Brian Evans
University of Alabama
Department of Art and Art History
103 Garland Hall, Box 870270
Tuscaloosa, Alabama 35487 USA
brian.evans@ua.edu

Foundations of a Visual Music

> In the moving image we have a synthesis—the spatial counterpoint of graphic art, and the temporal counterpoint of music.
>
> —Sergei Eisenstein

> There is a form of film that is trying to evolve that area of thinking which I call "moving visual thinking." And it is intrinsically a visual music . . .
>
> —Stan Brakhage

Abstract animation, often called *visual music,* can have a structural base similar to that of absolute music. In this article, I propose a groundwork for a practical theory of visual music composition. The approach builds from the simple premise that the resolution of tension moves us through time. The article will review some fundamentals from art and design and apply them to temporal structure. To establish the idea of visual consonance, the design principle of proportion will be presented. From this, I will develop and codify the idea that it is possible to resolve visual dissonance to consonance, and so move a viewer through time in a way similar to tonal harmony in music. We will then briefly consider problems of color. The article will review some traditional approaches to color harmony and suggest a simple hierarchical approach to working with color in time.

Expanding to a more comprehensive system of time-based design, the writings of Russian filmmaker Sergei Eisenstein will be mined for raw material (Eisenstein 1942, 1949). Eisenstein believed that "art is always conflict" (Eisenstein 1949, p. 46). For this article, conflict is defined as the opposition of forces that motivates and shapes action. The opposed forces are dissonance/consonance or tension/release. Eisenstein's writings contain many ideas that are useful in the development of a theory of visual music—especially his definition of cinematic montage. From the concepts of montage and time-design, a tentative but useable grammar for structuring time-based visual art can be developed and eventually applied to the integration of sound and image.

For this article, I created some simple video examples as illustrations. Where useful, I will also refer to the works of some visual music pioneers and other filmmakers. (Pieces were selected based on relevance and availability, and the list is by no means exhaustive.)

Computer Music Journal, 29:4, pp. 11–24, Winter 2005

Visual Music and its History

Visual music can be defined as time-based visual imagery that establishes a temporal architecture in a way similar to absolute music. It is typically non-narrative and non-representational (although it need not be either). Visual music can be accompanied by sound but can also be silent.

One can imagine the origins of visual music going back to the discovery of fire and dancing shadows on cave walls, or perhaps reflections of clouds when first seen on the surface of rippling water. In more recent history, color organs (instruments that projected colored light) were seen by the public as early as the 18th century with Castel's Ocular Harpsichord. In the 1920s Thomas Wilfred toured the United States and Europe performing on his Clavilux, an early electrical instrument that created clouds and streams of continuous color. This history intensified over the last century, parallel with the development of cinema.

While the Hollywood-style narrative dominated (and dominates) cinema, pioneering filmmakers were working to develop a non-narrative language of light. Typically non-academics, these pioneers worked in an experimental tradition similar to that of music composers such as Eric Satie, John Cage, and Karlheinz Stockhausen. Among the early filmmakers were Germans Walter Ruttman, Viking Eggeling, and Hans Richter. Oskar Fischinger, also from Germany and moving later to the United States, worked over thirty years creating abstract

animations (Moritz 2004). The career of New Zealander Len Lye, working mainly in London, spanned decades. American pioneers included John and James Whitney, Mary Ellen Butte, Stan Brakhage, and Jordan Belson. Norman McLaren, Evelyn Lambart, and their colleagues at the National Film Board of Canada also created significant works. There were many others (Russet and Starr 1976).

They all struggled against financial, institutional, and technical barriers, yet they left a body of work that provides a base from which to construct a theory of visual music. This work is finding renewed interest as evidenced by recent exhibitions. The *2003 Sonic Light Festival* in Amsterdam featured many of the pioneering animations, as did *Sons & Lumières, Une histoire du son dans l'art du 20e* siècle at the Centre Pompidou in Paris, closing in January 2005 (Cruse 2004; see also www.sonicacts.com/03). Also in 2005, the Los Angeles Museum of Contemporary Art opened the exhibit *Visual Music: Synaesthesia in Art and Music Since 1900,* which then moved to the Hirshorn Museum in Washington, D.C. (Brougher et al. 2005). Much of this work is becoming easier to find on video and DVD at online sources such as the Iota Center (www.iotacenter.org), the Center for Visual Music (centerforvisualmusic.org), and even Amazon.com.

Visual Music/Absolute Music

The concept of *absolute music* can be constructed from Igor Stravinsky's statement, "Music means nothing outside itself" (Stravinsky 1956). This illustrates a common mindset held by many composers for centuries. Form and content were often one in instrumental music. (This of course does not include programmatic music and especially the occasional attempts to imitate real-world sounds, such as bird calls, thunderstorms, etc.) Today, acoustic and electronic instruments can be used as generators of abstract sounds without a referent in the real world beyond the instruments themselves. These nonreferential sounds are used to create abstract temporal structures, from traditional musical forms to those of a more experimental nature. (Many of these new forms also find correspondences in visual music.)

A simple definition of absolute music is the structuring of time with the materials of sound patterns (excluding again program music and literary forms such as song, opera, and theater). Stravinsky's comment refers to music composition as the making of "art for art's sake." A musical work seeks from the listener an aesthetic response to the perception of sonic pattern—the appreciation of "significant form." This was a primary focus of modernism and its formalist leanings (Bell 1914).

When visual artists discuss composition, they are generally referring to static design, the formal distribution of objects in the picture plane or in three-dimensional space. (This article will deal exclusively with two-dimensional space.) In the 20th century, artists sought to express design in the abstract—significant form in visual space. They began to consider seriously their work in a musical way, and the abstract form of the work was the content. Kandinsky, considered by many to be the father of abstract painting, said:

> A painter . . . in his longing to express his inner life cannot but envy the ease with which music, the most non-material of the arts today, achieves this end. He naturally seeks to apply the methods of music to his own art. And from these results that modern desire for rhythm in painting, mathematical, abstract construction, for repeated notes of color, for setting color in motion. (Kandinsky 1914)

Using music as a basis for visual work was a concept explored by many artists in the early 20th century. Notable is the work by Sergei Eisenstein, a Russian avant-garde filmmaker who put these ideas into practice and, fortunately for us, wrote down many of his theories of cinema. These ideas still have potency and are especially relevant in discussions of visual music, time design, and the combining of sound and image.

Visual Consonance

For the visual artist, composition is "the arrangement of elements and characteristics within a defined area . . . a grouping of related components that

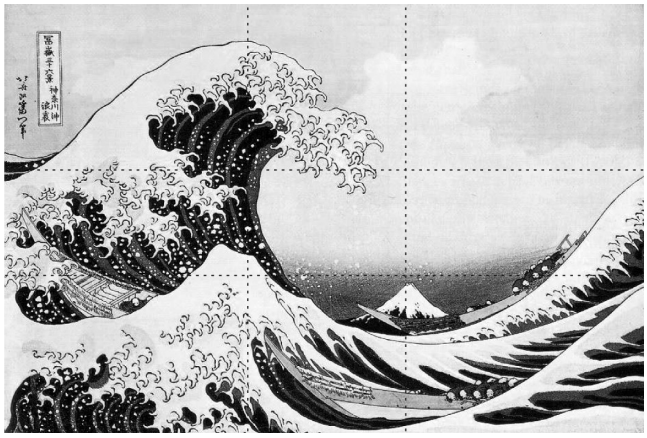


Figure 1. My harmonic analysis of Hokusai's The Great Wave off Kanagawa *(ca. 1831).*

make sense together . . . balanced by an overall appearance of continuity" (Bowers 1999). Good visual composition achieves harmony. Art theorist Rudolph Arnheim said, "One of the basic visual experiences is that of right and wrong" (1966). This idea forms the basis of design foundations as taught in most art schools. First-year art students study two-dimensional design and learn the principle of visual "rightness" and how to achieve it with the conceptual visual elements of point, line, plane, and volume. These elements manifest themselves as objects in a defined space, expressed as size, shape, color, and texture, establishing dynamic relationships as these elements interact.

Visual "rightness" is visual consonance. This becomes an axiom from which we can build a grammar of visual music. From this premise, composing visual music is a simple process. If rightness is codified and understood, wrongness is easily defined by not being right. We might call this wrongness *visual dissonance,* that is, visually active moments of tension in a temporal design. Progressing from visually wrong to visually right moves us from dissonance to consonance or tension to release, just as in music a dominant harmony resolves to a tonic (e.g., V–I). As in music, much of the interest and energy is in the moments of tension. (In tonal music theory texts, how many more pages are devoted to the dominant chord and all its variations versus the tonic chord?)

Using the tension/release construction, we can move dynamically—some might say musically—through time with visual materials. We can establish visual cadence points. With cadences, we can articulate units of time and so develop larger temporal units such as motifs, periods, and phrases.

Proportion

Consider a well-known design principle, the golden section (usually symbolized as the Greek letter ϕ or phi). A proportion of roughly 1:1.618, it is prevalent in the visual arts worldwide and throughout history (Huntley 1970). Figure 1 shows my brief analysis (what is sometimes called an harmonic analysis) of Hokusai's famous *The Great Wave off Kanagawa* (ca. 1831). The cut of the golden section both horizontally and vertically illustrates its utility as a compositional device. Many important points of attention lie on these formal lines: the highest wave is placed on a vertical, the ocean level on a horizontal, and Mount Fuji at an intersection of a horizontal and vertical.

The visual composition has a dynamic balance or symmetry in the picture plane. We see elements of repetition. (Note the wave at left-middle matching the shape of Mount Fuji, and the foreground shape of the big arching wave mirrored with the background sky.) We are not concerned here with the subject matter, but rather the underlying principle of composition or design.

Figure 2 shows my analysis of a drawing of an elk from the walls of the Caves of Lascaux in France. (Dating from approximately 13,000 BCE, it is doubtful that the golden section was applied consciously, although the drawing of the rectangle on the wall is curious as it predates Euclid by over 10,000 years. Even more unusual is the fact that it is a golden rectangle!) The animal seems to balance on the lower left intersection of phi as it divides the bounding rectangle on both the horizontal and vertical.

Whether by instinct or design, use of universal proportions is a common and effective formal device in visual and music composition. Proceeding from simple symmetry to a balanced but active composition is a useful skill in creating visual music. Phi is just one of many proportions that has been explored over the centuries. Jay Hambidge's studies in *dynamic symmetry* showed the applica-

Figure 2. My harmonic analysis of a drawing of an elk from the walls of the Caves of Lascaux in France (ca. 13,000 BCE).

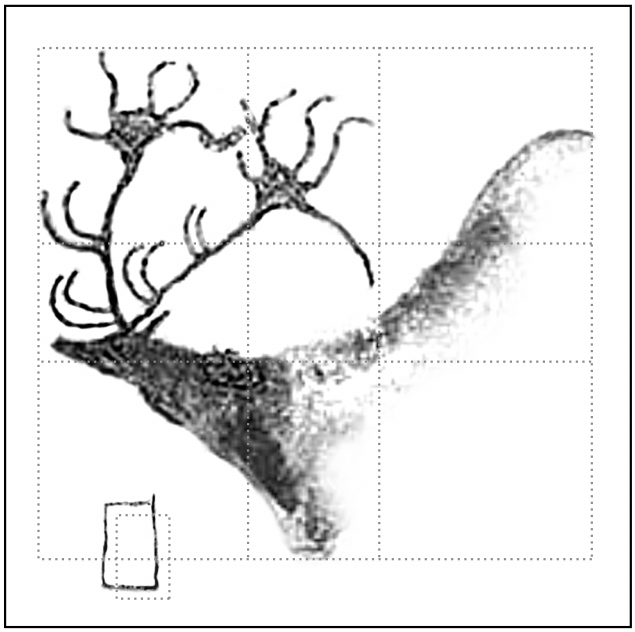
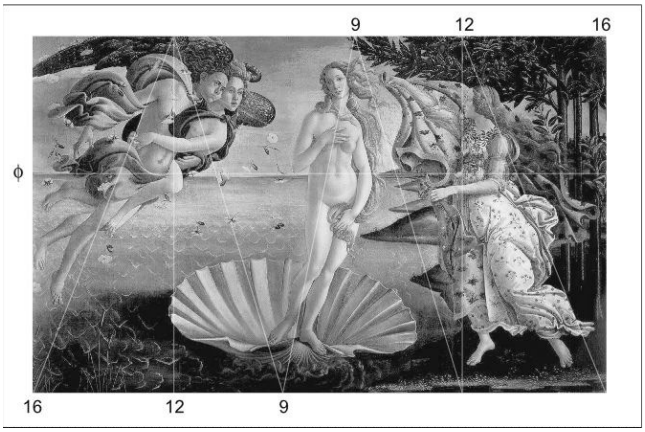


Figure 3. From a harmonic analysis of Botticelli's Birth of Venus *by Charles Bouleau (1963).*



tion of many mathematical constants, including √2, √3, and various other geometric constructions, illustrating their application in everything from Grecian urns to furniture to typography (Hambidge 1919, 1932). The "secret" of the golden section was revealed during the Renaissance in Luca Pacioli's book *Da Divina Proportione,* which was widely read by artists of the time.

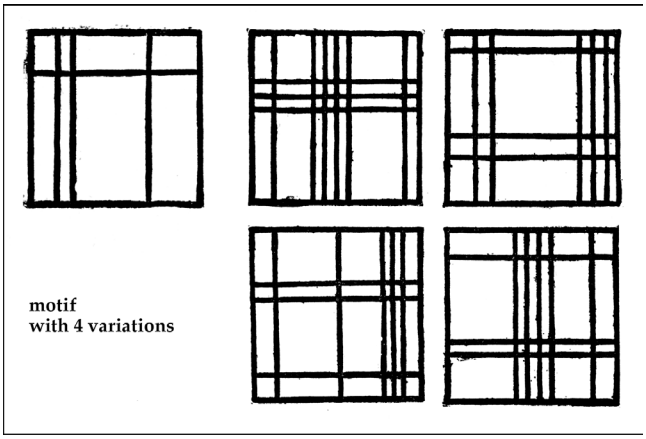
Also at this time, Leon Battista Alberti published his writings on painting and architecture. He described the system of one-point perspective and the application of the Pythagorean proportions of music to visual construction. Figure 3 is taken from a harmonic analysis of Botticelli's *Birth of Venus* by Charles Bouleau (1963). Here, we see an application of phi on the cut of the horizon line along with a composition based on the simple harmonic proportions of 9:12:16. Botticelli here used Alberti's musical consonances. For Alberti, proportion was a visual music. In Botticelli's application of these proportions, we see the interval of the perfect fourth (9:12, or 3:4) and what Alberti called the *duple diatessaron,* or double fourth (9:16, or 3^2:4^2) (Alberti 1955).

Simple-number proportions are a foundation of Western music intervals, but we can find application of proportions such as phi in temporal constructions as well. Bartók's and Debussy's uses of phi are well documented (e.g., Lendvai 1971; Howat 1983). Stockhausen often used the Fibonacci series (a number series from which phi can be derived) in several of his early works (Griffiths 1981). Evidence of phi can even be found in the music of Mozart, which should not be surprising as he was a Mason, and the divine proportion was part of the Masonic "secret knowledge" (Evans 1993).

It should be stressed here that these proportions are not a panacea for the creation of good composition or visual consonance. They are, however, undoubtedly useful and often used. Good visual composition requires sensitivity to all elements in the picture plane. Anyone interested in composing visual music should study two-dimensional design in all its various aspects, much as any first-year art foundations course will provide.

From an understanding of proportion and composition in the picture plane, ideas can expand into variations, and larger structures can build on relations of design elements. Figure 4 is from a design lesson in a book by Arthur Wesley Dow written over a century ago. The figure illustrates variations of a single design, which he calls the *motif.* Dow said, "Little can be expressed until lines are arranged in a Space [which] I shall call PRINCIPLES OF COMPOSITION . . . all being dependent upon a great general principle, PROPORTION or GOOD SPACING" (Dow 1899).

Figure 4. From a design lesson in a book by Arthur Wesley Dow (1899).



Temporal Design

In temporal design, we move the viewer from visual dissonance to cadences of visually balanced, well-composed moments. Time passes musically through patterns of tension/release. Filmmaker Ken Burns's use of slow pan and zoom of a camera over old still photographs is a simple technique that allows exploration of this basic approach to time design. Using only camera motion, we can see this phrasing in films such as *The Civil War—A Film by Ken Burns* (Burns 1990). Now a common technique in historical documentaries, camera motion over still images can also tell stories. See for example French filmmaker Chris Marker's 1962 short film *La Jetée,* an early and effective application of this device along with montage and visual composition (Marker 1962).

I created simple illustrations of camera movement with still images, included here as Video Examples 1 and 2. [Editor's note: video figures for this article can be found on the DVD accompanying this issue.] A virtual camera moves over digitally processed photos of flowers. Key frames are seen in Figures 5 and 6. Each example starts in balance with a dynamic or simple symmetry. Once motion starts, a tension is established that seeks resolution in stillness and a return to a balanced composition.

Each example can be considered as a phrase. We can eventually combine phrases, seeking points of commonality and difference in the visual dimensions of motion (direction and speed), composition, and color. We see repetition. The two examples are of equal duration of 8 seconds. They zoom in at roughly the same rate, with a central focus of two flowers moving to a single flower, focused left of center. The examples feel similar, but there is variation. They contrast in their vertical motion, angles of rotation, and colors; there is repetition, contrast, and variation. We have the beginnings of a syntax from which to construct larger forms and compose visual music.

The construction of a pattern of tension/release moves a viewer through time, and so we seek to build tension. The desire to resolve the tension naturally follows. Time passes dynamically. Video Example 3 is an exercise illustrating a single shot with elements entering and exiting the picture plane in an overall rounded binary form (AABA). The example has four phrases that land on well-composed, visually "right" moments. Phrase 1 introduces Mondrian-like graphic materials with red and blue squares, shifting in color and moving into a composed cadence, seen as a small yellow square spinning to combine with two thick black lines. Phrase 2 follows with a repetition of blue and red squares spinning in from the opposite direction, cadencing by coming to rest on a balanced design. A new, contrasting phrase follows with thin lines rotating around an axis and a return of the small yellow square and thicker black lines. This overlaps into a final phrase that again repeats the spinning red and blue squares while echoing the rotating lines that finally articulate closure, coming to rest on another complete and balanced design. The example illustrates how visual time can be musical, using motifs and phrases with the "camera" locked in place and the whole work unfolding in an single "shot."

Developing a visual music piece in a single shot with little or no editing is a common practice in visual music composition; many of John Whitney's films do this. His multi-movement work *Arabesque* is built on the development of 360 points of light moving in the picture plane (Whitney 1975). The points move independently through various numeric processes, in and out of visual resonances where the points resolve into circles or symmetries in simple numeric proportions. Whitney called this *differential*

Figure 5. Key frames from Video Example 1. Dots show points of focus, and lines indicate the basic trajectory of camera motion. The camera moves up and rotates counterclockwise.

Figure 6. Key frames from Video Example 2, illustrating points of focus and basic camera trajectory. The camera is moving down and rotates clockwise.

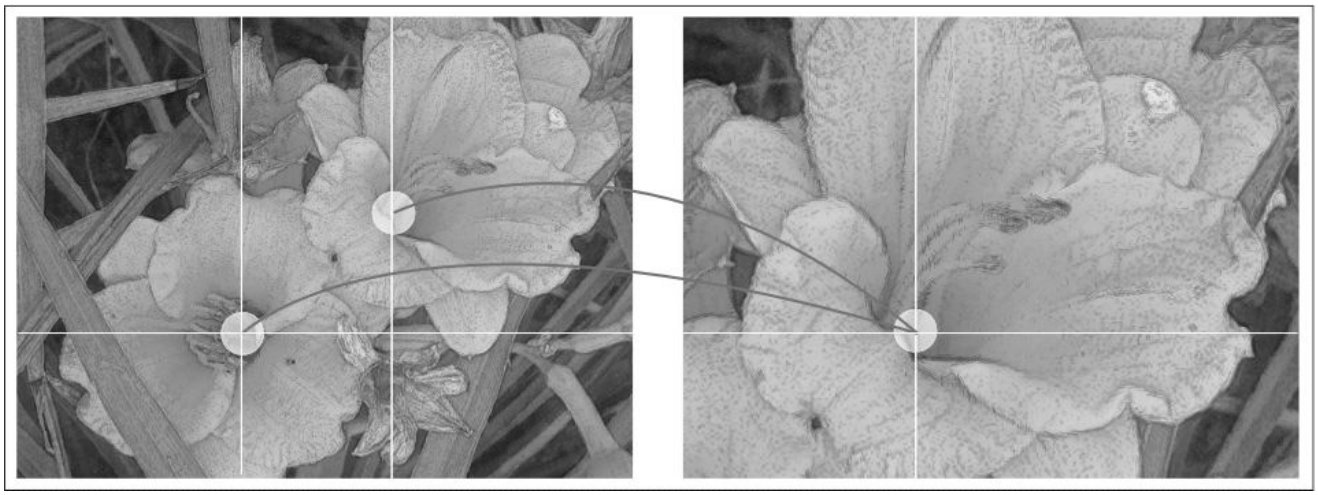


Figure 5

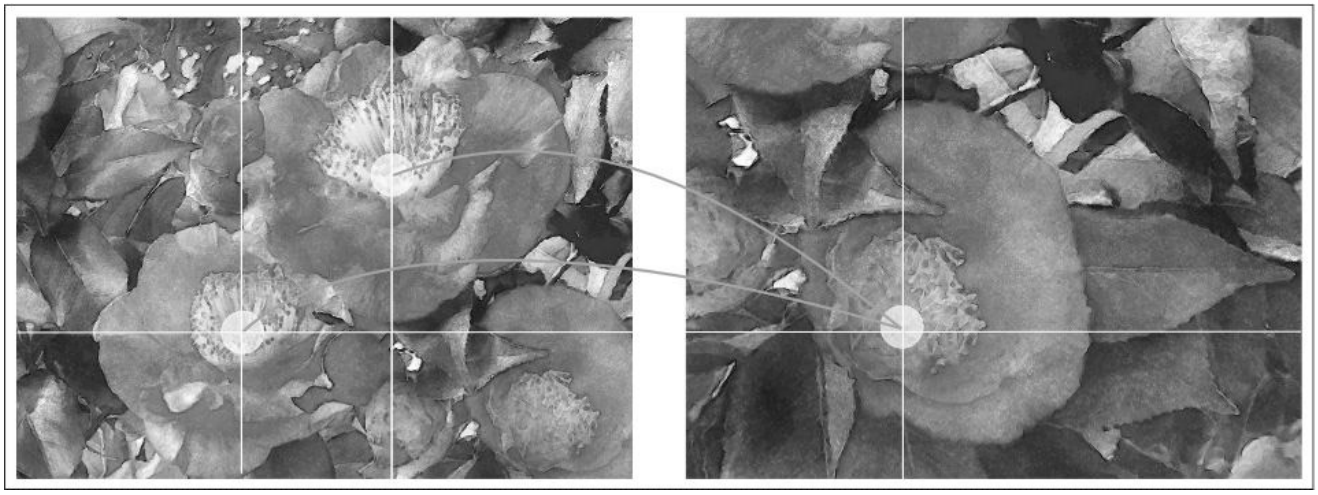


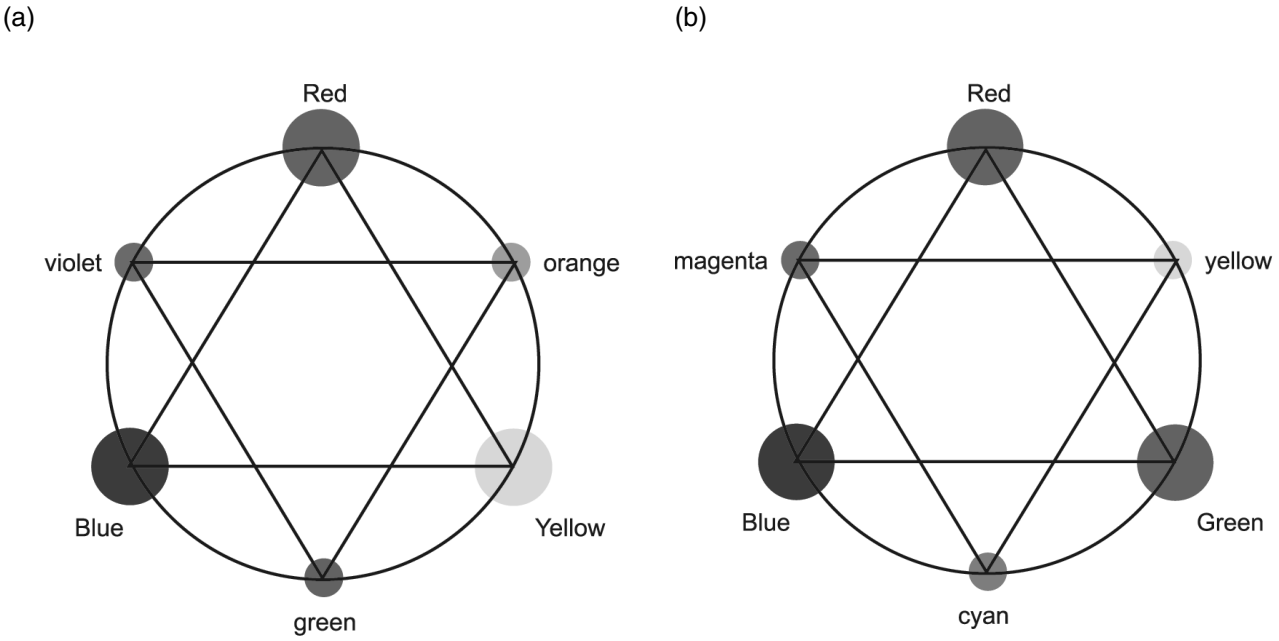
Figure 6

dynamics or *harmonic procession* (Whitney 1980). These moments of resonance provide points of visual consonance. Moving away from and back to the consonance creates tension and resolution.

Norman McLaren's *Pas De Deux* (1968) shows mostly single-shot action using time-lapse filming of two dancers. The piece presents another approach to Whitney's idea of differential dynamics. Definite points of resolution occur when all the time-lapsed movements of the dancers resolve into single figures, reminiscent of a musical canon at the cadence when all voices come together.

Oskar Fischinger's *Study No. 10* (1932), in black and white, is a wonderful example of this single-shot approach with movement through moments of visual consonance. In this piece, Fischinger creates abstract materials that directly follow and express the phrases and feeling of the music.

Figure 7. Two color circles: (a) Itten's ("painter's") RYB color circle; (b) additive ("computer") RGB color circle.



Considering Color

Color can also be used to create a tension that seeks resolution, although it is not an exact science. Artists have developed many theories of color over the past centuries. A common idea in these theories is harmony. In discussing harmony, theorist Luigina De Grandis said, "All theories have in common an assumption that the sensation of harmony, or concord of parts of a whole . . . results exclusively from the relationships and proportions of its chromatic components" (De Grandis 1984).

The problem lies in how those chromatic components should be organized, as there is no single universal categorization of color. Margaret Livingstone, Professor of Neurobiology at Harvard University, studies how the eye and brain process color. She notes that "much of what has been written about color in art is nonsense." She points out that the brain processes hue and luminance (also called value) separately. Some aspects of vision, such as depth perception and the separation of figure and ground, are colorblind, meaning that we get our visual cues from value alone. And the human perception of color, especially hue, can vary widely based on genetics and experience (Livingstone 2002).

At the very least, color in art is confusing. Take for example the two color circles seen in Figure 7. Both are used to determine complementary colors. A basic tenet of color theory is that complementary colors (those directly opposite on the circle) create color harmony. The circle on the left shows the standard color primaries as understood by artists over the centuries: red, yellow and blue ("RYB"). Johannes Itten formalized this color circle in developing the color theories he taught at the Bauhaus. This approach to color is still taught in many art schools today (Itten 1973).

The circle on the right is the one used in computer graphics displays as they project light. It uses the additive primaries of red, green, and blue (RGB). When added together, these primaries will make white light. The complementary colors in this circle are cyan, magenta, and yellow (CMY), also called the subtractive primaries, which theoretically produce black (the absence of light) when combined. As one mixes them, more and more light is absorbed or subtracted. Commercial and digital printing pro-

cesses use the subtractive primaries with black ink added (denoted K), as real-world pigments cannot create a true black, but rather a somewhat muddy brown.

In Itten's system, red and green are complementary and concordant in combination. In additive color, red and green are both primary and so are discordant. The two systems will create contradictory harmonies. What is the measure we use to decide which is right? Color continues to be problematic in seeking visual harmony. Furthermore, the perception of color is inexact, culturally influenced, and personal. I might use a variation of green that reminds me of the zest of a key lime. The same color might remind someone else of rotting meat, triggering a very different response. One must be careful with color.

There are some generalizations we can draw, however. First, color space is a three-dimensional construct. Familiar color spaces, all hue-based, include RGB, RYB and CMY. Some color spaces include other visual dimensions. One artistically useful space maps a hue-only approach to one where color is defined as hue, saturation, and value (HSV) (Smith 1978). There is no perceptual hierarchy of hue, but there are of value (bright to dark) and saturation (full chroma to grayscale). Movement through these hierarchies provides areas of exploration.

Chemist and colorist Maurice Chevreul in the 19th century suggested that the eyes seek the balance of grayscale, as evidenced through the visual phenomena of simultaneous and successive contrast (Chevreul 1854). A sensation of harmony or "chromatic equilibrium" should occur with complementary colors, as they would mix to create grayscale, a sort of visual tonic. Gray is balanced, whereas emphasis on one hue or another creates tension. Chevreul also points out the difficulty in practice using pigments or dyes as "we know of no substance which represents a primary color."

Newton showed us that we see white when all colors of the spectrum of visible light are combined; lower the value or brightness and we see gray. Digital technology simulates this with the additive RGB color space. We can measure the equilibrium of color combinations without concern for color circles or complementarity. Computer monitors work in an RGB space designed to create grayscale when the RGB component values are equal.

With digital images, it is possible to calculate sums of the amounts of red, green, and blue in an image. The closer the RGB sums are to being equal, the more "balanced" the image is with respect to color, as equal means gray. Color can move us through time by resolving an emphasis of a single hue to a coloring that balances or equalizes RGB summation values (Evans 1990). Tension/release with color can be established in this way, but it requires a detailed measurement of digital color that may not be practical. What is more practical is a general application of the perceptual hierarchies of brightness resolving to dark, and high saturation resolving to low saturation or gray.

A familiar example is mentioned in Bordwell and Thompson's analysis of the 1939 film version of *The Wizard of Oz* (Baum 1939; Bordwell and Thompson 2001). The larger structure of the film is an ABA form based on the settings Kansas-Oz-Kansas. Dorothy is home in Kansas in the beginning, filmed in grayscale. Through her adventures in Oz, where conflict builds, the film is in full color. Conflict resolves back to grayscale when she returns home. Shifting from color to grayscale is now a common device on television, used to articulate or simply end the action. A transition to grayscale or simple desaturation often takes us to a commercial.

Working with color in HSV space, it is useful to think of brightness as visual amplitude and saturation as spectral intensity. Bright, full, saturated colors can be tiring to a viewer when used in excess, just as loud, full-spectrum sounds can become tiring to a listener.

Oskar Fischinger's *Motion Painting 1* (1947) is a good example of a general approach to using saturation and brightness structurally. In the first section especially, the movement of color is important as it builds large phrases. In the film, we see a single painting develop layer upon layer over eleven minutes. We feel tensions resolve in short phrases as we see shapes developing as motives, begun and then brought to completion. Spirals coil and unfold; mosaics of diamonds appear and cover what is underneath. Larger structures are perceived as repetition and contrast of hue choices. Tension resolves as

high saturation, and bright colors give way to darker colors and lower saturations.

Montage and the Musical Phrase

Cinematographer Joseph Mascelli lists the "Five C's" of cinematography as camera angles, close-ups, cutting, continuity, and composition (Mascelli 1968). All but the cutting require some degree of preproduction planning. Through the practice of storyboarding, we see this planning in action (Begleiter 2001). With storyboards, filmmakers illustrate and articulate many of the decisions to be made when shooting a scene. The principles of composition used in support of the story are key to this process.

A director plans a scene as a series of moments, composed as a series of drawings. The drawings guide the eventual motion, framing, and design of the picture plane over time. Continuity, camera angle, and composition in a scene delineate phrases that cadence with moments of visual rightness. Time passes musically.

With visual time passing under compositional control, a larger temporal structure can be built. Visual phrases are combined in film in a process called *montage,* the cutting or editing of camera shots into a sequence. Eisenstein's theory of montage can be valuable to our developing grammar of visual music. He believed that it was montage that placed film at the "pinnacle of the arts." The unit of montage, what Eisenstein called the *cell,* combines with other units and creates a gestalt, expressing more than just the sequence of images. The power for Eisenstein lay in the heights of expression possible. "To determine the nature of montage is to solve the specific problem of cinema" (1949, p. 48). More simply, he states "montage is conflict" (1949, p. 38). Montage is a multi-tiered construct of tension and release.

We can think of each cell as a musical phrase. From good visual composition, we can build phrases and expand into a larger musical form. Temporal design at the level of the shot moves us forward to cadence points. Through montage, we construct relationships from shot to shot, building a larger form and expressing our ideas, formally or otherwise. Eisenstein lists five basic types of montage: metric, rhythmic, tonal, overtonal, and intellectual. In discussing the five types, we will focus here on the first three.

Metric Montage

In defining *metric montage,* Eisenstein writes that "pieces are joined together according to their lengths in a formula scheme corresponding to a measure of music" (1949, p. 72). We are only concerned with the duration of each shot, cutting each shot to a beat without regard for the content of the shot. We establish meter by making each cell the same length—the same number of frames. Each cell is like a phrase of music. With shots of short duration, cells are like bars of music—sometimes even beats within a bar.

A time signature can be loosely established through accent of some visual dimension. Occasionally, the filmmaker will actually put an all-black or all-white frame at the end of each cell, making the visual meter obvious. Sometimes cells are edited to match musical phrasing—a common device in music videos. Because the frame rate of film or video is constant, we can accelerate or slow down the meter by changing the length of the cells as they sequence.

Metric montage follows the "law of simple numbers," and it is applied for clarity when based on a musical time passage (3/4, 2/4, 4/4, etc.), where it can "bring into unison the 'pulsing' of the film with the 'pulsing' of the audience" (Eisenstein 1949, p. 73). Using an irregular measure, "16/17, 22/57, etc. . . . brings a degeneration of the method" (Eisenstein 1949, p. 72). From the extremes of simple to irregular is a middle ground where lengths are varied, alternating between two different contexts, juxtaposed temporally back and forth. What he refers to here are irregular-to-metered durations, such as *rubato* to *a tempo,* which is itself a basic process of tension/release.

Norman McLaren and Evelyn Lambart's film *Begone Dull Care* is an enjoyable example of metric montage. Strips of clear celluloid were painted, scratched, textured, and processed in a variety of ways. They cut the strips into cell lengths to match

with music by the Oscar Peterson Trio (Lambart and McLaren 1949).

Rhythmic Montage

In *rhythmic montage,* the length of the cell is determined by what we see. "Abstract determination of the piece-lengths gives way to a flexible relationship of the actual lengths" (Eisenstein 1949, p. 74). Durations of phrases support the content of the cells, and most specifically "movement within the frame impels movement from frame to frame . . . , objects in motion, or of the spectator's eye directed along the lines of some immobile object" (Eisenstein 1949, p. 75). This is apparent in basic continuity editing in which a long shot, usually containing lots of image detail, lasts longer than a close-up. The logic is that it takes longer to see all the details in a long shot. The passage of time becomes more psychological than strictly measured; experiential time is a consideration. Events and images within the picture frame determine duration.

Video Example 4 takes the simple camera motion with the flowers seen earlier and extends it into four phrases. The camera motion zooms in with cells 1, 2, and 8, slowly panning and rotating from and to a balanced composition. Consider these cells as phrases, cadencing with a similar design structure and a focus left-of-center.

Cells 3–7 contrast this and together make up a contrasting phrase. They are much shorter in duration, with centered images and the camera zooming in slightly. We see a basic rounded binary form: A (cell 1), A (cell 2), B (cells 3–7), A (cell 8). The duration of each cell reflects the amount of camera movement within the cell.

Metric montage is illustrated with cells 1 and 2 (each 8 sec long) and cells 3 through 7 (each 2.4 sec long). As the camera moves over more space in cell 8, it is consequently slightly longer in duration than cells 1 and 2, yet similar in camera motion and final composition. This keeps the perception of speed more or less constant in these cells. Rhythmic montage is developed so the durational feel of cells 1, 2, and 8 is similar, and the total duration of the B section (cells 3–7), feels in balance.

Camera motion and zoom come to rest at the end of each cell. Cells ease-in and ease-out at endpoints in the larger phrases. Cells enter with cross-fades, the cross-fade itself a visual equivalent to musical elision.

Tonal Montage

Eisenstein calls his third montage type *tonal montage.* Continuing his borrowing of musical function, he describes tonal montage as "melodic-emotive." We must be careful here to avoid misunderstanding his use of the terms tonal and dominant, as they do not relate very well to the musical definitions of the terms. The tone of a work is its expressive tone. The dominant is the primary emotional feel of the shots. He here begins to bring the thematic content of the film and the meaning of the images into the discussion. He attempts to maintain an objective measure by focusing on the graphic or plastic aspects of the moving image: color and shape. He calls it "graphic tonality"—the design elements of color, line, and shape, and what they can represent.

The time design begins to consider the subject matter and its signification. As Eisenstein explains his ideas, he presages developments in semiotics and structuralist thought. This goes beyond our discussion, as we are focusing on the plastic elements alone in composing visual music. The aim is to develop a simple design syntax built from a basic design idea. We are sidestepping issues of content that go beyond formal concerns. Bordwell and Thompson (2001) write, "Since a camera movement consumes time on screen, it can create an arc of expectation and fulfillment." This restates the premise: as camera movement alone fulfills an expectation, the viewer is moved through time to a resolution of tension.

Thus, while in tonal montage, color and shape do begin a move into connotation, we can reduce an image-in-motion to its design elements. From this distilled level of montage, we articulate phrases with movement in color (saturate to neutral, bright to dark) and movement from and to a harmonious distribution of materials on the picture plane.

In the flower example, the graphic tonality is seen in the repetition of points of visual focus and in rep-

etition or contrast of direction, rotation, and speed of camera zoom and pan. Centered images of the B section, with little motion, are shorter, and they set up expectations for a return to rotation and a longer phrase. This is another iteration of the basic rounded binary structure, this time at the level of tonal montage.

Many of Stan Brakhage's works, usually with no sound score, are good examples of tonal montage. *Mothlight* (1963), for example, is an interesting work that uses dead moths and other organic debris he collected from light fixtures. He used no camera but instead collaged twigs, blades of grass, dust, and moth parts onto sticky tape, and then he printed it to celluloid for viewing through a film projector. Cells were then cut to create a fast-paced and haunting montage.

Chris Marker's *La Jetée* and Ken Burns's *A Civil War* are good studies of tonal montage with minimal sources. For example, battle scenes in *A Civil War* may have shots of short duration, in contrast to longer duration pans over battlefields in the aftermath. Both films show how much can be done in developing large temporal structures with simple composition and effective editing.

Overtonal and Intellectual Montage

The last two types of montage are not really montage in the strict sense. Eisenstein calls his fourth type of montage *overtonal montage.* He begins to develop the relationship between the previous three types and the cinematic work as a whole. He sees the elements of montage metaphorically as harmonic partials of a musical timbre. Conflict develops and resolves as these elements wrestle and then agree. Eisenstein himself admits that overtonal montage is "not an altogether exact term" (1942, p. 86), but he is concerned with the larger structures of the film and its impact on the audience. From this, he takes us to his fifth level, *intellectual montage,* which will eventually build a "completely new form of cinematography" (1949, p. 83). In later writing, he calls this *intellectual cinema* and is really considering the movement of thought rather than moving light.

Returning to our simple, syntactical point of view, we can think about montage working at the three basic levels when building phrases of moving images. First is metric montage, which deals with the simple duration of shots. This is expanded to rhythmic montage, in which the duration of cells supports the plastic elements in motion within the shot. Lastly, the graphic elements such as color, line, and shape and the dominant expression of the shot build phrases in tonal montage. Each shot or phrase articulates cadences through a construct of tension/release. Phrases are combined and grow to a larger form through repetition, contrast, and variation.

Visual Counterpoint

An artwork can be defined as a complex relationship of elements; it is multi-dimensional. In developing the elements particular to an image in motion, it is possible to build the unfolding of these elements in a contrapuntal structure (Evans 1992). Various graphic elements can work independently in support of a larger structure. Eisenstein calls this *visual counterpoint.*

Montage is conflict. Cinematic conflicts unfold in the picture plane. These conflicts work within the cell or shot and in "colliding" cells or montage. Eisenstein lists the conflicts (which deal mainly with the compositional elements) as follows (the italics are his):

> *conflict of graphic directions (lines either static or dynamic); conflict of scales; conflict of volumes; conflict of masses (volumes filled with various intensities of light);* and *conflict of depth.* With a bit more "intensification," there can also be conflicts of *close shots and long shots; pieces of graphically varied directions (pieces resolved in volume with pieces resolved in area); pieces of darkness and pieces of lightness; and conflicts between an object and its dimension (and conflicts between an event and its duration).*

[In this final conflict, he refers to optical distortion, stop-motion, or slow-motion.]

What we have here is a comprehensive list of the design materials. With an understanding of two-dimensional design and tension/release, we can develop all of these materials independently within a shot and integrate them through many shots into a "unified system" of montage, the basis of a "tentative film syntax" (Eisenstein 1949, p. 55)

Again, consider the first two video examples. They repeat the horizontal direction of focus from left to right. They contrast in vertical direction (one with the camera going up, the other with the camera going down). They repeat the relative scale of zoom, and they repeat the duration. They contrast in rotation angle (clockwise vs. counter-clockwise). Different elements with different temporal structures—some in agreement and some opposed—create a counterpoint.

Vertical Montage (And the Sound Image)

Eisenstein expanded his counterpoint to combine the "visual image with the sound image, in the process of creating a single, unifying sound-picture image" (1942, p. 73). He considered this a new kind of montage, naming it *vertical montage.* Again, he uses the term loosely, referring to the interactions of various elements as they are synchronized in time.

He saw vertical montage as an orchestral score with several staves of activity unfolding on the horizontal yet with all elements interrelating on the vertical, all moving forward with "intricate harmonic musical movement" (1942, p. 74). The montage structure of the visuals is just one more line on the musical score. The varied techniques of music composition now can be applied to build a sound score that fits the overall contrapuntal texture of the work.

For Eisenstein, it was important that this be a true, polyphonic texture. He was not interested in direct synchronization of sound and image, but rather the "inner synchronization between the tangible picture and the differently perceived sounds . . . the 'hidden' inner synchronization where the plastic and tonal elements will find complete fusion . . . [where] we find a natural language common to both—movement" (1942, p. 81). This inner connection is among the images, the sounds, and the meaning of the work. For the audience it is the gestalt of the work as a whole, with all elements contributing to the overall meaning. In moving beyond the visual music foundations, much can be learned from the techniques of opera, musical theater, film scoring, and sound design.

When we use digital tools to create time-based work, layers of material are usually visualized on a timeline. Events are placed graphically like notes on staves in a musical score. Events unfold horizontally, but vertical relationships are also important—multiple lines moving independently yet in harmony. Eisenstein's image of the orchestral score is realized in the user interface of current graphics software designed for time-based work. (Good examples include Adobe After Effects, Digidesign Pro Tools, Macromedia Flash, and even Apple's iMovie.) Horizontal layers represent video, audio, still images, camera motion, interpolation paths, and more. These lines can interact and mix both horizontally and vertically.

Toward a Visual Music Theory

Eisenstein believed that every artwork is conflict according to its social mission, according to its nature, and/or according to its methodology. The methodology of cinema is montage. At the graphical level, montage can move us in time through the visual "rightness" one learns in two-dimensional design. With experience and sensitivity to design foundations, the artist, filmmaker, or visual music composer can explore the higher levels of artistic conflict, its nature and social mission. Even with studies in composition and movement using static images, a design sense can be exercised. Sensitivity to musical time can be developed and practiced with visual materials. A sound score can be added and musical sensibilities applied to a unified sound-image expression.

Literacy these days means speaking the language of new media, which is in many ways the language of cinema. As it is a temporal medium, we can find at the foundations of cinema the concept of music.

The tenets of musical expression give us a method for moving the viewer/listener through time. With visual composition, we create consonance and dissonance. The ability to resolve dissonance to consonance provides a syntactical basis for temporal design and the composition of visual music. Our compositions can be a single shot moving through phrases. We build phrases through repetition, contrast, and variation of the many visual dimensions including direction, speed, shape, size, and color. Eisenstein's montage provides a conceptual framework for putting the phrases together into larger forms. New media may require a new literacy, yet traditional ideas still have much to offer.

Acknowledgments

This work was done with the support of a Research Advisory Council Grant, University of Alabama.

References

Arnheim, R. 1966. "A Review of Proportion." In G. Kepes, ed. *Module, Proportion, Symmetry, Rhythm.* New York: George Braziller, pp 218–230.

Alberti, L. B. 1955. *Ten Books on Architecture.* London: Tiranti.

Baum, L. F., director. 1939. *The Wizard of Oz.* 2004. (1939) Motion picture. Hollywood, California: Warner Studios, ASIN: B00000JS62.

Bell, C. 1914. *Art.* London: Chatto and Windus.

Begleiter, M. 2001. *From Word to Image: Storyboarding and the Filmmaking Process.* Studio City, California: Michael Wiese Productions.

Bordwell, D., and K. Thompson. 2001. *Film Art: An Introduction,* 6th ed. New York: McGraw-Hill.

Bouleau, C. 1963. *The Painter's Secret Geometry.* New York: Harcourt, Brace, and World.

Bowers, J. 1999. *Introduction to Two-Dimensional Design: Understanding Form and Function.* New York: Wiley.

Brakhage, S. 1963. *Mothlight.* Recorded on *By Brakhage—Anthology—The Criterion Edition.* DVD. New York: Criterion DVD CC1590D (2003).

Brougher, K., et al. 2005. *Visual Music: Synaesthesia in Art and Music Since 1900.* New York: Thames and Hudson.

Burns, K. 1990. *A Civil War—A Film by Ken Burns.* Alexandria, Virginia: PBS Home Video.

Cheuvreul, M. E. 1854. *The Principles of Harmony and Contrast of Colors and their Applications to the Artist.* West Chester, Pennsylvania: Schiffer Publishing.

Cruse, V. 2004. *Sons & Lumières, Une histoire du son dans l'art du 20e* siècle. Paris: Centre Pompidou.

De Grandis, L. 1984. *Theory and Use of Color.* New York: Harry N. Abrams.

Dow, A. W. 1899. *Composition.* Garden City, New York: Doubleday.

Eisenstein, S. 1942. *Film Sense.* New York: Harcourt.

Eisenstein, S. 1949. *Film Form.* New York: Harcourt.

Evans, B. 1990. "Temporal Coherence with Digital Color." *Digital Image—Digital Cinema, SIGGRAPH '90, Art Show Catalog,* published in *Leonardo* supplemental issue. Oxford: Pergamon Press, pp. 43–49.

Evans, B. 1992. "Elemental Counterpoint with Digital Imagery." *Leonardo Music Journal* 25(4):13–18.

Evans, B. 1993. "Number as Form and Content: A Composers Path of Inquiry." In M. Emmer, ed. *The Visual Mind.* Cambridge, Massachusetts: MIT Press.

Fischinger, O. 1947. *Motion Painting No. 1.* Recorded on *Oskar Fischinger: Pioneer of Abstract and Advertising Animation.* LaserDisc. Carson, California: LaserDisc Corporation (1986) Pioneer Special Interests, PSI-88-012.

Fischinger, O. 1932. *Study No. 10.* Recorded on *Oskar Fischinger: Pioneer of Abstract and Advertising Animation.* LaserDisc. Carson, California: LaserDisc Corporation (1986) Pioneer Special Interests, PSI-88-012.

Griffiths, Paul. 1981. *Modern Music, the Avante Garde Since 1945.* New York: Braziller. p. 85.

Hambidge, J. 1919. *The Elements of Dynamic Symmetry.* New York: Dover.

Hambidge, J. 1932. *Practical Applications of Dynamic Symmetry.* New Haven, Connecticut: Yale University Press.

Howat, R. 1983. *Debussy in Proportion: A Musical Analysis.* Cambridge: Cambridge University Press.

Huntley, H. E. 1970. *The Divine Proportion.* New York: Dover.

Itten, J. 1973. *The Art of Color.* New York: Van Nostrand Reinhold.

Kandinsky, W. 1914. *Concerning the Spiritual in Art.* New York: Dover.

Lambart, E., and N. McLaren. 1949. *Begone Dull Care.* Recorded on *Norman McLaren: The Collector's Edition.* DVD. Harrington Park, NJ: Milestone Film and Video (2003) ASIN: B0000AOV4F.

Lendvai, E. 1971. *Béla Bartók: An Analysis of His Music.* London: Kahn and Averill.

Livingstone, M. 2002. *Vision and Art: The Biology of Seeing.* New York: Harry Abrams.
Marker, C. 1962. *La Jetée.* Recorded on *Short 2-Dreams.* DVD. Hollywood: Warner Studios (1997) ASIN: B000031VPS.
Mascelli, J. 1968. *The Five C's of Cinematography.* Hollywood, California: Cine/Grafic Publications.
McLaren, N. 1968. *Pas De Deux.* Recorded on *Norman McLaren: The Collector's Edition.* DVD. Harrington Park, New Jersey: Milestone Film & Video (2003) ASIN: B0000AOV4F.
Moritz, W. 2004. *Optical Poetry: The Life and Work of Oskar Fischinger.* Eastleigh, UK: John Libbey Publishing.
Russet R., and C. Starr. 1976. *Experimental Animation: an Illustrated Anthology.* New York: Van Nostrand Reinhold.
Smith, A. R. 1978. "Color Gamut Transform Pairs." *Computer Graphics* 12(3):12–18.
Stravinsky, I. 1956. *The Poetics of Music in the Form of Six Lessons.* New York: Vintage Books.
Whitney, J. 1975. *Arabesque.* Recorded on *The World of John Whitney.* LaserDisc. Carson, California: LaserDisc Corporation (1984) Pioneer Special Interests, PSI-88-003.
Whitney, J. 1980. *Digital Harmony: On the Complementarity of Music and Visual Art.* New York: McGraw-Hill.