interpolation between states.

Generating the space curve using embedding had both strengths and weaknesses. The main strength was the economy of only needing to generate a single one-dimensional signal to produce a three-dimensional shape. This economy, however, led to a subtle kind of perceived symmetry of the space curve. The cause of the symmetry is due to a simple phase relationship between the three signals. Since the signals are delayed versions of one another, any change in the delay amounts leads to box-like sweeping motions. This essentially breaks the spatial symmetry of the space curve as the sweeping motions are always axis-aligned. For example, it is not clear how certain shapes like a sphere or torus could be generated using embedding. For this reason, alternative sound to curve mappings, such as through complex or other higher-dimensional numbers, should be investigated.

During the process of composing *S Phase* many relationships were discovered between sound timbre and graphical curves. These relationships, while being largely implicit in the signals used, demonstrated a rich variety of possibilities and proved to be useful audiovisual material for composing on the meso and macro scales and giving "life" to an inanimate form. The approach used here for audiovisual art holds much promise and demands further investigation as it deeply connects areas of science and mathematics with the study of form and motion in the arts.

References

[1] Adriano Abbado: *Perceptual Correspondences of Animation and Synthetic Sound*. Thesis (M.S.). Massachusetts Institute of Technology, School of Architecture and Planning, Program in Media Arts and Sciences. 1988.

[2] Ben Laposky: *Electronic Abstraction: A New Approach to Design*. Exhibition catalog. Sanford Museum, Cherokee, Iowa. 1953.

[3] Ben Laposky: "Oscillons: Electronic Abstractions", *Leonardo*, pp. 345-354, 1969.

[4] Daphne Maurer, Thanujeni Pathman, Catherine J. Mondloch: "The Shape of Boubas: Sound-shape Correspondences in Toddlers and Adults", *Developmental Science*, pp. 316-322, 2006.

[5] Giovanni Battista Suardi: *Nuovi istromenti per la descrizione di diverse curve antiche e moderne*. Gian-Maria Rizzardi, 1752.

[6] Gordon Monro, Jeff Pressing: "Sound Visualization Using Embedding: The Art and Science of Auditory Autocorrelation", *Computer Music Journal*, pp. 20-34, 1998.

[7] Herbert W. Franke: *Artist and Computer*. Harmony Books, 1976.

[8] John Holt Ibbetson, Bern Dibner: *A Brief Account of Ibbetson's Geometric Chuck, Manufactured by Holtzapffel and Co.: With a Selection of Specimens Illustrative of Some of Its Powers*. A. Hancock, 1833.

[9] John Whitney: *Digital Harmony: On the Complementarity of Music and Visual Art*. Byte Books, 1980.

[10] Jules Antoine Lissajous: "Mémoire Sur L'étude Optique Des Mouvements Vibratoires", *Annales de Chimie et de Physique*, pp. 147-231, 1857.

[11] W Köhler: *Gestalt Psychology*. Liveright, 1929.

[12] Nathaniel Bowditch: "On the Motion of a Pendulum Suspended from Two Points", *Memoirs of the American Academy of Arts and Sciences*, pp. 413-436, 1815.

[13] V. S. Ramachandran, E. M. Hubbard: "Synaesthesia —- A Window Into Perception, Thought and Language", *Journal of Consciousness Studies*, pp. 3-34, 2001.

[14] Thomas Sebastian Bazley: *Index to the Geometric Chuck: A Treatise Upon the Description, In the Lathe, of Simple and Compound Epitrochoidal or ``Geometric" Curves*. Waterlow and Sons, 1875.

[15] Thomas Young: "Outlines of Experiments and Inquiries Respecting Sound and Light", *Philosophical Transactions of the Royal Society of London*, pp. 106-150, 1800.

S Phase

http://www.fupress.net/public/journals/18/2017-2018/s_phase_1080p.mp4