# EyeMusic: Real-Time Eye-Controlled Multimedia Performance

ANTHONY HORNOF

University of Oregon

TROY ROGERS

University of Virginia

JEFFREY STOLET and TIM HALVERSON

University of Oregon

---

A novel opportunity to transform eye movements into useful knowledge is to use the eye movement data to perform musical and multimedia compositions. Eye-controlled performance interfaces require new multimedia-based communication, such as between between eye tracking devices and multimedia authoring environments. This paper presents EyeMusic, an integrated environment for developing eye-controlled musical and multimedia compositions. In EyeMusic, a commercial eye tracking device communicates with a multimedia authoring environment to sonify both prerecorded and real-time eye movement data. A range of visual stimuli are explored to motivate the performer's eye movements, and the eye movements are used to trigger a range of musical elements. Compositions are created in which the performer visually interacts by alternatively reading text, studying paintings, and interacting with graphical objects that respond to the gaze. During real-time performances, the performer's eye movements also respond to the musical sounds created by previous eye movements. The project demonstrates that the eyes have a unique set of expressive capabilities that are different from, for example, the hands. Using the eyes to interact with a specialized multimedia environment can capture this expressive potential and result in compelling musical and visual performances. Two major challenges in designing eye-controlled multimedia software are addressed: (a) composing pieces that use the noise and error in eye movements and eye trackers for a positive expressive and artistic outcome and (b) finding a balance between having the performer directly control a specific musical outcome versus having the composition proceed somewhat on its own, responding in meaningful and interesting ways to the performer's eye movements. The EyeMusic system is implemented so that any composer using established composition software can incorporate eye movement data into their musical compositions. The usability of the system is demonstrated through user testing by intermedia composers. EyeMusic compositions have been selected by multiple juries for live performance at major computer music conferences.

Categories and Subject Descriptors: A.m [**General Literature**]: Miscellaneous; H.4.3 [**Information Systems Applications**]: Communications Applications; H.5.2 [**Information Interfaces and Presentation**]: User Interfaces - input devices and strategies, interaction styles; J.5 [**Arts and Humanities**]: Fine arts, performing arts.

General Terms: Design, Experimentation, Human Factors

Additional Key Words and Phrases: Human-computer interaction, computer music, eye tracking, new media art, performance.

---

## 1. INTRODUCTION

Eye tracking continues to hold great promise, not yet fully realized, for human-computer interaction, both to analyze and understanding how people interact with visual displays, and to provide an alternative means of interacting with a computer in real time. Eye movement data are useful to a variety of disciplines. Cognitive psychologists study eye movement data to understand human information processing capabilities. Human factors practitioners employ eye tracking to understand how people interact with devices and to improve usability. Accessibility researchers write software to enable physically disabled people to communicate by controlling the computer with their eye movements. Jacob and Karn [2003] provide a good overview of eye tracking research in human-computer interaction. Similar to how the scientific and expressive potential of eye tracking is not yet realized, computers in general hold great promise for reshaping our notions of multimedia performance [Wardrip-Fruin and Montfort 2003], but this is also currently a fledgling enterprise. Massive exploration of computer-mediated art, music, and performance is yet to come. This article discusses what may be the current "state of the art" in using eye movements for real-time computer-mediated creative expression.

This project emphasizes musical and multimedia compositions that can be performed live by a person using an eye tracker in front of an audience. Collaborators for this project include (a) eye tracking and cognitive modeling researchers and (b) computer music composers and performers. The project emphasizes the creation of a compelling artistic and musical performance, though the findings also have implications for adding useful and aesthetic sounds to other interfaces that use eye-input, and for sonifying eye movement data for analysis purposes.

There are a number of reasons that eye movements are useful and interesting for musical composition: (a) A performer could alternate between a higher-level visual task such as studying a painting, in which eye movements are programmed subconsciously, and another task in which he or she deliberately controls the music with his or her eye movements. (b) There is an inherent musical or at least rhythmic quality to eye movements that lends itself to composition. (c) People with physical disabilities who already interact with their computer by moving their eyes could enjoy new opportunities for musical expression. (d) From the perspective of scientists who analyze eye movement data to understand patterns of human visual processing, there may be patterns in the data that become most salient when the data are sonified rather than visualized.

## 1.1 Previous Eye-Controlled Music

We are aware of only one body of previous work in which eye movements direct musical compositions, work by the digital artist Andrea Polli [1999]. Her musical composition with eye tracking entitled Intuitive Ocusonics has been performed internationally. Excerpts are online at <www.andreapolli.com>. In this work, the eyes directly control aspects of the composition as it is performed. The pieces are striking, filled with haunting electronic sounds and digital samples of the human voice, sometimes singing and sometimes screeching. The compositions tend to be sparse, with just a few instruments or voices playing at a time.

There are three ways that our project differs from the work of Polli. First, our system provides more accurate eye movement data, and thus better control of the composition and performance. Polli's compositions respond to video images of the eyes by using STEIM's BigEye software <www.steim.org> to parse and process video images twelve times per second, but without the benefit of specialized algorithms for translating the video images into screen coordinates. EyeMusic uses a commercial eye tracker, the LC Technologies Eyegaze System, which reports the screen coordinates of the gaze position with higher spatial and temporal accuracy.

The second way that our approach differs from that of Polli is that EyeMusic works with both the gaze position as well as the fixations that the eyes make on visual stimuli. These are the most relevant psychological phenomenon for observing and understanding human visual processing [Findlay and Gilchrist 2003]. The resulting compositions and data sonifications relate more closely to the human processes that are at work.

The third departure is that EyeMusic puts more emphasis on the scene that is viewed during the performance, and how the gaze moves around and interact with the scene, and less emphasis on the image of the eye moving. Polli's live performances include a large video image of the eye as it moves. Our performances include a video image of the gazepoint superimposed on the scene viewed by the performer. Professor Polli is currently collaborating with us to compose using EyeMusic.

## 2. EYEMUSIC SYSTEM ARCHITECTURE

This section introduces the technical components of EyeMusic system, and explains how the components interact. In that the EyeMusic system architecture interacts directly with the human physical "architecture," a few terms pertaining to eye movements and eye tracking must be defined. The gaze is the vector that goes from a person's eye to the gazepoint, which is the point in a scene where he or she is looking. The eyes (and thus the gaze) examine a scene with a series of quick jumps called saccades, each of which lasts roughly 30 ms. Between saccades, the eyes (and the gazepoint) stay at the same location (with a slight tremor) for a fixation that lasts roughly 100 to 400 ms. People acquire visual information during fixations, not saccades. The eyes move so that people can put items of interest into the high resolution vision at the center of the gaze.

Individual eye movements are usually made subconsciously in the service of a higher-level strategy to accomplish a visual task, such as reading. However, people can also make deliberate, conscious decisions to move their eyes to a specific location, such as to fixate this letter *X* for two seconds and to then pass the control back to the higher-level reading strategy. For a more detailed explanation of how and why the eyes move, see Rosenbaum [1991].

Eye movements are monitored using an *eye tracker,* which typically incorporates a camera that sits below a computer video monitor and is focused on the eyes of the person using that computer. The video images are transformed, via software algorithms, into the (*x, y*) coordinates of where the person is looking on the screen.

Figure 1 shows the major software and hardware components in the EyeMusic system. Arrows indicate the flow of data. A *scene generator* displays a visual image on a video display. A person, the *eye performer*, moves his or her gaze around the scene. A video camera captures an image of the eyes and, in the *eye tracking computer*, converts it to the corresponding (*x, y*) coordinates of the video display. EyeMusic currently uses the LC Technologies Eyegaze System (www.eyegaze.com), which monitors the change in the spatial relationship between the pupil-center and a corneal-reflection as the gaze moves across the screen, and reports the gaze position sixty times per second, once every 16.67 ms. In our experience working with this eye tracker, once a good calibration is acquired for a performer, the accuracy is within 1° of visual angle. The sampling rate and the accuracy of the eye tracker are more than adequate for capturing the physiological phenomena as well as the unique personality and characteristics of eye movements. The signal-to-noise ratio is more than enough to clearly see and hear the eye movements translate into the intended music.

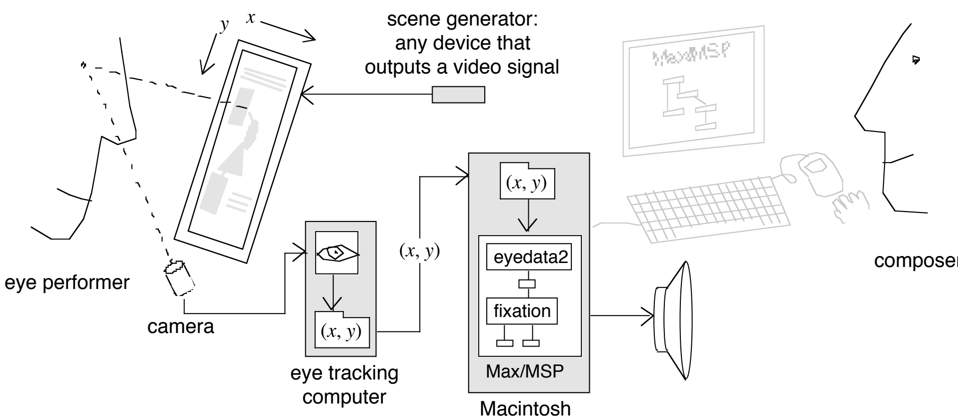


Figure 1. An overview of the major components in the EyeMusic system. Arrows indicate the flow of data.

EyeMusic makes the eye movement data available within Cycling 74's Max/MSP, a graphical environment for creating music and multimedia (www.cycling74.com). Max/

MSP is one of the most widely used software applications for electronic music composition. Max is typically used in conjunction with the Max Signal Processor (MSP) and thus referred to as Max/MSP. It is often used with an object library called Jitter that supports real-time manipulation of graphics and video. The eye movement data are made available in Max/MSP by means of an *eyedata* external object and parsed into fixation data by means of a *fixation* external object, both of which were created for this project and are discussed in the next two subsections.

EyeMusic works in two different modes: *playback* and *performance*. In playback mode, prerecorded eye movement data are read from a file into Max/MSP. In performance mode, the eye movement data are reported in real time, and the performer plays the music by moving his or her eyes. Alternative configurations to those shown in Figure 1 are possible. For example, the eye performer and the composer could be the same person, and the Macintosh could be used as the scene generator, with the scene changing as a function of where the eye performer looks.

## 2.1 The Eyedata External Object in Max/MSP

The basic building blocks in Max/MSP are data processing elements called objects. End-users can add customized objects called external objects that are written in the C programming language. In EyeMusic, eye movement data are read into the Max environment by means of an external object called eyedata (technically eyedata2 for version tracking). Figure 2 shows the eyedata object as it would appear in the Max/MSP authoring environment. This is where the eye tracker connects to Max/MSP. The gaze position (and the pupil radius) flows out of the eyedata object at a rate of sixty samples per second.

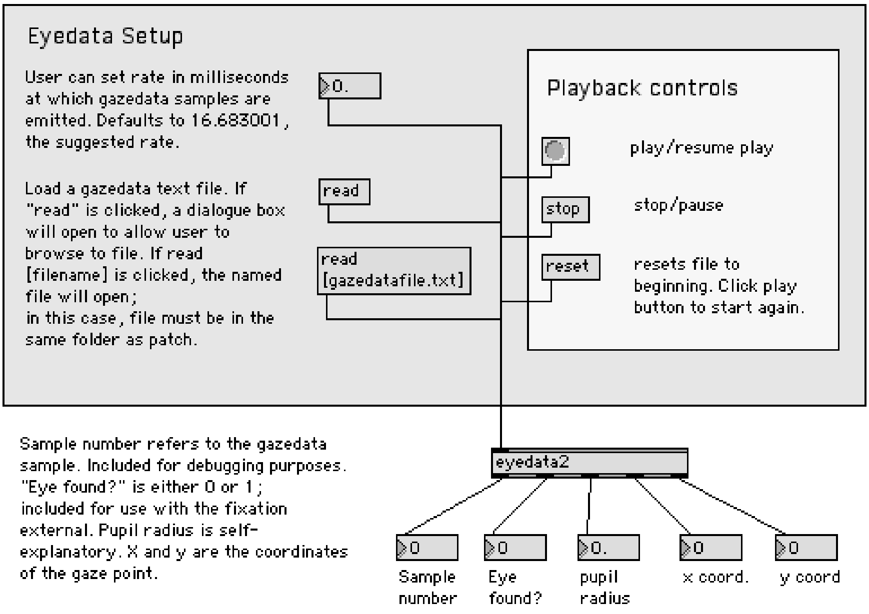


Figure 2. The *eyedata* external object in the Max/MSP multimedia authoring environment. In this visual programming environment, data and commands flow down the "patch cords."

The eyedata object is used in playback mode. It reads the eye movement data from disk, and outputs the data to any Max object that is connected to its outputs. Table 1 shows some sample data output from the external. *Sample #* increments 60 times a second. *Eye found?* indicates whether the eye was tracked for that sample (1 = yes, and 0

= no). The *x* and *y* are the screen coordinates of the gaze, in pixels, with (0, 0) at the top left of the screen. *Pupil radius* is reported in mm.

The data shown in Table 1 were collected while a person was reading. On Sample #390, the eyes finish a fixation at roughly (553, 112). On Sample #391, the eyes make a horizontal saccade to the right. On Sample #392, the eyes start a fixation at roughly (652, 109).

Table 1. Sample data from the eyedata object.

| Sample # | Eye found? | *x* | *y* | Pupil radius |
|---|---|---|---|---|
| ... | ... | ... | ... | ... |
| 388 | 1 | 553 | 112 | 1.31 |
| 389 | 1 | 552 | 112 | 1.32 |
| 390 | 1 | 554 | 112 | 1.31 |
| 391 | 1 | 576 | 111 | 1.32 |
| 392 | 1 | 634 | 108 | 1.3 |
| 393 | 1 | 663 | 108 | 1.31 |
| 394 | 1 | 659 | 111 | 1.3 |
| ... | ... | ... | ... | ... |

### 2.3 The Fixation External Object in Max/MSP

The fixation external object converts the gaze samples into fixations. The samples flow out of the eyedata object at a rate of sixty samples per second. The rate is an artifact of the eye tracker used, and different trackers have different rates. The relevant human physiological phenomenon is not the sample rate, but where and when the gaze lands for a fixation. The fixation object parses the gaze position samples, identifies where the fixations occur, and outputs each fixation and its location. This way, a composer can work directly with the eye movement data that eye tracking specialists believe to be most relevant. Figure 3 shows the fixation object as it would be used by a person creating a multimedia composition in Max/MSP.

The fixation object uses an established dispersion-based algorithm [Salvucci and Goldberg 2000]. The algorithm is based on the fact that eye positions sampled during a fixation tend to cluster in a small region for a minimum amount of time. Two parameters must be set in the algorithm: (a) the deviation threshold, which is the size of the region in which the fixations must cluster to be identified as as fixation, and (b) the duration, which is the minimum amount of time that samples must be recorded in a cluster for a fixation to be reported. The fixation object defaults to commonly-used settings for both parameters: (a) a deviation threshold of 20 pixels, which corresponds to 0.5° of visual angle at normal viewing distances, and (b) a duration of 100 ms, which corresponds to 6 samples (at a rate of 60 samples per second). Both of these parameters can be modified in the composition permitting, for example, one fixation object to respond to 100 ms fixations, and another to 500 ms fixations.

The EyeMusic externals, documentation, eye movement data, and audio-visual recordings can be downloaded at <www.cs.uoregon.edu/research/cm-hci/EyeMusic/>. An eye tracker is not needed to use EyeMusic in playback mode.

## 3. COMPOSITIONS USING PRERECORDED EYE MOVEMENT DATA

We created several compositions using the EyeMusic system in playback mode. In this mode, the system reads from a file that contains the (x, y) coordinates of the gaze positions recorded in an earlier session in which the composer viewed a scene. To compose a piece in which the music responds to where the eyes were looking in the scene, a visual recording of how the eyes moved through a visual scene is also needed. Without such a record, it is not possible for the composer to plan or the audience to observe a correlation between the eye movements and the sounds created by those movements. For example, the composition could be designed to play a bass note when the gaze lands on a frog, but both the composer and the audience must be able to see the gaze land on the frog. To solve this problem, when collecting the (x, y) data, we also record a video that captures the contents of the display that the performer is looking at, and superimpose the gazepoint on the scene using a hardware video mixer and the video output from the eye tracker (not shown in Figure 1). The video can be played back within Max/MSP simultaneously with the (x, y) data.

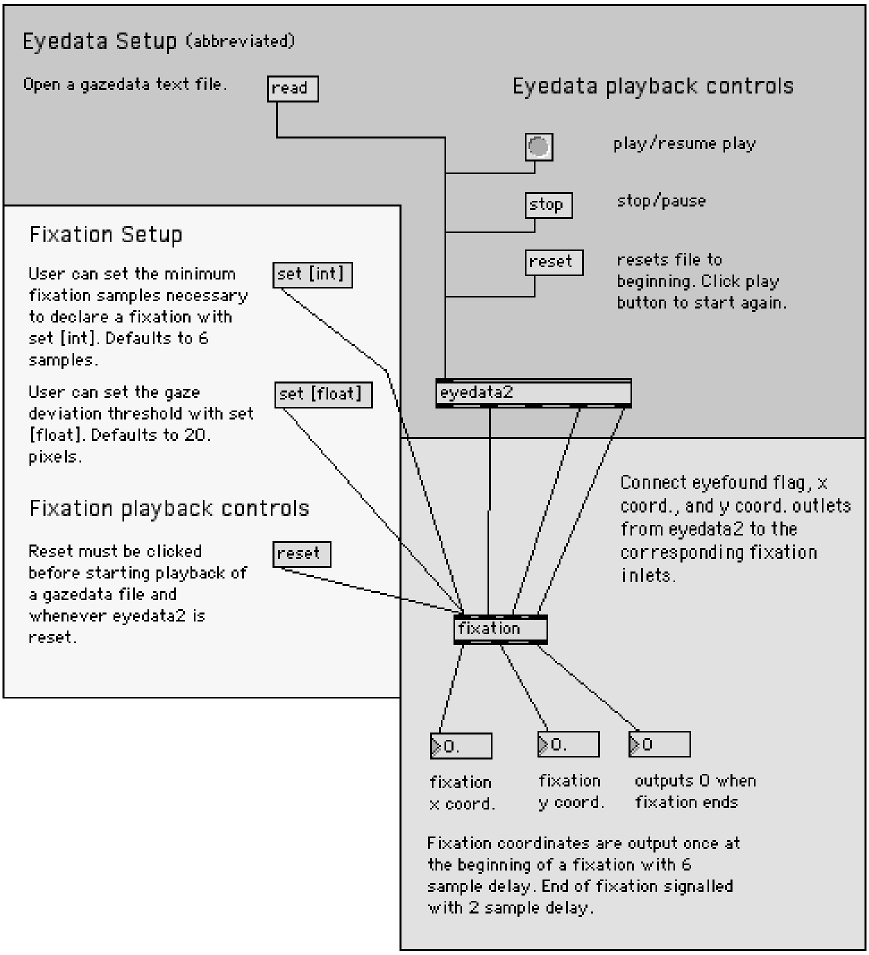


Figure 3. The output from the eyedata object flows into the *fixation* external object in Max/MSP.

## 3.1 The Reading/Typing Composition

One early compositional sketch is entitled Reading/Typing. The composition is based on the fixations made while reading a text that describes the fixations that are typically made while reading. For Reading/Typing, we recorded the eye movements of a performer as he simply read a passage of text displayed on the screen. To create a certain self-referential aspect to the piece, the text passage is from Rayner and Pollatsek [1989] and describes the characteristics of eye movements that people make, without realizing it, when reading. In addition to collecting the eye movements, a videotape was also made that superimposed the gazepoint on the reading material on the page.

The eye movement data collected for Reading/Typing are sonified as follows. Every time a fixation occurs, a typewriter keypunch sound is played. Every time the eyes move back to the start of a new line, an old-fashioned typewriter carriage return and bell sound is played. The fixations sound remarkably like typing because the fixations occur at the same rate that a slow typist would press the keys, about four per second, and because the slightly varied rhythm resembles that of typing. The sounds are played at the same time that the videotape is played back, and you can both see and "hear" the performer read the text. Reading/Typing is a rhythmic visualization. An audio-video recording of "Reading/Typing" is available on the EyeMusic web site.

### 3.2 The EyeMusic v.0.9b Composition (Viewing Kandinsky)

We next set out to produce a more musical playback of eye movement data, and produced EyeMusic v.0.9b. For this composition, the composer (the third author) watched a series of Kandinsky paintings that were slowly zoomed and panned across the video display. Data collected included the (x, y) coordinates of the gaze samples, and a video of the fixation point superimposed on the Kandinsky paintings.

For the composition of EyeMusic v.0.9b, the eye movement data are output from the eyedata object at the usual rate of sixty samples per second, and also selectively sampled at slower rates within Max/MSP. The selective sample rate is roughly once every 500 ms, but for musical purposes is varied slightly during the piece and ritards at the end. The fixation object was not used in this composition. EyeMusic v.0.9b produces a primary melody and a pointillistic counterpoint. The primary melody is derived from the horizontal dimension of the gaze location, with the pitch of the melody increasing as the gaze moves from left to right, as on a piano keyboard. The counterpoint is derived from the vertical dimension, with the slowed-down samples cycling through eight different MIDI (Musical Instrument Digital Interface) instruments, and each sound triggered roughly once every four seconds. The vertical dimension is also mapped to pitch, with higher gazepoints triggering higher pitches. The resulting music, influenced in important ways by the Max processing, displays clear parallels with the original data and produces a mysterious and lyrical ambiance. The piece is a little under six minutes long.

EyeMusic v.0.9b was presented at a Future Music Oregon concert on November 15, 2003. The audience watched the gaze move across the paintings and listened to the music generated by those eye movements. Max/MSP processed the data in real time during the presentation. An audio-video recording of the piece can be viewed on the EyeMusic web site.

The compositions Reading/Typing and EyeMusic v.0.9, capture two possible mappings of prerecorded eye movement data to sound. There are infinite other possibilities. The next section discusses how EyeMusic has been used for live performances.

## 4. COMPOSING FOR REAL-TIME EYE-CONTROLLED PERFORMANCES

After establishing the connection between the eye tracker and the multimedia authoring environment and creating several pieces in playback mode, we explored how the eye tracker might be used for real-time eye-interactive multimedia performance. This requires both the creation and the performance of the multimedia work. It also involves an interactive process of determining how the eyes can interact with the multimedia environment, a process akin to designing a new musical instrument.

Figure 4 shows how the EyeMusic system is configured for live performance. Figure 4 is derived from Figure 1, but the details of the eye tracking and Max/MSP systems are removed. The video image that the performer watches during the performance is no longer an arbitrary video signal, but is instead created by the Max/MSP multimedia system, which is now extended with the Jitter object library for real-time manipulation of graphics and video. Note that a feedback and control loop has been created, such that the performer can now visually interact with objects on the screen, such as by looking at objects and having the objects change based on the gaze. The auditory and visual feedback loop creates an experience analogous to playing a physical instrument. The performer's video monitor is also shown to the audience (though flipped horizontally) so that the audience can observe the visual elements of the performance, and see how the performer is interacting with the display. The image needs to be flipped horizontally (easily accomplished with most video projectors) so that the performer's physical gaze corresponds to the projected gaze shown to the audience. The output from the camera that is used by the eye tracker is run through a video splitter so that the audience can view, in a separate video monitor, the same image of the eye that is used by the eye tracker. This provides a close up view of what the eye is doing.

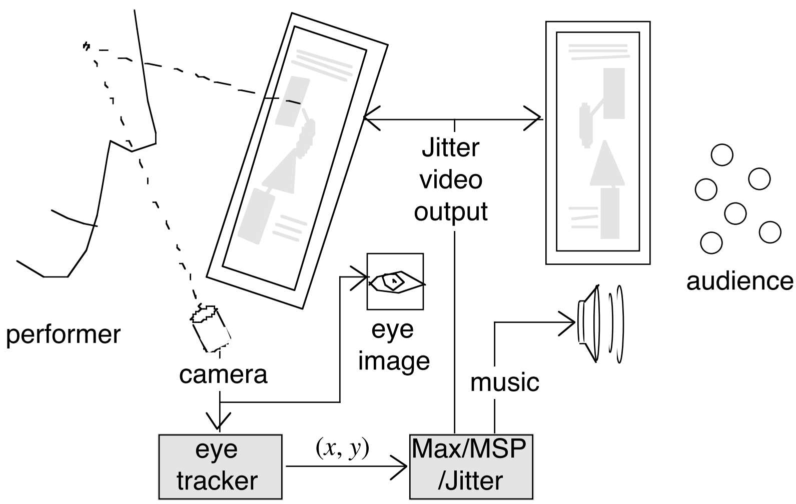


Figure 4. The EyeMusic system configured for live performance. Arrows indicate the flow of data and control.

The Max/MSP/Jitter composition is created in advance of the performance, and this preparation is analogous to both the construction of a musical instrument and the composition of a musical score. The next subsections discuss compositions created to explore the real-time expressive potential of the EyeMusic system, and our discoveries of how eye movements can and cannot be readily used for real-time creative multimedia expression. Our design exploration included the creation of multiple case studies, including an eye-controlled piano, discussed next.

### 4.1 Eye-Controlled Piano

To explore one possible means of playing music with the eyes, we created an eye-controlled piano. The eye piano functions analogously to eye typing, a well-established communication technique in which a key is pressed by holding an eye-controlled cursor on that key [Majaranta and Räihä 2002]. In our eye piano, a piano keyboard replaces the typing keyboard. Figure 5 shows part of the Max/MSP/Jitter interactive composition that we wrote to create an eye piano. When actually played, however, the octave of piano keys filled the entire computer display to give the performer larger targets and thus greater control and ability to land on the desired key with a single eye movement. Small dots were also placed on the keys as peripherally-salient targets for eye movements. The eye piano played a note, sonically and visually, when a fixation was detected on the key, using a dispersion-based fixation-detection algorithm with a 100 ms minimum fixation time. Salvucci and Goldberg [2000] discusses fixation-detection algorithms.

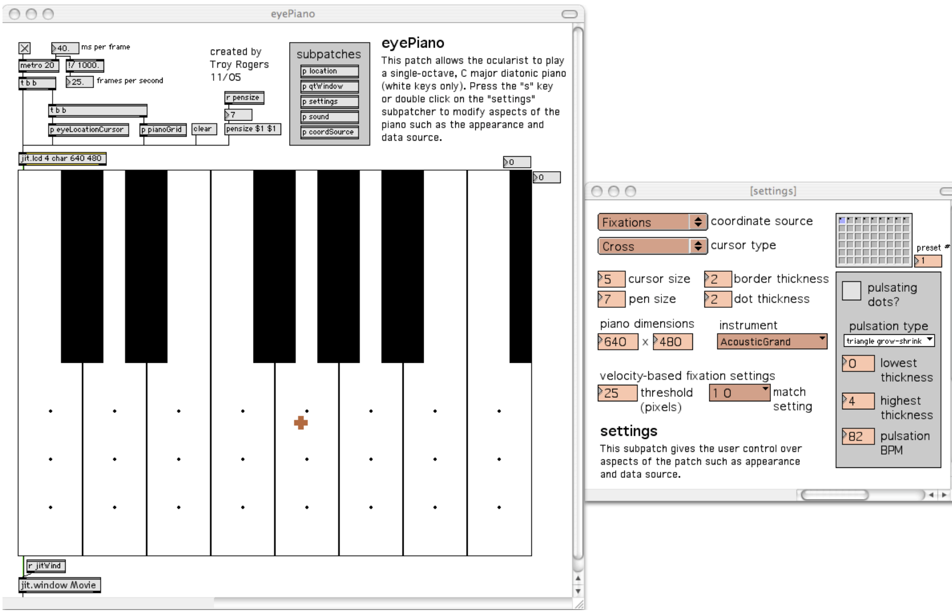


Figure 5. An eye-controlled piano in Max/MSP. The small plus sign on the keyboard is the gaze position.

In short, an eye-controlled piano based on a dispersion-based fixation-detection does not work. The second author, an experienced analog and computer musician, practiced playing the fixation-based eye-controlled piano for roughly an hour a day for two weeks, and intermittently for another month after that, and decisively concluded that (a) he was only slightly able to improve his ability to move to the intended piano keys and (b) he was not at all able to improve his rhythmic accuracy. He specifically worked on trying to improve his and the instrument's ability to play a steady beat, such as by practicing alternating between two keys, but with no improvement.

It is important contribution to report that an eye-controlled piano did not work. Numerous interface designers, computer musicians, and researchers (familiar with the success of eye typing) suggested that we create an eye-controlled piano. Based on the characteristics of eye movements (such as roughly four per second), we did not expect the piano to work very well, but we did expect it to help us to better understand the capabilities and limitations of the eyes in a musical context. This exploration contributed substantially to the design and discovery of the new idiomatic interface that was ultimately constructed.

## 4.2 Rhythmic Control

Since rhythmic control is critical in musical performance, and since it could not be achieved with an eye piano using a dispersion-based fixation-detection, subsequent instrument design was made working almost exclusively with the "raw" gazepoint data that arrives in Max/MSP/Jitter every 17ms. This created an interesting tension between the eye tracking researchers and the electronic musicians collaborating on the project. At least one of the eye tracking researchers insisted that fixations (when the gaze stays at the same place for roughly 100 to 400 ms) are the relevant psychophysical phenomena, and so the composition should be designed around the fixations, not the raw data from the eye tracker, which is merely an artifact of the instrumentation. The computer musicians, accustomed to working with the noise and idiosyncrasies associated with a wide range of sensors (accelerometers, infrared, etc.), preferred to work with the data from the sensor in its most native form, in this case the (x, y) gazepoint samples supplied every 17 ms. Over time, the musicians successfully demonstrated how using this raw eye movement data not only gave the performer more direct control and expressive opportunity with their instrument and the performance, but also permitted more interesting and varied compositions.

By starting with the raw sample data rather than the fixation data, we further determined that, despite its initial failure, the notion of an eye-controlled piano should not be entirely dismissed. There were two ways that it did work somewhat well. First, when the piano was modified so that a key would play when a single sample landed on it, rhythmic control improved, though at the expense of melody accuracy since every movement between piano keys produced a sloppy glissando. Second, a recent discovery we made is that an eye-controlled piano that uses a velocity-based (as opposed to dispersion-based) fixation-detection algorithm works better for rhythmic control with the eyes. When we tried this, rhythms could be performed better. This makes perfect sense in retrospect because the velocity-based algorithm can more quickly indicate the start of a fixation. The discovery came after the following compositions were created, however, and so the remaining compositions work primarily with the raw gazepoint data.

### 4.3 Visually Interacting with Objects

Besides working with the raw gazepoint data, the exploration moved away from having the eyes press buttons to make sounds, which is basically trying to use the eyes to control devices that are designed for the fingers. The eyes may have other better-suited expressive control capabilities. We moved to an interaction paradigm in which the performer moves an eye-controlled cursor around on the screen, and makes the cursor come into direct visual contact with other visual objects on the screen, producing an instantaneous visual and sonic reaction. We also sonified the eye cursor in a variety of ways so that it continually emits sound based on its position and movement. These interactions produced a variety of feedback loops, as with analog musical instruments, and perhaps a loop that is tighter and more responsive than the eye-controlled piano.

### 4.4 The EyeMusic v1.0 Composition and Performance

The most recent major milestone of the project is the creation of EyeMusic v1.0, an eye-controlled musical and interactive visual composition that builds on the ideas discussed in the previous subsections. Storyboarding was used to help design the content and flow of time-based, multimodal, musical compositions. Storyboarding is an established means for describing human-computer interactive behavior [Rosson and Carroll 2002]. Though this design technique was new to the computer music composers, they quickly saw its value in collaborating on the design of an interactive composition. Figure 6 shows an excerpt from the cartoon-based storyboard, which was augmented with a vocal track that sonically sketched out some basic ideas for what the various interactions might sound like. Perhaps the most important point of the storyboards is that an eye-controlled multimedia composition needs to be designed as an interactive experience, and not as instrument that waits idly for the performer to take control and play an independently scripted (such as memorized) composition.

Figure 7 shows the physical staging of EyeMusic v1.0, with the performer facing the audience, a 17” video display showing the video image of the eye, and a large video projection displayed behind the performer that matches what the performer sees.

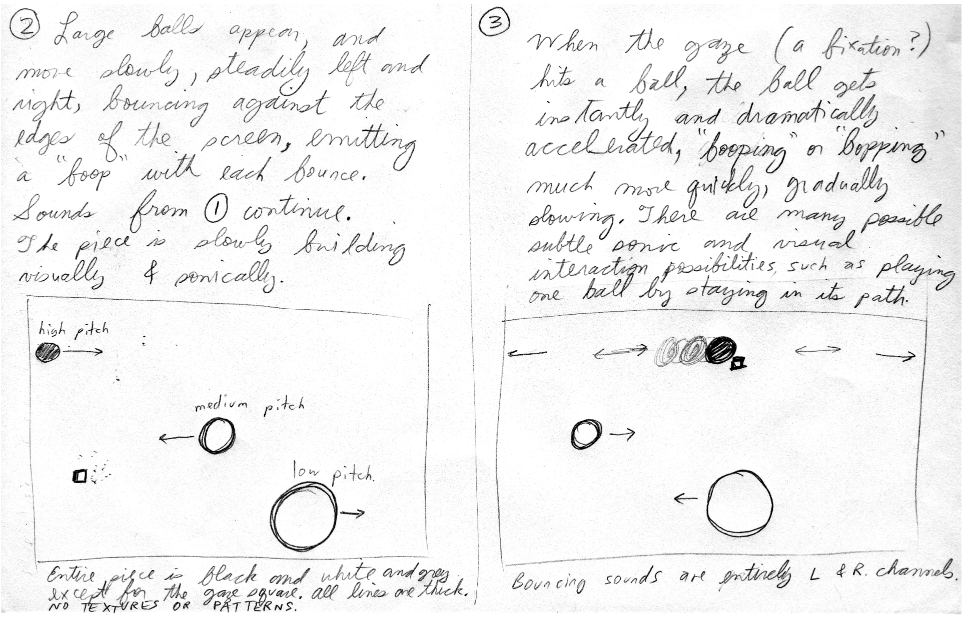


Figure 6. Two frames from a storyboard created for the EyeMusic v1.0 interactive composition. A recorded spoken track of sound effects accompanied each frame.

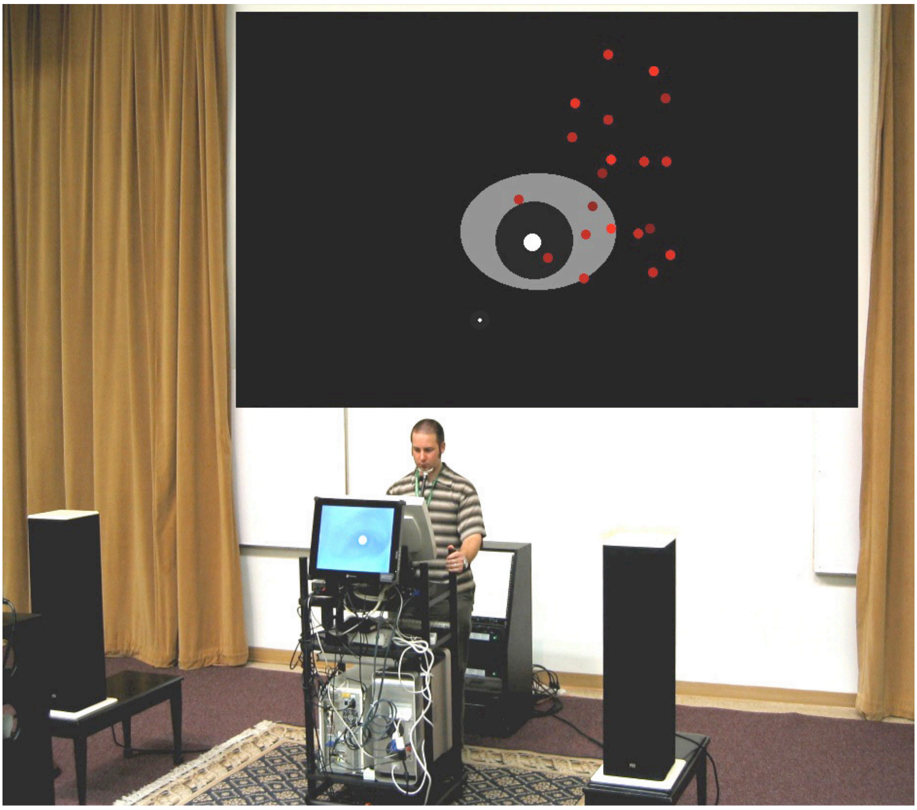

Figure 7. The second author performing EyeMusic v1.0 at SEAMUS 2006, and (inlaid above him) a screenshot of the video image that was projected behind him during the performance.

One of the goals of the composition and staging was to insure that the audience could easily and directly understand that the composition was performed with eye movements. This was accomplished by using extremely simple visual effects, and with great consideration of the physical staging. Showing the video image of the eye used by the eye tracker, for example, helped a lot. This was inspired by with what we learned in our accessibility research, in which the video image helps users stay in range of the camera [Hornof and Cavender 2005]. Directly asking the audience what they just saw following an early performance of EyeMusic 1.0 demonstrated that the audience was clearly able to understand that the eyes were controlling the performance.

EyeMusic v1.0 is roughly five minutes long, and starts with (a) the opening of the eyes which fills the small monitor with the eye image captured by the camera and (b) the appearance of a single gazepoint eye cursor appearing on the rear-projected screen. The eye cursor is a large white dot, drawn on a black background. Throughout the performance, quiet, scratchy, jittery noises play near the gazepoint, using stereo panning to move the sound from left to right as the eyes moved left and right. The composition moves forward using the closing and opening of the eyes, each of which is expressed as sudden, loud, slamming sound. Sometimes the eyes stay closed, during which the eye-closing sound decays, and the rear-projection is blank.

As the composition progresses, red circles appear on the display. The performer moves the eye cursor close to, but does not touch, the circles. The near misses create some tension. Eventually, the eye cursor touches a circle, which activates the circle to bounce between the left and right edges of the display, slowly decelerating over time. The performer "plays" the circles like an instrument, first tentatively and in a somewhat subdued manner with occasional glances, and eventually in a frenetic, over-energized manner such as by keeping the gaze directly in the left-right path of a circle, insuring continual re-activation. Each bounce produces a rubbery bouncing sound, with smaller circles creating higher-pitched sounds.

A loud crashing blink of the eyes replaces the bouncing balls with a cartoonish eye in the center of the screen. The eye mimics the video eye image of the eye that is used by the eye tracker and shown to the audience. The eye cursor moves around the screen and plays with a swarm of red dots that follow the boids model of bird blocking [Reynolds 1987]. The screenshot in Figure 7 illustrates the cartoonish eye and swarm of dots. The swarm initially avoids but is eventually attracted to the eye cursor, like moths around a light bulb. A small blip sound is played when the eye cursor touches a red dot. The entire piece is punctuated with the loud, sudden blinks, which dramatically increase in frequency near the end of the performance.

EyeMusic v1.0 was selected by juries for performance at SEAMUS 2006, the annual conference of the Society for Electro-Acoustic Music in the United States and NIME 2007, the international conference on New Interfaces for Musical Expression. A video of EyeMusic v1.0 is available on the EyeMusic website.

## 5. CONCLUSION

This paper discusses the exploration of new interaction techniques for real-time eye-controlled music and multimedia performances. Though designing any human-computer interface requires attention to the intended context of use, the user's expectations, and the intended outcome of the interaction, these requirements are perhaps even greater for eye-controlled interfaces. The basic input-output and human-computer control mechanisms that work at the 100 to 500 ms grain size must redesigned for each task. For example, the fixation-detection algorithm generally used for eye-typing does not work well for eye-piano-playing because it does not provide good rhythmic control. Further, some people with severe motor impairments have deliberate control of their gaze but not their blinks, and so for these performers the blink control in EyeMusic v1.0 would need to be replaced, such as with an off-screen glance.

Two of the problems of eye tracking that continue to keep it from its great promise are issues pertaining to noise and control. Working with computer musicians who routinely deal with these issues with a wide range of sensor technology provides a refreshing perspective. The musicians readily embrace the noise and sometimes even monopolize it as a feature, as did Jimi Hendrix with feedback and distortion on the electric guitar. When working with eye tracking data, rather than trying to remove all the noise, whether it comes from the eye tremor or the eye tracker, the musicians embrace and want to work with the noise. Regarding control, the musicians are comfortable with creating a composition that cannot necessarily be played the same way every time. In his closing plenary talk at CHI 2005 (the major annual ACM conference on human-computer interaction), Michel Waisvisz, the director of the Studio for Electro-Instrumental Music (STEIM), explained that, when designing electronic and computer musical instruments, the goal is not always control. Some materials such as a violin string have a life of their own, and the goal is to bring those materials to life. EyeMusic continues as an exploration into bringing to life the musical properties of the eyes.

EyeMusic will continue to be developed on a number of different fronts. We will continue to explore possibilities for live performances, to see how a musician can control a musical passage with his or her eyes. We will also explore the sonification of eye movements for data analysis purposes. Eye movement analysis for scientific purposes is difficult and time-consuming. Sonification of eye movement data will likely enhance current techniques. There are characteristics of human audition, such as a slower decay in working memory, that may make sonification of eye movements useful for displaying certain data trends. Lastly, we will explore opportunities for EyeMusic to open doors to musical composition and performance for people with severe mobility impairments, who interact with the world via eye tracking. We have already developed software that enables disabled children to draw with their eyes [Hornof and Cavender 2005]. Making music with the eyes would follow nicely.

## ACKNOWLEDGMENTS

Ying Tan provided invaluable aesthetic critique. Linda Sato programmed the Max/MSP externals. The project was supported in part by the National Science Foundation (NSF) through Grants IIS-0308244 and IIS-0553117. The eye tracking lab used for these projects was supported by the Office of Naval Research (ONR) through Grants N00014-02-10440 and N00014-06-10054. These grants were made to the University of Oregon with Anthony Hornof as the principal investigator. Any opinions, findings, conclusions, or recommendations expressed in these materials are those of the authors and do not necessarily reflect the views of NSF or ONR.

## REFERENCES

FINDLAY, J. M. AND I. D. GILCHRIST. 2003. *Active Vision: The Psychology of Looking and Seeing,* Oxford University Press.

HORNOF, A. J. AND CAVENDER, A. 2005. EyeDraw: Enabling children with severe motor impairments to draw with their eyes. In *Proceedings of ACM CHI 2005: Conference on Human Factors in Computing Systems,* ACM, New York, 161-170.

JACOB, R. J. K. AND KARN, K. S. 2003. Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In *The Mind's Eyes: Cognitive and Applied Aspects of Eye Movements,* J. HYONA, R. RADACH, & H. DEUBEL, Eds. Elsevier Science, Amsterdam, 573-605.

MAJARANTA, P. & RÄIHÄ, K-J. 2002. Twenty years of eye typing: Systems and design issues. In *Proceedings of ETRA 2002: Eye Tracking Research & Applications Symposium 2002,* ACM, New York, 15–22.

POLLI, A. 1999. Active vision: Controlling sound with eye movements. *Leonardo 32*(5), 405-411.

RAYNER, K., & POLLATSEK, A. 1989. *The Psychology of Reading.* Prentice Hall, New Jersey.

REYNOLDS, C. (1987). Flocks, herds and schools: A distributed behavioural model. *Computer Graphics 21*(4), 25-34.

ROSENBAUM, D. A. 1991. *Human Motor Control.* Academic Press, New York.

ROSSON, M. B., & CARROLL, J. M. 2002. *Usability Engineering: Scenario-Based Development of Human-Computer Interaction*. Morgan Kaufmann, San Francisco, CA.

SALVUCCI, D. D. AND GOLDBERG, J. H. 2000. Identifying fixations and saccades in eye-tracking protocol. In *Proceedings of ETRA 2000: Eye Tracking Research & Applications Symposium,* ACM, New York, 71-78.

WARDRIP-FRUIN, N. & MONTFORT, N., Eds. (2003). *The New Media Reader.* MIT Press, Cambridge, Mass.

Submitted for Review to ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP) on April 6, 2007.