Synchronized Multi-Channel Video and Music

Don Ritter

aesthetic-machinery.com

Abstract. This paper discusses a strategy for synchronizing music with a multi-screen, interactive video configuration. The paper describes a hypothetical configuration of four interactive video systems and four video projection screens that are arranged around an audience. This configuration enables sound-image relationships to be enacted within single screens and, also, in the movement of imagery between screens. The configuration associates different categories of music with different types of video imagery and screen movements, enabling audiences to be encircled with screens of synchronized sound and image. Video imagery can move between the four screens in different directions, at different speeds and in different orders, all according to any sound or music input. The unique aspects of composing a multi-channel work are discussed in detail by explaining its compositional elements and by outlining a notation system for creating an interactive, multi-channel video work that is controlled by live music.

The ideas presented in this text were originally written in Montréal during 1992 as part of a research project funded by the Canada Council. A small amount of editing was undertaken on the grammar in 2009, though the core ideas and the diagrams of the original document are the same. Consideration should be given to certain statements within this paper that pertain to interactive media and multi-screen projections, which were not as common in 1992 as they are today.

1. Introduction

When viewing a film or video, audience members usually remain stationary with their eyes directed towards a screen. Similarly, an audience at a live performance will usually look towards the stage of the event. Because the content of film, video, theatre, and dance are usually presented at a singular and specific location in space, an audience usually has no reason for looking in other directions. Within art galleries and museums, however, viewers are more kinesthetically involved as they walk around and look in different directions at paintings, sculptures, photographs, or whatever. In these spaces, viewers select objects for observation by walking into rooms and selectively gazing at the works. Viewers of film, video, dance, and theatre typically have little control over their visual experience, while viewers at museums and galleries can select and create personal experiences according to where they direct their attention. Non-interactive time based media typically provide little experience control for a viewer, while media containing static spatial elements, such as sculpture, offer viewers selective control according to the point-of-view. The difference between these two forms of media represent whether the experience of an artistic event is determined solely by its creator, or if a viewer can also modify it. McLuhan designates these two forms of media as being “hot” or “cold.”[1]

Ritter, Don (1992, 2009) “Synchronized Multi-Channel Video and Music.” Berlin: aesthetic-machinery.com

In addition to providing different amounts of control to an audience, media often differ in which human senses are being stimulated. When viewing a film, one typically sits in a chair for approximately ninety minutes and experiences a collection of sounds and images. When viewing sculptures in a museum, however, a person usually walks into a different room every few minutes and gazes at works through a variety of neck, head, and eye movements. The viewing of works in a museum provides a multi-sensory experience comprised of image, kinesthetic movement and perhaps sound.

In 1927, Abel Gance used three film projectors and three projection screens to present a film on the life of Napoléan. [2] Sixty-five years later, however, the use of multi-screen film projection is still uncommon as most films are presented using a single screen of imagery. The use of multiple screens is more common in the area of contemporary video art, such as works by artists Nam June Paik and Gary Hill that use multiple video projections or video monitors. In most cases, however, contemporary multi-channel video installations use multiple single channel systems having little or no synchronization between each channel.

In the field of audio recording, the synchronization of multiple audio channels has become so popular that monophonic recordings are rarely distributed because of the unique qualities provided through stereo sound. A pair of precisely synchronized video channels that are displayed through a pair of monitors or projectors may also provide unique visual qualities. An obvious reason for the existence of multi-channel audio is the accessibility of stereo and multi-channel recording equipment. I am unaware, however, of any video recorder capable of recording multiple video channels onto a medium that would allow synchronized playback of recorded channels. Although synchronized playback of multiple video channels is possible through a video controller and numerous playback units, these configurations provide little versatility in delegation of video images to different screens. In the area of audio, the delegation of different audio channels through various speakers is easily possible using an audio mixer.

The research described in this paper was undertaken to explore synchronized multi-channel video for use within video installations and live performances. Specific objectives included the development of a video installation that could provide viewers with a more physically involving experience than is provided with single channel video, and also to determine if synchronized control of multiple video channels could provide unique visual or sonic qualities that are not available within unsynchronized, multi-channel installations.

2. Synchronized Multi-Channel Video Installation

My involvement with multi-channel video installations evolved from various performances using interactive video that I began presenting in 1988. During those events, I presented single channels of interactive video controlled by music through single or multiple video projections. None of these performances used synchronization of video between multiple screens.

Exploration of a synchronized multi-channel installation was undertaken using the hypothetical configuration shown in Figure 1. Comprised of four projection screens arranged around an audience, the configuration enables an audience to be completely surrounded by video imagery. Because video imagery can be presented on one, two, three, or four screens at any moment in time, audience members must turn their heads to view a particular screen. This choice provides a viewer with some control of the visual experience, similar to what is available when viewing paintings on the different walls of an art museum. During a performance using this configuration, an audience would be seated in the center area on chairs that face different directions. As an installation, no chairs would be provided and viewers would walk in and out of the installation through the open areas between the four projection screens.

The multi-channel synchronization capability within this configuration is available through a control system that permits each video screen to be activated or deactivated relative to each other. This synchronization permits the illusion of video imagery to jump from one screen to another, created by rapidly switching each screen on or off in succession.

Audio is provided within this configuration through speakers located behind each of the four projection screens. This arrangement allows spatial synchronization of sound and video imagery within a three dimensional space. For example, a video image can be presented on *screen A* at the precise moment that sound is played through the speaker located behind this screen, thereby providing a spatial association of sound and image.

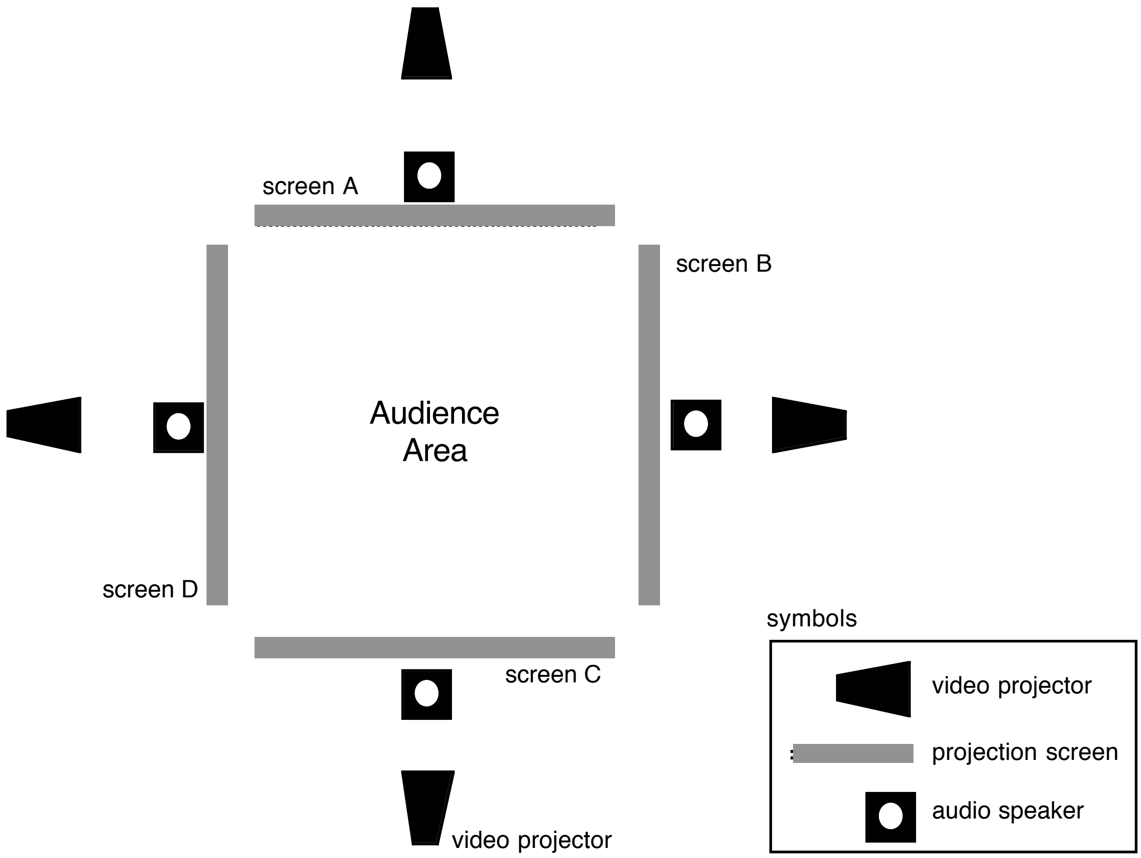

Figure 1. Top View of Synchronized Multi-Channel Video Installation

Potential synchronization between sound and image within this installation extends beyond presentation on a particular screen. This arrangement permits the association of different musical structures with specific images, cinematic effects, and image movements between the different screens. For example, the installation could be adjusted to present video imagery in a clockwise direction around the installation in response to high-pitched music, and in a counterclockwise direction in response to lower pitched music. Because of the synchronization available, the speed of the clockwise or counterclockwise movement could be controlled by the tempo or any other characteristic of the music. Thus, an audience could be surrounded with a circle of video imagery flying around them at different and directions under the control of live music.

The remainder of this paper discusses this multi-channel video installation in more detail by explaining the configuration of equipment and methods for synchronizing multi-channel video with music. The unique aspects of composing a multi-channel work are also discussed by explaining its compositional elements and by describing a notation system for creating a synchronized multi-channel video that is controlled by live music.

2.1. Synchronized Video and Music

The major component used within the multi-channel video installation is the author's interactive video system, *Orpheus*. Under development since 1987, this system provides real time control of video imagery though live music. This system allows precise synchronization of video imagery with improvised or composed music, and has been used in the creation of videotapes, video installations, and video performances. Collaboration performances with musicians have been presented with video controlled by voice, trombone, trumpet, saxophone, electric guitar, electric bass, acoustic bass, percussion, and keyboards. [3]

During operation, *Orpheus* analyzes live music and provides a musical categorization based on pitch, dynamics, note duration, rest length, tempo, rhythm, intervals, note density, and measure. Each categorization of music can evoke an associated video action consisting of specific video images and a cinematic effect. For example, the system could be programmed to present a video clip of a face laughing in response to a particular musical category, while another category would initiate a video clip of the same face crying. As the music alternates between these two categories, the face would alternate between laughing and crying up to a maximum speed of 30 times per second. The musical categorization can be viewed as an aesthetic judgement created by the system to allow association of music with corresponding imagery. If dissonance is desired, associations can also be made between music and imagery having opposing aesthetic qualities. Different analytical methods of categorizing music are available to allow selection of appropriate strategies for different music styles and musical instruments.

The major criterion in developing software for this system was complete flexibility in selection of video imagery, music type, and selection of correspondence between music and imagery. An overview of the system is provided in Figure 2.

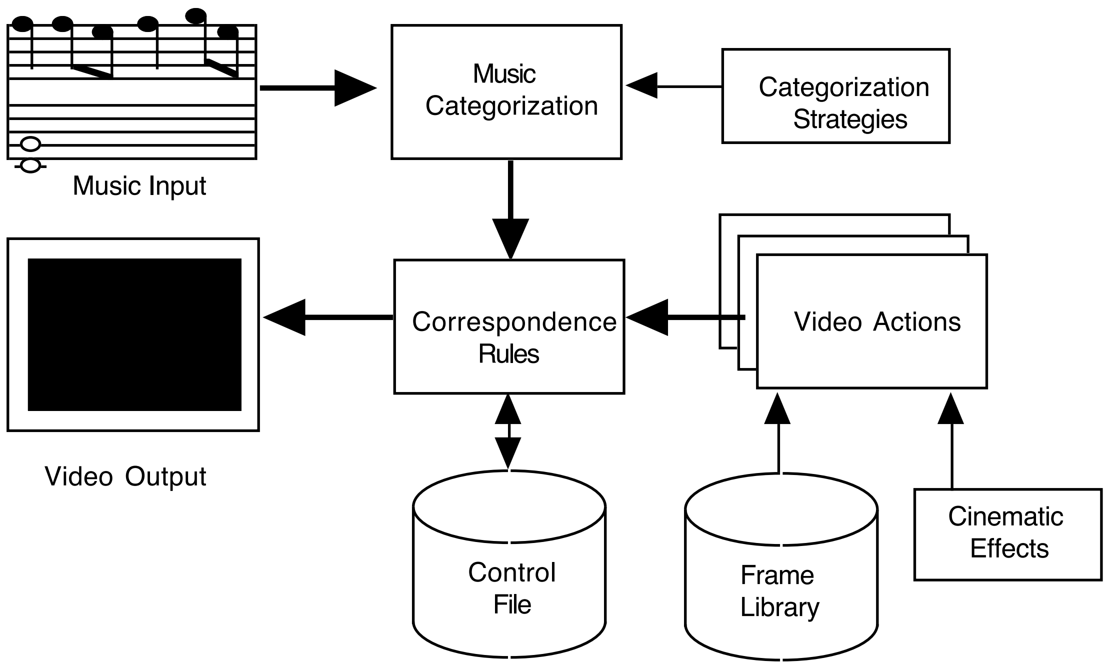

Figure 2. Overview of Orpheus Interactive Video System

This system interprets music in the form of MIDI data and it can be used with any instrument providing that form of output. For synchronization with acoustic instruments, sound can be translated into MIDI data using a microphone and a pitch follower. The video imagery presented and manipulated by the system is stored as digital frames within a computer's memory, or it can originate as video signals from tape, camera, or videodisk. A detailed description of *Orpheus* is provided in the author's paper, A *Flexible Approach for Synchronizing Video With Live Music (1992).* [4]

2.2 Multi-Channel Configuration

Four interactive video systems are used to create the multi-channel configuration described in this paper. Each of these four systems contains *Orpheus* software, a collection of video frames, and correspondence rules designating each system's response to MIDI input. A fifth computer controls the activity of each of these four systems independently and according to an analysis of incoming music. A diagram of the multi-channel configuration is provided in Figure 3.

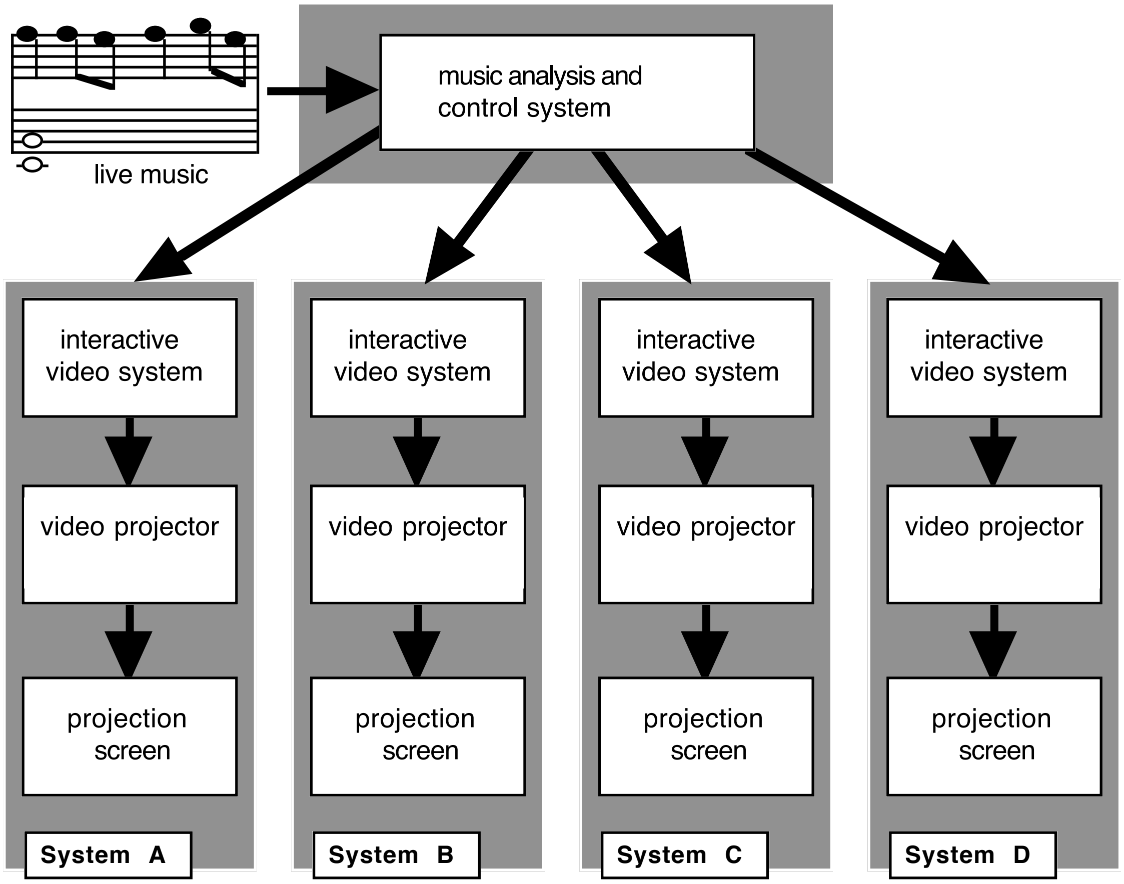

Figure 3. Multi-Channel Interactive Video Configuration

For a multi-channel configuration, each of the four interactive video systems operates in a similar manner as used with a single channel configuration. When using a single system, music enters directly into a system; with a multi-channel arrangement, music first enters a control system that independently analyses the music before selecting the activity for each of the four interactive video systems. Various visual effects become possible by using a single system to analyze music and control the four interactive video systems, such as video imagery jumping from screen to screen.

3. Compositional Elements

A detailed description of compositional elements within this installation is provided to demonstrate the versatility of a four screen configuration. The configuration has been designed to accommodate the presentation of different compositions, like a specific musical instrument being used to play different musical works.

Although the obvious advantage of multi-channel video over single channel video is an ability to display multiple screens of video imagery, a multi-channel interactive system provides other compositional elements. Because fast and intricate control of video imagery is available within this system, the illusion of movement can be created by having adjacent screens turn on and off in sequence.

Within the four screen configuration, an audience positioned in the middle of this arrangement would view a circle of video imagery flying around the screens. Because this motion can be presented at different rates and in different directions, screen to screen movement becomes a compositional element that is not available to single channel video or film.

In addition to screen movement, compositional elements within the installation include the actual video imagery, cinematic effects for presenting the imagery, the number of active screens, the spatial locations of active screens, changes over time, and the sound-image relationships associated with these elements.

The remainder of this section provides a detailed description of the various compositional elements available in the multi-screen installation. These elements are being discussed in detail because they represent the unique perceptual experiences made available through multiple screens of video, synchronized control, and video controlled by music.

3.1 Video Imagery

The compositional element of video imagery available in the installation is similar to film or video, referring to frames, shots, and sequences of objective or nonobjective imagery. Within this installation, imagery must be stored as digital frames within the memory of the interactive video systems, or they can originate from a video source. These frames can be created through various 2 dimensional and 3 dimensional animation software, or by digitizing existing imagery from videotape, videodisk, or video camera. An alternative approach for controlling imagery is available by using a single system to manipulate and process a video signal directly from tape, disk or camera.

3.2 Video Actions

When video frames are stored in a digital format, cinematic effects are available which are not possible with film or videotape. For example, frames can be presented in a forward direction, backward direction, oscillating from beginning to end, at different speeds, in a random order, or with colour changes. The term *video action* is used within the *Orpheus* system to describe the cinematic or digital effect used for presenting frames of selected video imagery. The system provides approximately 70 different video actions, each containing a number of variables for precise control, such as speed. An example of a video action is the *forward-speed-loop* action. When activated, this video action will loop selected video frames in a forward order, and the pitch of the music will determine the speed of the frames: higher pitches cause the clip to play faster.

3.3 Active Screens

The term *active screens* refers to the number and locations of video screens that are displaying video imagery at a particular time. Dependent on a viewer's orientation, a single active screen can be located in front, behind, to the right, or to the left. Thus, four different active screens are available. Sixteen different combinations of screen activity and location are provided within the four screen configuration, ranging from all screens off to all screens on. Table 1 lists these combinations using the letters A, B, C and D to represent the four video screens.

No Screens Active	**1 Screen Active**	**2 Screens Active**	**3 Screens Active**	**4 Screens Active**
A+B+C+D	A	A+B	A+B+C	A+B+C+D
	B	A+C	B+C+D	
	C	A+D	C+D+A	
	D	B+C	D+A+B	
		B+D		
		C+D		

total combinations = 16

Table 1. Active Screen Combinations

Because the video screens within the installation are oriented around an audience, both the location and number of active screens become compositional elements. Figure 4 contains some examples of active screen combinations. In these examples, the installation is viewed from the top, with white rectangles representing inactive screens and black rectangles representing active screens that contain imagery.

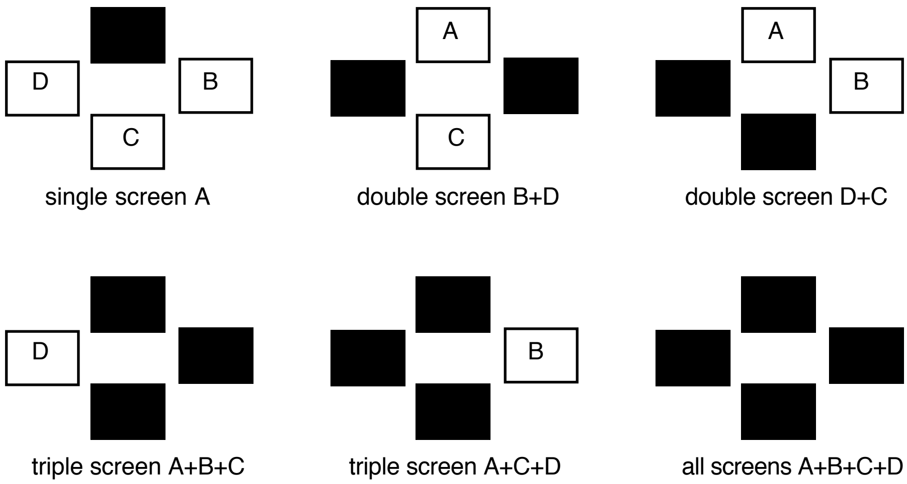

Figure 4. Examples of Active Screens and Locations

3.4 Screen Movements

Screen movements refer to a change in activity of a specific screen over time. Using the four screen arrangements shown in Figure 4, imagine screens A through D changing between active and inactive states in a sequential fashion. When viewed from the center of the installation, imagery would appear to move from one screen to another. The speed and direction for this form of movement is adjustable. Listed below are the different types of screen movements available within the four screen arrangement.

stationary: relative locations of active screens remain constant
clockwise: relative locations of active screens change in a clockwise direction
counterclockwise: relative locations of active screens change in a counterclockwise direction
random: relative locations of active screens change randomly
adjacent: relative locations of active screens alternate with adjacent screens
across: relative locations of active screens alternate with oppositely located screens

The effect created by a screen movement is obtained by having the number of active screens remain constant over a selected duration of time, and by changing their locations in a selected order. Figure 5 provides examples of different screen movements changing over time using various configurations of active screens.

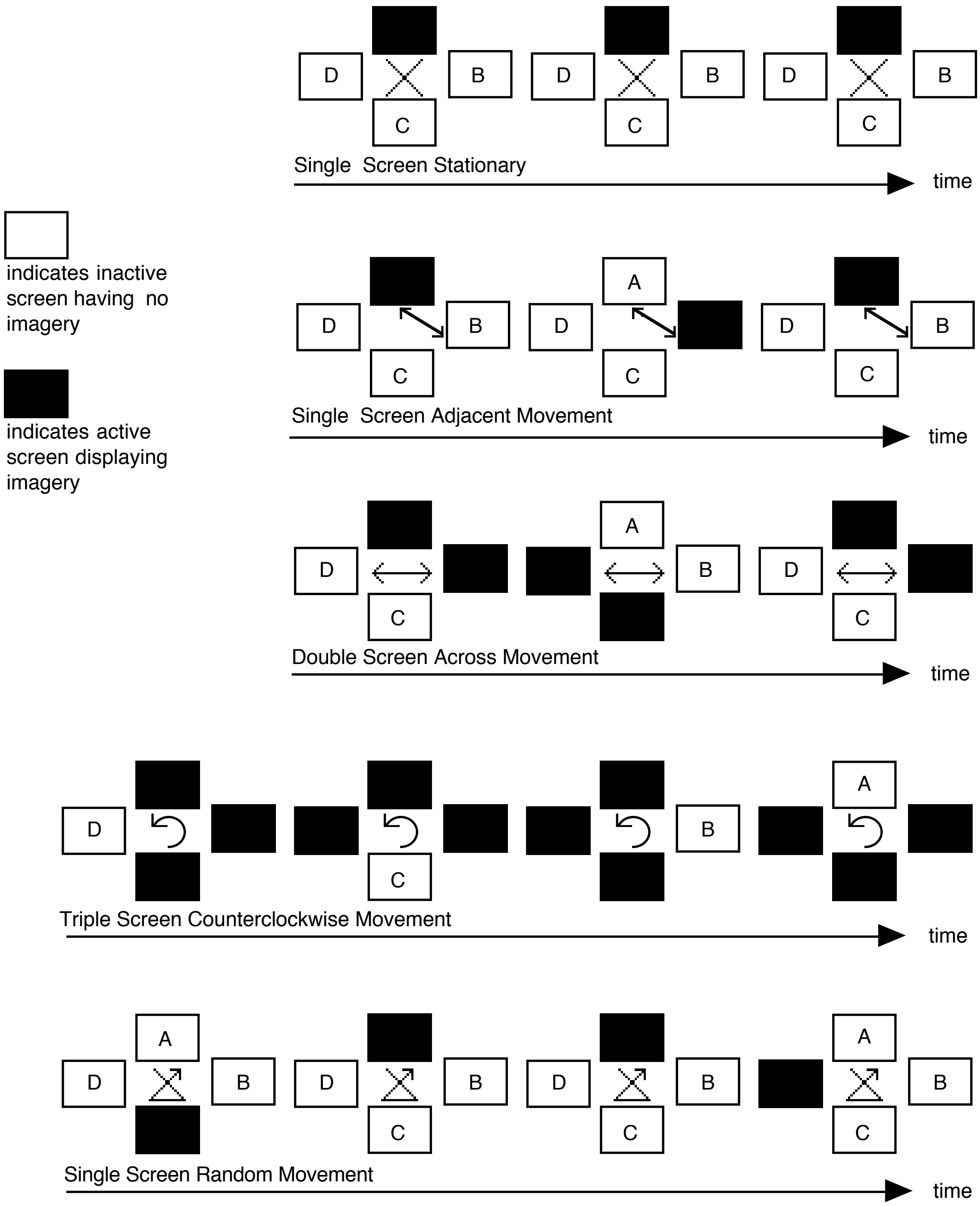

Figure 5. Examples of Screen Movements with Multi-Channel Video

3.5 Associations With Music

All of the previously described compositional elements can be directly associated to various characteristics of music. For example, the installation can be programmed to respond with a different collection of video frames for each octave of music, accompanied with a video action determined by the loudness of music. Similarly, the number of active screens and screen movements could be related to the voice being played, and the duration of a screen movement determined by the duration of an activating note.

4. Composing a Synchronized Multi-channel Video

Composing a work using this system involves the selection of associations between music, video imagery, video actions, number of active screens, locations of active screens, and the types of screen movements. Due to the large number of manipulatable components and potential decisions, a collection of symbols was developed to represent the various compositional elements. The symbols, shown in Figure 6, were created to aid composition on a score sheet resembling a music score. A multi-channel score contains the staffs and measures used with music composition, and also a variety of symbols representing activity of the image screens.

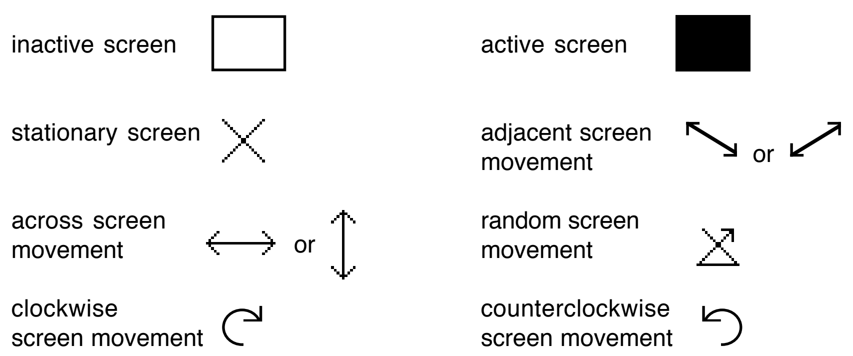

Figure 6. Symbols for Composing Synchronized Multi-Channel Video and Music

Figure 7 contains a portion from a completed score for synchronized multi-channel video. Located above each music measure in the score are symbols representing four video screens. Each combination of music measure and screen symbols are separated with a vertical line to their left, and by a horizontal line above the screen symbols. The combination of a single music measure and associated screen symbols are designated to be a single measure. The score example contains 8 measures, numbered 1 through 8 in the top left corner of each measure.

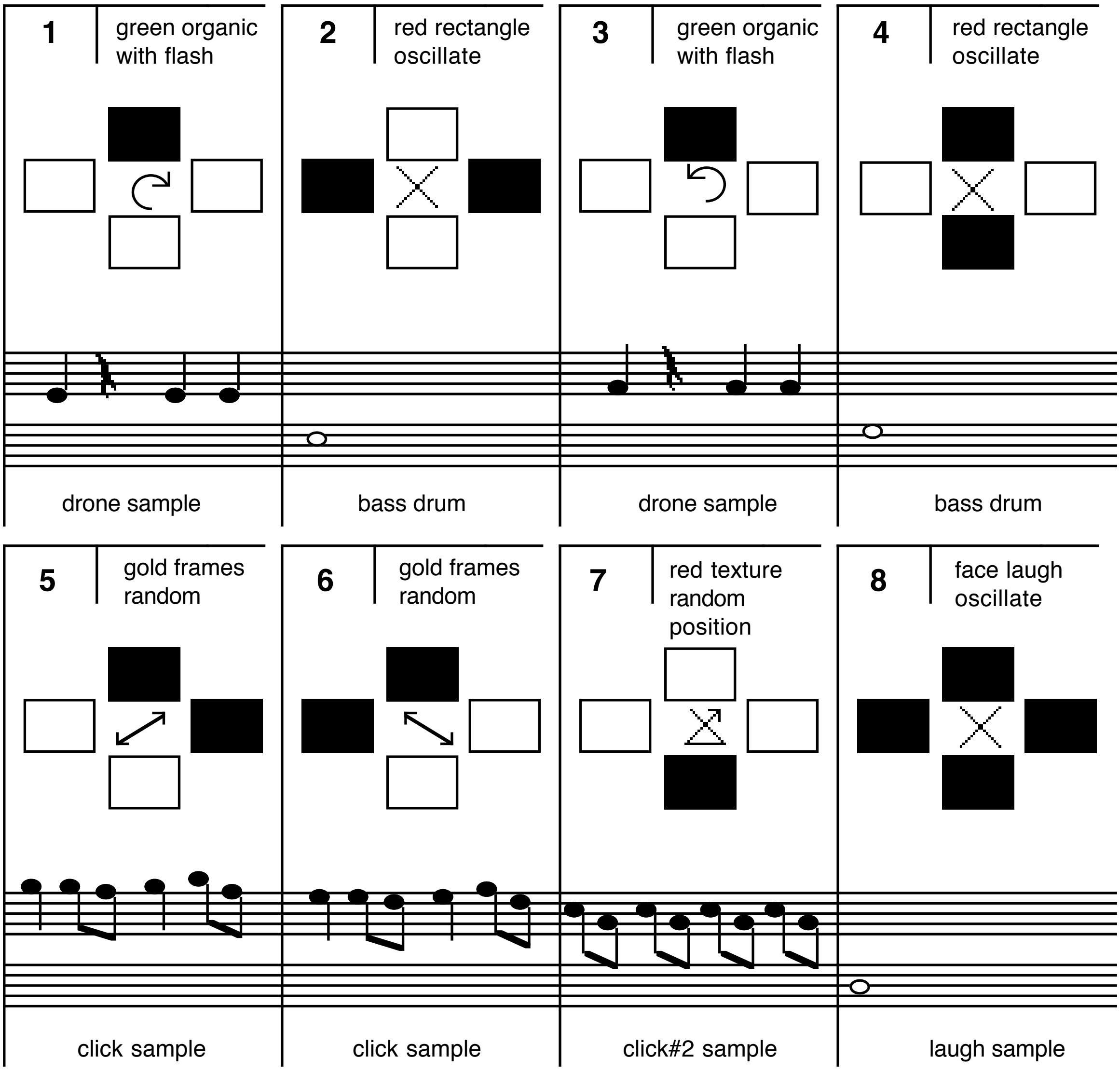

Figure 7. Score Example for Synchronized Multi-Channel Video and Music

Although complex music analyses are available within the system, music is categorized in this example according to pitch: 8 different pitch ranges are associated with 8 different music categories. Each measure in the example contains a unique range of pitches, with music type corresponding to measure number.

The score is not necessarily read from measure 1 to measure 8. Like all interactive media, this multi-channel system allows creation of a responsive environment that is nonlinear in nature. The system will react appropriately to different musical forms regardless of their occurrence in time. Therefore, all measures in the score can be played in any order or repeatedly. Measure numbers are indicated only for identification purposes.

The sole determinant for presentation of video imagery contained within a specific measure is the associated pattern of pitches contained within a measure. For example, playing of the notes in measure 1 will produce the video imagery, video action, active screens and screen movements indicated in that measure. More specifically, the playing of the pitches represented by the quarter notes will invoke the video frames *green organic* with a *flash* video action. The flash action will display a single frame of video for a brief moment, followed by a screen of black. When this effect is used with music of fast tempo, a strobe effect is produced. The name of the video frames, indicated in the top right hand corner of each measure, allows identification of different frames. An alternative is to draw small sketches of frames within the screen symbols for each measure. For measure 1, the symbols indicate only one active screen containing the *green organic* frames. The screen movement symbol indicates a clockwise rotation from screen to screen. Altogether, this measure indicates that playing the specified notes will cause the *green organic* video frames to flash around the screens in a clockwise direction. The number of revolutions produced by a single note would depend on the tempo of music, and the speed setting within the video action. The system could be adjusted to provide a complete revolution for each quarter note. The final piece of information contained within the measure is the voice for the music notes, which is a *drone* sample.

Measure 3 is similar to measure 1, having the same voice, video frames, video action, and number of active screens. They differ, however, in their screen movement: measure 1 uses a clockwise movement while measure 3 is counterclockwise. As indicated in the music measure, the video of measure 3 is produced when the music is one step higher than the music of measure 1.

Measure 2 is quite different from 1 or 3. This measure uses 2 active screens that are produced in response to a *low bass drum* note. These two screens will display an image of *red rectangle* frames using an *oscillate* video action. This effect involves playing various frames repeatedly from first to last to first, an effect resembling a pendulum swinging back and forth. Measure 4, however, is similar to measure 2, differing only in which pair of screens are active in response to a *bass drum* note that is 2 steps higher than the note in measure 2. If measures 2 and 4 were played in a repeated sequence, the pair of *bass drum* notes would produce alternating pairs of screens in synchronization with the duration of the notes. A detailed analysis of measures 1 to 4 played in sequence is presented in Table 3.

	Music Note:	**Video:**
measure 1	1: quarter note drone	single screen of green organic frames flashes in a clockwise direction around the screens for duration of note
	2: quarter rest	no response
	3: quarter note drone	single screen of green organic frames flashes in a clockwise direction around the screens for duration of note
	4: quarter note drone	single screen of green organic frames flashes in a clockwise direction around the screens for duration of note
measure 2	1: whole note bass drum	two screens containing red rectangle frames oscillate for duration of note
measure 3	1: quarter note drone	single screen of green organic frames flashes in a counterclockwise direction around the screens for duration of note
	2: quarter rest	no response
	3: quarter note drone	single screen of green organic frames flashes in a counterclockwise direction around the screens for duration of note
	4: quarter note drone	single screen of green organic image flashes in a counterclockwise direction around the screens for duration of note
measure 4	1: whole note bass drum	two screens containing red rectangle frame oscillate for duration of note

Table 3. Analysis of Measures 1 to 4 for Score Example

The remainder of the score is read in a similar fashion. Measures 5 and 6 use the same *gold* frames that are alternately displayed on a pair of adjacent screens in response to the *click* sample voice. Measure 7 uses a different voice, *click#2* and its music produces a single screen of *red texture* to move between the screens in a random order. And measure 8 presents an image of a *face laughing* in response to a *laugh* sound sample.

When creating a composition, I start with a collection of video frames from my frame library, or I create a collection of new frames. These frames are organized into various shots and sequences to be accompanied by a cinematic effect or video action. Once completed, sounds are selected that correspond aesthetically with the selected video sequences. The next step involves the creation of musical phrases on the score sheet, accompanied with selected voices and video sequences. A musical collaborator is typically used at this stage. According to differences between the phrases, a method of categorizing music is selected. This method would then be incorporated into the control system, while the video frames and selected response patterns are programmed into the four interactive video systems. At this stage, associations between different musical phrases and video responses would be initiated and modified if required. Once modifications are complete, the installation would be ready for presentation.

5. Conclusion

Although the installation being described in this paper was not actually constructed, the work was completely designed and all software was developed. The research were undertaken using a sampling keyboard, sequencing software, and a computer simulation of the installation. This configuration required two computers, one running the simulation and the other running sequencing software. Sampled music and the simulation were controlled by the sequencing software through MIDI. The simulation was created by rendering the installation using 3D animation software, and then controlling these rendered frames with the same software that would be used in a complete multi-channel configuration. The audio and video of the simulation were recorded onto videotape.

The conclusion of this research is that viewing a synchronized multi-channel video installation would provide visual experiences that are not available through single or unsynchronized installations. More physical involvement and perceptual selection would be available to a viewer by following video images travelling between the different screens,. Pursuits are currently being undertaken to acquire the resources for constructing the entire installation: five computers, four video projectors, four projection screens, two sound systems, video and MIDI interfaces, many cables, and a large space.

References

1. McLuhan, Marshall (1966) *Understanding Media: The Extensions of Man.*
New York: McGraw-Hill Book Co.

2. Cook, David (1981) *A History of Narrative Film.* New York: W.W. Norton & Co. p. 324-326

3. Ritter, Don (1993) “Interactive Video as a Way of Life." *Music Works,* vol 56, Toronto.
http://aesthetic-machinery.com/articles-way-of-life.html

4. Ritter, Don (1992) *A Flexible Approach for Synchronizing Video With Live Music.*
Berlin: aesthetic-machinery.com 2009
http://aesthetic-machinery.com/documents_pdf/Ritter-flexible-soundimage.pdf