Toward a New Model of Performance
Mon-chu Chen
Institute of Applied Art
National Chiao-tung University
u9342514@c.c.nctu.edu.tw

Abstract
This paper describes a performing model in which composers/performers manipulate aural output by controlling visual sources. The performance model is divided into four parts for detailed discussion. The functions of these four parts and the relationship between audio and video will be stated. A simple prototype of supplementation that adapts the performing model will be presented later. Finally, several transformations, extensions and alterations of this performance model will be summarized.

Background
Shooting a film or arranging a sequence of spotlights on a stage is a different procedure from composing a piece of music. There will be different intentions and meanings in different works, regardless of whether the works are categorized into musical arts or visual arts or other art forms. Are there any relationships between these dissimilar genres of arts? Are there any connections between these methods in which people create thousands of various forms of works? In other words, can music be composed in the same manner in which other art works are created?

The main objective of this paper is to attempt to compose and perform music in the manner in which lights for a play are designed, a picture is painted, or a home video is produced, etc. It can be viewed as making a black box, putting visual materials into it, shaking it, then attaching one's ear to listen to the sounds the box generates. Hence, performers and composers can control music and sound characteristics such as melody, harmony, timbre, frequency, amplitude, etc., by arranging, combining, and adjusting the visual materials such as light, shapes, colors, etc., to create music.

Overview
There are several ways to build such a black box, but two requirements are necessary. Our requirement is that the box is able to catch all the visual sources. The other is that it is able to generate sounds. Besides these two requirements, how the visual sources are combined and how the visual materials are transferred into aural materials have to be considered. The above considerations form the main parts of the performing model: the VIDEO SOURCE MODULE, the VIDEO-AUDIO TRANSFORMATION MODULE, the VIDEO-AUDIO TRANSFORMATION MODULE, the SOUND GENERATION MODULE. The performing model can be implemented with different technologies, but the computer is no doubt the best tool to realize the performance. The discussion below will be from, but not limited to, the viewpoint of computers.

VIDEO SOURCE MODULE
This is the most important part of the performing model. Composers/performers arrange all the visual sources to control the resulting sound. The sources can be video tape players, laser disc players, television programs, or cameras. The contents of the video tape or the laser disc can be any commercial film or motion pictures pre-produced by the composers. These films can be a period of computer-generated animation, sights of from Mars or Pluto, or even a simulation of a play's motions. A camera is a powerful tool in the performing model. Composers can set up a stage and use a camera to catch all the actions occurring there. Performers can wear a special kind of suit that is highly reflective or decorated with luminous objects while dancing, acting, or otherwise moving on the stage in the manner that the composer desires. Composers can also put a screen on the stage, use special masks to filter the lights, and project different colors and shapes onto the screen. Slides and projectors can enrich the varieties of the final works. Moreover, composers can turn the camera around and face the audience, thereby allowing the audience to decide how the music will develop.

Besides the single video source described above, a composer can sit in front of a video mixer and video special effect machine, manipulate a few video inputs such as several cameras from different angles, and apply special effects to them such as cross-dissolving, wipe effects, etc. All the techniques which are used in movies and TV programs can be considered in this module. Until now, composers have dealt with
visual sources without connections to aural subjects. Thus, composers have composed visual art works instead of musical works. How can visual material be transformed into aural material? This question will be discussed in the following sections.

**VIDEO DIGITIZER MODULE**

After working with all the visual sources, there should be only one video signal, which will be digitized here into data accessible to a computer. A video signal analog-digital converter is required. It is better for the AD converter to have color-conversion capability. The video AD converter should accept as many video formats as possible, such as NTSC, PAL, SECAM. Usually computers are not originally equipped with this kind of hardware; therefore an expansion card for capturing motion pictures should be added. There are two kinds of ideal computers which are suitable for this type of works. These are the 386 INDY and the Apple Macintosh Quadra 540av/650av (or PowerMac av series). Both of these have built-in video digitizing hardware. The frame rate on this hardware can support is an important consideration. A faster sampling frame rate will produce a better speed of response. Usually the video signal will be digitized as a series of still images, so-called frame grabbing. Other digitizing methods are possible. A special AD converter which provides a stream type of output digitized data can be designed.

In this case the computer will be dealing with the continuous data flow type of digitized video material instead of consecutive pictures. How the video digitizer module digests the incoming video signals will affect how the video-audio transformation module functions.

**VIDEO-AUDIO TRANSFORMATION MODULE**

The incoming data from the will be processed in this module. Composers have to carefully consider the relationship between the video sources and the intended resulting sounds. After the visual part of the entire art work has been arranged, this module will be in a crucial position. How the transformation module is constructed will definitely determine the resulting sounds, which can be considered both in the macro and the micro-scope. Composers should also consider the relationship between the characteristics of the visual sources and those of the output sounds, such as a climate or a calm. Such considerations should also include the way in which the colors, shape, brightness and other characteristics of the video source will represent appropriate timbre, frequency, amplitude, and other characteristics of the output sounds. A simple rule is to isolate the impressions from the input video sources and determine which properties each visual impression will represent.

There are several strategies for evaluating the characteristics of the incoming video signals. The red green, and blue components of each pixel in each frame, or the hue, the saturation, or the lightness components of each pixel are elements that can be processed. Averaging each frame or differencing between each two frames are helpful techniques. If the computer is fast enough, or if a performance in real time is not desired, image processing/analyzing techniques can be applied. There are thousands of mature skills which support the creation of even more amazing works.

How to determine which properties each characteristic of the visual material will represent will depend on the composer's personality and preference. The choice will be the spirit of the final work. One can simply view the incoming video raw data as the waveform of the output sounds, or one can apply several functions such as chaos or fractal to the input data. The result from this module is not the exact sound, it is merely the abstract information concerning the output sound. This data includes several properties, such as when to play and how to play the sound, etc. But, the information need not be entirely intuitional. It may encompass the parameters of the carrier in the FM synthesis for the next sound generation module.

**SOUND GENERATION MODULE**

The actual sounds are produced in this module. There are several existing techniques to produce sounds, such as DSP, MIDI. Other special composer-compatible instruments can also be considered. In this module, the information from the video-audio transformation module is to be sent selectively to the instruments. In the DSP scheme the information is explained as the sound waveform or the parameters of the user-defined wave functions. In the MIDI scheme the information is explained as the note, velocity, program change, etc. One should remember to filter out the unreasonable values. Thus, the work is completed.

**Implementation**

One prototype was implemented. A program running on an Apple Macintosh Quadra 840av was finished. The video source module is a pre-produced 4-minutes VHS video tape. The visual material is a computer-generated video produced by the Adobe Premiere 2.0 on an Apple Macintosh, which shows some of the Vincent van Gogh's paintings. The video digitizer module is the Apple Macintosh Quadra 840av's built-in video analog-to-digital converter which provides S-Video.
and RCA inputs and NTSC, PAL, SECAM video formats, and captures incoming video signals at a 50fps frame rate. This program currently provides only gray-scale conversion. In the VIDEO-AUDIO TRANSFORMATION MODULE the program was aimed to produce four voice parts. There are four hypothesized lines, two vertical and two horizontal, moving around the screen. Each line represents one voice and has a virtual point, called leading point, which determine where the line will move to. The gray-level of each pixel on the line and of the leading point will be sent to the SOUND GENERATION MODULE. In the SOUND GENERATION MODULE the wave table scheme in the Apple Sound Manager was used. The gray-level of each pixel on the entire line will be put into the wave table. The gray-level of the leading point will be the amplitude, and the difference in the gray-level between two connected lines will determine the interval between the current pitch and the next one. Thus, there will be four voices, and these four voices are limited to the right channel with mid-pitch, left channel with mid-pitch, both channels with high-pitch, both channels with low-pitch, respectively.

Finally, the work was sounded like someone murmuring and was named “Van Gogh’s Murmur!” The intention of naming it such a title is to give audience a motive to think about who is really murmuring. Is Vincent himself? Are the confused audiences? Or is the computer which created the sound? The results of this performing model are interesting, especially when you fix the VIDEO-AUDIO TRANSFORMATION MODULE and change the VIDEO SOURCE MODULE. For an implemented example, keep all the modules used in the “Van Gogh’s Murmur” except the VIDEO SOURCE MODULE, and take the Shakespeare’s Othello video tape as the source material. The results seem like a silence movie dubbed with an amusing narrative.

Possible Practices

Peculiarity: Put a hyaloid fishbowl on a projector instead of the transparencies. Drop, scatter, or dump water into the fishbowl. Add coloring into the water. Blend and let it pervade. Knock, shut the bowl. Moreover, put couples of goldfishes and tropical fishes into the bowl. Use a camera to capture all the actions which the projector projects on the screen and send the video signals to the computer.

The Pianist: Use the optical music score recognition techniques in the VIDEO TRANSFORMATION MODULE, but make the SOUND GENERATION MODULE like a beginner who usually plays piano timidly and presses the wrong keys.

Further Works

The current algorithm is quite straight forward and inefficient. It captures one frame, processes the frame, then generates sounds. The frame rate is about 10fps. It means that the resulting sounds are controlled by the input sources no more than 10 times per second. There are several ways to eliminate the limitation. One is the double buffers technique. That will let the image capturing and sound processing/generating be done simultaneously. Porting the program to a Power Macintosh, which provides calculation power of the i960 processor PowerPC, is also considered.

The MAX is a really satisfied environment for composers to experience. Writing a MAX external object which provides the functions of the VIDEO DIGITIZER MODULE is planned. Thus, composers can apply existing techniques on this performing model and explore more ways to create.

Refine the “Van Gogh’s Murmur” and realize some examples listed in the previous section.

Conclusion

This performing model provides a fascinating experimental opportunity to explore the world. Existing techniques from different art form offer wide variety of possible ways to explore.

The main idea of this performing model is to create art works in the manner in which other works of different form are created. From this view, there is not only the video-to-audio model but also many other transformations.

The frame rate is an important issue of this performing model. The dynamic of the resulting sounds will be better if the system can provide faster processing frame rate.

Acknowledgements

The author also wishes to express appreciation to both Dr. Arun Chandr and his teaching me DSP knowledge and his kindness, and Dr. Cherry Budgell for her editorial assistance in the preparation of this paper.