

# EyeMusic: Making Music with the Eyes

Anthony Hornof and Linda Sato

Department of Computer and Information Science

University of Oregon

Eugene, OR 97403 USA

{hornof, lsato}@cs.uoregon.edu

## ABSTRACT

Though musical performers routinely use eye movements to communicate with each other during musical performances, very few performers or composers have used eye tracking devices to direct musical compositions and performances. EyeMusic is a system that uses eye movements as an input to electronic music compositions. The eye movements can directly control the music, or the music can respond to the eyes moving around a visual scene. EyeMusic is implemented so that any composer using established composition software can incorporate prerecorded eye movement data into their musical compositions.

## KEYWORDS

Electronic music composition, eye movements, eye tracking, human-computer interaction, Max/MSP.

## 1. INTRODUCTION

An *eye tracker* is a device that reports where a person is looking. Eye trackers usually incorporate a camera that sits next to a computer screen 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.

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 [1] provide a good overview of eye tracking research in human-computer interaction.

This report discusses EyeMusic, a system that transforms eye movement data into musical compositions and data sonifications. The system enables musicians and performers to control a composition, as it is being performed, with their eyes.

## 2. HOW THE EYES MOVE

To understand how EyeMusic works, a few terms pertaining to eye movements and eye tracking must be defined. The *gaze* is the vector that goes from the eye to the *gaze point*, which is the point in a scene where a person is looking. The eyes (and thus the gaze) move

around a scene with a series of quick jumps called *saccades*, each of which lasts roughly 30 ms. Between saccades, the eyes (and the gaze point) 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 shift from higher-level strategy execution and 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.

## 3. WHY MAKE MUSIC WITH EYE MOVEMENTS

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.

## 4. PREVIOUS MUSIC WITH THE EYES

We are aware of only one body of previous work in which eye movements direct musical compositions, work by the digital artist Andrea Polli [2]. Her musical composition with eye tracking entitled *Intuitive Ocusonics* has been performed internationally. Excerpts are online at <[www.andreapolli.com](http://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 identifies fixations, which is a highly relevant psychological phenomenon. The resulting compositions and data sonifications relate more directly 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 eyes move around 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 gaze point superimposed on the scene viewed by the performer. Professor Polli is currently collaborating with us to compose using EyeMusic.

## 5. HOW EYEMUSIC WORKS

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 from 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. The screen coordinates (as well as the pupil radius) are delivered to Max/MSP at this rate. Eyegaze runs on a Microsoft Windows computer.

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.

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.

The eye movement data are made available in Max/MSP by means of the *eyedata2* external and parsed into fixation data by means of the *fixation* external, both of which were written by the authors and are discussed below in greater detail.

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.

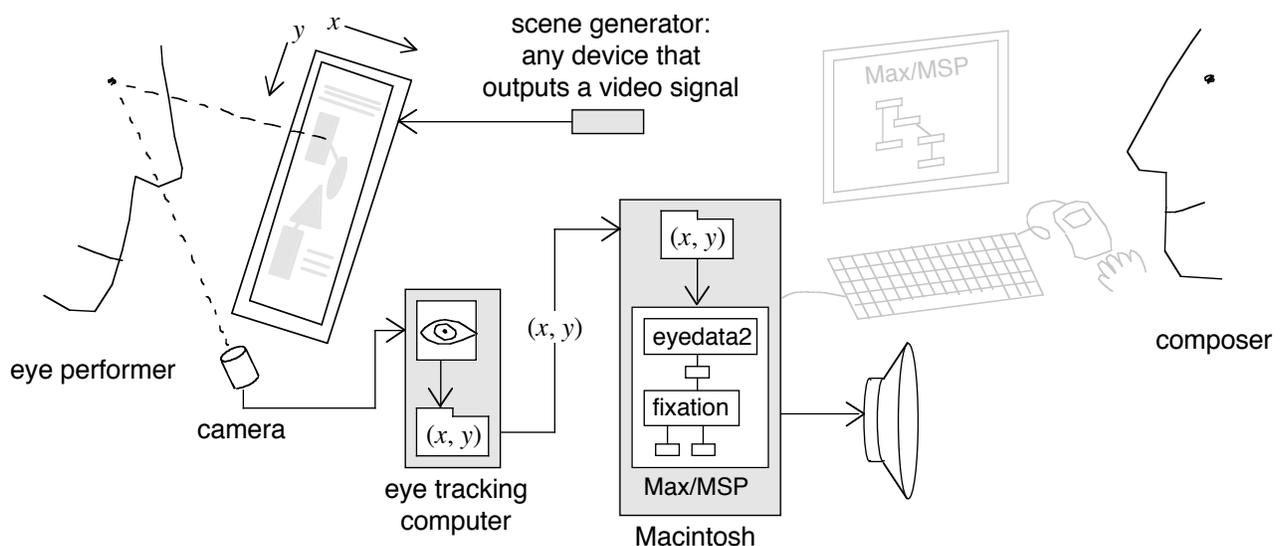


Figure 1. An overview of the major components in the EyeMusic system. Arrows indicate the flow of data.

EyeMusic works in two different modes: *playback* and *performance*. In playback mode, prerecorded eye movement data are read from disk 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.

### 5.1. The Eyedata2 External in Max/MSP

The basic building blocks in Max/MSP are data processing elements called *objects*. End-users can add customized objects called *externals* using the C programming language. In EyeMusic, eye movement data are read into the Max environment by means of an external called *eyedata2*. Figure 2 shows the eyedata2 external in use.

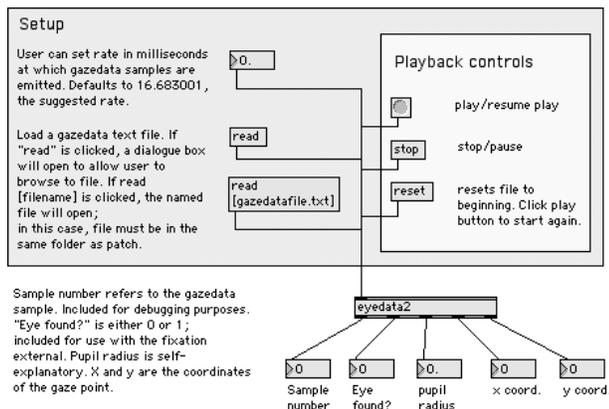


Figure 2. The eyedata2 external in Max/MSP

The eyedata2 external 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).

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
...	...	...	...	...

Table 1. Sample data from the eyedata2 external.

### 5.2. Video Playback

In playback mode, the *eyedata2* external outputs the recorded gaze position as (x, y) coordinates. This is adequate to play the musical composition but not for the composer or the audience to observe a correlation between the eye movements and the pitches created by the eye movements. For example, the composition might be designed to play a bass note when the gaze lands on a red blob, but there will be no way for the audience to see this happen. To address this problem, we record a video during the eye tracking session that shows the eye-performer's gaze point superimposed on the scene that he or she is viewing. (The video mixer and recorder are not shown in Figure 1.) The video is then played back within Max/MSP simultaneously with the data output from *eyedata2*. The composer and audience members can thus see that the bass note is played when the gaze moves to the red blob.

### 5.3. The Fixation External in Max/MSP

The fact that the data flow out of the *eyedata* external at a rate of sixty samples per second is an artifact of the eye tracker used. The relevant human physiological phenomenon is not this sample rate, but where and when the fixations occur. The *fixation* external parses the eye movement sample data, identifies where the fixations occur, and outputs each fixation and its location. This way, a composer can work directly with the relevant human phenomenon. Figure 3 shows the fixation external in use.

The fixation external uses an established *dispersion-based* algorithm [3]. 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

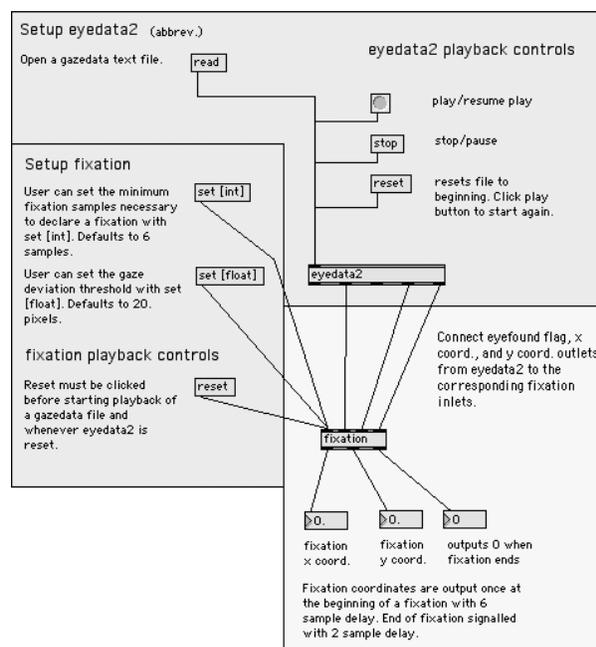


Figure 3. The fixation external in Max/MSP

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 external 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.

## 6. RECORDINGS AND PERFORMANCES

Jeffrey Stolet, our collaborator in the School of Music, created an electronic composition using EyeMusic entitled “EyeMusic v. 0.9b” that was performed (in playback mode) at a Future Music Oregon concert on November 15, 2003. The composition uses the eyedata2 external to play back Stolet’s eye movements that were recorded as he studied Kandinsky paintings at an earlier eye movement recording session. A video of the fixation point superimposed on the stimuli was played during the performance. An audio-video recording of “EyeMusic v. 0.9b” can be viewed on the EyeMusic web site. The piece is a little under six minutes long.

In “EyeMusic v. 0.9b,” the eye movement data are output from the eyedata2 object at a rate of sixty samples per second and selectively sampled at slower rates within Max/MSP. The rate is roughly once every 500 ms but, for musical purposes, varies slightly during the piece and ritards at the end. “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 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 composition “EyeMusic v. 0.9b,” described above, captures one possible mapping of the eye movement data to sound. There are infinite other possibilities.

The fixation external was developed after “EyeMusic v. 0.9b” was composed. To demonstrate how the external works, the authors composed a composition 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 a text. 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 bell and carriage return sound are played. The fixations sound remarkably like typing because the fixations occur at roughly the same rate that keys are pressed while typing, about four per second, and because the varied rhythm of the fixations resembles the rhythm of typing. An audio-video recording of “Reading/Typing” is available on the EyeMusic web site.

The EyeMusic externals, documentation, eye movement data, and audio-visual recordings can be downloaded at <[www.cs.uoregon.edu/research/cm-hci/EyeMusic/](http://www.cs.uoregon.edu/research/cm-hci/EyeMusic/)>. An eye tracker is not needed for playback mode.

## 7. FUTURE WORK

EyeMusic will be developed on a number of different fronts. We are exploring possibilities with performance mode, to see how a musician can control a musical passage with his or her eyes. We will continue to collaborate with Jeffrey Stolet and Andrea Polli to develop innovative musical compositions.

We will 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.

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. Making music with the eyes would follow nicely.

## ACKNOWLEDGMENTS

The authors thank Jeffrey Stolet for creative collaboration and technical support and Tim Halverson for assistance with the eye tracking. This work is supported by the National Science Foundation under Grant No. IIS-0308244. The eye tracking lab used for these projects is supported by the Office of Naval Research through Grant N00014-02-10440.

## REFERENCES

1. Jacob, R. J. K., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises (Section commentary). In J. Hyona, R. Radach, & H. Deubel (Eds.), *The Mind's Eyes: Cognitive and Applied Aspects of Eye Movements*. Oxford: Elsevier Science.
2. Polli, A. (1999). Active vision: Controlling sound with eye movements. *Leonardo*, 32(5), 405-411.
3. Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye-tracking protocol. *Proceedings of the Eye Tracking Research and Applications Symposium*, New York: ACM Press, 71-78.