EyeCon -- a motion sensing tool for creating interactive dance, music and video projections

Robert Wechsler  
Doncaster College School of Intermedia and Performance Arts  
Doncaster, England  
robert@palindrome.de

Frieder Weiß  
Palindrome Inter-media Performance Group  
Nürnberg, Germany  
frieder@palindrome.de

Peter Dowling  
Independent Composer  
Glasgow University  
Glasgow, England  
pdowling@fish.co.uk

Abstract

EyeCon is a video-based motion sensing system which allows performers to generate or control music and projected images through their movements and gestures in space. It was developed Frieder Weiß as an application-driven project of the Palindrome inter-media performance group and is thus more a tool for artists than for researchers. EyeCon does, however, allow motion sensing according to a variety of movement parameters and lends itself to experimentation with media mapping.

Unique and largely unexplored problems face composers, graphic artists and choreographers as they collaborate on interactive performance works, not the least of which is settling on schemes for mapping the various parameters of human movement to those within the world of sound and image. Among the myriad mapping alternatives, good design choices are paramount to creating effective interactive performance works. The authors have made progress in understanding the special issues involved and are devising strategies for these choices.

1 Introduction

The authors' work is an investigation into the perceptual relationships of human motion to sound, and, to a lesser extent, those of human motion to video projection art in which the motion of the performer determines or influences the secondary medium. Through nine years of collaborative work creating interactive performances, the authors are able to draw certain conclusions concerning the nature of interactive performing, i.e. what makes it artistically tenable and what does not. Some of these conclusions relate to the choices of media, some to the process of collaboration and still others to the choices of mapping. These are summarized in the form of tips for performers wanting to work with interactive systems.

Mapping is the process of connecting one data port to another, somewhat like the early telephone operator patch bays. In our case mapping has a very specific connotation—it means the applying of a given gestural data, obtained via a sensor system, to the control of a given sound or video synthesis parameter. The dramatic effectiveness of a dance, however, invariably depends on myriad factors—movement dynamics of body parts and torso, movement in space, location on stage, direction of focus, use of weight, muscle tension, and so on. And although sensors may be available to detect all of these parameters, the question remains: which ones to apply in a given setting, and then to which of the equally numerous musical or visual parameters they should be linked.

2 Motion Sensing and Analysis

The primary motion-sensing system the authors use is the EyeCon software. It is a camera-based motion sensing system developed by the two principle researchers for this project, Robert Wechsler and Frieder Weiß, especially for stage and installation art work as an on-going project of the Palindrome Inter-Media Performance. The first version of EyeCon appeared in 1995.

EyeCon permits movement to control or generate sounds, music, text, stage lighting or projectable art. It is adaptable to an enormous number of applications, lending itself to experimentation in genuinely new directions in performing and installation art. It offers a way to create interactive video environments without the need to get into graphical or script-based programming. This means, that people without special computer skills are able to use it.
Although EyeCon has been used by a number of dance and theater companies, singers and performance artists, Palindrome is its dominant user. For a overview of some of the work Palindrome has created with EyeCon (including video samples) visit Palindrome's homepage at www.palindrome.de.

2.1 Terminology

The terms motion tracking, motion capture, motion recognition, motion analysis and motion sensing are used variously and with overlapping applications. Motion capture is a technique developed by the film industry for creating more realistic movement of animated characters and while the technology has found numerous other uses, it has generally not been possible or practical to use it in live performance settings. The data collection process involves cameras completely surrounding the performance area and the performers must wear highly noticeable reflective markers. Even if a performer were willing to perform with these accoutrements, the fact remains that motion capture systems have, until very recently, not been capable of rendering the data into graphic images in real time. The frist major are, however, now underway to do this by Paul Kaiser, Trisha Brown, Bebe Miller and others at Arizona State University's "Intelligent Stage" lab, as a part of the "motone" project.1

The term motion tracking also fails to hit the mark for the authors' work. Or, better said, it would seem to imply only limited aspects of human motion, namely where a person is located on the stage and the speed and direction they are traveling. This information is not only limited in its expressive potential -- audiences are much more concerned with what a performer is doing than where he or she is located -- but it is also a particularly poor choice of parameter for reasons of transparency. Because a performer cannot jump instantly from one part of the stage to another, the transitions involved are slow. This makes the interactivity far less convincing since it could easily be simulated (faked, as in a technician following the performers motion with mouse movements). Indeed, it may simply appear that the performer is following the media in instead of the other way around!

2.2 How the Motion sensing Takes Place

A video signal is fed into the computer and the video image appears on the computer screen. Using the mouse, lines or fields in different colors are superimposed onto the video image. When a performer, through their video image, touches one of these lines, or moves within a field, a media event is triggered or modulated, for example, a certain sound might be heard. Eyecon works by comparing the individual pixels of two different video frames and analyzing them for differences in brightness or color. The difference between these two frames is time. That is, they are the same scene, but one is the present time, and one in the past. In some cases, there may be only 0.04 seconds between the two, but when applied to a body part in motion, this is easily sufficient time to allow the determination of motion. Generally speaking, not all of the pixels in the two images are compared, but rather only those marked off by EyeCon's elements.

2.3 EyeCon's Sensing Elements

EyeCon provides an assortment of different ways to sense motion. To understand the need for this diversity, let us look for a moment to the expressive potential of movement. Human movement in general, and dance in particular, is perceived as a collection of movement features -- parameters if you will. That is, human movement is not one quantity, but rather a great many, which are weighted very differently in our perception depending on their artistic intent. In one case, it may be overall speed of the mover which speaks to us, while in another, we are concerned with the precise positioning of the limbs. Thus if a motion sensing system is to be of use to artists, it must be capable of sensing a variety of different movement parameters. Without this palate, and the skill to use it effectively, the dance becomes a slave to the system. To be effective, the movement must be choreographed to fit the technology, instead of the other way around.

The systems by which EyeCon senses movement are represented through graphic structures called Elements. The Elements are superimposed onto the live video image in the Video Window where they can be scaled and manipulated. The Elements are then assigned their individual properties such as the volume control of a particular sound or changes in a projected image.

TOUCHLINES. Touchlines are lines drawn on the video image in the computer, these act as triggers for the presence or absence of body parts or objects, but can also be scaled, so that different places along the line can have different effects. In this sense they provide an easy way to track position along a line.

DYNAMIC FIELDS. Dynamic fields are boxes which represent fields and respond to movement dynamic. They can be used to trigger sounds and images or to control the volume and pitch of sound files so that, for example, the faster the dancer moves, the louder the sound or the higher the pitch.

FEATURE FIELDS. Feature fields begin with the same boxes as Dynamic Fields, but provide formal analysis of objects within the field. For example, a Feature field might measure the dancer's overall size (how expanded or contracted they are), or how close two dancers are to each other. They can also analyze shape (width compared to height), height or degree of left-to-right symmetry. Finally, there are a set of controls for sensing the direction of movement, so that a step to the left, for example, will sound different than a step to the right, reaching up different than reaching down, etc.

1 http://ame2.asu.edu/projects/motone/.
**Position Trackers.** Position trackers track the mean position of one or multiple persons as they move around the video image. This means, if you have an overhead camera, you can track the location of persons as they move within a room and thus the environment can be made to respond to each person differently. Touchlines and Dynamic fields can be attached to the tracker, so that a given array of controllers can move with the performer as they move around the space. Finally, this feature permits color-specific tracking, presenting the possibility of distinguishing between dancers by the color of their costumes.

EyeCon is often used to control secondary programs either running on the same machine as EyeCon, in another computer via data link or network. The additional computer does not, of course, need to be a PC. External software and hardware which can be controlled by EyeCon include software and hardware synthesizers, as well as programs like Director, MAX/msp, Reactor, and Isadora. The data link can be via Ethernet and use MIDI or OSC protocols.

### 3 The EyeCon User Interface

Figure 1 shows the user interface of EyeCon 1.60.

[Image of the EyeCon user interface]

In contrast to many interaction-oriented digital media systems, EyeCon is not freely programmable, i.e. does not work like a programming language. Rather it is based on the fixed architectural elements described above. Rather than creating entire environments with highly accurate and specialized properties, EyeCon is conceived more in terms of offering dynamic control -- both spatially (movable shapeable elements) and temporally (sequencer functions). The result is a relatively intuitive system for performers and by its nature, relatively easy to operate. Artists with little or no computer experience can use it.

The EyeCon software can be divided in two completely different parts: motion sensing and multimedia action(s). The concept of EyeCon is that you create motion sensing elements which are in many cases graphically represented as lines and rectangular fields. The Element Editor is the main mapping panel where you assign how movement controls media.

EyeCon has two operating modes, Simulation and RUN mode. In RUN mode that actual video analysis is performed and media is played. In simulation mode you can use tiny boxes and move them with the mouse to emulate moving objects in the video image. The Control window allows you to control the main parameters of the video processing like camera selection, threshold for movement detection, basic timing etc.

The Video window allows you to watch the live video and see how the motion sensing elements react.

To allow the creation of complex interactive setups, EyeCon offers a scene management function. Each sensing element is functional in a selectable level. The Sequencer Window allows you to arrange these interactive setup levels.

### 4 Interaction as an Artistic Phenomenon

Regarding interactive performance works, one could speak of four levels of interactivity: One is the purely technical level, that is, if the system successfully accomplishes the conversion of one media to another (for example, if a dancer's movement does indeed generate a sound). The second level is whether the performers are aware that their movements are affecting their environment. The third level comes when a system functions in such a way that the audience is aware of the interaction (without having it pointed out). The fourth level is when the audience members themselves cause changes to occur in media, for example, through their movements in their seats.

Some have argued that a successful interactive performance work requires only that the second level interaction be reached. The reasoning is that since the performer is aware of the effect they are having on their environment, then this will invariably affect the way they perform. And thus the audience, even with no clear understanding of what is happening, nevertheless feels the art work in a different way. While believing that this point has some merit, the authors of this paper rather see this as a cop out. If an art work and its technology are prepared carefully enough, then the interactivity can be accessible to an audience directly and this the greatest artistic potential.

#### 4.1 The Psychology of Interaction

There is another aspect to interaction which, in a sense,
supercedes the question of who, from a technical or perceptual standpoint, interacts with whom. Like wolves and other primates, humans are an extremely interactive species. They tend to clump together in groups, spending inordinate amounts of time speaking, gesturing, touching and otherwise communicating with one another. Its what we are not doing now. If you were to write us back a letter, we may touch on this quality, but we don't really start interacting until we sit down together and hash it out. That we do this with such relish belongs to the most primitive of human instincts.

Human beings have been dancing and making music for 10,000 years. During most of this long history, performances were highly interactive -- much more so than they are today. The distinctions of "performer" and "audience", and even those of "musician", "dancer", etc. were far less clear. Everyone was part of the event. There are still today examples in Africa of traditions for which the same word is used for both dance and music\(^2\). Participants fed off of each other's energy in a way which is seen today only in such settings as night clubs (the good ones). Jazz music provides perhaps a last bastion of interactive performing in the West.

Beginning with the predominance of the bourgeoisie, theater in Europe saw a closing off of interaction between performer and audience. With bright lights on one side of a proscenium, and a darken area with seats on the other, the audience's role was pretty much reduced to sitting quietly and then clapping before going home. This has not changed much in the last two hundred years.

Ironically perhaps, modern technology is a major culprit. Recording and sampling techniques have meant that musicians, for example, often work separately. Pop music relies heavily on sampling ("stealing with respect", as it is known among DJs) rather than creating sounds from scratch or playing musical instruments.

Dancers and musicians, meanwhile, rarely work directly with one another anymore. Only a tiny fraction of dances performed today directly involve musicians in any part of the actual stage production: creation, rehearsal or performance.

But the biggest interactivity-buster of all is surely the projection screen. Not only did video further reduce the need for dancers and composers to work together creatively, but of course one doesn’t even need to be part of an audience today to watch a performance. Just turn on the television.

Palindrome has a piece called Publikumsstück ("audience piece") in which 10 audience members are brought back-stage during the intermission. There, they are taught ingredients for a structured improvisation -- essentially given tasks to accomplish with one anther -- as well as some examples of how interaction works. It is very much a "feeling thing" -- a subjective, rather than objective phenomenon. In some cases, small amounts of participation by an audience can utterly change their experience of a performance event. On the other side, giving the audience many things to do may have little effect on their sense of interactivity. I.e. their sense that they were part of the event depends on special factors.

Thus, whether we are speaking of interaction between artists, between artist and audience or between a person and a computer system, the same basic principles apply. They share psychological roots and in practice function in a similar way. In all cases, it is dependent on the performer being relaxed enough in their role to be able to respond genuinely, in a sense innocently, to what they are experiencing. Note, this may and may not involve improvisation on the part of the performer, at least not in the sense that the word is generally used by dancers and musicians. There must however be some degree of play in the performance. Each time must be different. If a performance is utterly fixed, there can be no interaction. It is the paradox of all good performing; it should look and feel spontaneous, even when it is carefully prepared.

5 Two Current Motion Tracking Based Projects Of The Authors

Two current and very different examples of applications are presented (with live demonstration). The first is an excerpt of a motion-to-sound piece, and the second, an excerpt from a motion-to-projected video image. These and similar examples are available at www.palindrome.de.

5.1 "ICE 9" (2003)

The real-time sound synthesis environment was designed in MAX/MSP. A PC running EyeCon is linked to a Macintosh PowerBook running MAX/MSP, sending the gestural data gathered by EyeCon to the real-time sound synthesis parameters.

The MAX/MSP program for "ICE 9", is a musical synthesis environment that provides many control parameters, addressing a number of custom-built DSP modules that include granular sampling/synthesis, additive synthesis, spectral filtering, etc. All mapping is accomplished within the MAX/MSP environment, and changes throughout the work.

\(^1\) The Awa of the Dogon tribes (Guinea Coast) offers one example of many.

\(^2\) The Awa of the Dogon tribes (Guinea Coast) offers one example of many.
Control of the musical score to "Ice 9" is accomplished through a cue list that enables/disables various EyeCon movement sensing parameters, mapping and DSP modules to be implemented centrally. Both EyeCon and MAX/MSP software components are organized as a series of "scenes", each describing a unique configuration of video tracking, mapping, and DSP. Scene changes for both computers are synchronized and can be initiated by a single keystroke from either station.

"Ice 9" is a music-dance performance/research work that applies choreographic evaluation methods and motion sensing technology to music composition issues. Of particular interest are certain qualities which may be applied to both music and dance and the application of these to multi-modal expression. For example, we have developed a motion sensing (motion tracking) system which responds to the direction of the dancers' motion. That is, whether the impulse, or overall movement tendency is to the left (vs. right), upwards (vs. down-wards), and downstage (vs. upstage). In this way, 3 fundamental bi-polar tendencies can be identified and, importantly, easily perceived by the observer. Each of these directional parameters are mapped to a different bi-modal acoustic model which is applied to the real-time generation of sound.

The application represents the essential paradigm for us. Needless to say, the music will be influenced and inspired by the process and style in both gesture and in the transparent mapping of movement to music. However, the composer’s role must not only be seen as concentrating on the composition of sound sources and the live processing of triggered materials (and of course their creative implementation with relevance to the choreographic situation), but also as determining ways in which the form of implementation -- how we go about doing this -- can be creative in its own right. The selection of mappings, for example, is in itself a complex issue; choices must be made between multifarious possibilities with the criteria for selection being of a cross-disciplinary, and largely unexplored nature, i.e. involving both parameters and expression of human movement (dance) as well as those of sound (music). It must be understood that these choices are neither cosmetic nor trivial, but are crucial to the transparency, ergo the effectiveness of the complete sound/movement experience.

In short, the interactive (and technological) process is on a par with the composing of the sounds and the creation of the choreography when thinking conceptually and artistically about the invention of the piece. Our performing experiences have shown that, generally speaking, when only part of a piece is transparent and convincing in its interactive relationships, then audiences tend to accept additional more complex relationships. They become 'attuned', as it were, to the functionality of the piece.

This instills in us a wish to negate the obvious and the transitory notion of mimetic dialogue. If we take electro-acoustic musical discourse as being constituent of aural discourse – abstract musical content – and mimetic discourse – a complex of auditory, visual and emotional stimuli, we have a basis for multi-layering perceptions.

For example, timbral mimesis is the direct imitation of the timbre of a sound, whereas syntactic mimesis is the imitation of relationships between natural events, like the orchestration of speech rhythms (which Palindrome have approached creatively on numerous occasions in the past). In composition (and to an extent in choreography), both [aural] discourse and mimesis are always present in some form – a continuum exists between the two. So, in aural discourse we can always extract ‘pure’ musical elements, even from directly recorded natural sounds. In mimetic discourse we can always perceive (or imagine) some source of or – importantly in this collaborative case – cause of the sound(s).

To cross-reference these concepts with the interdisciplinary issues at stake in mixing the dance (the visual stimulus of movement) with the music is essential. They are refracted through the interface of the technology and the processes of artist and audience interaction. In this way, there is generated a fascination with the creative technique as well the more obvious potential for multi-layering of perception generated from work of this kind.

5.2 "Ich, mich und mir" (2004)

The work applies the age old theater technique of a shadow play, only here it is combined with digital media - specifically, an infrared light source, an infrared camera, motion sensing and real time image and audio signal processing.

What would happen if we could release our shadow image, follow it like in a dream, amplify it or face it?

Since the beginning of human consciousness, people have known their "virtual" companion: the body shadow. It follows us quietly. You can neither catch it, nor step across it. And while belonging to our body, it is in no way a part of it. It is always darker than ourselves. Some cultures believe it belongs to an emotional world. The shadow was thought to be the home of the soul. Who ever does not have a shadow is regarded as dead.

After many years of a hype about the virtual experience, most of us are left with disillusionment. We are learning that we can’t live in a world that is disconnected from our physical and emotional realities.

Our performance piece is meant as a reminder of the organic connection between body-image and body-reality. Our theme is the shifting border between body and mediated virtual body image.

Figure 2 shows a moment in the dance.
5.2.1 The Technology

Our shadows gain freedom from their source in the following way: In one corner of the space we put a light source which is throwing the shadow image on the large projection screen. This shadow, however, is of a type which the human eye cannot see. Because all of the visible light has been filtered from the light, only the infrared light is reaching the screen. A special infrared camera picks up the shadow image from the screen and feeds the digitized image into the computer where it is processed. A connected video projector makes the processed images visible, in fact they are projected on to the exact surface where the invisible shadow is located. (Video projector light, being low in infrared, does not interfere with the shadow we are filming).

The digital processing includes continuously variable delays, multiplication, transposition, coloring, reversing, accelerating/decelerating, freezing and dissolving.

Brain research has found that what humans experience as ‘now’ is actually a time band of up to three seconds. Our brain is constantly trying to sync our different senses, making predictions about what is most likely to happen, and then integrate the whole into a perception we call ‘now’. How much can a shadow be delayed before we loose the sense that it is still "connected" to its owner? The computer shadows seamlessly shift between what we are used to and the unexpected. It is this play along the borderline between the known and the surprising makes the piece fascinating to watch.

6 Conclusion: Some Specific Suggestions for Interactive Performers

- Performers must be relaxed and open enough on stage so that their performance can be informed by the influence they are having on their environment.
- Map to multiple outputs. For example, you may wish to link a movement to a particular sound as well as to a visual element such as a stage lighting change or video projection element. Although this may seem obscure the correlation, because the mappings are parallel, it will have the have effect of reinforcing the connection.
- Think about camera angles. Choose one which helps the movements to be accurate and repeatable.
- The dancers (or moving performers) should tell the choreographers or technicians what they feel they need, instead of the other way around.
- Link media events to particular gestures. The way a movement becomes marked in the mind of the audience. I.e. use memorable, movements, those with character, even though technically there may be no advantage to doing it that way.
- Trigger or control the same events from the same stage positions or from the same body posture even though, again, this may be irrelevant to the system you are using.
- Look for intuitive mappings (higher body level-to-higher pitch, faster-to-louder, busier movement-to-busier sound, heavier movement-to-heavier sound, etc.). Artists tend to have great reluctance to do this for reasons which are not entirely clear to us.
- Near the beginning of the piece, or at least during the piece, use the system in a clear and transparent way. In this way, it can explain itself to the viewer. Having done this, the audience becomes sensitized to the interactive experience. They will then be attuned to and accepting of later, subtler mappings.
- Either before, or after the piece, explain to the audience what the technology is and how it works. There are as many good reasons to do this as there are not to, but it is an option. Some pieces don't need it, some don't want it, and others simply love it. Either way, whether you do it or not, I will guarantee you one thing: someone will come to you after the show and thank you from the bottom their heart for doing it, just as the next person in line will castigate you for the same thing.
- Program notes are not nearly as communicative as an announcement, yet may be far less intrusive to a work of art.

Acknowledgements

The authors would like to thank the dancers and choreographers with whom the works described in this paper were created: Emily Fernandez, Helena Zwiauer and Larry Hahn. Important background work concerning DSP was conducted together with Butch Rovan.
References


