

7-6-2007

## Oculog: Playing with Eye Movements

Juno Kim

*University of Wollongong*, [juno@uow.edu.au](mailto:juno@uow.edu.au)

Greg Schiemer

*University of Wollongong*, [schiemer@uow.edu.au](mailto:schiemer@uow.edu.au)

Terumi Narushima

*University of Wollongong*, [terumi@uow.edu.au](mailto:terumi@uow.edu.au)

Follow this and additional works at: <https://ro.uow.edu.au/creartspapers>



Part of the [Arts and Humanities Commons](#), and the [Social and Behavioral Sciences Commons](#)

---

### Recommended Citation

Kim, Juno; Schiemer, Greg; and Narushima, Terumi: Oculog: Playing with Eye Movements 2007.  
<https://ro.uow.edu.au/creartspapers/11>

# Oculog: Playing with Eye Movements

Juno Kim  
Department of Psychology  
Faculty of Health and Behavioural  
Sciences  
University of Wollongong, Australia  
+61 2 4221 5450  
oculog@juno3d.com

Greg Schiemer  
Sonic Arts Research Network  
Faculty of Creative Arts  
University of Wollongong, Australia  
+61 2 4221 3584  
schiemer@uow.edu.au

Terumi Narushima  
Sonic Arts Research Network  
Faculty of Creative Arts  
University of Wollongong, Australia  
+61 2 4221 5450  
terumi.narushima@gmail.com

## ABSTRACT

In this paper, we describe the musical development of a new system for performing electronic music where a video-based eye movement recording system, known as Oculog, is used to control sound. Its development is discussed against a background that includes a brief history of biologically based interfaces for performing music, together with a survey of various recording systems currently in use for monitoring eye movement in clinical applications. Oculog is discussed with specific reference to its implementation as a performance interface for electronic music. A new work features algorithms driven by eye movement response and allows the user to interact with audio synthesis and introduces new possibilities for microtonal performance. Discussion reflects an earlier technological paradigm and concludes by reviewing possibilities for future development.

## Keywords

Eye movement recording, video, MIDI, algorithmic composition, Pure Data, microtonal tuning.

## 1. EXPRESSIVE CONTROL INTERFACES

Musical expression has conventionally been dependent on and conditioned by the muscular responses of vocalists or performers playing musical instruments that are usually held or touched. In recent times the conventional concept of expressive control has been expanded to embrace electronically sensed choreographed movement. Examples of this include early electronic instruments such as the theremin and terpsitone [2], or more recently developed systems such as The Hands [19] or SensorLab [18]. Expressive control offered by these electronic instruments requires a new kind of musicianship to develop a more deterministic relationship between the choreographic action of the performer and the musical outcomes produced.

Eye movement has been an important expressive feature of choreography in certain dance traditions such as south Indian Kathikali and Bharat Natyam, and these traditions played a role in the early electronic dance experiments of Philippa Cullen, a choreographer and collaborator with the second author [6]. The ability to monitor eye movement offers electronic performers an alternative kind of expressive control. Eye movements can be

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

*Nime '07*, June 07-09, 2007, New York University, USA.

Copyright remains with the author(s).

many times faster than the choreographic movements involved in playing standard musical instruments. Expressive control of music using eye movement responses involves finding a balance between voluntary deterministic movements and involuntary autonomic movements.

In the 1960s, analogue technology made it easier for experimental composers such as Lucier, Rosenboom, Teitelbaum and others to create music using biological signals [14]. The use of biofeedback techniques was sometimes enlisted to bring about greater expressive control. One example of this is the control of eye movement using amplified biological signals such as electrocardiogram to activate flash lamps in various positions [3]. Digital technology has gradually given the sophistication required to process neural signals in order to create and control more complex music [14].

*Gape* by Andrea Polli, performed in 1997, was one of the first works in which eye movement response was used to control digital sampled sounds [12]. The principal focus of her creative approach was the use of saccades, or abrupt, rapid eye movements which occur when the eyes fix on one point after another in the visual field. More recently, researchers at the EyeMusic project have developed a digital system that allows a performer to play music using fixation-detection algorithms [4]. This resulted in a public performance of *EyeMusic v1.0* by Troy Rogers presented by the composer at the annual SEAMUS conference at the University of Oregon in March 2006.

## 2. EYE MOVEMENT DETECTION AND RECORDING TECHNIQUES

### 2.1 Electrophysiological Methods

A common electrophysiological technique for measuring eye movements in the clinic involves recording difference potentials generated between electrodes placed either side of the corneoretinally polarised eye, a method also commonly known as EOG, or electro-oculogram [20]. EOG systems provide an indication of unidirectional saccadic and tonic eye movement along an axis from one electrode to the other, usually horizontal. The corneoretinal potential is not robust but has been found to vary over time, and is affected by background activation of eye muscles [5]. Although EMG (electro-myography) recording of electrical activity from eye muscles may be more reliable, it is difficult to obtain information about eye movement in more than one direction.

### 2.2 Infrared Reflectance Methods

Infrared reflectance systems are known to have inherent non-linearities, but one advantage is that they are not influenced by

background electromuscular noise as with EOG systems. Additionally, eye movements can be detected in both vertical and horizontal directions with greater reliability than EOG. While infrared reflectance and EOG offer high sampling rates (>5kHz), they do not provide an effective indication of torsional eye rotation.

### 2.3 Magnetic Search Coils

Extremely accurate systems for recording three-dimensional changes in eye position have been developed using systems based on magnetic search coil techniques [13]. Typical scleral search coil systems are extremely fast, scanning at rates of up to 100kHz. Such systems require a coil to be placed on the locally anaesthetized surface of the eye in order to relay voltages generated by a surrounding cube magnetic field.

Search coil systems have three principle disadvantages: the mobility of the user is restricted in that the head must be restrained from moving while the system is in use (unless a head coil is used to measure offset of the eye coil in the magnetic field); its use involves quite invasive procedures; and it relies on hardware that is relatively expensive (ca. US\$40,000). These problems make it more suitable for a laboratory environment rather than for use in performance.

### 2.4 Video-based Eye Movement Recording

Increasingly over recent years, video eye movement recording systems have offered a solution to the limitations of non-video methods of eye movement recording. New high-speed digital cameras allow three-dimensional changes in eye position to be reliably determined in real-time at rates upward of 250 frames per second (fps). The major limitation of these systems is the throughput of current digital camera technology and the need for greater intensity of infrared illumination to allow adequate passing of light from the eye to the CMOS sensor with each open-close cycle of the camera shutter.

Video eye movement recording systems allow reliable tracking of the pupil in both the horizontal and vertical directions as well as tracking of the iris as it rotates torsionally around the optic axis [9, 23]. Three-dimensional planar rotations are determined in terms of their angular displacement from an arbitrary reference vector computed from an initial reference bitmap image at the time of hardware initialization and calibration. The reference vector is obtained from the position of the pupil within the bitmap image and the angular offset of the iral signature (the spectral arrangement of striations around the pupil of the eye unique to each individual). This reference acts in much the same way a detent position provides a mechanical default position for the pitch wheel on a MIDI synthesizer.

## 3. TECHNICAL IMPLEMENTATION

The system described in this paper is a modified version of the Oculog Recorder software <<http://www.oculog.com>>, an ongoing research and development project of the first author [7]. It is a three-dimensional eye movement recording utility to track the pupil and iral signature in real-time during digital video acquisition. Tracking can be configured to perform at rates of up to 120fps from images partially scanned over a hardware area consisting of 320x240 8-bit monochrome pixels. In the current implementation a slower rate of 30fps was used in order to reduce

the processing cost associated with running eye movement analysis.

Digital video images were acquired in real-time using a Fire-i IEEE-1394 (Firewire) board camera assembly (Unibrain). This was secured to a pair of modified snow goggles fitted with an aluminium insert as shown in Figure 1.



Figure 1. User wearing the eye camera.

A terminal from an infrared LED emitter was soldered in contact with the input from the 12V Firewire bus-powered terminal of the 6-pin Firewire cable originating from the host machine. The other terminal of the LED was soldered to the end of an 820 Ohm resistor, the opposite end of which was soldered to the ground pin on the 6-pin Firewire cable. Alternatively, an isolated battery power supply can be used. The effective constant-current power driving the LED in either case was regulated below 50mA, resulting in an intensity of infrared emission within suggested safety limits [9].

The modified camera goggles were contoured to the face with a soft foam contact seal and may be further secured using an adjustable elastic head band. When the goggles were positioned on the wearer, the side-mounted camera was adjusted so that it captured a view of the eye on an infrared reflective hot mirror. The bus-powered infrared LED emitter was mounted on a flexible wire guide attached to the frame of the goggles. The eye was bathed in infrared illumination emitted by the LED, as shown in Figure 2. A small piece of infrared-pass film was fitted over the back of a 16mm lens to block natural light wavelengths. This prevented contamination from natural room light, thus allowing the camera to capture a stable infrared view of the eye.

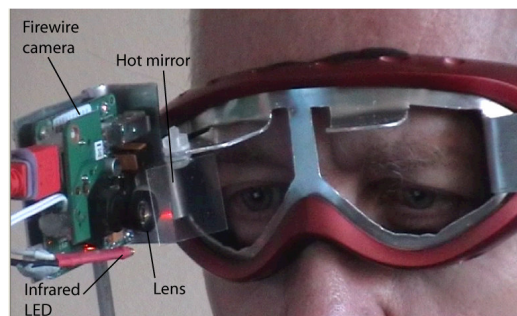
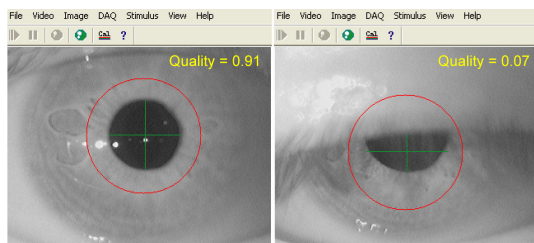


Figure 2. Eye camera detail.

The system was calibrated for each user to establish the radius of the eye in pixels. This was achieved by measuring the change in horizontal position of the pupil's centre (in pixels) over a known angular change in horizontal eye position [9]. Once the radius of the eye was known, horizontal eye position in subsequent video images could then be determined using simple trigonometric

functions. Real-world video aspect ratio in millimeters (measured using a ruler) was taken into account when determining vertical angular eye position [7]. The intensities of pixels falling along the points forming a circle (an annulus) on the bitmap image at a configurable radius around the pupil were stored in memory for subsequent correlation in order to determine the precise torsional eye position in a given video image. Geometric compensation was used to deform the sampling annulus in order to closely fit the camera's perspective view of the ellipsoid iris during horizontal and vertical eye rotation [9].

In order to register blink activity of the eye reliably, the iral signature was stored in memory to provide a constant reference as the image changes. The level of similarity between the initial and current iral signature is high when the current video image closely matches the initial image, and low when comparison is degraded, primarily due to partial or complete covering of the iris by the eyelid (see Figure 3).



**Figure 3. Image analysis of iral signature: high quality (left), low quality (right, blink).**

In total, our eye movement recording system can provide up to five channels of information that characterize various attributes of eye movement: horizontal position, vertical position, torsional position, blinking, and pupil size. Control of these attributes can be either voluntary or involuntary: horizontal and vertical eye positions tend to be voluntary, while torsional eye movement and pupil size tend to be involuntary. Blinking may be either voluntary or involuntary. Combinations of these physiological characteristics are algorithmically encoded and mapped to MIDI messages using embedded calls to routines provided in STK [1].

We chose to focus on the transmission of three of these characteristics as MIDI data. Horizontal eye position was mapped to note number and vertical eye position was mapped to key velocity. Torsional eye movement was reserved for varying the delay between note-ON and the subsequent note-OFF message: clockwise torsional eye movement lengthens the delay while counter-clockwise torsional eye movement shortens it. Alternatively, torsion may be implemented by transmitting a MIDI Controller Change message. Currently the delay between note-ON

and note-OFF is a preset value. The fourth characteristic, blinking, is used to trigger the transmission of MIDI data packets.

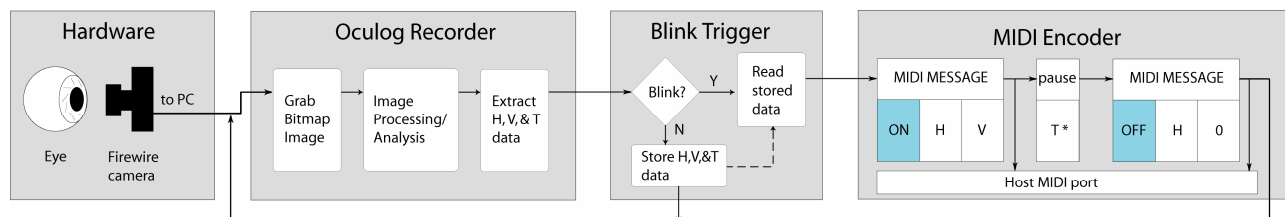
A vital feature of the mapping algorithm is the ability to trigger MIDI events. For example, when a blink is registered and the torsional quality threshold is no longer exceeded, eye position information obtained immediately prior to the blink is encoded as a MIDI note-ON message. MIDI messages extracted in this way are then sent to the default MIDI port on the host PC.

This relationship between eye movement processing and MIDI is summarized in the flowchart shown in Figure 4. The first block on the left is the Firewire video acquisition hardware which is mounted on the goggles. The second block is the Oculog Recorder software. The third block is the blink detection algorithm triggered on a thresholded reduction in iral signature quality. The fourth block is the MIDI encoding of eye camera data. It shows a MIDI note-ON message followed by a delay and finally a MIDI note-OFF message. In the case of the MIDI note-ON message, horizontal and vertical eye positions are transmitted as the 1<sup>st</sup> and 2<sup>nd</sup> data bytes respectively; in the MIDI note-OFF message, horizontal eye position is transmitted as the 1<sup>st</sup> data byte followed by 0. Transmitting 0 for the 2<sup>nd</sup> data byte complies with the MIDI protocol for note-OFF. Finally, the transmission of the MIDI note-OFF message is also a cue that synchronizes scanning and storing of image data in the first block.

MIDI was chosen for its compact message packet size and its applicability across a broad range of electronic music applications. The MIDI note-ON message was used to transmit horizontal and vertical position coordinates in a single message packet triggered by the blink detection algorithm. 7-bit precision available in MIDI data bytes supports the degree of precision achievable by performers using position coordinates.

MIDI was used to communicate between the Oculog Recorder software and the musical application, both running on the same processor (AMD Athlon 64 3800+ CPU 2.41GHz). On slower machines the operations associated with eye movement processing compromised the performance of the musical application, usually resulting in the break up of audio. Under this circumstance the problem was overcome by running each application on separate machines interconnected using external MIDI communication.

Currently a pair of eye position coordinates is sent from Oculog to the target musical application as a single message packet using a MIDI cable. In future we envisage a wireless version of this connection in which MIDI is used as a sub-layer to transmit eye movement coordinates to a music program running on a remote machine. Doing this is an efficient way to address problems of communications bandwidth associated with wireless networks.



**Figure 4. Overview of Oculog eye movement system.**



Horizontal eye position sent to the Pd patch as a MIDI note number interacts with the cyclic sequence to vary the melodic contour further. This is achieved by storing the received note number in the first counter of the cyclic sequence generator thereby resetting the sequence and making it jump forwards or backwards to a different point in the cycle. This gives the melodic contour a characteristic of saccadic eye movement, proceeding in a manner that may appear to be random and non-sequential, yet nevertheless coherent. This variation algorithm extends techniques for interaction between a live performer and an automated procedure developed by the second author between 1986 and 1993 [16]. In the context of *Saccadic Variations*, this algorithm embraces an aspect of noise that one might expect from a biological performance interface.

### 4.4 Chorusing

The cyclic sequences play melodically by controlling the pitch of four tone generators. Each generator consists of five sine wave oscillators configured in the manner of Risset’s ‘chorusing’ instrument. The Pd patch in Figure 9 shows one tone generator consisting of a centre frequency (310Hz) flanked on either side by two higher tones (311.8 and 310.9Hz) and two lower tones (309.0 and 308.1Hz). Each tone generator uses a different set of tuning offsets to achieve other rates of chorusing than the ones used here.

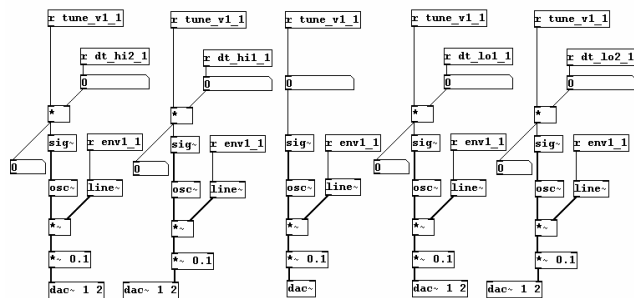


Figure 9. Pd patch for chorusing.

In addition to fluctuations in amplitude that result from chorusing, amplitude envelopes were also used to shape the sound for each note event. In Figure 8, the output of mod 3 was used to select three different envelope shapes and in a separate embedded Pd object (not shown) one of three preset attack and decay times was selected. In Figure 9, these parameters enter the patch via the five ‘r env1’ objects and, using the ‘\*~’ and ‘line~’ objects, vary the output levels of each oscillator over time.

### 4.5 Microtonal Tuning

The process of auralizing eye movement responses in *Saccadic Variations* is part of our ongoing investigation of just intonation tunings. The MIDI protocol allowed the eye movement system to control music. However MIDI was originally intended to support music composed only in 12 equal divisions of the octave.

We decided that a new gestural interface should not be limited by such musical assumptions. The degree of control a performer may exercise manually is significantly different from the kind of control people exercise using eye movement, a view supported by Hornof, et al [4]. The key question is: what kind of music might be played on the new interface that cannot be played on any other

interface? Would the new interface, for example, allow us to play microtonal music that cannot be performed on standard instruments? Both the second and third authors have produced microtonal works in which presentation format has been restricted by the lack of a suitable performance interface [17, 10]. Even though Oculog uses MIDI for communicating information related to eye movement, we saw it as potentially usable for playing microtonal music. This was synthesized in Pd.

For *Saccadic Variations*, a table of pitch data was used to implement microtonal tuning. The system of tuning used is a 15-note just intonation scale called a pentadekany, one of a genre of tunings called Combination Product Sets (CPS) developed by contemporary theorist Erv Wilson [22]. Table 1 shows a pentadekany used in *Saccadic Variations*. This was generated using harmonics 1, 3, 5, 7, 11 and 13. Degrees of the scale are formed by multiplying combinations of pairs of these harmonics, reducing them so they fall within the range of one octave and sorting them in ascending order. The table shows the pitches of the pentadekany, their just relationship with the implied tonic, the generators and their historical interval names. Because a harmonic is never multiplied by itself in a CPS scale, such a scale has the interesting musical property of not having a tonic, yet many of the intervals sound consonant. Tuning is implemented within the ‘pd pitches1’ object in Figure 8 where vectors produced by the cyclic code generator are used to access the table.

Table 1. 15-note scale used in *Saccadic Variations*.

Pitch	Ratio	Generators				Cents	Historic interval name
0	1/1					0.000	unison, perfect prime
1	65/64		5		13	26.841	13th-partial chroma
2	33/32	3			11	53.273	undecimal comma, al-Farabi's 1/4-tone
3	35/32		5	7		155.140	septimal neutral second
4	143/128				11 13	191.846	
5	77/64			7	11	320.144	
6	39/32	3			13	342.463	39th harmonic, Zalzal wosta of Ibn Sina
7	5/4	1	5			386.314	major third
8	21/16	3	7			470.781	narrow fourth
9	11/8	1			11	551.318	undecimal semi-augmented fourth
10	91/64			7	13	609.354	
11	3/2	1	3			701.955	perfect fifth
12	13/8	1			13	840.528	tridecimal neutral sixth
13	55/32		5	11		937.632	
14	7/4	1		7		968.826	harmonic seventh
15	15/8	3	5			1088.269	classic major seventh
16	2/1					1200.000	octave

## 5. CONCLUSION

This work is ongoing and has yet to be fully evaluated. The impetus for this work has been the eye movement detection system developed by the first author. In this collaboration we have taken this system from the clinic into the concert arena. Our musical experimentation so far has been limited to vertical and horizontal eye position combined with blink detection. In the process we have encountered issues of control and noise similar to those reported by the EyeMusic team [4]. The interest in saccadic eye movement expressed by Polli [12] also resonates with our approach although there are noticeable differences, the most obvious being that she uses sampled sounds whereas we synthesize live audio.

The feature of our eye movement system that perhaps best distinguishes it from systems developed by others is its ability to monitor torsion. This offers the possibility of monitoring



involuntary torsional eye movements such as those that occur when observers move their heads from side to side while gazing at a fixed point of reference. Even though the Oculog Recorder software already monitors torsion, we have not yet implemented a musical application of this feature. This will be the next step in our project. Wireless connectivity offered by new developments in Ultra-Wideband technology indicates that the torsion control scenario is now a practical reality [11]. If Wibree technology fulfils the expectations of developers in the Ultra-Wideband community, smart spectacles with embedded optical sensors may become a commonly used item of wearable computing [21].

Finally it is also worth noting that the eye movement system in its current implementation is monoscopic. By adding a second camera, stereoscopic eye movement detection can be achieved. This potentially offers a system with two receptors that monitor eye movement independent of neural activity. Whether this can be a useful musical feature will only be determined once musicians have played with it.

## 6. REFERENCES

- [1] Cook P. R. and Scavone, G. P. The Synthesis ToolKit (STK). In *Proceedings of the International Computer Music Conference* (Beijing, China), 1999.
- [2] Galayev, B. M. Light and Shadows of a Great Life: In Commemoration of the One-Hundredth Anniversary of the Birth of Leon Theremin, Pioneer of Electronic Art. *Leonardo Music Journal*, 6 (1996), 45-48.
- [3] Eaton, M. L. Bio Music. *Source*, 5, 1 (1971), 28-36.
- [4] Hornof, A. J., Rogers, T. and Halverson, T. EyeMusic: Performing Live Music and Multimedia Compositions with Eye Movements. Submitted to ACM CHI 2007 <<http://www.cs.uoregon.edu/~hornof/publications.html>>.
- [5] Iwasaki, S., McGarvie, L. A., Halmagyi, G. M., Burgess A. M., Kim, J., Colebatch, J. G. and Curthoys, I. S. Head taps evoke a crossed vestibulo-ocular reflex. *Neurology*, accepted 4 December 2006.
- [6] Jones, S. Philippa Cullen: Dancing the Music. *Leonardo Music Journal*, 14 (2004), 64-73.
- [7] Kim, J. A simple pupil-independent method for recording eye movements in rodents using video. *Journal of Neuroscience Methods*, 138 (2004), 165-171.
- [8] Kumar, A. and Krol, G. Binocular infrared oculography. *Laryngoscope*, 102, 4 (1992), 367-378.
- [9] Moore, S. T., Haslwanter, T., Curthoys I. S. and Smith S. T. A geometric basis for measurement of three-dimensional eye position using image processing. *Vision Research*, 36, 3 (1996), 445-459.
- [10] Narushima, T. *Composing for Carillon: Exploring the Relationship Between Tuning and Timbre*. M.Mus. Thesis, University of Sydney, NSW, Australia, 2003.
- [11] Nekoogar, F. *Ultra-Wideband Communications: Fundamentals and Applications*. Prentice Hall, Upper Saddle River, NJ, 2005.
- [12] Polli, A. Active Vision: Controlling Sound with Eye Movements. *Leonardo*, 32, 5 (1999), 405-411.
- [13] Robinson, D. A. A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Trans Biomed Engineering*, 10 (1963), 137-145.
- [14] Rosenboom, D. The Performing Brain. *Computer Music Journal*, 14, 1 (Spring 1990), 48-66.
- [15] Schiemer, G. *MIDI Tool Box: An Interactive System for Music Composition*. Ph.D. Thesis, Macquarie University, NSW, Australia, 1998.
- [16] Schiemer, G. Improvising Machines: Spectral Dance and Token Objects. *Leonardo Music Journal*, 9 (1999), 107-114.
- [17] Schiemer, G. *Tempered Dekanies: Chorus Effect Using Microtonal Intervals Based On Just Intonation*. In *Proceedings of 7th International Conference for Music Perception and Cognition*, Causal Productions, Adelaide July 2002, 300-303.
- [18] Tanaka, A. Musical Technical Issues in Using Interactive Instrument Technology. In *Proceedings of the International Computer Music Conference*. International Computer Music Association, San Francisco, 1993, 124-126.
- [19] Waiswicz, M. THE HANDS: A Set of Remote MIDI Controllers. In *Proceeding s of the 1985 International Computer Music Conference*. Computer Music Association, San Francisco, 1985, 313-318.
- [20] Werner, J. A method for arousal-free and continuous measurement of the depth of sleep in man with the aid of electroencephalo-, electrooculo- and electrocardiography (EEG, EOG, and EKG). *Z Gesamte Exp. Med.*, 134 (1961), 187-209.
- [21] Wibree. <<http://www.wibree.com/>>, Oct 3 2006.
- [22] Wilson, Erv. D'alessandro, Like a Hurricane. *Xenharmonikôn*, 12 (1989), 1-38.
- [23] Zhu, D., Moore, S. T. and Raphan, T. Robust and real-time torsional eye position calculation using a template-matching technique. *Comput. Methods Programs Biomed.*, 74, 3 (2004), 201-209.