5-18-1988

A study in music visualization

Charles Marsh

Follow this and additional works at: http://scholarworks.rit.edu/theses

Recommended Citation

This Thesis is brought to you for free and open access by the Thesis/Dissertation Collections at RIT Scholar Works. It has been accepted for inclusion in Theses by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.
A Study in Music Visualization
by
Charles D. Marsh

May 18, 1988
Approvals

Adviser: James Ver Hague _________________________________________
Date: __5/20/88__

Associate Adviser: Robert Keough _________________________________
Date: __5/20/88__

Associate Adviser: Dr. Charles Warren ______________________________
Date: __5/20/88__

Special Assistant
to the Dean for Graduate Affairs: Philip Bornarth __________________
Date: __5/27/88__

Dean, College of
Fine and Applied Arts: Dr. Robert Johnston _________________________
Date: __6/9/88__

I, ________________________________, hereby grant permission to
the Wallace Memorial Library of RIT, to reproduce my thesis in whole or in part.
Any reproduction will not be for commercial use or profit.
Date: __5/20/88__
Contents

Introduction ................................................................................................................. 1
History ......................................................................................................................... 3
Development ............................................................................................................. 6
Music Selection .......................................................................................................... 9
The Studio Work ......................................................................................................... 11
Figure 1 .................................................................................................................... 12
Figure 2 .................................................................................................................... 13
Figure 3 .................................................................................................................... 15
Figure 4 .................................................................................................................... 17
Equipment ................................................................................................................ 19
Conclusion ............................................................................................................... 21
Bibliography ............................................................................................................. 22
Introduction

Music visualization, as perceived by this writer, is any creative visual effort whose intent is to convey an understanding of a piece of music; it is a creative translation of ideas from an aural to a visual medium. I use the word creative to make the distinction between an artistic endeavor and simple coding which is what written music constitutes. This paper will serve as a documentation of my master's thesis project in which I attempted to meaningfully translate ideas from already existing musical works, three of Frederic Chopin's Preludes for the piano from the composer's Opus 28, into a visual medium, in this case, two-dimensional computer animation.

My intent at the outset of this project is best expressed, (and rightfully so), in my thesis proposal.

The purpose of this thesis will be to develop several different approaches to music visualization, and to create a series of computer animations utilizing each of these approaches.

In addition to being a personal artistic expression, I wanted the thesis to be an investigation of the visualization process in a larger sense; I intended to consider some of the limitless number of approaches to the subject so that when the studio work was complete I would be able to place it in some sort of context. I feel that there are many observable, universal parallels between the two forms of communication; parallels, similarities which, if considered and classified, would constitute a sort of knowledge base, a set of rules or guidelines by which music visualizations could be created and evaluated. To develop such a comprehensive body of knowledge is certainly beyond the scope of this project; I contented myself to try to create artwork which would satisfy me both as an artist and a musician. Having studied both art and music, I have always been interested in the possible parallels one might draw between the two disciplines. This project is the first that has stimulated both my artistic
and musical selves.

The topics in this thesis report will be covered in roughly the same order in which I considered them in the past few months; thus, this paper will serve as a documentation of my creative and organizational processes.
History

Music visualization is a subject that has been given extensive treatment by many prominent painters such as Paul Klee and by animators such as Norman McLaren and Walt Disney. Architecture has even served as a medium: Castle Neuschwanstein in Bavaria was reputedly built as a tribute to the music of Wagner.¹ While not intending to devote a great amount of time to research, (I wanted my ideas to be relatively uninfluenced by outside sources), I nevertheless found much of interest in the work of these artists with respect to music visualization in general.

Paul Klee, in his notebooks, writes extensively concerning the translation of musical forms, particularly rhythmic patterns, into visual images. Not concerned so much with specific compositions, his writings and artwork refer to musical forms in a more general sense. His Three-Part Time/Quartered, a gouache from 1930, graphically represents three part rhythm using regularly spaced patches of color in three values: dark, medium, and light. Many of Klee's other paintings also reflect this translation of regular temporal rhythms into spatial ones.

Klee de-emphasized the importance of lines as 'shape-bounders', stressing their individuality and importance as discrete entities. He likened lines to melodies and compared their interaction to that of melody lines in counterpoint. According to author Andrew Kagan, Klee tried to develop a science for creating pictures just as there was a science to define and create music.² Fugue in Red, a 1921 watercolor pictures that repetitive musical structure as stacks of similar forms that change subtly in shape and color. It was Klee's scientific, exacting approach to such a traditionally subjective area that intrigued me; his attempts to find one-to-one correspondences between aurals and visuals made me consider applying a more empirical spirit to my project.


Where Klee's work inspired me to place my work in a carefully considered context, that of Norman McLaren seemed to suggest the opposite, namely, to rejoice in each work for its own sake. McLaren's animated films are wonderfully lighthearted and free-form. His *Begone Dull Care*, 1949, which I had the great good fortune to view last year is a visual accompaniment to music of the Oscar Peterson Jazz Trio. The images have a spontaneous quality, owing in part to the fact that they were drawn and scratched manually onto pieces of blank motion-picture film. In an interview with author Maynard Collins, McLaren says of his film: "It's the movement in the music that captures my imagination first,...the images, the particular shapes are relatively unimportant." The overwhelming sensation while viewing the film is one of movement. The music is quick, rhythmic, and light, and McLaren's visuals complement it perfectly, with abstract shapes jiggling and swooping around the frame. Still, McLaren does seem to be concerned with shape to a degree; a continuity can be seen in his treatment of different sounds. Lines representing notes vary in their weight according to pitch, and high brilliant notes tend to be represented with small circles or dots.

*Canon*, 1964, is a more analytical film, akin to Klee's *Fugue in Red* in that it is an analysis of a musical form rather than a specific piece of music. The soundtrack for the film was composed simultaneously with the development of the visuals which is apparent when one sees the finished work. Child's blocks, human actors, cats and a butterfly, each representing different melody lines, act out a complicated, intertwined series of movements that describe the melodies' interactions. It is an interesting use of 'real' objects to relate an abstract idea.

In 1971, McLaren took advantage of the fact that soundtracks were optically read to produce *Synchromy*, a film where the sights and sounds truly are one. Cards printed with precisely measured series of lines were photographed so that they appeared in both the image and sound areas of the film, with the result that the image the viewers see is actually producing the sounds that they hear. To me, *Synchromy* is interesting in that it shows that an

---
3 Maynard Collins, *Norman McLaren*, (Toronto: Canadian Film Institute, 1976), p.61
aural/visual translation, in this case pitch height/line width and spacing, could be physically demonstrated or proven. The wide, heavy lines produce low tones, while the thin, delicate ones produce high tones.

The 'Toccata and Fugue' segment of Walt Disney's Fantasia was the only part of that film that directly addressed the problem of finding pictorial parallels to musical sounds. Small, ethereal shapes in pastel colors shimmer in time to the notes in the upper register, the violins and woodwinds, while huge objects in dark, saturated hues represent the bass tones, the brass and string basses. An abstract film, like Synchromy and Begone Dull Care, 'Toccata' uses similar devices; some of the same parallels seem to be in operation. The relationship between object size and color, to the notes' pitch and loudness is strikingly consistent between the work of the two artists.

The term 'music video' would seem to mean 'visual music'. For a few videos it is an accurate name but not for most. Generally, music videos are concerned more with the performers than with the music being performed, and if they aren't miniature concert films, they are usually dramatizations of the lyrics being sung. In spite of their promising name, the music videos that I have seen offer little in the way of music visualization.
Development

The main body of work, the most time consuming part of this project was the consideration, evaluation of various organizational approaches to the subject. Ideally, the studio work should consist of animations which would enlighten the viewer both by their individual content and their mutual contrast. The task, then, was to break down the subject in different ways, evaluate the various classification schemes, and to choose one or more that could be effectively illustrated using three of four short animated sequences.

All the music visualizations that I had seen seemed to fall into one of three basic categories: narratives, music and visuals that tell some sort of story, what I call 'mood' paintings, non-narrative images that simply evoke the mood of a musical work, and abstract images whose elements represent specific musical forms. Examples of all three can be found in Disney's Fantasia. 'The Sorcerer's Apprentice' is a visual narrative based on a piece of music which itself describes a narrative. 'Ave Maria' presents tableaux that simply express the solemnity of Schubert's music. As described in the previous section, the 'Toccata and Fugue' creates sound-specific abstract parallels.

Another, similar classification creates four groups: narrative/objective, telling a story with 'real objects'. narrative/abstract, non-narrative/objective, or total abstraction, neither narrative nor representational as in 'Toccata'. This is slightly more flexible, accounting for some of the crossovers possible in the first structuring. A non-narrative/objective visualization could be represented by McLaren's Canon, where everyday objects act out the various intertwining melodies in the soundtrack. Non-narrative, it complements the puzzle-like structure of the music, playfully revealing parts, then the whole. It would be neither narrative nor objective if the 'actors' were non-representational objects.

This method of classification, a continuum between the referential and the abstract may be useful in classifying existing works but not, I thought, as a guide to creating new ones. At least not on the relatively small scale of this project. It is a scheme that encompasses far too many possibilities. It would require realistic rendering for the objective sequences, something difficult to do
on the Genigraphics system, (see the section titled ‘equipment’). Furthermore, in a discussion with Dr. Charles Warren, a member of my thesis committee, we questioned the value of working with narrative material at all. Music based on a literary or artistic work such as The Sorcerer’s Apprentice by Charles Dukas or Danse Macabre by Camille Saint-Saens, was created with specific imagery already in mind. To visualize it would be, to an extent, redundant. Finally, to create the animations around that structure would in itself be redundant. Disney already did it; the point has been made.

In his interview with author Maynard Collins, Norman McLaren says of his film Begone Dull Care: "The movement is the common denominator between the picture and the sound." Since the medium for this project is animation and since animation is movement by definition, an organizational framework based on the comparison of different types of motion would be worth considering. At its simplest, motion in animation consists of two kinds of movement: subject motion and viewer motion. If one considers these to be basic units, there are four possible combinations: static subject/static viewer, static subject/moving viewer, moving subject/static viewer, and moving subject/moving viewer. These classifications apply to live-action films as well, of course, with the moving viewer being simply a camera move. An example of a static subject/moving viewer situation would be where the camera pans across an empty landscape. Of course, these situations are not mutually exclusive; in a film the situation usually changes often depending on the camera work and the action being filmed.

The viewer of a film or video tends to identify strongly with the moving camera, perceiving it as being his or her own point of view. This gives each of the four situations listed above a very different effect. A static viewer and a moving subject yield a theatrical effect; the audience is passive, all interest is focused on the subject. With a changing point of view, the viewer takes an active part in the proceedings; with a static subject, the experience can seem like a journey or a tour through the scene; with an active subject the experience is even more participatory. The fourth situation, that where both the

viewer and subject are static is important in that it strengthens the other cases by opposing them. Just as black, the absence of light, can enhance color compositions by providing contrast, so the absence of motion can be used to punctuate activity.

It was at this stage in the development of the thesis that I turned to the second part of the process, the selection of the music. The viewer/subject motion framework I found satisfactory as a groundwork on which to create the final animations. It was neither too restrictive, it didn’t preclude working with any other variables, color for instance, nor did it require the investigation of matters that lay outside the topic at hand, story and character development, for example.

By this time I had completely dismissed the idea of including narrative elements or characters in the final animation. Instead, I resolved to follow the example of Paul Klee and work only with abstract visuals.

To me, the timing of the transition from preliminary to studio work was fairly critical; too early and the studio work may have been too subjective, based entirely on instantaneous impressions from the music, not on a carefully considered framework. Too late, and the visuals might owe too much to preliminary research and the work of other artists and not enough to the selected musical compositions. As it was, I had developed a workable structure into which I could place my ideas as I received them.
Music Selection

The organization of this paper would suggest that the music selection process began after the organizational work had been completed. In fact, it was simultaneous to a degree. As the different systems were considered, narrative, objectivity, motion, etc., so were various pieces of music that might fit into these systems. The music which at first seemed most attractive was the first to be dismissed. Several programmatic works came to mind including the 'Bydlo' from Mussorgsky's *Pictures at an Exhibition*. In the animation I would have highlighted the dynamics of the piece, which is really one long crescendo-decrescendo from start to finish. I had in mind a scene which would be viewed in close-up initially then revealed gradually with a zoom-out as the crescendo progressed, thus parallelling the dynamic level and orchestration with the viewer's proximity to the subject. It was rejected, not so much because it was 'pre-visualized', (the imagery that I had in mind was unrelated to the painting that inspired the music), but because anyone familiar with the piece might not be able to clear their mind of the original visuals long enough to appreciate my new ones.

I considered the possibility of comparing the works of two or more composers or the music of different periods. That idea was dropped almost immediately as it would, like the treatment of narrative material, swing the focus of the thesis away from the abstract visualizations that I had envisioned towards something more like a music history project.

It was while I was developing the viewer/subject framework of organization that I considered two works for the keyboard: Chopin's prelude no. 10, and a two-part invention by Bach. Both had musical forms that strongly suggested visuals to me. The main theme of the Bach is a miniature canon, a run upwards in an harmonic minor scale followed by a run downwards which is a repetition of the first run, but in reverse: a mirror image. I planned a polygon whose configuration suggested the first half of the theme; the shape would reproduce and mirror itself as the melody doubled back and other voices joined in. The Chopin prelude has a striking repetitive structure. A set of downward
runs followed by chords is repeated four times. The rhythmic structure of each section is identical which suggested to me an image consisting of four identical parts, perhaps oriented different ways to reflect the subtle changes in character between the sections. This piece, one of the ones featured in the studio work will be discussed in greater depth in the next section.

A look at the rest of the twenty-four Preludes revealed two others whose structure seemed particularly complementary to the tenth. It was the relationship between these three pieces that led me to select the viewer/subject motion framework as the guide for creating the animations.

The Preludes, on the whole, are ideal material for this project. They are short, concise musical statements. Their brevity allowed me to use them in their entirety; to have to edit a piece of music was something that I wanted to avoid. Leopold Stokowski's attempts at streamlining the music for Fantasia was a major source of criticism for the film. What is more, the Preludes are very much a series, each written in one of the twenty-four major and minor keys in the circle of fifths. The idea that my series of animations could later be expanded to twenty-four amused me. Finally, the Preludes, like virtually all of Chopin's music, are abstract, owing nothing to visual or literary material. Descriptive names attached to his compositions, the titles 'Raindrop' Prelude, 'Butterfly' and 'Ocean' Etudes, were not Chopin's, but were added after the music was published.

The Studio Work

As I have previously mentioned, the striking feature of Chopin's Prelude no. 10 in C-sharp minor is its modular organization. The Prelude consists of the quadruple repetition of a series of runs and chords. Of the three preludes examined, it is the one whose structure is the most apparent to the listener. The rhythmic pattern of each of the four sections is virtually identical: a brilliant cascading run downwards of thirty notes, (six groups of five), ending with six somber chords. The consistency of the Prelude's components, the straightforwardness of its exposition, and its modest length all suggested to me the construction or assembly of a simple object. Since the listener is so strongly impressed with the overall structure of the work, I felt that the viewer of the visualization should be as well.

The music, having four equal parts each containing thirty-six notes, (six sixes, a perfect square), practically demands that a square be used to represent it. In the animation, the notes of the runs are pictured as circles which fall (as the melody descends), into rows corresponding to their quintuple groupings in the music. Circles were chosen to represent individual notes (the simplest unit in music), because they are arguably the simplest shape, having a single, unbroken edge. Squares depict the chords, both as miniature echoes of the Prelude as a whole and as complex sounds. The chord-squares appear in a row next to the completed grid of notes from the runs, closing off each module of the piece as the viewer-camera rotates towards the next. The fifth group of chords is the last thing that the listener hears, informing him that there will be no fifth module and that the piece is over, (see figure 1).

As implied in the last paragraph, the listener cannot perceive the piece as a whole, see the music as a square, until the end is reached. Thus, the animation only reveals the entire image at the end. To accomplish this, both the subject, (the grid), and the viewer have motion; as the grid composes itself, the viewer turns above the image, slowly moving farther away until the square is entirely visible, (see figure 2).
Figure 1
Figure 2
In the animation for Prelude no. 10, both the subject and viewer move, but in the visualization for Prelude no. 3, the viewer is stationary which makes for a less involving experience for the viewer. In no. 10 there is a feeling of participation in the construction of the image, but in no. 3 one stands back at a distance to watch the action. The reason for this is that, for me, what is most striking about Prelude no. 3 is not its repetitive structure, (it is rather uniform), but the contrast between the two voices as divided between the two hands. The left hand in no. 10 is no more than an accompaniment, a grounding for the intricacies of the right hand voice. In no. 3 the left hand competes with the right for the viewer’s attention. The visualization for this prelude, then, accentuates this competition and contrast.

The accompaniment is a rapid series of notes that starts from the tonic of the current chord, races up through almost two octaves, then down through roughly the same notes back to the tonic. It's a feverish and persistent motif that implies verticality in its wild race up and down the scale, and circularity by its consistant return to the tonic of the chord. Representing the various notes in the left hand is a circular bundle of rods which begin at 'ground level', a plane representing the base note of the chord, and rise to the heights reached by their corresponding notes in the scale. The base to which they return changes with each modulation into one of the piece's four keys. The right hand motif is pure verticality; single clear notes or simple chords (ellipses in the animation), play out the melody in huge vertical leaps in pitch, some almost an octave. As in the animation for Prelude no.10, spatial height represents pitch height, but here it is much more precisely organized. As the interaction between the two voices was to be emphasized, great care was taken to make the spatial increments between the heights of the rods and ellipses correspond with the true intervals in the music.

Color is used similarly here as in no. 10. Brilliant upper register notes are brighter, while lower notes are darker, more subdued. In no. 10 the minor key is reflected by the predominance of blue. The blue-purples of the G-major Prelude give way to bright oranges and yellows in the high notes (see figure 3).

The viewer is in motion once again with the visualization to Prelude no. 12 in G-sharp minor, but this time the subject remains stationary. Number 12
Figure 3
offers the listener the least information as to its overall structure and the
directions its theme will take next. Neither of the other two preludes in this series
stray very far at all from their home keys, but G-sharp minor twists and turns
every which way. This unpredictability suggested to me a road where the
traveller never knows which way the next turn will take him. In the animation the
viewer flies above an abstract landscape upon which an object (identified with
by the viewer), moves along a road lined with tree-like green disks. The turns in
the road coincide with the chord changes and modulations in the music.
Moving along the road, the object exactly marks the melody's progress through
time; the 'trees' correspond to the percussive bass notes in the accompaniment
which sound as the travelling object passes them by (see figure 4). The notes
of the melody itself, rather, the pitch height of the melody is indicated by the
viewer's height above the landscape. At the opening of the prelude, the
melody is relatively low in pitch; as it rises and falls in the scale the viewer rises
and falls above the landscape reaching his or her zenith (along with the
melody), just before the restatement of the main theme, halfway through the
piece. Any section of the music occurring in the home key of G-sharp minor is
accompanied by a lower to upper (south to north), movement by the
object/viewer.

It is during this animation that absence of motion is used to its greatest
advantage. As discussed in the section titled 'development', motionlessness as
a compositional device can be used to punctuate or accentuate motion. The
forward, driving nature of the melody in no. 12 is depicted by the motion of the
object/viewer. When the music pauses, the object and viewer both come to a
complete standstill which accentuates the loss of forward impetus in the music.

Where the music is repetitive, so is the 'scenery'. During the
development section, the viewer follows a course which brings him or her back
to the same ground for the restatement of the main theme. Thus, this
visualization attempts to create a logical, physical space to represent the
musical space heard in the soundtrack. Unlike the tenth prelude, however, one
never sees the entire space because that space is never fully described in the
music; the twists and turns are abrupt, the modulations unpredictable, and the
repetition minimal.
Figure 4
The series as a mode of expression has always held a great deal of interest for me. A single entity is, to me, always made more understandable when placed in the context of a series. Before these animations were conceived I had already prepared one context in which to place them; the viewer/object motion series. Other frameworks could be said to apply to them as well. With respect to formality they range from very formal: the first animation, rigidly organized, through moderately formal: the second, concerned less with structure as with texture and color, to apparent informality with the third, which, though carefully organized, is calculated to match the wayward, meandering spirit of the prelude in G-sharp minor.

As types of investigation the animations form another sequence. The first is compositional or constructional, revealing information about the music through the assembly of its representative image. To picture the instantaneous quality of the music, its total configuration at any one time is the aim of the second. The third animation translates time into distance; the music's passage through time becomes the viewer's passage through space.

My decision to use only abstract imagery was followed shortly by a resolution that I would include no visual information that did not arise out of musical ideas. Extraneous matter, I thought, would only be confusing and would dilute the effectiveness of the project. Even so, there are some details whose presence is questionable. The diagonal shadows cast by the vertical rods in the second animation have no direct link to the music. I included them to help define the ground plane which is important as it depicts the tonic of the home key: G-major. The shadows of the falling circles in the first animation are present for the same reason. The fountain-like structures that are seen at several junctures in the third animation were originally included as landmarks, markers to orient the viewer, particularly in the beginning where the object goes around the loop in the road. It was originally intended for the angle of view to be much narrower, in which case the direction of movement may have been unclear. They were made redundant when the viewing angle was made wider, but I decided to keep them anyway as they provided a focus around which the turns in the road could be made.
Equipment

All the animations were created on the Genigraphics 100V computer graphics system. Little of interest can be said about the actual animation process; the system was used conventionally without any attempt to push it beyond the capabilities for which it was designed. The capabilities of the machine were, in fact, taken into consideration when making certain aesthetic decisions. Representational artwork was dismissed not only for the reasons discussed previously, but also because of the Genigraphics' inadequacy in rendering animated objects convincingly.

The most challenging technical aspect of the project was the synchronization of the visuals with the music. Once the music was transferred to three-quarter inch videotape from digital compact disc, my original intent was to carefully time the music with a stopwatch, generate the animation using those figures, record the animation onto a separate videotape, and then edit the music and visuals together onto a third tape. This process promised to be very time-consuming and even expensive if done commercially. It eventually occurred to me that a much more efficient way to go about it would be to record the animation section by section onto the tape that already contained the music, correcting for any inaccuracies during the process. This is what I did; the results were satisfactory and a videotape generation loss was avoided.

For those familiar with the Genigraphics, here is a more detailed explanation of the process. First, the position on the tape where the music begins should be located. This can be done simply by playing the tape and resetting the counter to '0' at the moment the music starts. Accuracy at this stage is not important. Then, Genigraphics frame code should be placed on the tape for ten or so seconds before and after the place where the music begins. This can be done by recording a black still-frame from the controller. The tape should then be rewound to a position several seconds before and after the place where the music begins. If the counter is at '-8' then eight seconds should be recorded; (recording
the ten or so seconds of black still-frame first prevents timing problems due to the four-second title frame inserted by the controller for new sequences). When the tape is played, the still image should disappear roughly when the music starts. By recording additional frames of the still, or over-recording some of the still with black frames, the exact frame number of the music's beginning can eventually be located. Once it is located, the rest of the animation can be appended, section by section. If the stopwatch timing was accurate, synchronization should be maintained. Any inaccuracies can be corrected by adding frames or starting new sections early.
Conclusion

To close this report it remains to address two points. The first deals with the independence of the visuals in a project such as this. Should the visuals be required to stand alone? In this medium, animation, and in a project whose intent is to create artwork that is to be viewed over time, I think it justifiable to consider the music and visuals as inextricable. The temporal nature of music, its rhythm is the reason for choosing animation as the target medium in the first place. The animations I have created are derivative; their content was drawn from the music and was not meant to stand alone.

Finally, what is next? What is the logical extension of this project (apart from visualizing the other twenty-one preludes)? A more classically experimental approach could be tried, for example, a series of visualizations of the same musical composition, each using different sorts of imagery. One could express pitch height using color intensity, another with viewer proximity, another with some other visual variable. It is this sort of experimentation that might lead to a standard vocabulary of aural to visual translations, if such a thing could exist. For my part, I think that the success of any music visualization will always depend heavily on the particular tastes of the viewer.
Bibliography


