

2007

-- online since 1993

Home > Journal > Essays > LEA-ACM Special

ISSN NO : 1071 - 4391

The MIT Press

QUICK LINKS : [LEA E-JOURNAL](#)[TABLE OF CONTENTS](#)[INTRODUCTION](#)[EDITOR'S NOTE](#)[GUEST EDITORIAL](#)[ESSAYS](#)[ANNOUNCEMENTS](#)[:: GALLERY](#)[:: RESOURCES](#)[:: ARCHIVE](#)[:: ABOUT](#)[:: CALL FOR PAPERS](#)**Human Movement and Clear Affordances Promote Social Interaction**v o l 1 5
i s s u e
05

SEARCH LEA

 GO

Advanced Search

natra haniff

editor, leonardo electronic almanac

http://lea.mit.edu natra [at] leoalmanac [dot] org\n

-----\n\n",0)); D(["ce"]); //-->

Human Movement and Clear Affordances Promote Social Interaction[Click here to download pdf version.](#) (429kb)

by Andrew Webb

awebb [at] cs [dot] tamu [dot] .edu

Andruid Kerne

andruid [at] cs [dot] tamu [dot] .edu

Eunye Koh

eunye [at] cs [dot] tamu [dot] .edu

Interface Ecology Lab

Computer Science Department

Texas A&M University

College Station

TX 77843

U.S.A.

Keywords

choreographic buttons, computer-supported cooperative play, social interaction, gesture, movement, Laban notation, iterative design, mappings, affordance

Abstract

This research connects methods from computer vision, human computer interaction, choreography, design, and art. We used human movement as the basis for designing an embodied collaborative aesthetic design environment. Our intention was to promote social interaction and creative expression. We employed off-the-shelf computer vision technology. Movement became the basis for the choreography and recognition of gestures and the development of imagery and visualization. We developed a new type of affordance, the choreographic button, which integrates choreography, gesture recognition, and visual feedback. Jumping, a quick movement, and crouching, a sustained gesture, were choreographed to form a vocabulary that is personally expressive, and which also facilitates automatic recognition.

For evaluation, we held an integrated exhibition, party, and user study event. This mixing of events produced an engaging environment in which participants could choose to interact with each other. The study results demonstrate that our movement-based design environment promotes social interaction.

Introduction

Research on human computer interaction involving computer vision-based motion tracking needs to start with the body, and then consider technology. The body is the site of experience. We sense the environment, form understandings, and actuate responses. Movement is an actuated essence of people's everyday experiences. Movements articulate locomotion, and also convey information regarding emotions and intentions. A gesture is a linguistic form of movement that communicates without the use of vocal articulation or written language. Thus, gesture is social. Gesture can be utilized to

develop expressive vocabularies that form a basis for the interaction between humans and machines. Choreography is the creation or specification of movement in order to affect a specific intention and convey meaning. In this research, human movement serves as the basis for human computer interaction, social interaction, and visual imagery. Expression through movement forms the basis of an interactive environment in which participants collaborate to create design.

LEA - ACM**Morsels of
Multimedia
for the Taking**ACM Multimedia
Interactive
Arts Program
Special**Unyazi****Welcome the year
with a new groove**Check out LEA's
Supplement to
Leonardo
Music Journal 16

We develop a collaborative aesthetic design environment with a movement-based interface. The movement-based design environment consists of five structural components: (1) superimposed choreographic button duos; (2) a human Choreography Grid and a visual Imagery Grid; (3) a Movement Imagery Collection; (4) temporal visual structures; and (5) interactive affordances.

This article begins by presenting the design process of the five components. Then, we will develop an evaluation method that mixes a social event with an experimental study. We close with discussion about the role of the body and movement in human computer interaction and the application of principles of interaction design in movement-based interfaces.

Choreography and Recognition

Our design process focused on exploring possibilities for human expression, and the social function and utility of video motion tracking technologies [0]. We utilized the Max signal processing and integration environment, along with the Jitter suite of video processing objects, and the Cyclops video analysis plug-in [0]. A program in this environment is known as a Max/Jitter "patch." We developed gesture-recognition algorithms as a Max/Jitter patch.

We sought to choreograph gestures that were on the one hand expressive, and also not difficult to develop reliable recognition algorithms for. As we developed prototypes iteratively, we discovered that it is important to graphically convey the state of the gesture recognition process. Otherwise, the participant cannot clearly understand what state the environment is in, and what interactive movements make sense. Further, the physical space in which the participant can effect the environment also needs to be clearly defined, in order to make the possibilities for interaction clear. Norman associates with the term affordance, "the perceived and actual properties of a thing ... that determine how it can be used" [0]. Thus, we developed a tight binding between the demarcation of physical space, gesture and its recognition, and graphical representations of state. We call this affordance the choreographic button. Designing a choreographic button involves three interwoven stages: (1) defining (choreographing) and recognizing the movement; (2) designing graphical indicators that represent the state of the movement; and (3) mapping changes in the state of the movement to changes in the graphical affordances and the system.

Choreographing Recognizable Gestures

Laban provides a structural approach to understanding movement and choreography [0]. He developed six elementary schematic structures of effort in correlation with the six fundamental directions in space: up and down, left and right, backward and forward [0]. Based on Laban's schematics of effort, we define a choreographic button duo: one button uses a quick upward jumping movement, while the other uses sustained downward crouching. Each was chosen because of the simplicity provided in computationally monitoring single-axis unidirectional movement (up and down). We recognize jumping movements with a single wide-angle video camera, and crouching movements with another such camera. Modern computational resources enabled our Max/Jitter patch to apply commodity recognition algorithms to both video streams concurrently in real time using a single computer.

These two cameras are located in the corner to the northeast of the active choreographic area of physical space, providing a wide-angle profile view (Figure 1a, right). Camera A is used to recognize jumping is located at floor level and provides a view of the feet (Figure 1b). By placing the camera parallel with the plane of the floor, only one algorithm is necessary to

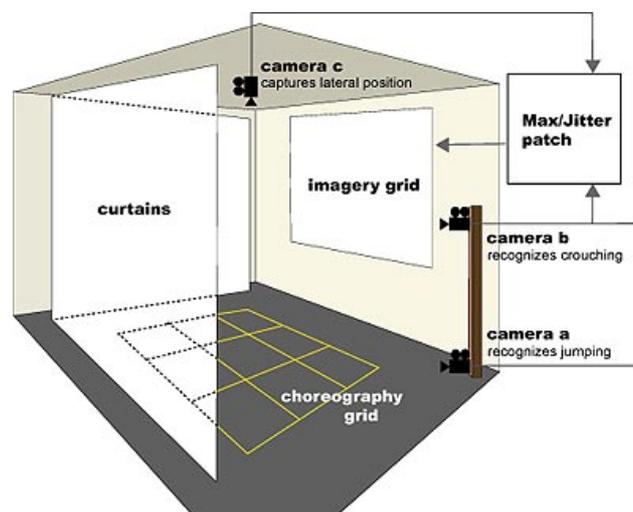
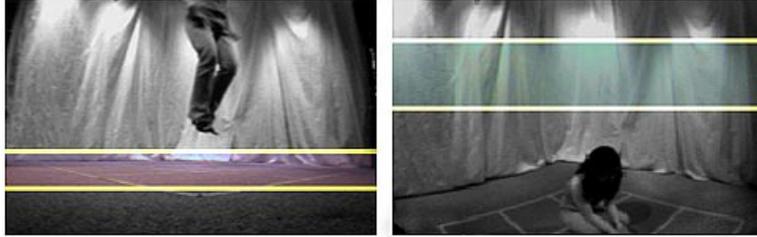


Figure 1: System Diagram with camera views
(a) System Diagram: Choreography Grid, Imagery Grid, and their relationships to the space and technology
 Copyright © Andrew Webb, Andruid Kerne and Eunye Koh

recognize jumping, because only a single rectangular area (junction between the floor and wall) needs to be analyzed regardless of the position of the feet along the floor. Camera B, which recognizes crouching, is located directly above and functions similarly to the floor camera, but with a view of the upper torso (Figure 1c). A curtain running just outside the south and west edges of the active choreographic area provides a static backdrop to reduce noise for motion tracking.



b) camera a view, recognizes jumping

(c) camera b view, recognizes crouching

Copyright © Andrew Webb, Andruid Kerne and Eunye Koh

This recognition mechanism is sufficient to support a single choreographic button duo corresponding to a horizontal area that spans the range of the cameras. Lateral spatial resolution requires employing a third camera to track the position of participants. The camera (Camera C) that differentiates lateral positions is placed overhead, facing the floor, orthogonal to the other cameras. Lateral space can be discretized. Upon subdividing the feed from the overhead camera into nine discrete regions (cells) arranged in a 3x3 grid, a choreographic button duo is assigned to each region. By putting together the profile and overhead views, we are able to recognize distinct movement forms, in association with specific lateral positions.

A Tale of 2 Grids

The nine choreographic button duos of our movement-based design environment are manifested not just in physical space, but also in an associated visual space. The grid in the physical space, where movement recognition takes place, is called the Choreography Grid, while the projected space of visualization and feedback is the Imagery Grid (Figure 2, below). The two grids are characterized by a direct one-to-one positional mapping of cells. The top-left corner of the Choreography Grid maps to the top-left corner of the Imagery Grid.

The boundaries for the cells of the Choreography Grid are marked out on the floor of the physical space with yellow lines. The Imagery Grid (Figure 2) projected on the forward facing wall is the arrangement of the imagery states for the choreographic button duos.



Figure 2: Imagery Grid Example. The bottom center square is selected and locked.

Copyright © Andrew Webb, Andruid Kerne and Eunyee Koh

The 3x3 grid structure provides a collage creation structure with degrees of freedom designed to support an experience that is aesthetically evocative and creatively open.

Movement Imagery Collection

Visual representations rendered in each cell of the Imagery Grid are based on the Movement Imagery Collection, a set of 29 images that we chose in order to represent forms of movement in a physical or emotional sense, developing a movement vocabulary based on the gestures of the choreographic buttons: jumping and crouching.

Temporal Visual Structures

We designed three temporal visual structures for selecting elements from the Movement Imagery Collection and rendering them: cross-fading, fast-forwarding, and still. Each of these structures consists of rules for selecting an element, an option for compositing with the previous element, and an option for sequencing over time.

The first and default structure is cross-fading. When a grid cell is cross-fading, the visualization continuously transitions between pairs of images from the Movement Imagery Collection. Each pair consists of a source image and a destination image. Each successive destination image is selected randomly. Cross-fading progressively renders source and destination images with inverse levels of translucence, as per these equations:

Following each cross-fade, there is a five-second pause before the next cross-fade occurs with a new destination image. The pause provides a participant with time to discern individual images in the collection, as well as their blended states.

At initialization, every cell in the Imagery Grid is invoking the cross-fading temporal structure. The initialization times of the cross-fading structures in different cells are offset, in order to desynchronize the fades. This conveys a sense of mutual independence, rather than synchrony.

The second temporal visual structure is fast-forwarding. When this structure is invoked, the imagery transitions between discrete images from the Movement Imagery Collection in a fixed sequence. There is no compositing. A pause between transitions allows the participant time to observe each successive image, and make decisions about whether to select it with a movement.

The simplest structure is still. This is the identity element. In the still structure, the imagery visualization is fixed to a single image from the Movement Imagery Collection. The still temporal visual structure does not transition between images.

Affordances and Mappings

We choreographed the gestures of quick jumping and sustained crouching as the basis for each button duo. Each duo is associated with a square cell of physical space in the Choreography Grid, and a corresponding cell in the projected visual space of the Imagery Grid. In the physical space of the Choreography Grid, while the choreographic buttons of a duo are superimposed

laterally, their positions are differentiated along the vertical axis. Figure 3 (below) defines the mappings by specifying a state diagram of associated movements, affordances, temporal visual structures, and transition logics.

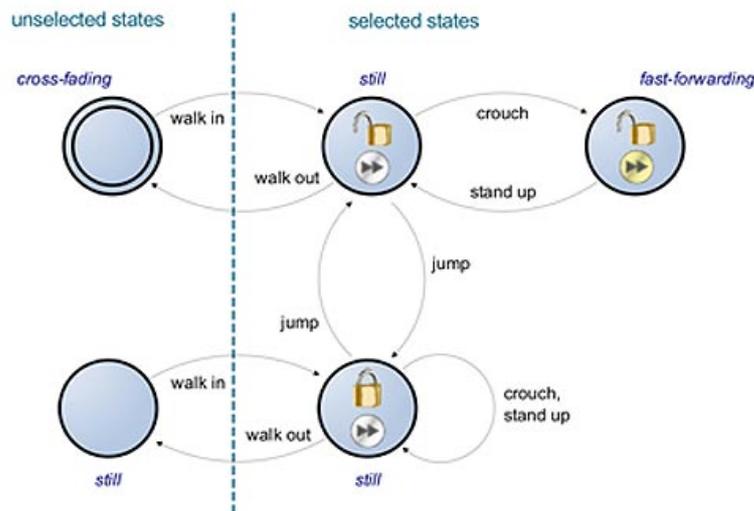


Figure 3: State Diagram for a Choreographic Button Duo. The state in the top left outlined by doublelines is the initial state for all duos.
Copyright © Andrew Webb, Andruid Kerne and Eunyee Koh

Selection of a choreographic button duo occurs when a participant moves laterally into a particular Choreography Grid cell. This is indicated in the visual space of the Imagery Grid by a yellow selection border that frames the corresponding cell.

Selection also results in the presentation of indicators reflecting the state of each button in the duo. The button state icons are positioned vertically, corresponding to the movements required to trigger each of them. The jump button state icon is located in a top corner since upward movement triggers it. The crouch button state icon is located in a bottom corner because downward movement triggers it.

In developing the design for iconic visual feedback and mappings to temporal visual structures once a choreographic button duo is selected, we were guided by physical constraints and the characteristics of movement. Jumping is quick and sudden, and thus well suited for a momentary state-changing operation.

Locking is a toggle operation that transpires instantaneously. Thus, jumping is mapped to the toggle of a lock state. Locking is indicated by a padlock that is open when not activated or closed when activated. Once it has been locked, a choreographic button duo invokes the still temporal structure, whether it is actively selected by the participant, or not. Locking also inhibits the affects of crouching. Crouching is a sustained gesture that can last for an indefinite amount of time. When a button duo is unlocked, crouching is mapped to the sustained selection of the fast-forwarding temporal visual structure. The fast-forwarding icon is a circular button with double forward arrows that is silver when not activated and yellow when activated.

After interacting with the button duo within a Choreography Grid cell, the participant will eventually move laterally. The cell becomes unselected, and the yellow selection border is removed in the Imagery Grid. When a grid cell enters the unselected state, it either invokes the cross-fading temporal visual structure if it is unlocked or the still structure if it is locked.

The interactive semantics of locking allows a person to create a collage that is entirely static (all cells locked) or a heterogeneous dynamic composition in which some cells sustain cross-fading, while others are locked to still. This enables the participant to create patterns (see Figure 2). Changes made in the grids persist even after a person leaves the design environment space. A subsequent participant can enter the space and make changes to the previous creation, consequently making the new collage a collaborative work between the current participant and prior participants.

Evaluation

We need to discover methods for evaluating interactive systems with playful and social design intentions. We deployed the choreographic buttons design environment in the context of the exhibition / party / user study event with the intention of provoking physical activity and social interaction, and developing understanding of emerging practices of use. Amalgamating these

three different types of events allowed us to use HCI evaluation techniques in a realistic (semi-controlled) environment while giving the participant the opportunity for an experience less like that of a typical experimental subject, and more like a participant at a social art event.

In order to structure the evaluation process, we wanted to be able to compare two “usage” scenarios. Thus, we constructed a mouse-based version of the aesthetic design environment. In order to select a square, a participant places the mouse cursor over the desired square. Locking/unlocking a square entails clicking the padlock icon for that square. Fast-forwarding a square is accomplished by holding the mouse button down over the fast-forward icon and releasing the mouse button when complete.

Participants voluntarily took responsibility for engaging in the structured activities of using the two systems, and answering the questionnaire.

Results

Nearly all the participants (18) found the movement-based version more entertaining to watch other participants use [$\chi^2(1) = 15.21, p < 0.001$]. Most (15) felt that the movement-based version involved more social interaction. This result shows that the participants were more socially involved and interacted more with each other using the movement-based version than the mouse-based version [$\chi^2(1) = 8.00, p < 0.01$] (see Figure 4). Participants who selected the movement-based version said that it involved movement of their entire body, produced good exercise, or was fun to play with, like a game.

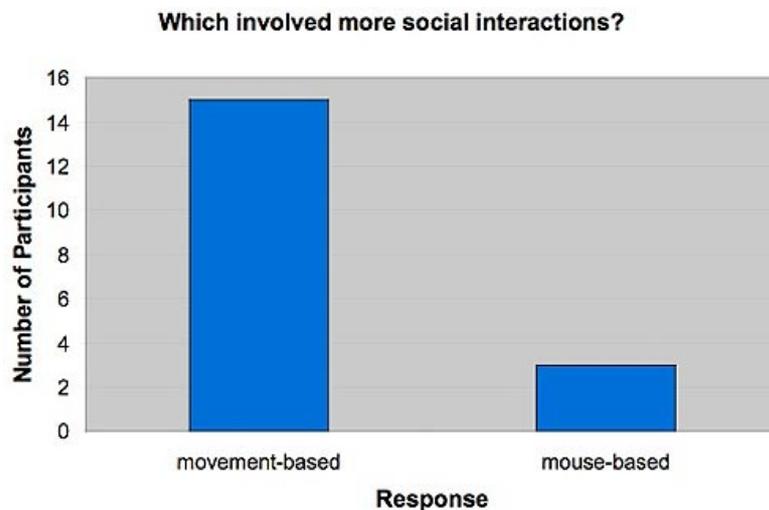


Figure 4: Quantitative Experiential Results
Copyright © Andrew Webb, Andruid Kerne and Eunye Koh

The results showed that computer-supported cooperative play in the movement-based design environment was reinforced through the system’s collaborative and ludic nature as expressed by these participants.

“It’s funny to see people jumping and concentrated in looking at the screen. It seems more natural to talk and interrupt people in the movement-based than in the mouse-based.”

“Moving is very different. This difference causes interesting questions to arise. It feels easier to question and comment about someone who is rigid and socially stiff. The vulnerability of the composer makes initiating conversation easier.”

Discussion

The experience of human computer interaction should be no less than other experiences in human life. As Merleau-Ponty develops, the sensation is the unit of experience, and the body is the site of sensation [0]. We humans are physical creatures. Thus, to bring human computer interaction onto a par with the rest of life, we need to discover new modalities that directly involve the body. Choreography is the field that has developed knowledge of expressive human movement. In framing this research, we set out to involve principles of choreography, of human movement, of effort and shape, in interface and experience design.

Computer-vision based motion-tracking technology served as the basis for recognizing human movement. Expressive movements and imagery were co-designed with gesture recognition algorithms. The role of technology was essential, yet not central in the interface design process. Further, while we used choreographic principles to design movement, we also needed interface design principles to develop a coherent experience for participants. The need for perceptible affordances, consistency, feedback, and clear mappings was if anything intensified,

rather than obviated, by the involvement of computer vision. Choreographic buttons were found to produce a clear, intelligible user experience. The association of movement and visual state feedback was effective.

One of our primary goals for the choreographic buttons system was to promote human to human social interaction. This manifests a value system in which developing human social relationships is considered worthwhile. As the theater director and performance studies scholar Richard Schechner said, "Process — a term used often in environmental theater — means 'getting there' rather than 'getting there' emphasis on the doing, not the done... to be alive to the here and now, to express oneself here and now. What an immense risk that is! Those who love products value things and make things of all living beings. Those who love process value living and make living beings of all things" [0]. This research represents an activity-based rather than task-based, a process rather than product oriented goal structure.

We observed social interaction. One participant would create a clear pattern, using repetition and symmetry (i.e., Figure 2). Subsequent participants would create related patterns. Social interaction developed through two stages. The grids afforded the creation of patterns; the movements and the display afforded human to human social interaction, in which the grid-based patterns functioned as vocabulary. The result was the emergence of design style. Compared to a mouse-based interface, choreographic buttons promote social engagement. Non-artists made art together. Non-dancers got involved in moving in public. The pure physicality of the interaction and the exchange of roles in reciprocal turn taking are factors. Clear affordance design contributed to the efficacy of the movement-based interface.

References

Cycling '74. Cycling '74: Max/MSP for Mac and Windows, <http://www.cycling74.com/products/maxmsp.html> (2004).

Laban, R., *The Language of Movement: A Guidebook to Choreutics*, (Plays Inc., First American edition, 1974).

Merleau-Ponty, M., *The Phenomenology of Perception* (New York: Routledge, 1995).

Moeslund, T.B., Granum, E., "A Survey of Computer Vision-Based Human Motion Capture", *Computer Vision and Image Understanding* Vol. 81 Issue 231-268 (2001).

Norman, D., *Design of Everyday Things*, (Basic Books, 2002).

Schechner, R., *Environmental Theater*, (NY: Applause, 1994).

Author Biographies

Andrew Webb is a M.S. student in Computer Science at Texas A&M University. He is a member of the Interface Ecology Lab [<http://ecologylab.cs.tamu.edu>], where he researches human computer interaction. His research at the Interface Ecology Lab employs human-centered computing principles to the design of creative and expressive interactive systems. His areas of interest include interface design, graphic design, computer vision, and information visualization. He is involved in the development of *combinFormation*, a mixed-initiative creativity support tool that integrates processes of searching, browsing, collecting, mixing, organizing, and thinking about information. Andrew holds a B.S. in Computer Science from Texas A&M University. He has published papers in ACM Multimedia and Joint Conference on Digital Libraries (JCDL).

Andruid Kerne is an assistant professor of Computer Science at Texas A&M University, where he directs the Interface Ecology Lab [<http://ecologylab.cs.tamu.edu>] and the Perceptive Sensor Networks Lab, and serves in the Center for the Study of Digital Libraries. The Interface Ecology Lab investigates human-centered computing support for expression, creativity, and social engagement. Andruid is the principal investigator of *combinFormation*, a mixed-initiative creativity support tool that integrates the everyday activities of browsing, searching, and collecting information. He holds a B.A. in applied mathematics / electronic media from Harvard, an M.A. in music composition from Wesleyan, and a Ph.D. in computer science from NYU. Kerne's output has been presented by the Guggenheim Museum (New York), ACM SIGGCHI, SIGGRAPH, Multimedia and Document Engineering, New York Digital Salon (New York, Spain, London, Beijing), ISEA (Paris, San Jose), Computational Semiotics in Games and New Media (COSIGN), the Milia New Talent Competition (Cannes), the Ars Electronica Center (Linz), the Boston Cyber Arts Festival, the Pan-African Theater Festival (Ghana), and the town square of the village of Anyako (Ghana). The National Science Foundation, the Rockefeller Foundation, Dance Theater Workshop, the Spaulding-Potter Fund for Innovative Education, and the Texas A&M Department of Computer Science and Humanities Informatics Initiative have supported his work.

Eunyeek Koh is a Ph.D. candidate in the computer science department in Texas A&M University. Her research area includes human computer interaction and information retrieval. Her recent research work focuses on development and evaluation of the information system named *combinFormation*. The research results in publishing papers in Joint Conference on Digital Libraries (JCDL), European Conference on Digital Libraries (ECDL), ACM Multimedia, ACM

Computer Human Interaction, SIGIR, and ACM Document Engineering. She was a reviewer for IEEE transactions on Human-Centered Computing and ACM Computer Human Interaction. In addition, she is actively involved in volunteering and mentoring work in the women in computer science group, and also she was a student volunteer in various conferences such as the JCDL, ACM CHI, and Grace Hopper Conferences.



Citation reference for this Leonardo Electronic Almanac Essay

MLA Style

Webb, Andrew; Kerna, Andruid; Koh, Eunyee. "Human Movement and Clear Affordances Promote Social Interaction." "LEA-ACM Multimedia" Special Issue, Leonardo Electronic Almanac Vol. 15, No. 5 - 6 (2007). 10 May 2007 <http://leoalmanac.org/journal/vol_15/lea_v15_n05-06/WebbKerneKoh.asp>.

APA Style

Webb, Andrew; Kerna, Andruid; Koh, Eunyee. (Apr. 2007) "Human Movement and Clear Affordances Promote Social Interaction," "LEA-ACM Multimedia" Special Issue, Leonardo Electronic Almanac Vol. 15, No. 5 - 6 (2007). Retrieved 10 May 2007 from <http://leoalmanac.org/journal/vol_15/lea_v15_n05-06/WebbKerneKoh.asp>.