Gesture Mapping for Interaction Design: An Investigative Process for Developing Interactive Gesture Libraries

Thesis

Presented in Partial Fulfillment of the Requirements for

The Degree of Masters of Fine Arts in

The Graduate School of the Ohio State University

By

Lane Marie Kuhlman

Graduate Program in Industrial, Interior and Visual Communication Design

****

The Ohio State University

2009

Thesis Committee:

Professor Alan Price, Advisor

Professor Maria Palazzi

Professor Liz Sanders

ABSTRACT

Gestures play important roles as facilitators of language development, temporal-spatial learning, and non-verbal communication. Gesture-based interaction design seeks to capitalize on this natural method of human communication by using gestures as a means of interfacing with digital content. While technological factors address important issues related to sensing gestural input, design factors are the most critical factors relate to developing useful and approachable gesture-based interactivity. The goal of this research is to articulate more clearly some intrinsic characteristics of gesture that are significant to gestural interface designers, while providing methodologies that designers can use to gather and implement this information in a fashion that suits their unique design processes.

Gesture researchers have published a great deal of research that has significant implications related to gestural interface design, but most research in the field of gesture studies relates to gestures that are produced in combination with speech or in place of speech. Directly applying this research to visual interface design is difficult because many of the examples of gestures provided by these researchers analyze gesture in terms their linguistic characteristics. Because interface designers are seeking gestures that can be incorporated into interactive scenarios, there is a need for

example of gestures produced in response to visual-spatial cues. The aim for this study and beyond is to create a library of gestures that can serve as a reference to designers who are seeking visual-spatial representations of a broad range of gestural expression. This study presents methods of visual and spatial contextualization that can be applied or expanded upon by gestural interface designers who are seeking to build unique gestural vocabularies on a project-by-project basis. This document outlines a pragmatic approach to gestural interface design that aims to inspire designers toward further investigation.

This thesis documents the development processes for several interactive prototypes. Each of these prototypes helped to define specific research questions that may be important issues as gesture-based interaction design moves forward as a field of research. Discussion of interactive prototypes is followed by documentation of a user centered research study. This study presents new strategies for provoking, documenting, analyzing and contextualizing gestures within specialized visual-spatial scenarios.

The results of this study include documentation of an approach that can be used to generate libraries of interactive gestures. Several categorical patterns of gestural expression emerged from this research study, which reaffirms the potential for standardization of gestural interaction. Because gestures have recognizable visual and formal relationships to the things that they represent, their interpretation is closely tied to the context in which they are used. Through a process of contextualization, interaction designers can create visual-spatial frameworks for understanding the intended meaning of the gestures that a user produces. This thesis discusses best practices for applying gestures within interactive scenarios by defining many characteristics of gesture that represent a broad range of gestural expression.

DEDICATION

Dedicated to my family, my husband, and Ruthanna Rule who have greatly supported me throughout my pursuit of this degree

ACKNOWLEDGEMENTS

I owe a great deal to the faculty and students at the Advanced Computing Center for Arts and Design and the faculty and students in the Department of Industrial, Interior and Visual Communication Design at the Ohio State University. They have been an inspiration to me and created a community in which creativity can truly thrive.

I would like to give thanks to my advisor, Alan Price, who has been a phenomenal mentor to me. Thanks Alan for sharing your brilliant ideas and giving me a sense of direction through thought provoking conversations and insightful criticism. I am also thankful to have had Maria Palazzi and Liz Sanders who were also on my committee. Maria, although your criticism often presented me with difficult challenges, I am thankful to have received it because it forced me to address important problems and helped me to be a better artist, designer, and communicator. Liz, thanks for keeping things fun and interesting. Your playful approach to design research has helped open my eyes to a new ways of thinking about and approaching design.

To all the students at ACCAD, thank you for being an inspiration. Your work and dedication to your art has always energized me and encouraged me to go further. Thanks for sharing all of your creativity, friendship and laughs. Thanks Fran for the Chocolate Martinis and great advice. Thanks Beth and George for letting us turn your awesome San Francisco apartment in to a crazy youth hostel. Thanks Sucheta and Nitz for amazing Indian dinners. Thanks Jane for inviting me on that awesome trip to Vegas and for encouraging me to come along to San Francisco. Thanks Josh and Mary for being a super cute duo and for all the laughs and stories. Thanks Boris for helping me understand the need for balance in grad school. A special thanks to everyone who was willing to be my guinea pig during this research studies – Paulo, Andy, Ryan, Jae, Iuri, Vita, Asheer, George, Beth, Jane, Mary, Josh, Iuri, Jon, Annie, Louis, Matt, Ashley, Catalina and Elliot.

Mom, thanks for everything. I can always count on you to help me put my head back on straight and keep moving in the right direction. Thank you Dad for always being there and for going out of you way to make my life a bit easier whenever you can. Thanks Ruthanna for your unwavering support. I really would not have come this far without you.

Thank you Matt. You have been the most supportive element in my life during grad school. Your unconditional love and kind encouragement have helped me overcome a lot of my fears and anxieties. With you by my side I know I can get through anything. I look forward to many new adventures with you.

VITA

May 2003...............................	George Clements Memorial Scholarship for International Travel, Cleveland Institute of Art
June 2005...............................	Student Exhibition, Fachochshule Schwaebisch Hall, Germany
February 2006.........................	Student Independent Exhibition, The Cleveland Institute of Art
April 2006..............................	Honorable Mention for *Lane's Brain*, West Virginia Flash Animation Festival
May 2006...............................	*Listening Labyrinth*, BFA Thesis, Cleveland Institute of Art
September 2006....................	Story Development and Interaction Design, *Chaco Interactive Project*
November 2006......................	Microsoft Female Academic All-star, Serious Games Summit, D.C.
May 2007.................................	Graduate Teaching Assistant, Electronic Media for Artists and Designers
August 2007...........................	User Interface Design Lead, AOL International, Columbus, OH
April 2008..............................	Geometry Drawing Table, Multi-user Interface with tangible tools
May 2008...............................	Information Architect, Colangelo Synergy Marketing
August 2008...........................	*Climate Model Interactive*, Design and Programming, Byrd Polar Research Center

Major Field: Design

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

I. Introduction to research

Gesture-based interaction design is a new paradigm of computing in which gesture is the primary communication modality that impacts design considerations. Designers of interactive media are currently seeking out research and evidence that can help them design new gestural interfaces. Gestural interface designers currently have relatively little published research, few precedents and few established research methods to call upon. When comparing this emerging field of inquiry to more established fields such as web or gaming design resources, the amount of documented research is scarcer and the complexity of reaching an adequate level of understanding far greater.

Currently, designers of gesture-based interaction must forge new ground and test unique assumptions when attempting to stimulate socially beneficial gestural interaction and communication. Working in this area can easily become overwhelming due to the complexity of sensing and interpreting the meaning of gestures and the difficulty of moving from technology-centered to user-centered concerns. Results of gestural interaction are often difficult to analyze, document and build upon because researchers have not always share unified goals or views related to the best practices for developing gestural interaction. As the field continues to develop, there is a growing need for research materials that clearly outline the benefits of incorporating gesture and illustrate design strategies that lead to enhanced user experiences. In addition, the benefits of gestural expression need to be more clearly defined and discussed, so the gesture-based interactivity is not viewed as a an gimmick or extravagant pursuit. This will help designers unify around specific design objectives and allow them communicate these objectives to technologists and business stakeholders.

This thesis suggests that for this area of research to evolve, designers must take particular care in the methods they use to test, develop, and document gestural interaction, as these issues are the first step toward establishing new conventions of gestural interaction that will take hold and be built upon. As designers form a greater consciousness of the role gestures play in day-to-day life, we can mimic the benefits of gesture we experience in daily life such as person-to-person communication and externalization of our spatial and temporal knowledge. It is important not to forget that gestures have played an important role in human communication far longer than computers. There is much to be learned from a more conscious view of ourselves as we search for “best-practices” for gesture-based interaction design.

As the designer attempts to isolate design factors related to gesture-based interactions, the problem is complicated further by many varying definitions of “gesture-based interaction”. Gesture is a widely researched form of communication yet there are nearly as many categorizations of gestures as individuals investigating its use. It is clear when surveying research in related fields such as gesture studies and gesture-based communication in technology-centered research that gesture research outlines unique dilemmas. These dilemmas are not presented in a manner that applies directly toward design concerns. Researchers in the field of gesture studies have explored gesture in terms of its connection to speech, language development, and social interaction. Most research in this area focuses on psychological or developmental factors related to gesture. Computer science has explored gesture-sensing in terms of technology-centered methods of analysis. Most technology-centered research address the question of “how” to recognize gestures, but overlooks the question of “why” and “when” technology should address gesture. The technology-centered approach often forgoes addressing the human-centered value of gesture-based interaction.

Transferring natural gestural expression into digital forms presents difficult challenges. Motion sensing technologies have advanced to a point where it has become possible to differentiate patterns of motion acted out by a person in real time. This creates many new possibilities, yet as industry and academia explore these new potentials they often re-iterate conventions from mouse and keyboard driven paradigms of computing. There is an innate desire to build upon existing infrastructures, yet the models on which these infrastructures were designed often contrast with gestural modalities of communication. Current systems are predominately oriented towards single-channel, linear interaction schemas, rather than the multi-channel, non-linear schemas that would be more conducive to gesturing. It is the position of this thesis that prevalent interactive schemas are not designed to support the full spectrum of information that people are capable of expressing through natural gestural communication. Clinging to design strategies with deep menu systems and indexical organization reinforces linear approaches to interaction design and stifles the expressive potential that is an inherent part of the embodiment of communication as found in physical gesture.

Current motion sensing technologies make seemingly possible the idea of breaking ties with conventional systems and starting fresh with an approach that incorporates a larger capacity for gestural expression. This task begins with difficult questions. How can we make the leap from our current understanding of interaction design to new systems that more effectively support gestural communication? Without understanding gestures full role outside of computer interaction design, we cannot clearly envision what a gesture sensing system should look like. A greater understanding of gesture is important because a gestural interface design should strengthen the abilities that people naturally possess. Designing information architectures that more effectively supports gesture will require an understanding of recognizable factors that combine to form communicative gestural phrases. To move forward, there needs to be focus, not on how gesture can conform to computing, but on how computing can conform to include the larger scope of information that is communicable via gesture. The goal of this research is to articulate more clearly some intrinsic characteristics of gesture that are significant to gestural interface designers, while providing methods that designers can use to gather and implement this information in a fashion that suits their unique design processes. Ideally, a greater understanding of gesture can lead to further thinking related to how gesture-sensing technology should be shaped.

Before approaching this problem, it is important to understand the breadth of gesture as it is considered herein. Gestures can take many forms. They have uniquely illustrative and descriptive qualities that can communicate meaning similar to written or spoken words. Gestures can take the place of nouns, verbs, adjectives or adverbs used in speech. Gesture is inherently effective at communicating temporal and spatial information. Studies of gesture produced during speech show that gestures are used to indicate the shape of referents or the motions of referents over time. In many ways spatial concepts can be more easily articulate through an “air-drawing” or by contorting the hands into a form. We can represent a broad palette of information about space and time using only our hands and arms. Gestures can describe shape, form, action, and even pacing. They can be expressed through a medium, used in conjunction with speech or expressed alone as a form of pantomime.

Gestural communication possesses unique qualities because of the spatial nature in which they are expressed. It can be argued that physical gesture is, in many ways, more efficient than speech as a mode of communicating temporal and spatial information. In addition, gestures can be used to integrate information from various senses. It is important to consider that gestures do not exist in isolation. Through their gestures, people can refer to other co-present beings, objects and environments. As people gesture they frequently refer to elapsing or elapsed time, showing that the meaning expressed in a gesture also has a context within time. They often define abstract space or form through gesture. Using gestures, they even establish symbolic references to the abstract spaces or forms they have generated previously using gestures. Gestures can also be used to develop codified languages such as those used in American Sign Language or other systems. All these topics will be explored in greater detail in chapter 1.

These broader descriptions of gesture cover many of the roles that gesture has played throughout culture and may in effect make it a daunting task to develop a roadmap for the design of gestural interfaces. It is important to consider that this is not a task that can be accomplished by one person or even just a few individuals. The position taken herein is therefore a call for designers to take action. This document outlines a pragmatic approach to gestural interface design that aims to inspire designers toward further investigation. Design research methodologies proposed consider research in the field of cognitive science as well as a broad view of emerging research in the realm of gestural interface design. When considering the relevancy of gestures to interaction design it is necessary to consider how the computer can sense a gesturer and respond to a gesture in some useful and meaningful way. This thesis will review current applications of do-it-yourself and off-the-shelf gesture sensing technologies and provide analysis of the possibilities and limitations that accompany these technologies.

What needs might gestural interfaces fulfill? What useful and desirable things can we do with a gestural interfaces that we have never imagined doing before? Inspiring development of new methods that help to probe these questions is the main focus of this investigation. Through this research I hope to bring new scope to the definition *of gesture-based interaction design*.

II. Overview

This thesis addresses the development of experimental prototypes used to test assumptions related to gestural interface design and inspire more in-depth investigation. Early prototypes led to more critical investigation of the role of gesture within cultures throughout the world. A detailed overview of research from the field of gesture studies will be presented. Significant findings relative to interaction design will be discussed. Several findings that have arisen in the field of cognitive science imply that there are very specific formal characteristics of gesture that can be recognized and categorized systematically. This research paper has attempted to incorporate relevant features of gestural communication that were found during an extensive literature review. An overview of this research may help to more clearly establish the benefits of incorporating gesture into computing experience. This research document proposes new directions based on these findings and documents my path towards incorporating these discoveries into the design of gestural interfaces.

Several case studies will be discussed in terms of their focused approach to specific gestural interaction design ideas. This will include a review of discoveries made during public exhibitions and testing sessions. It will also discuss research questions that developed and helped to inspire further research. The final project to be reviewed addresses what I see as a need for generative and participatory design research that relates to gesture-based interaction design. It involves exploration of methods used to collect and analyze a broad range of natural gestures produced in responses to visual cues. Early prototypes informed the development of a research framework for participatory research that will be discussed in greater detail in the final chapters of this thesis. Earlier stages of design experimentation made clear the need for a more formal illustration of design strategies that are supportive of gestural communication. The conclusions drawn from this research outline a strategy for documenting and analyzing gestures. I feel that understanding how to approach gestural interaction in the early generative phases of design research is crucial step toward the development of gestural interaction that aligns with natural gestural communication

III. Expected Outcomes

This research aims to comprehensively illustrate the implications that gesture has for computing when human interaction is sensed via motion sensing technologies and used as a means of interface control within highly designed interactive experiences. It addresses these potentials in order to emphasize the broad scope of gesture and the rich implications it has for interaction and user experience design. Results from this research may serve as a reference for individuals seeking to explore various schemas of gesture analysis through the focused application of technology.

For new interface designs to implement a richer array of gestures, designers need to be inspired by thoroughly documented precedents that exemplify a fuller range of possibilities available. The final results of this research will include a digital archive of information related to gestural communication. This archive will include a collection of videos demonstrating a large array of gestures collected during a generative research study. Numerous gestures will be recorded in a digital video format and be accompanied by analysis of the gesture's potential utilization within an interactive application. The aim will be a library of gestures that can be referenced and expanded by gestural interface designers seeking to build unique gestural vocabularies on a project-by-project basis.

CHAPTER 1 BACKGROUND

1.1 Gestural Communication

The decision to give gesture a place within interactive infrastructures creates many difficult challenges for designers, developers, producers and marketers of this technology. Making a transition toward computerized systems that recognize and interpret our subtle gestural expression and consider the intricacies of our kinesthetic experience will require considerable effort, time and money on the part of many people. It is therefore important to start by weighing the difficulty of this task with the potential benefits that can come from such an effort. It is important to mention that incorporating gesture is not the solution to every problem, but the task may provide more optimal solutions for problems of a certain nature.

We are often unaware of how significant a role gesture plays in our lives. As I have developed my understanding of gestures, it has led to greater awareness of its role in my daily life. Most of the time, production and recognition of gestures occurs unconsciously. I attribute this feature of gesturing to how unaware we as a society of the role it plays in communication. Many early discussions with students, faculty, family and friends have made clear that there is a lack of social awareness related to gesture. Most commonly, discussion of gestures begins with hand gestures like “thumbs-up” or “ok”. The discussion does not go much further. Yet, when asked whether more arbitrary hand motions enacted by another person were intentionally communicative, people come to consensus that certain motions are meant to convey specific meaning (Kendon, 2004). It is clear that gestures are meaningful, but less clear why this is so.

A great deal of research into gestural communication has arisen, which sheds new light on the role of gesture in society. Gesture has been shown to have a significant role in language development, facilitating pre-verbal skills such as sensory integration, the sensory apparatus required during speech (Dejean, 2009). Over time, researchers have been able to isolate the fundamental elements of gestural communication, which repeatedly combine together to create richly communicative gestural phrases. Gestural forms of expression often contribute meaning that is beyond our verbal capabilities (Goldin-Meadow, 2003). As the factors involved in gestural communication become more clearly isolated, more systematic methods of interpreting gesture may become possible.

Systematization of gesture makes clear features of gestures that can be more easily distinguished by a computer. Gesture research has made it easier to apply gesture-sensing technologies in a manner that corresponds to natural gestural communication. Gestural phrases can be analyzed in chunks and phrases with clearly distinguishable structures. By referencing documentation of these natural structures of gestural communication, designers have the potential to harness the power of these structures in order to improve communication, education, and user experience.

1.1.1 Features of intentionally communicative gestures

William Kendon, who has done extensive research on the role of gesture in communication, provides an in depth investigation of the definition of gesture in his book, *Gesture: Visible Action as Utterance* (2004). He defines gesture as, “visible bodily action”, and goes further to discuss gesture as utterance, or an ensemble of action that counts for another. He says it is, “any unit of activity that is treated by those co-present as a communicative ‘move’, ‘turn’, or ‘contribution’”. Through his studies of how people interpret and recognize gesture, he discovered that observed bodily movements are given differing levels of attention by observers. Movements were given varying status based on features of *manifest deliberate expressiveness*. Movements were seen as deliberately expressive when it was found that, “the movement had a sharp boundary of onset and offset that was an excursion, rather than that resulting in any sustained change in position” (ibid). In other words, when the body was moved and then returned to the position it started in, when a motion was repeated, or when it appeared to be done for its own sake, it appeared to be more deliberate, and people could therefore more easily attribute meaning to it (ibid). These movements did not include movement meant to create an appropriate setting for interaction such as moving from one position to another. They also did not include involuntary or habitual movements, such as readjusting clothing. Gestures that were seen as intentionally communicative were:

> "... composed of a nucleus of movement having some definite form and enhanced dynamic qualities, which is preceded by a preparatory movement and succeeded by a movement which either moves the limb back to its resting position or repositions it to the beginning of a gesture phrase."

When a gesture is understood to be intentionally communicative it has clear onset and offset. Gestures involve meaningful preparatory movements that can influence their interpretation.

This notion helped me begin to think differently about the relationship between computing and gestural expression. A great deal of the complexity of integrating gestural expression into computing is awareness of the numerous subtle elements that work together to form a gestural expression. This new definition presented the notion that gestures have clearly distinguishable features which can be sensed by a computer.

1.1.2 Gestural Systems of Communication

There are many gestural systems of communication invented to facilitate communication. Referees use signs to communicate to players and fans during athletic competition. Gestural signs are an ideal way for a referee to communicate because gestures are less easily distorted than sounds during the yelling and cheering that takes place during an athletic event. Large crowds who gather can be aware of a referee's decision from across a large stadium without being confused by inaudible speech. (See Figures 1).

For scuba divers it is extremely important to be able to communicate underwater. There is a set of underwater hand signals established by the Active Divers Association (ADA), which illustrate a few essential signals (see figure 2). There are a lot of important things that need to be communicated during diving to ensure diver safety.

American Sign Language is a form of gestural communication that primarily uses the hands and arms to communicate. There is a large vocabulary of specific movements for many of the words and phases used in the American language. It is a highly effective method of communication for people with auditory impairments. Due to the complexity of language it takes a lot of time and practice to develop fluency in sign language.

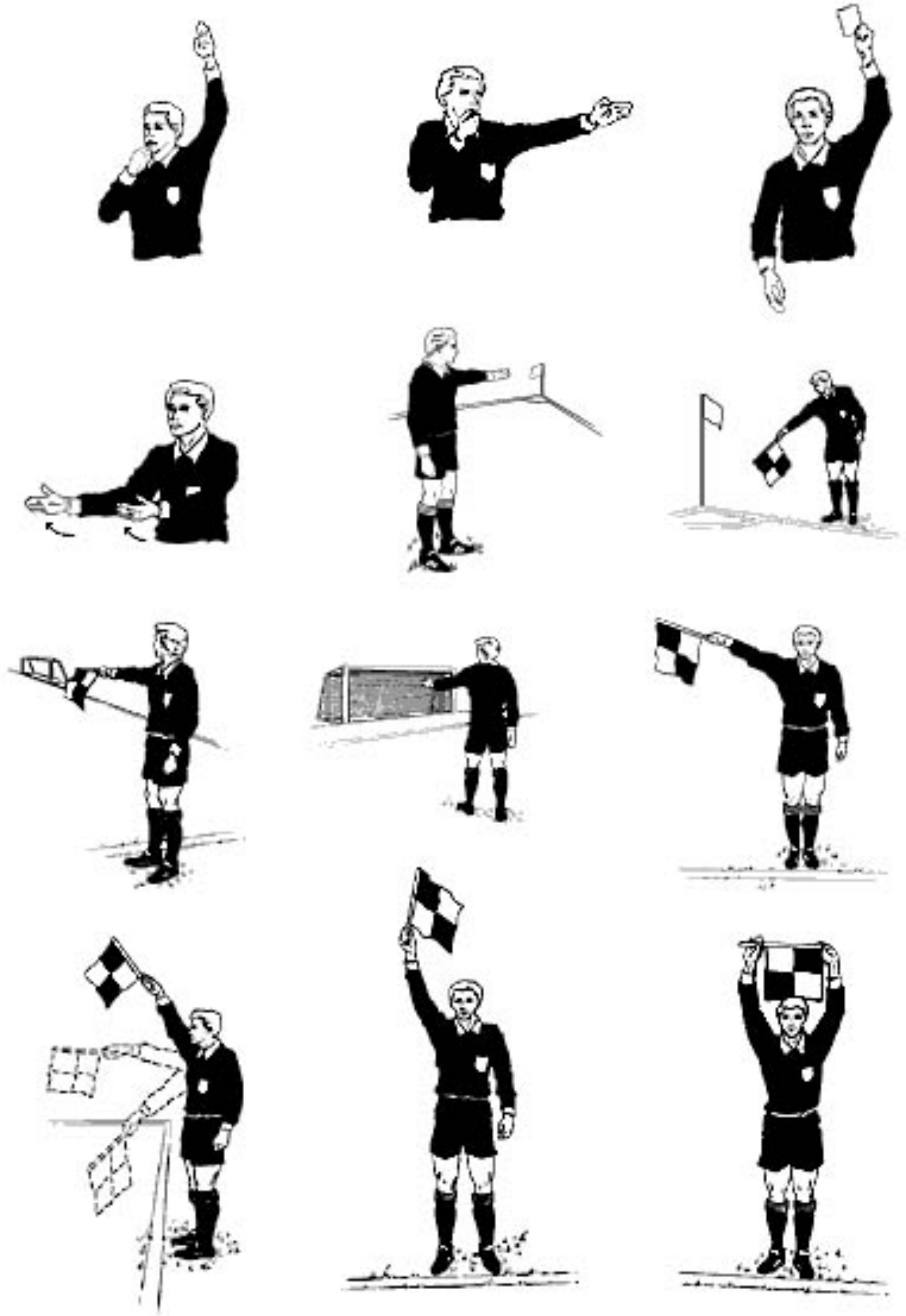


Figure 1 - Signs used by referee during a soccer game

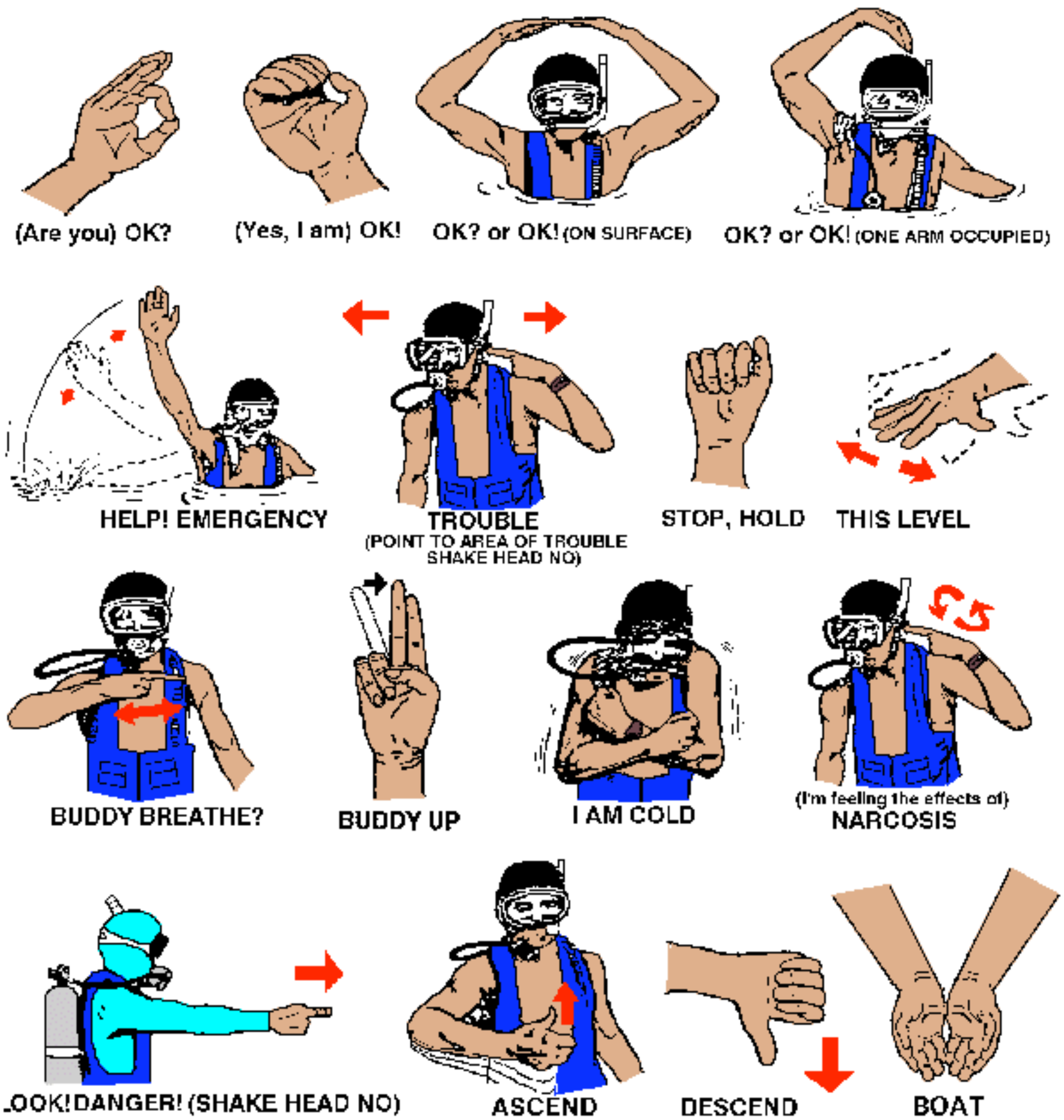


Figure 2: Scuba Diving Hand Signals

There are nearly as many signs as words in the English language. The limits of language over national boundaries apply to sign language in the same way. British sign language is much different than American as are many other national sign languages.

Many parents use baby sign, a form of sign language, to communicate with infants before development of the babies' vocal abilities. Research studies have shown that babies' cognitive development often precedes the physical development of their vocal-motor capabilities. Their ability to talk is preceded by the ability to communicate through hand gestures. There are several benefits to establishing this sort of communication before a baby develops the vocal-motor skills. Communicating through hand signs can minimize frustration for the baby, jumpstart literacy skills, and allow a baby to express basic needs more easily. (Handspeak, 2006)

We have discussed how gestural systems of communication are extremely helpful for the deaf and also very useful in situations where verbal communication is either impossible or impractical. In addition to these features, gestural systems of communication are much different than spoken languages because they require information reception to take place through vision and motion-processing systems. These systems also require communicative action to take place through the integration of motor systems (Emmorey, 2002). In spoken languages complex words can be constructed using concatinative processes like the use of prefixes and suffixes, but with sign languages these forms of meaning are constructed through concatinative processes in which a sign is nested within various movement contours and planes in space (ibid). For example, the word give can be represented with simple hand form and hand motion, yet variations in hand motion can change the meaning of give to indicate giving exhaustively or continuously. (ibid) (see Figure 3). Morphology of motion plays a role in the interpretation of the word, for example whether the same symbol refers to a verb or noun form (ibid). For example, the words for phone, paint and erase can be represented in noun or verb forms depending upon the size of the arching motion used during expression. (see Figure 4)

Another fundamental distinction between signed and spoken languages is the form meaning relationships. Spoken words, with the exclusion of onomatopoeias, possess arbitrary relationships between their form and meaning, while sign languages primary retain meaningful relationships to the things or actions to which they refer (Emmorey, 2002).

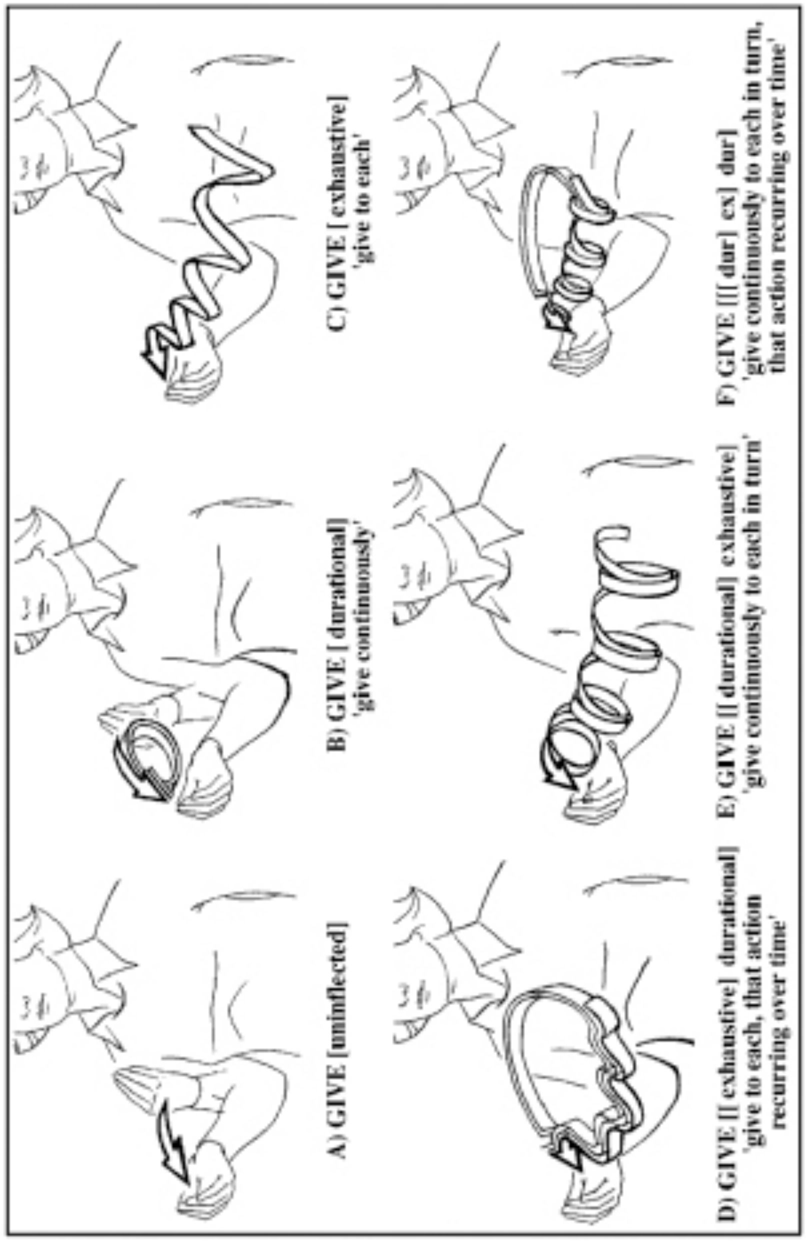


Figure 3: Sign Language Concatenation (Emmorey, 2002)

Figure 4: Verb and Noun forms of the same sign phrase (Emmorey, 2002)

1.1.3 Gesture's role in the Development of Language

When comparing gestures that are part of established communication systems to gestures performed during speech it can be seen that the meaning of established sign language systems is less dependent upon context. Codified systems, like American Sign, have less specialized repertoires of gesture. Specific sign-meaning relationships allow gestures to be used like words in speech. Gestures performed during speech, can be interpreted based on context. Yet, gestures can take on properties similar to verbal languages when used alone. With time gestures develop into unspecialized systems. It is the lack of specialization, and creation of singular meaning, that allows gestures to act as a language that can stand on its own without the aid of speech. (Kendon, 2004)

In his book, Origins of the Modern Mind, Merlin Donald (1991) discusses gestures role in the development of language. He argues that the earliest steps toward shared systems of communication and representation were developed through mimetic systems of representation. During this evolutionary stage, humans transitioned toward symbolic thought processes because they began to communicate through pantomime. Pantomime required an integration of sensory experiences that eventually allowed humans to evolve into the use of symbolic systems of representation.

Valerie Dejean (2009) has researched autistic children with deficiencies in the development of language. In the process she discovered a link to their development of pre-verbal skills. She found that children go through stages of non-verbal communication that are an important step towards their development of verbal skills. She notes:

> “Praxis involves ideation, or the creation of an "idea" of what one wants to do, followed by the organization and execution of a plan in order to do it. Praxis allows for intentional, purposeful communication and interaction with our world. Mimetic thought, the pre-verbal aspect of intentional communication, is in some ways the first step to intentional communication.” ..... “Pre-verbal children who go through a Sensory Integration combined with auditory training program, often develop intentional yet non-verbal communication prior to developing language”

This research suggests that gesture may play a more critical role in language development and sensory integration than we are commonly aware of. If designers gain greater understanding the components of gestural language this may lead to mediated experiences that are more effective in encouraging sensory integration. As a form of representation, gesture may facilitate integration because they require less abstraction than other forms of representation like language or other forms of notation. They provide a direct connection to the sensory apparatus that enable people to advance toward higher-level integration (ibid).

Several studies have investigated what happens when individuals were asked to retell a story through gesture-only communication. Several patterns emerge that resemble the structures used in spoken languages. Susan Goldin-Meadow characterizes these features as the *resilient properties of language* (2003, The Resilience of Language). Through her research she discovered that common language structures develop whether or not a verbal structure is imposed upon the gestural systems (ibid). Her research involved the study of children who were born deaf, and followed these children through the early stages of developing their personal sign languages (ibid). What she found was that systems of language structure, similar to those found in spoken languages, began to emerge. Often a predicate structure was used that involved 1) pointing to the topic 2) performing a characterizing gesture and 3) a gesture that predicates something of the topic. Simple syntax developed.

Gestures used in sign language take on different qualities than gestures used in the context of speech or during pantomime. While pantomime is always transparent and iconic, signs can be opaque and arbitrary (Emmorey, 2002). Sign languages have “a componential level of structure below the level of the morpheme (the smallest meaningful unit) that is akin to phonological structure in spoken languages”. (ibid.)

What all of this information related to the structure of gestures expression eludes to is the notion that gestural expressions have structures that are seemingly ingrained in us. Regardless of ones background, there are some features of gestural expression which seem to emerged repeatedly despite variations in culture, upbringing or limitation in sightedness or hearing. There are *resilient* patterns that are expressed and understood because they have a relationship to human cognition and sensory experience.

1.1.4 Functional Equivalence and Spatial / Temporal Expressiveness

Articulating certain rhythmic characteristics, such as syncopation or synchronization can often be more effectively articulated through *beat gestures* than though verbal cues. As their name suggests, beat gestures beat time and denote rhythm and pacing. They can be used to construct a temporal framework in which meaning unfolds, and therefore add meaning through their context instead of their content (Goldin-Meadow, 2003).

To see examples of the temporal expressiveness of gestures one need only look at the range of expression used by a conductor when communicating to an orchestra. An orchestra conductor uses hand gestures to shape the sounds of musical performers. The conductor indicates the beat with the right hand while cueing other members or the orchestra or signaling other changes in musical dynamics with the right. A conductor uses their left hand for cuing of individuals and to indicate dynamics, phrasing, expression and other musical elements. The conductor's role is

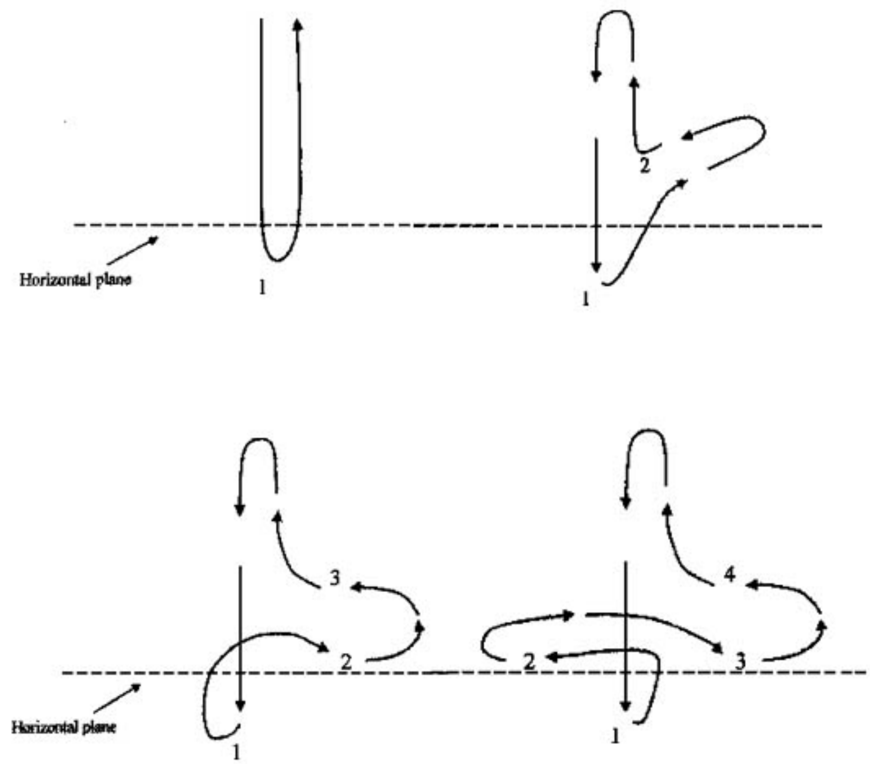


Figure 5: Patterns for conducting in 2/4, 3/4 and 4/4 time

important for ensuring that musical entries are made at the correct time and that members of the ensemble maintain a unified beat and expressive style. Figure 5 shows patterns of motion used to indicate time.

As a former member of the school orchestra, it is interesting to consider that most of my orchestra members responded to the conductor without explicit formal training as to how to interpret these motions. After having gathered information related to conducting, I recognize many of the patterns being discussed. As a musician I responded to these cues, but most of the information was understood without being explicitly stated. There is certain awareness associated with the use of temporal spatial metaphors with music.

Sounds produced by orchestra members are often related to movements (hitting, stroking, plucking, bowing). In a paper title, *Gestural Imagery in the Service of Musical Imagery* (2004), Rolf Inge Godoy presents the idea that gestures are effective in triggering images of musical sound because sounds have a *functional equivalence* to the gestures that are used to produce them (ibid). Godoy defines *functional equivalence* as the close resemblance between the neuronal apparatus involved in actual actions and/or perceptions and in imagined actions and/or perceptions (ibid). He argues that images of gesture seem to be efficient in evoking images of sound (ibid). Godoy suggests that mental simulations of movements used to produce music are preceded by ones ability to mimic sound producing movements (ibid). He writes:

> ".. what we often (or in most cases) think of as the meaning or content of music is actually a matter of gestural images: Images of effort, velocity, contours, trajectories, gait, etc, could all be understood as gestural phenomena, as gestural images transmitted by sound and decoded in listening back to a gestural language as in ballet or pantomime. "

Observations of neurological activity in the brain indicate that certain motor areas of the brain seem to be activated when imagining sound, indicating that action imagery is produced concurrently (ibid.). Brain imaging researchers provide related research through its discovery of *mirror neurons*. While studying neurological brain patterns of primates, Giacomo Rizzolatti and Michael A. Arbib (1998) discovered, ".. neurons that discharge not only when the monkey grasped or manipulated the objects, but also when the monkey observed the experimenter making a similar gesture". Similar patterns were discovered in the human brain, which indicates that there is further value in observing motions and gestures enacted by others. Simply by watching others motions we can share their spatio-motoric experience.

Kita & Ozyurek (2003) point out that gestures can arise from spatio-motoric functions. In other words, some gestures are developed from actions, and therefore, can provide kinetic information about actions as well as spatial information. When sets of gestures combine to form a large phrase, gestures that refer to action rather than subject can modify the meaning of a gesture phrase in a manner similar to adverbs or adjectives (Kendon, 2004). Characteristics of spatial movement can be expressed through gesture and be used to communicate not only "how", but "in what way". Gesture researchers often refer to these as *kinetographic* gestures, or gestures that depict bodily motion. For example, a pointing gesture might refer to a bird as the subject of a gesture, which might be following by a flapping gesture indicating the bird is flying. In this case, the flapping gestures would be considered a kinetographic gesture.

1.1.5 Shifting point of view and the gesturer

During a study in which participants were asked to retell stories, Alibali (2005) found that people were more likely to gesture when talking about spatial prepositions. People tend to gesture more frequently when discussing spatial information, such as when giving directions or describing the layout of a room (ibid).

Several studies indicate a high degree of gesturing occurs when discussing geometric relationships and spatial transformations (Goldin-Meadow, 2003 & Alibali, 2005). Several studies have been conducted that involve participants who were observed as they used gestures to retell popular fairy tales (Goldin-Meadow, 2003, Alibali, 2005 Emmorey, Tversky, & Taylor, 2000). Most of these scenarios revealed shifts in perspective that affect the interpretation of a gestures meaning. For example, someone might take on a bird's eye view while drawing out a map, or and "in route" perspective while describing directions along a street or walking path (Alibali, 2005).

The point-of-view adopted by a gesturer has some affect upon spatial interpretation of gestures (ibid). Considering the motion of a machine with many gears and moving parts. The spatial relationship between parts within the machine could be described from many angles. Perhaps a top down view would be optimal for viewing information about the placement and orientation of various parts. A side view would be more effective for viewing the vertical motion of moving parts. A bisection view of the machine might reveal the internal connections between various parts. When using a gesture to express a path of motion or the shape of an object it can be represented from many angles and may appear very different while being present from each angle. Some gestures play unique roles in that they define shifting perspective. By defining point of view they allow someone observing the gestures to know how consecutive gestures or verbal descriptions are to be interpreted.

1.1.6 Gesture as a Learning Modality in Collaborative Environments

Analysis of gestures used within classroom settings indicates that gestures are a powerful teaching aid. Gestures are particularly affective modes of expressing information related to temporal and spatial transformations. Several studies have been conducted indicating that students perform better on average in classroom where students and teachers can observe each others gestures (Goldin-Meadow, S., Kim, S., & Singer, M. 1999, Goldin-Meadow, S., 2004) .

There are several benefits to the use of gesture in educational settings. First, when students gesture they externalize information that teachers can use to assess a student’s current level of understanding (ibid). Gestures can often allow a student to express their knowledge through a visual/spatial modality, which can be a helpful outlet if, for example, they have not yet mastered the vocabulary necessary for verbalizing a concept (Church, R. B., & Goldin-Meadow, S., 1986). Gestures can be a helpful way for students to share their thoughts related to mental transformation tasks, such as how a process evolves at given increments and rates of changes (ibid.). Through gestures, students externalize their knowledge, which could perhaps make it easier for other students to visualize and animate a problem in their head, encouraging students to share their ideas. A student can observe the knowledge that other students possess when it is externalized through their gestures.

One of the most intriguing characteristics arising from gestural interface design is the implementation of multi-user computing scenarios. Gestural interaction is inherently collaborative in nature, which leads to unique patterns of learning and interaction. When comparing multi-user interaction to single-user interaction there are several clear advantages in terms of learning. People who interact with a multi-user interface are likely to learn more information at a faster rate than an individual acting alone (Kuhlman, Lane, and Price, Alan, 2008) People are more aware of issues, problems and questions and when these matters arise there is more opportunity to debate solutions and share pertinent information related to a topic. (ibid) Members of a group are more confident in the information they learned and are likely to take action with the knowledge they have learned more quickly.(ibid).

Given the potential of collaborative, multi-user interaction and the potential for gestural communication within multi-user scenarios, multi-user and gestural interactions seem to me to be a logical point of integration. An interaction language based on natural, human methods of gesturing could facilitate communication in multi-user interactive scenarios because it could capitalize on recurring patterns of gestural expression and create a more common language of gestural expression. In addition, users could more easily adopt these systems of interaction because they can learn by observing or mimicking the actions of others.

Unlike traditional mouse and keyboard interfaces, which have a more detached relationship between what a user is doing and what is observed on the screen, gestural interfaces can make the connections between a users actions and their affect on virtual scenarios more apparent. For example, try to recall an experience where you were observing a projected screen while someone standing at the back of the room, pressing buttons and moving their mouse, controls an interface. It may have been difficult to understand how someone else is influencing the digital scenario. Even when staring over the shoulder of another computer user, it can be hard to decipher which button or combination of buttons are being pressed in order to influence a piece of computer software. After watching a software demo, it can be difficult to repeat the steps that you saw another person take.

On the contrary, when we watch another person move their body, we can copy their motion (given their motions were too physically demanding). Gestural interfaces may be easier to learn to use because they make it easier to observe how other people are interacting with them. It is more difficult to copy subtle motions of the fingers (mouse-clicks and button pushes) than it is to copy larger and more general shapes and movements of the arms and hands. More significant changes in body shape and position are easier to distinguish from a distance.

The potentials that I see arising from larger scale, collaborative, and gestural devices relates not only to the interactive gestures that could be used to interact within virtual environments, but the resulting gestural vocabularies that could be used to communicate with other people. For example, after or during interaction with a gestural interface, users may be encouraged to use similar gestures when talking to other users across the table. They may be encouraged to use gesture to describe to others how they would like to affect content on the screen. They may be able to establish more effective methods of gestural communication outside of the computing environment as a result of the gestures they were encouraged to produce gestures within it. It seems the most interesting potentials of the technology would not be the new ways that we can interact with technology, but the ways that it can facilitate natural human communication. The most appealing scenarios I envision are those in which the technology no longer requires our focused attention and the attention is instead shifted toward interaction with other people.

1.2 Gesture Sensing Technologies

With so many gesture-sensing technologies it is hard to know where to begin. The greatest variations in the implementation of gesture-sensing technologies are arising from communities of bloggers, hackers and DIY techies. Many people have excitedly posted blogs about how to hack the Ninetendo Wii remote to make gestural music synthesizers (Endo, Ayaka and Kahura, Yasuo, 2008) or multi-touch surfaces (Lee, Johnny Chung, 2008). You can find volumes of exuberant technologists blogging on the Internet about ways to make everything from a DIY hand gesture interface for Google Earth (see Figure 8) (Phiffer, D. and Zer-Aviv, M., 2006) to a Gesture-based sugar dispenser (see Figure 9) (Suominen, J., and Keune, A., 2007).

Several websites have quite a following and provide many solutions from out-of-the-box hacks to total DIY projects that involve soldering your own circuitry. Several notable sites include NUI Group – Natural User Interfaces (NUI Group, 2009), Make Magazine, Processing (Processing, 2009) and OpenCV (OpenCV, 2009). Explorations of motion sensing technologies are emerging within several special interest groups that have gathered with the united interest of building unique projects with small microcontrollers such as Arduino (Arduino, 2009) and the Basic Stamp (Parallax, 2009).

Figure 6: DIY hand gesture interface for Google Earth

Figure 7: Gesture controlled sugar dispenser

With all this information it might be difficult to decide where to begin or what technology will best suit for a specific gesture-sensing task. What makes one gesture sensing technology better than another is tied to many factors such as the qualities and characteristics of motions to be tracked, the context in which interaction takes place and the nature of the audience who will use a given device. Technology should not overshadow a design purpose, which must always consider such factors. An overview of current technologies can spark many ideas but it can also quickly become overwhelming, leading to distraction or confusion. The pros and cons of various technologies will vary depending upon design objectives. Each project may possess unique challenges related to setting, durability, expense, budget, environment, space, technical skill, and time constraints. Experience with several technologies will reveals drastic and subtle differences that impact design strategies for gesture-based interaction. The following is an overview of several gesture sensing technologies that weighs some basic pros and cons related to gestural interaction.

1.2.1 Physical Computing

Physical Computing is an umbrella term that includes physical interaction that occurs when a computer recognizes, records or responds to stimuli in the physical world. Physical computing is a way of thinking about computing that may provide some relevant insights into my understanding of what constitutes a necessary approach to gesture-based interaction design.

Figure 8: "How conventional computers see us", an illustration from Physical Computing by Dan O'Sullivan and Tom Igoe

Physical computing focuses on how computers and humans sense each other. In their book *Physical Computing*, Dan O'Sullivan and Tom Igoe (2004) illustrate how a conventional computer sees us. They write:

> "A computer's image of human beings is reflected by its input and output devices. In the case of most desktop computers, this means a mouse, a keyboard, a monitor, and speakers. To such a computer we might look like a hand with one finger, one eye, and two ears. To change how the computer reacts to use, we have to change how it sees us." (see Figure 8)

Physical computer implies the use of various kinds of sensors that allow the computer to receive information from the physical world through transduction. Transduction occurs when a signal in the form of one type of energy is converted to a signal in another through a transducer. A transducer is a device for converting sound, temperature, pressure, light or other signals to or from an electronic signal. Examples might include photo sensors, accelerometers, or pressure sensors. In terms of computing, transduction is a conversion of analog information from the world into binary, or digital information. In a sense, all computing is physical. Hitting a button or moving a mouse requires the transfer of energy from the body into the computerized forms, but these physical movements are just a few simple examples of what computers can sense and how humans can communicate.

The signals that a computer produces affect the value and purpose it has in our lives. Meaningful encounters with computers provide sensory experiences because sensing is the way we understand the world. Computers operate based on binary information, which is a very simplified in comparison to the processes involved in human cognition. As a result, when information passes from a human to a computer, a process of simplification occurs. Computerized processing are performed on a low-level, binary scale and then amplified again they are sensed by people through vibrations, sounds and images as it is output and received by human senses.

In developing my perspective on human-computer interfaces, I developed an illustration that would show the relationship between human processing and computer processing as I see it (See Figure 9). Human processes like intuition and association have analogies to more abstract computerized forms of representation like indexical arrays or categories of information. While we operate based on the quality and meaning of information, a computer can only operate based on predefined quantifies and values. We produce complex and subtle behaviors and actions, while a computer can only produce predefined functions and forms of output. The nature of the transformation of information that occurs between humans and computers has implications beyond the notion that we must specifically tell a computer what to do before it can do it. I see computers as something more like a looking glass because what we tell a computer to sense, think and do reflects upon what we see as our deeper values and motivations. Because a computer can only do what we tell it to do, the things that we tell it to do and the systems that we construct with technology are a reflection of what we consider to be valuable human endeavors. This point of view has some level of atonement with the philosophy of physical computing.

The philosophy of physical computing addresses a flaw that underpins contemporary computing. Traditional computing arose from a desire to develop systems capable of mimicking human intelligence. This is the goal of artificial intelligence (AI). Alan Turing's original plans for a computer originated from the desire to duplicate human intelligence through a state machine (Bodanis, 2005). Here lies the main difference between traditional computing and physical computing. Physical computing seeks Intelligence Amplification (IA) as apposed to AI, which means it serves human intelligence instead of seeking to replicate it (O'Sullivan and Igoe, 2004). The purpose is to support people rather than duplicate them. Computers serve the people and not the other way around (ibid).

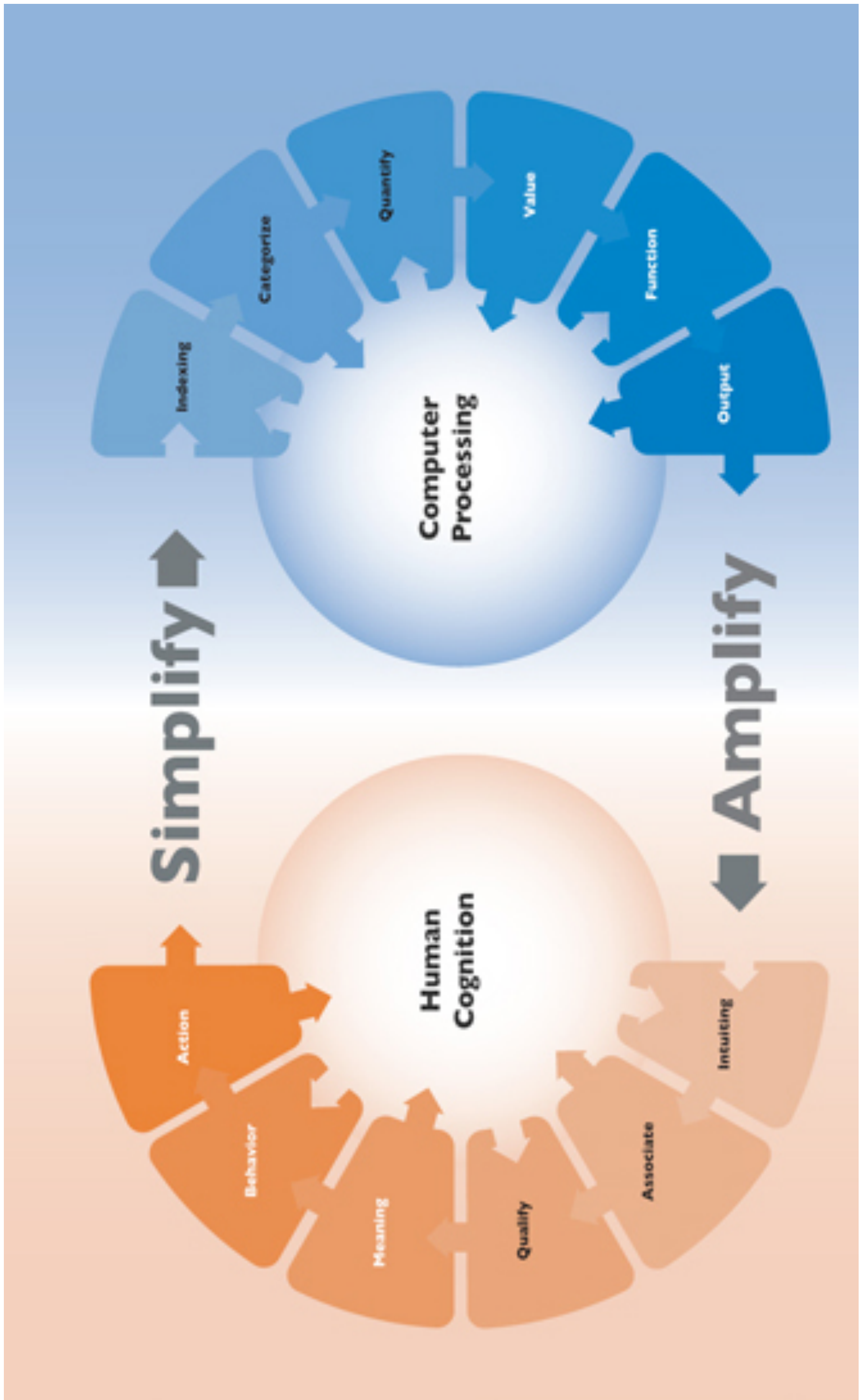


Figure 9 - Illustration of Human Computer Interface

Gestures and the spatial-temporal information that they embody represent a significant portion of human capacity for knowledge. This is knowledge that develops from physical interaction with our world. Theoretically, many gestures have found their origins from re-enactment or pantomime of bodily movements used to interact with our physical world (Kendon, 2004). As technologies have developed, many devices have been created which require people to conform or adapt their body motions to the computer. This counters the notion of intelligence amplification, and in so doing has had negative results.

In our contemporary world, many technologies have led to a decline in interaction with nature. This condition has come to be known as *videophilia*, the new human tendency to focus on sedentary activities involving electronic media *(*Pergams, O. R. W. and P. A. Zaradic. 2006 *).* There are many ways in which incorporating gestures into computing experiences could amplify and augment human abilities that traditional interfaces do not support. By including gestures, perhaps we are also allowing a greater portion of human experience to be incorporated into our technological endeavors, which may counter this effect.

To fully integrate the mind, body, and environment during human-computer interaction, it must not only sense information that is meaningful to us, but it must also convert the meaningful information being sensed into something that can be perceived as useful to peoples' existing abilities. These abilities include our predilection to be physically engaged with the world and to communicate our experiences of physical interaction with the world through gestures.

These modes of thinking about computing are helpful when considering approaches to gesture sensing technologies. Physical computing exemplifies one approach to incorporating gestures into human-computer interaction, but it requires thinking first about gestures and what they represent in our daily lives. Using spatial pattern recognition and the relative distance between postures changing over time, gesture can be used to communicate with a computer and to manipulate virtual space. These interactions can take on new and useful forms that help to educate, simulate, visualize, and communicate via gesture.

In order to construct a system that responds to human motions in useful and meaningful ways we must explicitly tell the computer what to "sense", "think", and "do" at any given moment. We must first know what features of a given gesture should be sensed, what transducers can sense these features, how relevant data can be interpreted and processed by a computer, and how it can be represented to a user through system response. Only after achieving this can we construct a system that captures meaningful information about gestures and transfers this information into something that is useful and meaningful to the people who use the system. This is a more difficult question than "which technology should I use?" The point in the development of gesture-based interactive experiences where designers must define their approach precedes decisions about technology.

The choice of technology should be a decision that arises from a desire to improve human potentials. The following are a few more specific examples of emerging technologies that are being applied to develop research and industry precedence.

1.2.2 Touch Tables and Surfaces

Touch tables and touch surfaces are a genre of motion sensing devices is primary geared toward recognizing the interaction of hand and fingers with a flat surface. There are several categories of touch screens on the market and the technology that have developed along various trajectories. Each method has unique characteristics that should be considered before choosing one approach. Each approach has its own unique capabilities and limitations. All allow tracking of multiple fingertip contact points on a two dimensional plane, but some allow more points of contact to be tracked simultaneously. Because these systems can only track gestures on a two-dimensional plane they significantly diminish the ability to incorporate some three-dimensional characteristics of natural gestural communication. These systems may be useful for recognizing two-dimensional features of hand gestures as they hover over or come in contact with a flat surface. These features include:

1) shapes or paths drawn on the surface

2) the shapes produced by hands in contact with the surface

3) temporal-spatial patterns of motion enacted over the surface

4) relationships of fingers and hands to the surface and each other

5) relationship between fingers and other visual elements in the interface.

1.2.1.1 Diffused Illumination

Many tables use a technique known as diffused illumination, or DI. Diffused illumination tables are typically comprised of a transparent glass or plastic surface onto which several diffused infrared lights are projected from either the top or underside. These tables involve a projector and a camera with a filter that sees only the infrared light spectrum. They typically implement a computer vision system, which processes imagery from a digital camera near the surface in order to determine the placement of multiple fingertips. Some development frameworks, such as Microsoft Surface, incorporate programmatic features that recognize logical grouping of hand features, such as finger sets that belong to a specific hand or a a person on a certain side of the table. Some of these systems can allow tracking of over 50 distinct fingertips.

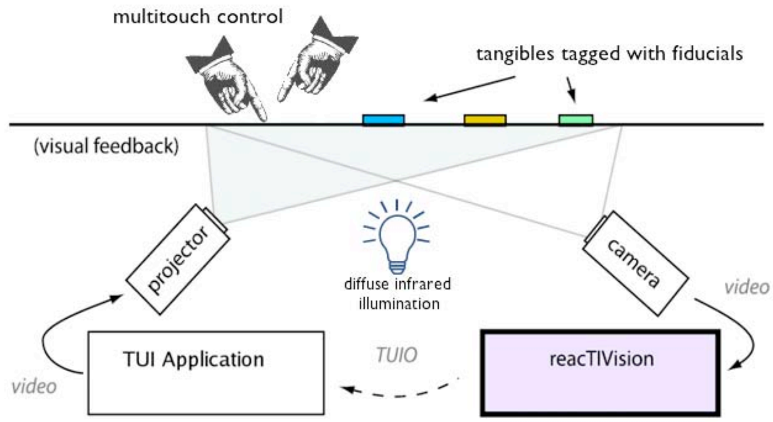


Figure 10: ReacTIVision diffused illumination system

The computer vision engines used for many diffused illumination multi-touch surfaces can also accommodate marker tracking. (Music Technology Group, 2009, Microsoft Surface, 2009) These systems typically implement a computer vision algorithm that processes video input. One example of a computer vision engine is *Reactivision* (Music Technology Group, 2009). This system is an opensource computer vision engine with printable markers and finger tracking capabilities that can be downloaded and used under a General Public License.

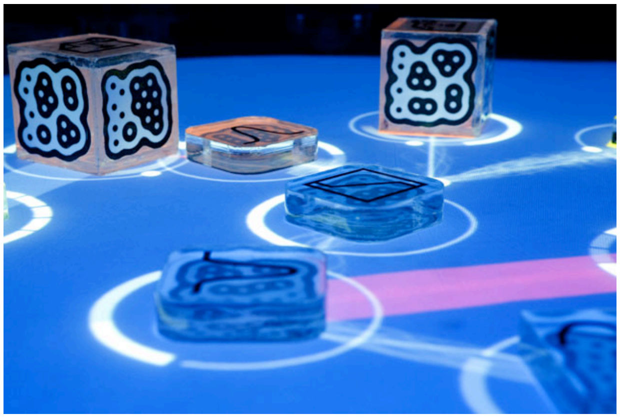


Figure 11: The reactable, a music synthesizer

Markers can be attached to objects that provide contextual interpretation of gestures performed with the objects in hand or near the objects that are placed on the surface. For example, a physical object with a marker attached could represent a specific tool with specific functionality. Gestures performed with a tool could have different meaning than gestures performed with fingertips or tools of a different type. Additionally, as with the Reactable (2009)(Figure 11), the placement of various marker objects can create a context in which gestures have new meaning. For example, the ReacTable utilizes beams of projected light, moving between objects, to represent music synthesizers. Each block generates sounds or affects music synthesis in different ways. The orientation and spatial relationships between sound generating block and sound filter blocks combine to create infinite music variations. Beams can be interrupted with fingertips in order to augment beats or tones in real time.

Unlike most touch surfaces, diffused illumination surfaces allow the shape of fingers that hover slightly above the surface to be seen. This allows some variation of motion sensing because hands and objects can be sensed when they are hovering above the surface, which allows some information about the distance between the hands and the table to be sensed. Many other forms of multi-touch surfaces only recognize the onset or offset of a gesture when fingertips touch the surface or are removed from the surface. Recognizing hovering provides some initial variation in the onset or offset of gesture recognition. It provides an additional layer of sensitivity within the framework that could be used to structure interactivity. For example, this may allow recognition of significant preparatory movements, which occur before a gesturer touches the surface. This could also be used to determine the direction from which the users hand is approaching or the general shape of the hand prior to contact the table.

1.2.1.2 Frustrated Total Internal Reflection - FTIR

Jeff Han and his associates at *Perceptive Pixel* have developed some remarkable multi-touch demos through a unique combination of technologies. These systems are capable of recognizing finger pressure and motion with high temporal and spatial resolution (Perceptive Pixel, 2009). Their technique is made possible in part by a phenomenon known as Frustrated Total Internal Reflection, or FTIR (See Figure 12, image from Scientific American). This technique uses light refracted from fingertips to deliver crisp images of fingertip shape to a camera located behind the touch surface. Perceptive Pixel incorporates engineered surfaces that magnified brightness of finger images based on finger pressure. This allows pressure sensitive manipulations that can be used to perform tasks such as change the spatial layer of virtual objects in the foreground, middle ground and background. The ability to recognize finger pressure may be more significant to direct manipulation tasks than it is to gestural interpretation of motion, but it has some significance to the physicality of interaction. This feature allows for greater tactile engagement with digital content. Therefore, the experiences of such physicality may stimulate gestural communication when reflecting upon experience of such an interface.

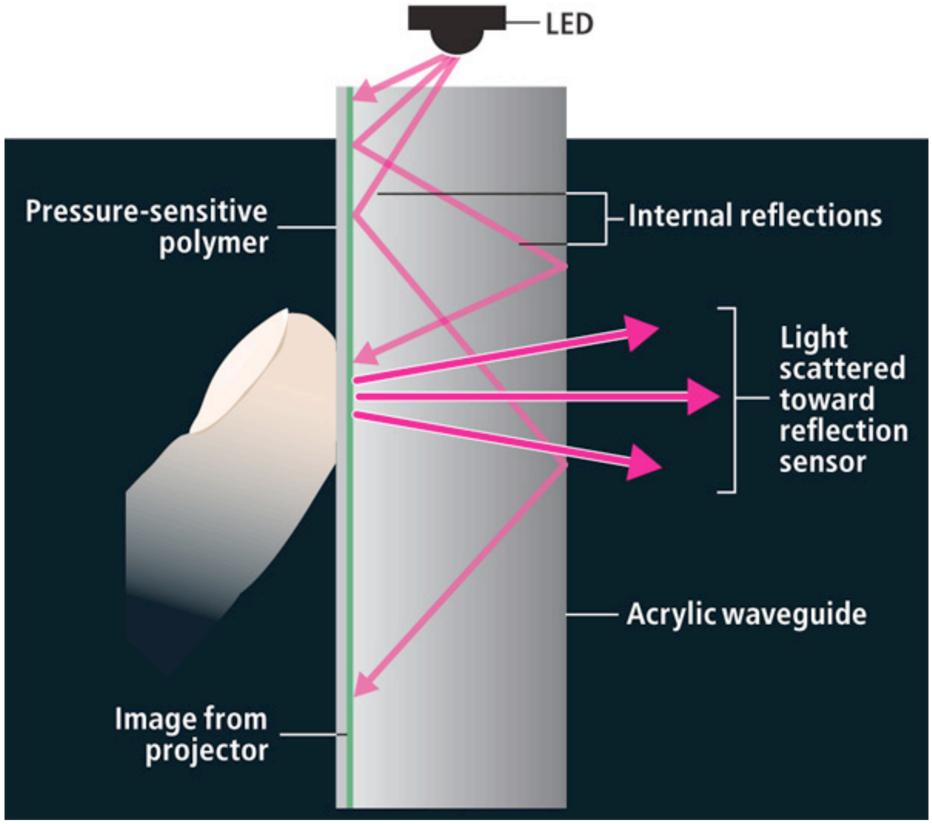


Figure 12: Illustration of light reflection in an FTIR multi-touch surface (Brown, Stuart, 2008)

A few drawbacks to an FTIR technique include the fact that it does not allow marker tracking, as diffused illumination techniques do. Most contact with the surface of this device would appear as a solid bright light, and therefore distinguishing black and white markers would not be possible. In addition, this system is not adept at recognizing “ghost fingers” or fingers hovering slightly above the surface.

1.2.1.3 Capacitive

Capacitive Touch Screens are typically glass surface with a transparent metallic conductive coating. They allow imagery from a LCD screen to pass through this capacitive surface with minimal visual disturbance. Because it does not require space behind the surface for a projector, this technique allows for the creation of fairly thin multi-touch surfaces, like the iPhone. When a finger touches the screen it draws a minute amount of current to the point of contact and creates a voltage drop that can be used to insolate the coordinates of the finger-touch location on the capacitive touch surface. This technique sometimes relies on the small amount of capacitance provided by human skin, so some of these devices may not operated when using gloved hands.

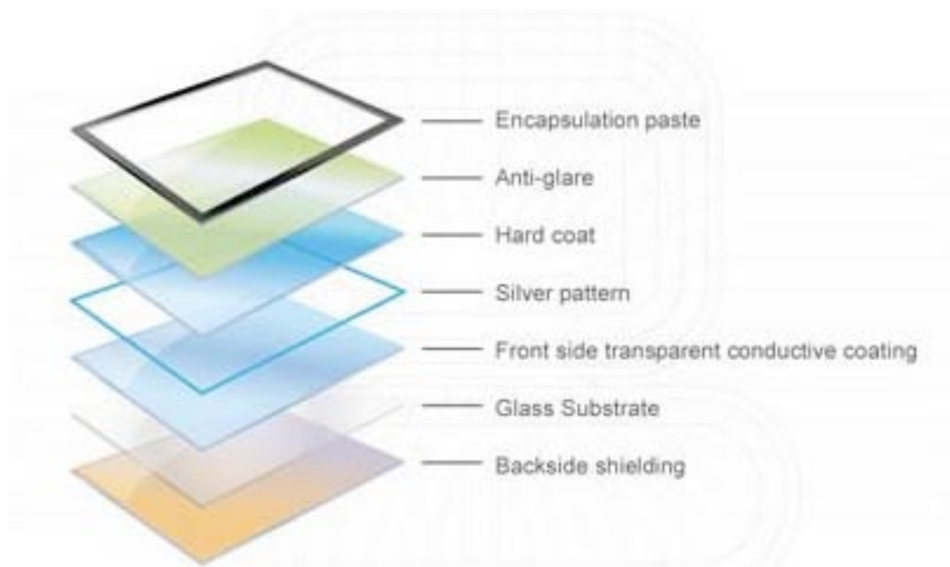


Figure 13: Layers of a capacitive surface

Unlike multi-touch surfaces discussed previously, most capacitive surface provide only a single interactive layer of recognition related to onset and offset of motion. While some devices recognize hover, a hover-touch is usually an equivalent signal to a touch signal. Therefore, the device only senses the presence or absence of a finger. Many capacitive touch screens are limited in the number of finger points that can be recognized at one time and do not allow as many simultaneous points of contact as do FTIR or diffused illumination touch-surfaces.

Figure 14: The iPhone utilizes a capacitive touch-surface

Projected Capacitance Touch technology is another type of capacitive technology that allows tracking of motion within an XY coordinate system. Projected Capacitance technology is also being used for external "through window" touch-screen applications (3M, 2009). These technologies can allow kiosks to be operated through a window, which provides some advantages in terms of security in public areas. As we speak new innovations are being made in the realm of capacitive touch screen. At the moment, developments in this area are hard to keep up with. One of the biggest advantages of capacitive multi-touch screens the potential to incorporate them into hand-held mobile and portability devices.

Innovations are being made on the level of microprocessor technology that will affect the feasibility of multi-touch ready operating systems. Mitsubishi Electronic Research Laboratories has explored many off these possibilities and published suggested directions for development in this area (Shen, C., Forlines, C., Wigdor, D. Vernier, F., 2007) . In an "Open Letter to OS Designers from the Tabletop Research Community" (ibid) individuals from Mitsubishi Electronic Research Laboratories wrote:

In order to properly enable a system for multi-user tabletops, each of these changes would need to be made to event delivery mechanisms:

1. Events need additional fields, such as 'touch area shape', 'touch area location', 'touch areas', 'user ID', 'device ID', 'touch pressure/strength'.

2. Focus mechanisms need to be redefined to recognize both multiple input areas by a user, and simultaneous input by multiple users

3. Widget behavior needs to be defined to properly react to input from multiple users

4. Widget behavior needs to be defined to properly react to a single user making simultaneous input to multiple points.

5. Multiple keyboard support, be they soft or hardware keyboards, needs to be integrated.

It is clear that new operating systems will be prepared to handle multi-touch screens. There are several multi-touch tablets and monitors being released on the market and there will surely be many innovations in this area in the not so distant future. Multi-touch computing has perhaps been the most widely adopted form of gesture-based interactivity. Further adoption will likely take place in the next few years, as both Apple and Microsoft will further incorporate multi-touch gestures into their operating systems.

Apple is currently seeking various patents for multi-touch input and has recently been awarded one (patent number 7,479,949 on January 20, 2009) which attempts to limit other companies from use of gesture-based interactions schemas which they have released. This raises a great deal of questions about the feasibility of developing cross-platform standards for multi-touch devices. Only time will tell how these scenarios will unfold, but is seems as though there is a potential that aggressive pursuit of patents may inhibit industry standards.

1.2.3 Video-based Gesture Recognition

Video-based motions sensing systems enable many interactive scenarios in which gestures can be interpreted from two-dimensional camera data. As discussed previously, both FTIR and diffused illumination tables often rely on video-based computer vision systems. Some video-based motion sensing systems can drive visual displays without requiring participants to contact a physical interface. Video-based motion analysis relies on computer vision algorithms that decipher key information related to position, direction, timing and velocity of movement.

Computer vision algorithms process arrays of pixels from video data is real time, so there are some complex processing and performance issues associated with these techniques. Certain movements are difficult to capture when occlusion occurs. To a certain degree, this effect can be countered by combining information from multiple cameras. Signals from video cameras provide a great deal of information, especially with higher resolutions cameras. The difficulty of the task relates to separating meaningful signals from noise. Lighting conditions are a major consideration because they can affect a computer algorithm’s suitability for interpreting body movement from a video feed. These scenarios often require controlled lighting scenarios.

Some motion sensing systems model the structure human body through the computing architecture, enabling accurate estimations of joint hierarchy and positioning (Cassel, R. and Collet, C. 2004). Some of the most advanced applications of these techniques have developed from researchers working in theatre and dance (Eyesweb, 2009). In each of these cases, computer models must address logic and processing of information from the cameras in an *a priori* fashion. Movements typically have a great deal of variations so logic systems that model body movements must address a great deal of contingencies.

Motion-based models have been implemented to track motion trajectories of human movement like walking, running or gymnastics (Rohr, K. , 1997). Coordinate systems have also been developed to interpret motion within 3-dimensional space (Ibid) Several common computer vision techniques can often be combined to draw out important information about human motions captured with video. Here are a few examples of techniques and their definitions:

1. Symbol Tracking – determining orientation of predefined black and white symbols based on algorithms that decipher shape distortions

2. Blob Tracking – determining movement based on areas of pixels with similar values or colors

3. Frame Differencing – determining the difference between consecutive frames in order to isolate areas of change or motion

4. Background Subtraction – requires starting with a static background images and determine figures based on the difference between the background image and the image with the a new figure included

5. Edge Detection – determining the outmost edge of a figure

7. Area of Motion – usually determined using a combination of frame differencing and background subtraction

6. Center of Motion – based on the center of the area of motion

8. Quantity of Motion – the number of pixels where motion is detected

9. Direction of Motion – based on the areas of motion and their relationship to the figure's previous position

10 Trajectory of Motion – using assumptions related to the direction, position, and velocity of motion a trajectory can be determined

In terms of gesture recognition, video-based motion sensing can be quite complex because it requires some assumptions about context, body form and limb movement to be made. A video signal is quite complex and often has noise in the signal that needs to be filtered out. Processing video input can be complex because it is difficult to sort out extraneous changes in pixel quality from meaningful information. These systems require application of explicit and highly descriptive systems of logic in order to recognize meaningful information about body movements and limbs. A large body of knowledge related to computer vision is developing, which reveals great potentials to come in this area.

Some high-resolution camera-based systems can capture motion using the combination of data from multiple view points in 3D space. Vicon Motion detection system use an array of camera, typically positioned in a circle around the subject whose motion is being captured. This system can track subtle movements because it captures movements of the limbs and joints at a high spatial and temporal resolution. This system has been utilized in the motion picture industry to map human or animal motions onto 3D animated characters. It has also been used for virtual puppeteering and for recording intricate movements within dance performances.

There are several drawbacks to high-resolution motion capture. These systems can have occlusion, but including cameras from many angles can counteract this effect. This system may require a great deal of work can go into filtering out unwanted noise in a signal. The data retrieved during motion capture can be quite large which poses problems for storing and processing of resulting data.

1.2.4 Data Gloves

Typically data gloves detect pitch and role of the hand as well as bend of the fingers. Data gloves can capture a large range of expressive information that can be combined to create complex hand-sign systems of interface control. Data gloves are typically expensive and are also typically only manufactured in one size, making it difficult to create an interface that accommodates various hand sizes. Performance can vary based on how the glove fits, which is another difficulty with this device. The device is often directly connected to the computer, which makes data gloves only suitable for use in constrained environments.

Datagloves have often been combined with high-resolution motion sensor placed on the wrist in VR simulations. This provides additional information about the placement of the hands and their relationship to each other as hand shapes change over time. By including these features, fairly high resolution sensing becomes possible. Datagloves provide a continuous analog signal that can often be difficult to work with because it requires the signal to be processed in real time. As with video-based motion tracking, it can be difficult to draw out important information and separate it from unwanted information or noise.

This technique provides the greatest amount of detail related to hand muscle contraction. Contraction of hand muscles is often more difficult to interpret from vision based systems, which are limited because they are trying to interpret muscle tensions from two-dimensional image. The kinesthetic experience of muscular tension is an important part of the experience of gesturing, which is often harder to pick up from other methods of gesture sensing. This technique is most useful for sensing subtle changes in the bending and flexing of fingers and wrists, which can be important points of interpretation for certain types of gestures. Other techniques require logical models that infer joint hierarchy, which typically possess a greater degree of inaccuracy related to subtle movements.

1.2.5 Electromagnetic

It is a well-known fact that the human body naturally produces electro-magnetic fields. Electrostatic motions sensing systems typically enhance this field so that it can be more easily recognized. Applications of this nature do not require physical contact with an interface (Smith, J. White, T. Dodge, C. Allport, D. Paradiso, J. Gershenfeld, N. 1998). They also do not require calibration for different users or environments. One example of this technique implements a frame that allows users to control the interface from a distance by pointing. (Yinlin, L., Groenegress, C., Wofgange, S. and Fleischmann, M. 2004).

Another technology that uses electromagnetic motion sensing is the Polhemus FasTrak. The Polyhemus FasTrak uses several small sensors that can be positions on the body and used to track subtle motions in three-dimensional space. The Polyhemus FasTrak system uses an electromagnetic motion-detection system that enables six degrees of freedom without occlusion. Electromagnetic motion sensing is very unique among motion sensing technologies when you consider that the body produces detectable electromagnetic waves. Systems that recognize the electromagnetic force (EMFs) of the human body do not rely on the visible light spectrum as do camera based systems.

1.3 The quest for “natural” interactivity

Among communities of technologists and interaction designers, it has become popular to taut the naturalness of certain types of interaction over others. Gestural interactivity is often attributed with more natural interaction. Natural User Interfaces, or NUIs, have become their own genre of computer interfaces. There is a desire to create more natural interactivity into computing experiences, but definition of the characteristics that might lead to a more natural experience for users are fairly elusive. So what is a NUI and what qualities do NUIs possess?

Natural User Interfaces are often referred to as a new paradigm of computing which was preceded by graphic user interface (GUI) and the command line interface (CLI). When comparing these paradigms, NUIs have several distinctive characteristics that indicate the results of movement toward more “natural” interaction. To better illustrate these principles, I will refer to the design principles for natural interactivity as applied by the Microsoft Surface design team (August de los Reyes, 2009):

1) *Principles of performance aesthetics* - Natural user interfaces are more evocative than CLI or GUI paradigms because they relate more closely to real-world kinesthetic experiences.
2) *Principles of direct manipulation* - NUIs allow more direct interaction than previous paradigms because meaning does not need to be communicated through translation into button pushes or mouse coordinates.
3) *Principles of Scaffolding* - NUIs should provide a less mediated path toward the execution of a task or command. A GUI typically provides several paths that enable the execution of a given task to be approached in various ways instead of a single, clear solution. A CLI would allow open access to all of a system’s functionality from any point within the system hierarchy. For this reason, CLIs requires foreknowledge of the scripts needed to execute a given command.

4) *Principles of Contextual Environments* – NUIs should present information to users based on context so that users are presented with what to do rather than required to figure out what to do. Graphical user interfaces typically provide exploratory experiences in which many options are accessible at any given time. GUIs do not typically limit users options based on context, but some interfaces, which are presumably more “natural” would present options based on context of use. CLIs allow all commands to be executed at any time, which requires direct and explicit action on the part of the user. CLIs provide no contextual information related to a command so users must rely on additional written documentation in order to write code with appropriate syntax and structure.
5) *Principle of the Super Real* – NUIs should allow users to quickly and intuitively utilize the system by mimicking and/or extending kinetic interactions from the real world. They should allow people to immediately understand how the system will respond. GUIs rely more heavily on icons and metaphors to create a sense of realism. CLIs utilize much more abstract textual representations.

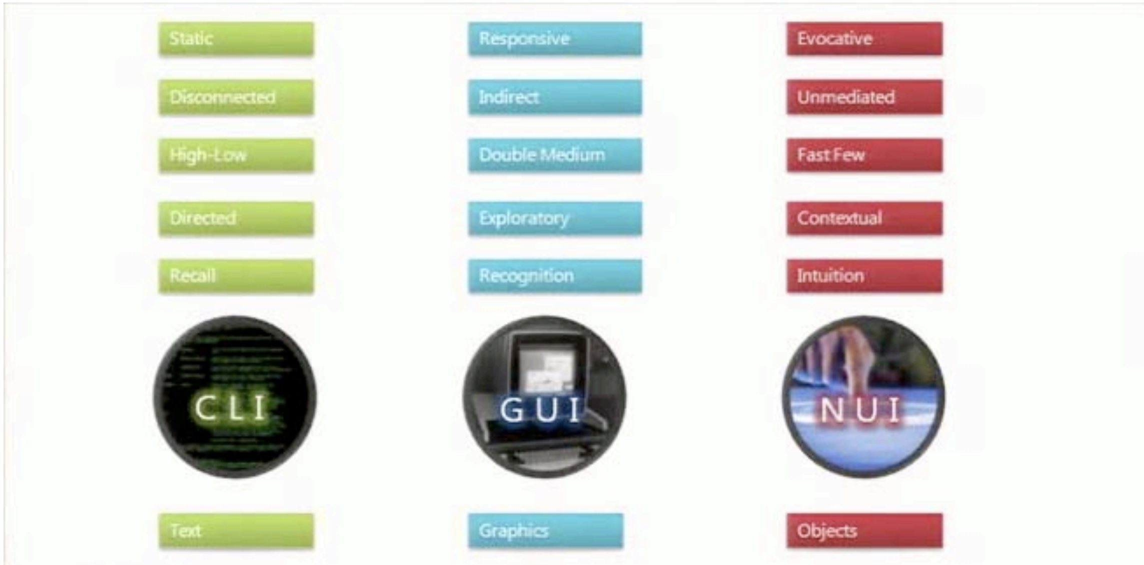


Figure 15: Comparison of Computing Paradigms,
from August de los Reyes, 2009 Microsoft Surface

Despite the best of intentions, many interfaces that claim to be "natural" and therefore "intuitive" are far from achieving this goal. When comparing CLIs to more contemporary form of human-computer interaction, it can be seen that we are moving in a direction that is more *like* natural experience. Yet, clearer definition of what constitutes "natural interaction" seems to be needed before they can claim to have created truly natural experiences.

Webster's definition of nature includes terminology that implies exclusion of things made by human hands. By this definition, all computer interfaces could be considered to be quite unnatural. Perhaps the characteristics of interfaces that designers and researchers have come describe as "natural" could more simply be stated as the effectiveness of one form of representation vs. another. Various forms of representation have developed throughout the ages with the result of more effective communication and self-expression. From numbers and letters to algorithms and languages, various forms of representation simplify cognitive and communicative process, which would otherwise be overly complex. Depending upon the situation or purpose, one form of representation may be more affective than another. Bill Buxton (1990) discussed this phenomenon saying:

> "Where all of this is heading is the observation that the notion of a universally understood natural language is naive and not very useful. Each language has special strengths and weaknesses in its ability to communicate particular concepts. Languages can be natural or foreign to concepts as well as speakers. A true "natural language" system is only achieved when the language employed is natural to the task, and the person engaged in that task is a native speaker. But perhaps the most important concept underlying this is acknowledging that naturalness is domain specific, and we must, therefore, support a variety of natural languages... If there is anything that makes a language natural, ... it is the notion of fluidity, continuity, and phrasing."

In truth, gestural interaction is no more natural than any other mode of interaction. The characteristics that make interaction seem more affective seem to be relative to the context in which a given mode of interaction is implemented. The relationship between the mode of expression and the information being communicated has a greater impact on its effectiveness within a given scenario. In one scenario, verbal communication might be best suited for a task. In another, gesturing may be the most effective form of representation. Verbal communication seems to be best suited for single-threaded communication while gestures seem to be most suited for multi-channel communication (Bill Buxton, 1990).

While technological approaches addresses gesture sensing in a more elaborate manner than design, I believe that what inhibits gestures from becoming more fully incorporated into computing experiences is a result of a lack of design strategy. The research that I have found, related to developing interactive gestures, has primarily focused on developing gestures based on predefined tasks (Nielsen, Michael, Storring, Moritz, Moeslund, Thomas B. and Granum, Erik., 2003) or the predefined "needs" of the users (Saffer, 2009). It has become a tradition in technology fields to begin by defining what a person can and should do before they even do it. Of course, to develop certain types of interaction, developers must begin with an understanding of what an application should enable a person to do. I would not argue this point. However, there is a flaw in this approach because it limits our view of human needs and capabilities to what technology is currently capable of doing. Perhaps the reason that technologically centered approaches lead to unnatural interactivity may be the result of several flaws in approach:

1) They start with the application
2) They focus on the technological status quo

3) They over segment interaction by focusing on predefined tasks

Technologists often focus on what is needed in order to programmatically respond to gesture. The problem with this approach is that it does not begin with either a view of how gesture is useful to communication or an understanding of how people comprehend gestures. For technologists it is clear that there is a justifiable need to address concurrency and parallelism. There is a great deal of useful discussion in technological fields that relates to the functional elements of gesture. Clear definitions of programmatic functionality have been around since the 80s, yet we still seem to lack directions in terms of useful applications. Rhyne discusses several functional elements quite clearly in his 1987 paper title "Dialogue Management for Gestural Interfaces". These elements include:

1) Scoping - Scoping is the act of selecting a set of objects to be acted upon
2) Elements within scope of may be defined by specifying a property, pointing, or by creating motions that enclose, traverse, or indicate spatial extents of an objects referred to.
3) Targeting – pointing to or isolation the element
4) Operations – gestures that operate on the meaning of objects or other gestures

He even goes on to address issues related to each:

1) Syntactic compress – several small strokes can connect together to form a single meaning

2) Contextual effects – Form and function of elements in the general vicinity of a gesture affect its meaning. Create the need for application dependent meta-rules which control the interpretation of gestural dialog
3) Closure – the event which signals the end of dialog phrase
4) Embedded Dialogs – multi-threaded, asynchronous dialogs

It seems that many of the innovations related to gesture-based interactivity have sprung from a very tech-centric focus. It is time for a reversal in terms of approach. Designers need to play their part in addressing the challenges related to creating intuitive and engaging experiences that incorporate gestures. This is not for the sake of technology, but for the sake of improving human experience. The discussion needs to begin with the useful role gestures play in human communication and in the embodiment and externalization of human knowledge. Stagnation in advancement is not due to technical inadequacies. Technical innovations are plentiful, but seem to lack directions in terms of usefulness, usability and applicability to human experience. Technologists need creative inspirations from artists and designers who can address the deeper human issue involved in gestural interactivity. Only then will the field truly advance.

Through the previously literature review I hope to have made clear the state of gesture-based interaction design. It seems to me there are three elements that are a crucial to discussions that will foster progress in this area. We have seen significant research in both cognitive science and computer science. The third element, which seems to be unequally represented, is the design element. It is my conviction that progress in this area is held back by a lack of design debate related to such topics. There is a need for communication between the design community, technological communities and cognitive science communities in order for this problem to be more fully addressed. Through this thesis I hope to encourage discussion within the design community that can broaden into collaborative dialogs with other research communities. It is time for designers to do their part. In the next section I will discuss several projects that helped to inspire and define my approach to building such creative dialogs.

CHAPTER 2 DESIGN RESEARCH METHODS

2.1 Development of Research questions through design exploration

Much of my body of research related to gesture developed in stages, inspired by experimental prototypes. During the first section of chapter 3, I will discuss this body of work and how it has led to the formulation of the primary research objectives addressed by this thesis of mapping schemas of gesture-based interaction design. Several projects that were important steps in developing the design methodology will be discussed in this section.

My process developed from a strong interest in design methods that could increase physical engagement in computing experiences. Ironically, the first project that was developed goes by the title “The Digital Seed Project”. Although this project did not deal with gesture directly, the ideas that it embodied were a major stimulus for subsequent projects, which focused on gesture. Section 3.1.1 addresses the development of this process and resulting outcomes from public exhibition of the project. Awareness gained from public exhibitions helped frame several key issues that would later informed the development of a new approach.

The second major project to be discussed explores gestural systems that could be used as a means of controlling virtual characters. The “Data Glove Character Controller” project involved the creation of a gesture system for context-related interpretation of hand gestures. This system recognized hand signals that could be use to trigger actions of a character within a virtual environment. Each gesture was interpreted based on the characters location in space and proximity to objects in the environment. Some limitations and potentials of “context-sensitive” interpretation of gestures were explored and the findings were applied in subsequent investigations.

The third major project was the “Geometry Drawing Table” project. For this project, I created a geometry-drawing interface that was controlled with tangible tools. Users of this interface could interact by using two tangible tools, which enabled different functionality depending on whether they were used together or separately. The development of this interface was a major turning point in the research, because it made clear the need for a more rigorous research process that could isolate formal models for gesture-based interaction design. I began to realize that multimodality was an important issue that could unlock what I saw as the greatest potentials of gestural interaction.

As development and testing of interactive elements of the touch table project continued, I felt that a definition of the true potentials of gestural interface design was inherently connected to the multimodal and multiuser capabilities that gesture-based media has the capacity to support. The model for the final “Generative Design Research” plan was informed by a research study conducted using “The DNA Workbench” (Price & Berezina-Blackburn, 2008). This study compared learner awareness in single-user and collaborative, multi-user scenarios in order to isolate the most salient differences between these scenarios. I discovered some interesting benefits that may also be relevant benefits for gestural interaction, such as the person-to-person communication and learning benefits that arise from observation of knowledge externalized by others through gestures.

The methodological approach of the primary “Generative Design Research” project will be introduced in this sections, but detailed analysis of the final results will be discussed in the following section. The final research study involved a participatory approach. Several test subjects viewed an array of video clips and were then asked to retell what they had observed using gesture. Many of the early prototypes discussed in this chapter led to the development of the research objectives that were the focus of the final study.

As will be discussed, much of the complexity involved in designing gestural interactions relates to the lack of research precedents available to individuals working in this area. Many of the projects in the section could be considered on the fringe of what I am beginning to define as gestural interactions, including areas of research such as tangible media and surface computing. Each of these types of interfaces possesses distinct areas of intersection with gestural interaction design through the multimodal and kinesthetic concerns involve. The final research study proposes a new way of thinking about interactivity in terms of kinesthetic experience that has applications beyond the realm of purely gestural interactivity.

Figure 16: "Digital Seed" containing waterproof RFID tags

2.1.1 Digital Seed Project

Prior to developing the “Digital Seed Project” I was inspired by several research projects in the area of research known as Tangible Media (O'Malley, C. and Stanton-Fraser, D., 2004). Tangible Media developed in part from the notion of *Tangible bits, which* were first discussed by Hiroshi Ishii (1997). In “Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms”, Ishii discussed “what we have lost in the advent of computing", namely the "richness of the physical world" (ibid). Unlike traditional, mouse and keyboard driven interfaces, tangible interfaces "attempt to bridge the gap between cyberspace and the physical environment by making digital information (bits) tangible."(Ibid.) There are several processes that tangible media applies in order to create a sense of directness (O'Malley, C. and Stanton-Fraser, D., 2004):

1) Externalizing computer processing that typically take place internally
2) Remapping the relationship between users actions and computers response by coupling computer output with physical metaphors that relate to real-world prehensile engagement & spatial reasoning
3) Representational mapping that symbolically couples an interaction with the resulting effects

As a designer, I often work by responding to physical materials. As I became more interested in the potentials of computing power I felt that my computing experiences did not provide adequate opportunities to explore through physical interaction. I felt that there was a great deal of potential to communicate to users through the physicality of designed objects, but was frustrated with the limitations of the mouse and keyboard as a mode of representation. I was drawn to the ideas presented by researchers working with tangible media because they demonstrated ways in which knowledge from kinesthetic experiences could be applied to create computing experiences with a stronger sense of physical engagement. I saw this as an opportunity to enable new forms of interaction that supported more diverse forms of sensory engagement.

I was interested in creating education media that would support the development of kinesthetic forms of learning. Prior to developing this project, I had also been researching the life and work of Maria Montessori (1962). Montessori believed in the importance of designing materials for educational environments that could model abstract concepts and allow student to independently explore concepts by physically manipulation of designed objects (ibid). After further research, I found researchers developing tangible media that encouraged Montessori-style education through the development of computationally enhanced versions of Montessori materials. These materials were given the name of digital MiMs or Montessori inspired-manipulatives (Zuckerman, O. , Arida, S. , and Resnick M., 2007).

These ideas inspired the first research project to be discussed, which I now referred to as the “Digital Seed Project”. This project focused on creating a situation in which learners could interact with computationally enhanced objects and discover further information through physical engagement. I started with a notion that information can be embedded into objects with unique identifiers, like RFID tags or barcodes. Retailers use these technologies to access and synchronize information from vast databases and map this information through indexing serial codes. I decided to apply this technology to learning materials. The concept of using a seed seemed appropriate, because recognition of a seed could be metaphorically coupled with information related to photosynthesis. Like seeds, many physical objects possess information that can be unlocked over time through interaction. Examples might include gears, levers, or simple machines. In the case of seeds, a world of different potentials is unlocked through its interaction with water, soil and sunlight. I was interested in embedding information about photosynthesis and the plant life cycle into physical seeds. My vision was that students could unlock this embedded information through activities analogous to real-world interactions with seeds. I began to consider how I could remap the relationship between the users' actions and the computers response by coupling the computer output (animations of plants growing or dying) with the users actions (watering and "planting" digital "seeds"). I created a digital experience in which a student could place a seed inside a pot, water it and watch it grow.

Each plant would have an accelerated life span, which could be dependent upon factors like the amount of water or the time passed since watering. I was able to sketch out initial plans based on foreknowledge of several technologies, including the Parallax Basic Stamp (See Figure X). With the basic stamps I was able to connect small sensors to a PC through a serial port connection. Using the DirectCommExtra for Macromedia Director (Berda, Tomer, 2004), I could receive signals from the basic stamp and then triggered animations of plants growing and dying. RFID tags could be embedded within sculpted, painted seeds. When the seeds were placed within a small pot situated near a RFID reader, the RFID reader could send the serial number related to each RFID tag to Director and trigger an animation that related to a each specific seed. A moisture sensor was also situated below the pot and connected to the basic stamp. When the seed was watered using a small watering can, the moisture sensor would trigger an animation of plant growth. Using the combination of several variables including presence of RFID tags, ID of sensed RFID tags, and presence of moisture, animations of 3 plant types could be triggered to grow or die. Three seed types were sculpted to resemble corn, bean, and sunflower seeds and waterproof RFID tags were placed inside each. When seeds were placed in the pot, then watered, a simulated of the corresponding seed projected above the pot could be triggered to grow.

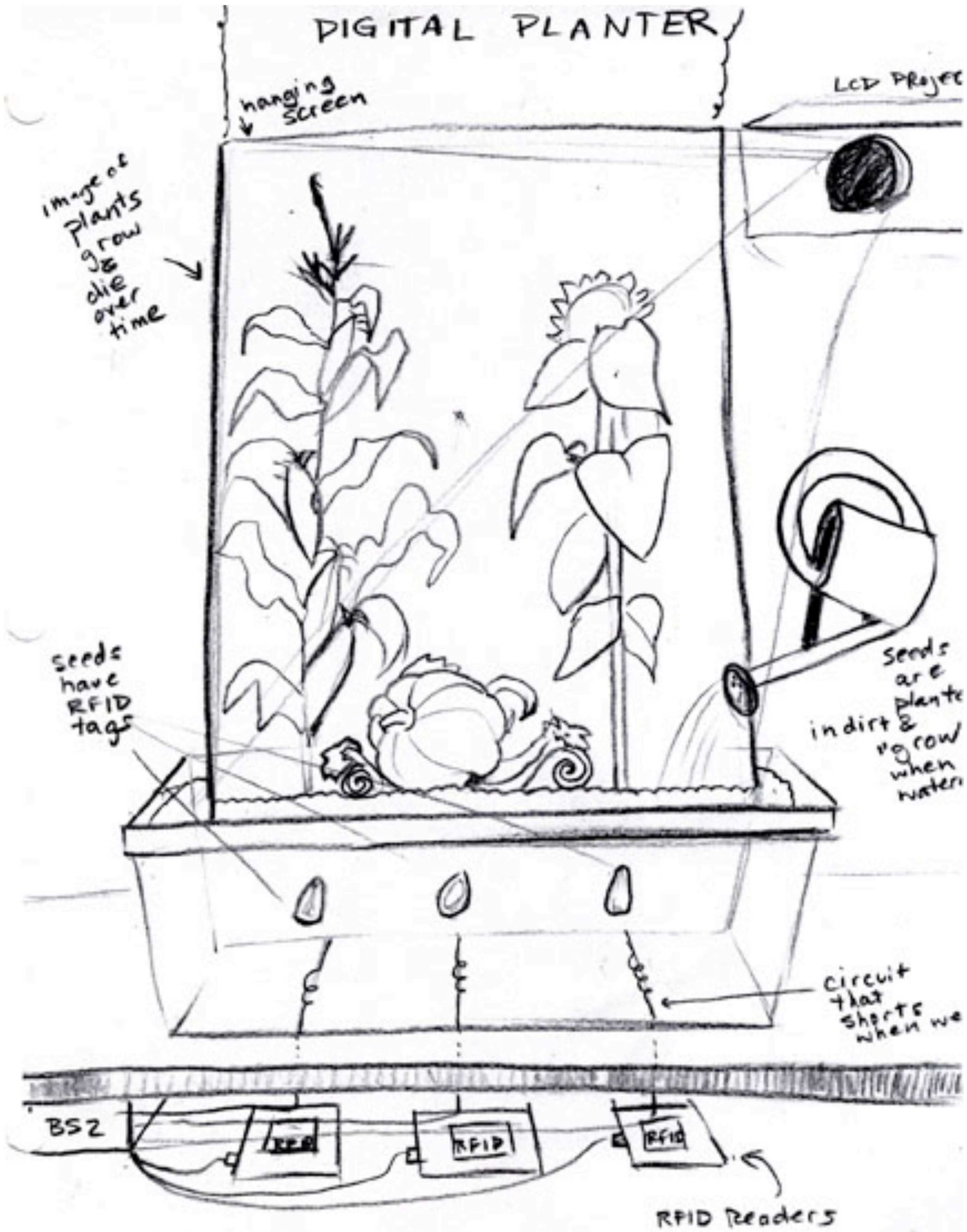


Figure 17: An early sketch of the digital seed interface

The project was completed for a public exhibition in autumn of 2006. Most participants were surprised by the interactivity, and expressed that they found the project to be very unique. Several individuals played with the interface and tested the growth of each seed with various applications of water. After realizing that some plants were living longer than others, some individuals began to experiment in order to see if the length of a plants life was related to seed type, time of planting, location of planting, or application of water. Several individuals became aware of the fact that the plants in the water were sensitive to the application of water. For example, excessive watering would kill the plant more quickly and so would not watering a plant at all.

When people were presented with the interface they often tried to understand the logic of the system through physical interaction. People attempted to alter their behaviors systematically in order to understand the deeper logic within the system. They attempted to trick the system to see if they could find an error. This suggests that there is some benefit to constructing systems with discoverable systems of logic that can be observed through varied patterns of interaction.

Public exhibition of the project brought several issues to the forefront of my mind. Firstly, I became more aware that it was possible to create opportunities for participants to discover information through self-directed physical engagement. People can learn about the real systems that a simulation represents when they can use their own unique kinesthetic strategies to explore related concepts within a simulation. Giving participants an opportunity to test the concepts kinesthetically can encourage kinesthetic problem solving. If simulations lack realistic physical reactions to kinesthetic interactions, the outcomes can easily misrepresent real systems or cause the people using the system to have doubts about what they have learned. Providing alternative kinesthetic entry points creates an important sense of realism for users.

The exhibition made clear the need for greater robustness. After the exhibition, the project was damaged due to inadequate water seals. While mixing electricity and water would send up red flags for most people, I felt that using water was a worthwhile experiment because it demonstrated that natural and technological systems could be integrated. Water and electricity can mix, although in the future I know that I will need to approach the problems differently and include strong seals. The initial interface inspired a large range of unpredictable reactions, which I had not fully anticipated. Many responses to the interface seemed to demonstrate a desire to test its boundaries. I feel that the use of water was successful in that it inspired people to test the boundaries of the interface in way they may have likely never imagined. In the future I need to be more prepared for the large range of explorations, which included large water spills, overfilling and even water flinging.

Looking back on the exhibition, I feel one of the largest shortcomings of the seed project was the restriction of movement created by the interface constraints. In order to use the interface participants were required to act out a specific pattern of motion, which was partly an undesired result of the limited range of the RFID reader. For this reason, the design was that it was not *discoverable* (Saffer, 2008). For an interface to be discoverable it must make clear to the people at the exhibition how to approach and interact with it. The seed interface was not discoverable for several reasons. Perhaps one reason was that the interface was uncommon and unfamiliar. When confronted with a digital interface, most people would hesitate to pour water anywhere in the vicinity. They are not likely to consider using water to interact with it unless they are encouraged to do so. Before anyone could use the interface or understand its capabilities I had to guide them through the process of using it and explain a seed placement and watering scenario. Participants had to place a seed directly into the pot and then water it. Perhaps this issue could have been addressed with signage, but ideally the design should invite these interactions. Restrictions created by the range of the RFID readers restricted the space in which users could interact and discover. I had internalized these concerns and would attempt to address them in future projects.

Shortly after the show, I became interested in the nature of the movement created by a physical interface, an issue that I would investigate in future works. I realized that several of the concepts such as *physical metaphors for computing* were very critical, not only in terms of the graspable objects used to control the interface, but the movements and gestural metaphors performed with the objects in hand. For example, I felt that the act of pouring water was an important element related to learning about the sensitivity involved in caring for plants. While the complexity of an action such as pouring is often overlooked in traditional education, in Montessori-style education it is understood that there is a great deal of sensory development involved in this task for learners in early stages of development. For example, pouring too much water into a pot too quickly could cause the pot to overflow and dirt to spill all over the floor. Developing the muscle control that enables one to avoid spills requires some practice because it require one to simultaneously observe and act. Successful gardening also requires sensitivity in terms of observation and physical experimentation.

It is often true that one gesture or action in the physical world can metaphorically represent a larger more complex action within a coupled digital environment. For example, a short-cut key can trigger a complex function. Toggling a button sets the system within an authoring environment, and can alter the meaning of consecutive actions in a state dependent manner. Computer interfaces tend to abstract gestures from the real world. For example, the “drag and drop” action that is performed with a mouse is not at all like a drag and drop action in the real world, yet we can somehow still abstractly equate the two. I began to question the implications of the mapping systems that connect physical actions to their digital representation.

With the seed project, I began considering alternative ways in which the action of “pouring water” could be represented with a tilt-switch. I wondered if for example, could the tilting of the water pot represent pouring water if the real water was removed from the scenario? I also began to consider the act of “planting a seed” and how to represent or signal that a seed had been planted. Is the seed planted when it is inside the pot or within a given proximity? It became clear that each scenario I considered began to present problems of logic. For example, it was not my goal to communicate that a seed must be in a pot or in a given location in order for it to grow, yet the spatial relationship between the RFID tag, the RFID reader, and the projected plant growth animation created unwanted restraints in terms of proximity. I realized that the technological constraints were beginning to create symbolism that contradicted several aspects of the natural cycles of photosynthesis. I had no way to resolve these issues at which point I choose not to pursue the project further.

After abandoning the Digital Seed project, I began to consider the symbolism that arose from the coupling of physical actions to digital reactions. I began to consider the potentials available to designers willing to explore the effect of this phenomenon upon users’ experiences. There seem to be a great potential for embodiment or engagement that I began to imagine as a result of my work on the Digital Seed Project.

Despite the many shortcomings of this project, it helped me to realize that there are many complex relationships between physical actions and their digital representation that combine to create a users experience. Believability, discoverability, and usability are all tied to virtual cues that provide information that presents a simulation’s capabilities, limitations and the current status. I realized that, through careful design, virtual simulations could present feedback that people could use to contextualize and internalize the significance of their gestures or actions in meaningful ways.

2.1.2 Data Glove Character Controller

Moving forward I began to explore various ways in which body movements could be used to signal or cue actions within virtual scenarios. I wanted to create a scenario in which the meaning of real-world physical actions could be re-contextualized in order to generate meaningful relationships to virtual scenarios. I considered how a manipulation of real-world humanoid armatures could map on to various virtual outcomes.

During this time I began to consider the expressive and communicative potential of the human body. I discovered the work of Paul Ekman, Edward T. Hall, and Ray L. Birdwhistell who were exploring codified and systematic approaches to the interpretation of gestures, body postures and non-verbal communication. Each of the aforementioned researchers contributed to my understanding of non-verbal expression in different ways and helped lead me towards related research. Paul Ekman has done extensive research in facial expression, working to develop a Facial

Action Coding System (FACS), which has been used to distinguish the meaning in facial expressions through discrete analysis of facial postures. His research was quite interesting because he breaks down facial expressions into small meaningful units that make clear the existence of consistent relationships between form and meaning. This helped lead me to the idea that there was a deeper logic within postures of the face and body that could potentially be systematized and made computationally accessible. I began to feel that a system of postures could be used to trigger actions within virtual scenarios. Like Ekman, Ray L. Birdwhistell has written extensively about the elements of non-verbal communication that reveal intentions and emotions. This helped me to consider the subtleties of non-verbal expression, like muscles tensions and postures. Edward T. Hall was the founder of a field of research called *proxemics*, which studied the impact of physical environments on human communication and behavior. This also helped me to develop the notion that environmental context impacts the meaning of a postures or gestures.

I began considering the contextual interpretation of body postures within virtual environments. I explored ways in which a physical skeletal armature could be used to control the movements or actions of a virtual character. Through research I discovered the Monkey 2, a commercially available armature designed to facilitate 3D character animation. (Inition, 2007) Knowing that an armature of this nature could be created, I conceptualized an interactive system in which poses of a real-world armature could act as triggers for a virtual character's activities (Kuhlman, 2007). I felt that poses of a physical armature could become signals that created a method of character navigation. Codified relationships could be established between the characters pose and his actions in the virtual environment. I considered the body poses used by drill sergeants to direct their troops and the signals used by air traffic controllers to direct pilots. I considered other postures that could act as signals. For example, when adopting a running pose with deep forward lean he would run or perhaps change direction. More significantly, I began to consider how certain body motions could have "context sensitive" meaning relative to the location of the character in the virtual setting. I thought, for example, perhaps extension of an arm on the physical armature could cause the virtual character to pick up a nearby object in the simulation or open a nearby door. If a virtual character had a specific object in hand, like a basketball, perhaps extending the arms could cause them to pass the ball.

During this project, I realized that the gestures and postures generated by the user could be interpreted based on the status of the virtual entity that the gestures are meant to represent. In this case, the position of the character within its environment and its relationship to other objects, characters and entities also affected how a gesture should be interpreted. The same pose or movement could have various outcomes depending on the context in which it was used. This allowed an otherwise one-to-one mapping of virtual character to physical posturing to be extended in a manner that would allow simple gestures to represent complex sets of movements. Rather than the non-representational methods used in puppeteering, where the wiggle of the finger equals the wiggle of a limb or the press of a button equals jump, a system of symbolic representation would be tied to the context in which it was used.

Developing an armature was complex given time and budget constraints, so I took the project a slightly different direction by implementing the 4D dataglove. It was much simpler to test these concepts with a prefabricated interface that was readily available. Using the dataglove, I was still able to test the basic concepts considered previously. I could enable virtual character control through hand-poses that acted as signifiers. The data glove output signals related to the pitch and roll of the hand and bend of the fingers. It also output gesture codes based on combinations of bent and extended fingers. To develop the hand signals I would use as cues, I reviewed several hand gestures used in American Sign Language (Michigan State University Communication Technology Laboratory, 1997) and considered hand poses that could be recognized using the data glove interface. I decided I would be able to implement modified forms of the signs for jump, walk & sit.

Figure 18: Data Glove

I developed the simulation to test these concepts in a computer program called Virtools. I created an environment using some of the existing 3D assets from the Virtools environment, including a 3D Environment and animations cycles from a virtual character named *Eva*. The programming environment in Virtools uses a schematic programming interface with building blocks that can be arranged and connected in order to create interactive structures. In addition to standard built-in blocks, I implemented a unique building block developed by Professor Alan Price, which received data output from the 4D data glove through a serial port connection (Price, 2007). Using the schematic interface, I programmed the structure to control Eva's actions using the data glove (Kuhlman, 2007).

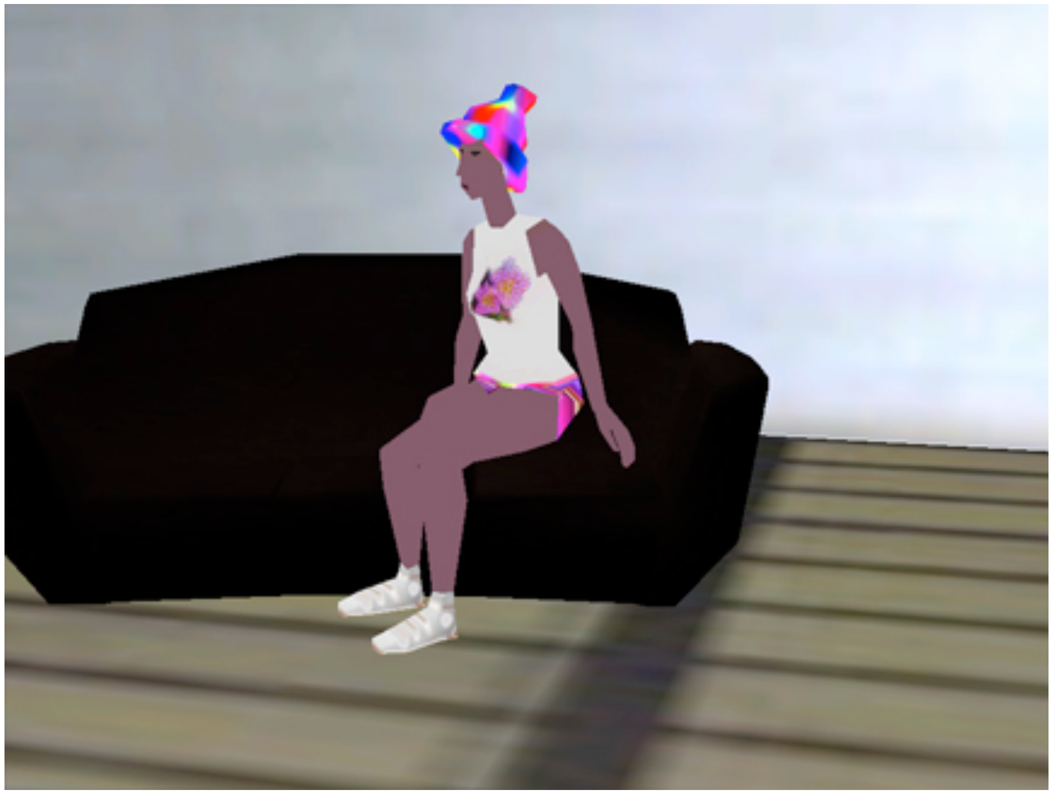


Figure 19: Eva, a virtual character, moves to the couch and sits down

When the index finger, middle finger and thumb are extended Eva walks. The rotation of the wrist determines the direction in which she walks. Extending the index finger and thumb simultaneously will trigger Eva's run cycle. Rotating the wrist makes Eva run to the left of right. When the thumb alone is extended and rotated upward Eva jumps. When the user makes a bunny ear or quotation mark gesture by bending the index and middle finger repeatedly, Eva sits down. The sit animation is only triggered if Eva is near a chair, in which case she will walk toward it, align with it, and sit down. (see Figure 19) Making a fist and then rotating the hand at the wrist to the right and left repeatedly will cause Eva to reach for door or drawer, but only if she is in close proximity to one.

This project was presented during an open house event at the Advanced Computing Center for the Arts and Design in the spring of 2008. Before operating the interface each person received a quick demonstration of all the gestures that could be used to control Eva. Most people were able to remember and repeat the gestures after watching a quick demo, but the quality of performance often varied depending on how well the data glove fit. Unfortunately, the size of the glove made it difficult to recognize the gestures of people with hands that were smaller than the glove. Despite that issue people could quickly understand how to walk, run and turn. Most were quickly engaged by the interface. Some people mentioned arm fatigue after extended use. After demonstrating the interface for several hours, I also developed very sore wrists from the subtle turning needed to control the character.

As mentioned previously, the gestures used to control Eva were developed from modified American Sign Language. I use the term “modified sign language” because when I was exploring the sensory capabilities of the data glove, I knew that the data glove did not have the capabilities to recognize all the spatial qualities of these forms of sign. It could merely recognize the pitch and roll of a single hand, as well as which fingers were extended or abducted. For example, the American Sign Language sign for sit involves curling the fingers of one hand over the fingers of the other hand.

After I showed people the American Sign Language sign for sit, many people found it easier to mimic my presentation of the full sign language form, even through system would only react to the double finger bend motion on one hand. When this motion was performed alone or as part of a larger motion, the system saw it the same way. What I came to find out during public exhibition of this project was that, although the glove did not sense information such as the spatial movement of the hand or the relationship between hands, many people found it easier to use both hands to mimic the sign language forms. Perhaps this is because it is easier to curl the finger into the desired shape by forming them around the other hand. This posed interesting questions that brought to my awareness the notion that unique hand shapes could arise as a result of one hand being influenced by the other or the shape of a hand being influenced by the surface of an object.

In addition to controlling the virtual character named Eva, another application was developed that allowed users to create a simple two-pose animation. By pointing up, down, left and right the users could navigate the joint hierarchy of a boxy figure. Opening and closing the hand would trigger the interface to switch between a joint-hierarchy navigation mode and a joint-rotation mode. In the joint-rotation mode, pointing up, down, left and right would rotate whichever joint was currently selected in the X, Y or Z axis respectively. Users could set the position of figure’s initial position, set the position of the final position and the trigger animation between the two states.

This pose-to-pose simulation and the Eva character controller discussed previously made several dilemmas related to hand pose navigation clear. This system proved to be difficult to navigate in some ways, and easy to use in others. When controlling the Eva character, the simplest interactions were walking and running, because both of these processes presented continuous feedback, signaling that the hand gesture had been recognized. Eva responded by either walking or running. It was very clear to the user whether their hand signal had been read because they could see the response immediately. Likewise, when rotating the joints of the block character it was clear whether or not the appropriate signal was being received from the data glove because the joint would rotate accordingly.

The more difficult tasks, and the tasks which generated the most fatigue and frustration for users, were the tasks that were triggered through sequences or combinations of gesture signals. For the Eva character controller, difficult gestures included those for actions such as jump and sit which require the hand to be positioned more precisely and follow a more specific pattern. For example, the jump gesture required that the users start at a horizontal thumb-extended position and then rotate their wrist upward and to the right while maintaining the thumbs-up position. Sitting required the index and middle finger to be bent and extended twice while the hand was upright and while Eva was within a limited proximity to the sofa. For the block character-posing interface, it was difficult to select and navigate the joint hierarchy. Selecting a joint was difficult because it required users to open and close their hand twice in a row. Often when a system response to the gesture was not initially recognized, people would begin repeatedly opening and closing their hands waiting for the intended response. Because the same gesture was used to switch back and forth between the joint-rotation and joint-navigation modes, continuous repetition of the gesture would cause the modes to switch and then quickly switch back again.

With the glove interface there were frequently dilemmas with the sequencing and timing that triggered the simulations response. If a gesture acted as a trigger or button push, as apposed to altering or adjusting a variable, there were frequently problems with continuity. For example, joint navigation was difficult because pointing up, down, left and right was at times recognized by the computer faster than the users could react, causing people to sometimes overshoot the joints they were targeting. If the user was enacting a sequence of gestures and one gesture of a sequence was missed, this would not trigger a system response. When these problems occurred, users had no way of knowing which part of their gesture was not recognized, which made it difficult for them to adjust their behavior and improve the outcome.

Through this project I learned of an additional layer of complexity involved in creating a system that interpreted the meaning of human gestures. Gesture is not simply a codified system of hand poses of body postures. Interpreting the meaning of gestures requires a system with sensitivity to subtle elements of timing and ordering. Within the realm of usability research, it recognized that it is important to give users opportunities to understand the status of the system and provide them with information that they can use to recognize and interpret the cause of system errors, which will inevitably occur at times. For traditional interfaces, this is a manner of addressing the point along a linear path at which a problem occurs. For gestural interfaces, addressing this issue is much more complex because of the multimodal and multi-threaded nature of gestural interactivity. Creating a scenario that involves sequencing and timing of subtle hand gestures means that there are many stages at which an error could occur. Developing strategies to detect these errors becomes much more complex. Simultaneously, it becomes even more important that users are aware of how the system sees them, so they can alter their actions in a manner that allows the system to see their actions as they intend them to be interpreted.

2.1.3 Geometry Drawing Table

Through the dataglove project, I began considering how the computer could be used to interpret expressive motion from human hands. I had learned a great deal from this project, yet I knew I still had a great deal more to learn. I began further research into sign language and gesture studies (see section 2.1.1.3 for a review of findings). I became convinced that gestural interactive systems could be beneficial when implemented within an educational setting. I had found some publications that discussed the benefits of using gestures to learn about various concepts in geometry and mathematics (Goldin-Meadow, S., Kim, S., & Singer, M., 1999) For the next interface I developed, I wanted to generate examples that demonstrated how gesture-based media could help students to learn kinesthetically. Geometry was an area of study for which I could imagine many beneficial applications of gesture-based interactivity.

I was particularly interested in developing interfaces that could work in classroom settings. This required a solution that could be effectively used by many different users. The dataglove interface was too limited for several reasons. First, the dataglove was very expensive, which meant that it would be difficult to find schools willing to invest in the technology, regardless of its educational potential. In addition, most datagloves have a set size, close to the size of the average male hand, which meant many students could not use the dataglove. I needed a solution that would not limit educators' access based on cost or complexity.

Moving forward, I began conducting interviews with teachers regarding their use of technology within their classrooms. I also discussed their views on incorporating kinesthetic learning into the classroom (see appendix). After interviewing several teachers, I learned about a technology that was increasing in popularity, called a Smartboard. *Smartboards* extend the role filled by the tradition chalkboard or whiteboard by allowing digital layering through a digital projection and digital pen recognition system. Teachers can draw over projected imagery and also record notes that can be accessed later or digitally disseminated to students and their parents for later review.

After talking to teachers from various schools, I learned that they had felt the technology would be useful but was out of reach due to budget constraints. I found one interesting solution to this problem when I discovered a demo on the TED website by Johnny Lee Chung. He had developed a simple multi-touch interface using the small digital camera inside the Nintendo Wii. His solution would allow teachers and students with access to a forty-dollar Wii-mote to download a free application that could transform a flat surface into a smartboard-like system.

I felt this was the level of financial accessibility that was required in order to create affordable and accessible solutions for teachers and students. I knew at this point that a video-based solution could potentially overcome some of the barriers that limit the use of gestural technologies in the classroom. Because teachers have very busy schedules and many educational standards that they must constantly struggle to achieve, classroom technologies need to be cheap, easy to setup and use, and be a significant benefit to students.

After discovering the potential benefits of camera-based interactivity, I began researching video-based motion tracking systems. With a video-based motion tracking system, I could avoid the dilemmas that I had faced with both the digital seed and data glove projects because interaction was no longer limited to a specific physical construction or tied to a wearable device. Video-based solutions would allow access to all participants who possessed a digital camera. Video-based motion tracking seemed like a more reasonable solution for schools because a digital camera could serve multiple functions that were not dependent on one specific application and was much more affordable than a smartboard.

My knowledge of computer vision was limited, but I had learned about several open source applications that I might be able to apply. The first computer vision engine I explored was called EyesWeb. (Camurri, Antonio, Mazzarino, Barbara, Volpe, Gualtiero, 2003) Eyes web is development platform designed to allow video analysis of expressive gesture. It interprets video data in real-time so that it can be used to control interactive applications (Infomus Lab, 2007). I ran into limitation related to my programming experience and eventually found another framework I could work with more easily, called Reactivision (Kaltenbrunner, Martin , Bencina, Ross. 2007). The final application utilizes Flash OSC (Chun, 2006) and Reactivision to create a table-top tangible user interface, or TUI.

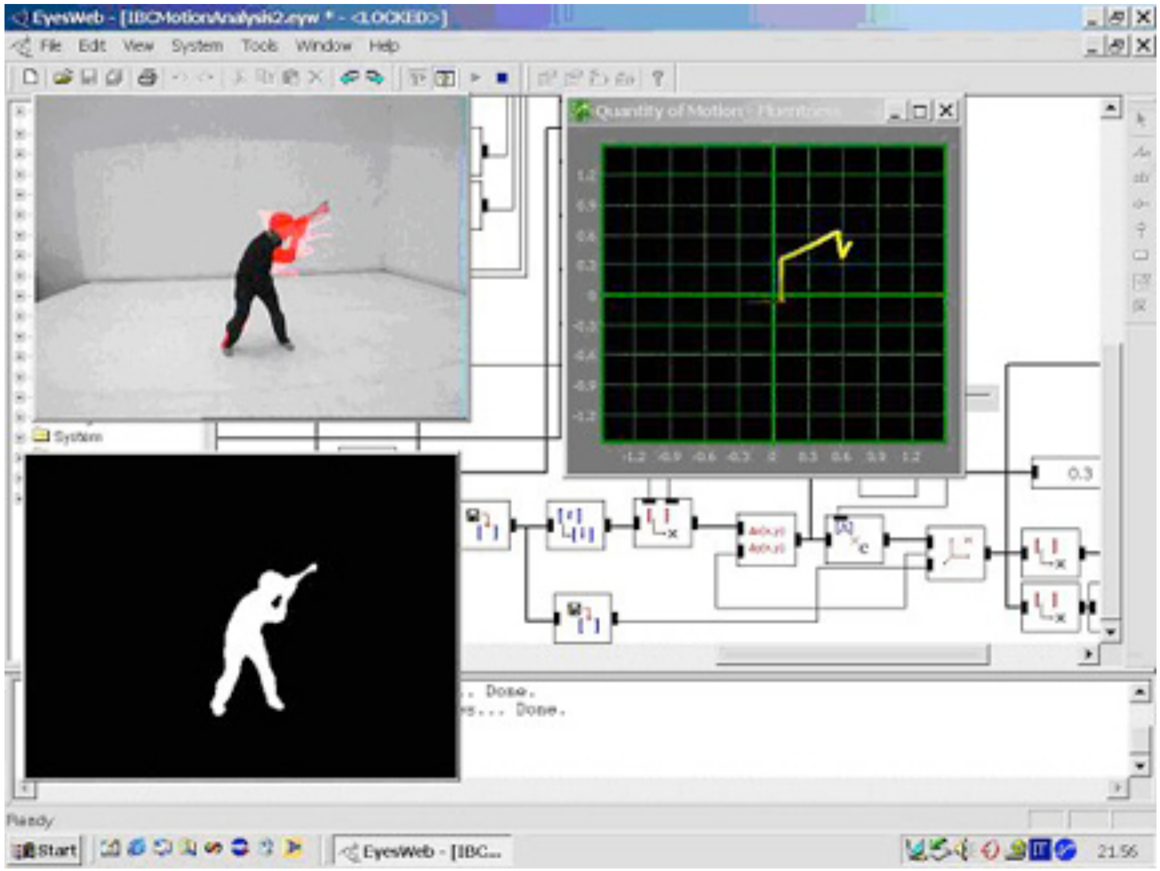


Figure 20: The Eyes Web development environment

Several early tests were conducted during the development of the project's code base. During these tests, I developed a framework that would allow data handling related to messages sent via the reactTIVision engine. The reacTIVision engine interprets information from a small video camera. It detects the location and rotation of special *fidicial* markers (see Figure X) relative to the 2-dimensional camera plane and then sends this data through FLOSC to flash. It also sends data regarding the placement of circular blobs, which it interprets to be fingertips. Each tool was associated with a unique ID, so presence or absence of each tool could be used to change drawing modalities and manipulate geometry assets.

After initial tests, I proposed a plan for the design and development of a geometry-drawing interface that could be used in education settings to teach geometry concepts through a tangible, kinetic interface. Several concepts were inspired by exercises available to users of the geometer's sketchpad software (KCP Technologies, 2008). This mouse driven interface utilized kinetic manipulations of points, lines and geometric relationships. I found that several of the concepts addressed through the mouse and keyboard interface could be addressed through this new interface and could act as a grounds for comparing multi-modal, multi-user interfaces to a more traditional mouse and keyboard driven system.

Due to the significant time involved in developing such an interface, mock-ups were developed to test the interactive concepts. This involved video recording hand motions enacted in front of a green-screen then compositing this video with motion graphics that illustrated potential plans for a graphic interface responses to movement performed with tangible tools. Through this process I discovered that some concepts were more successful than others.

The *Geometry Drawin*g Table has gone through several iterations. The first iteration was on display for the open house at the Advanced Computer Center for the Arts and Design in the spring of 2008 (See Figure X). At that point, its capabilities allowed several participants to interact with the drawing table simultaneously. Each person was given two tools, a *modifier tool* and a *drawing tool*. When the tools were used alone or in combination they could produce varying results. When the *modifier* was placed on the table alone, a menu appeared. Turning the modifier would highlight various icons on a menu that represented drawing modes. Lifting the tool from the table would select whichever mode was currently highlighted. The *modifier* could be used to select one of 6 drawing modes: Free-Hand, Line, Triangle, Square, Circle, and Erase. The drawing tool could be used to draw shapes, depending on which mode was currently set. In the Free-Hand mode, the drawing tool could be used to draw squiggles and curves. While in line mode, each time the tool was placed on the table and then lifted off, a new point was set and a line is drawn to any previously set points. While in shape drawing modes, placing and lifting the drawing tool on the table would draw circles, triangles, and squares. Rotating the tool while in a shape mode would rotate the shape currently being drawn on the table surface. Touching the drawing tool over the drawing while in *erase mode* cleared the entire screen. Placing the modifier tool on the table at the same time as the drawing tool allowed new kinds of interaction. With both tools on the table, rotating the modifier would set the drawing color instead of the drawing mode. While drawing shapes, increasing the distance between the two tools would scale the size of the shape up, and moving them closer together would make the shape smaller.

Figure 21 - The Geometry Drawing Table

After the first public exhibition of this project, I learned a great deal by observing visitors and listening to their feedback. The second iteration of the table included several revisions based on these discoveries. First, I discovered that several people had difficulty reading the imagery on the modifier menu. While developing I frequently tested the interface on a screen, but during the exhibition the imagery was project on a diffused transparent surface from underneath. This resulted in a lack of contrast and subtle color shifts that made it difficult for some people to see the icons and detect whether or not they were highlighted. New menus were designed in the second iteration with more optimal contrast and more readable status indicators for highlights.

During the exhibition, I found that *system status* was not clearly communicated to users. This is one important issue for an interactive application and an issue that Jakob Nielsen addressed this issue in his *Ten Usability Heuristics* (Molich, R., and Nielsen, J. 199o). He wrote:

> "The system should always keep users informed about what is going on, through appropriate feedback within reasonable time. "

The following version attempted to address this issue in several ways. First, different sound bytes were added to indicate when a tool was added or removed from the table and when the drawing mode had changed. This would help the users know whether or not the tool and its movement had been recognized. Secondly, visual menus were optimized to help provide feedback about system status, such as color settings and shape modes.

After the initial exhibition, several improvements were made to the reacTIVision engine, which improved the quality of the computer vision and finger tracking. This helped to simplify some of the difficulties that arose from in the earlier version. For example, in the first version of the table the system did not accurately read markers near the edges of the table because of distortion created by the wide-angle lens on the camera near the underside of the table. Background subtraction was added to the new version of reacTIVision, which counter-acted some undesirable affects that arose from uneven infrared lighting on the undersurface of the table.

Despite improvements in the computer vision engine, there still exists a potential for vision errors that need to be addressed in the design. For example, the initial design had several drawing modes that were selected each time the modifier tool was lifted off the table. This caused a dilemma when a tool moved to an area on the table where distortion, occlusion, or inadequate lighting cause the marker to be temporarily unseen, causing a drawing modes to be accidentally selected. On several occasions during the public exhibition, people were startled when the entire image they were creating was erased because they had unintentionally switch to erase mode. To account for this dilemma, the new menu system has changed the method in which erasing works, making it unlikely that the erase mode will be accidently selected. When a tool is first recognized it is always set to a null position by default so that other shape modes cannot be accidentally selected.

Another issues that become apparent is that some functions, such as line drawing, could be best utilized through finger touch. Shape drawing was difficult with the large tangible tools in the initial prototype because it blocked the view of line drawing operations.

Although gestural interactivity has been a major interest and influence at this stage in the research, the interface as described thus far does not fully address the definition of gesture that has been previously introduced. Although not purely gestural, this interface does take on many of the design problems that I believe are at the center of gestural interface design. From the beginning, development and planning of this interface has considered the potentials of an interface that would allow students to interact with multiple factors in a geometric problem simultaneously. Gestural interfaces must address similar concerns because gesticulation is a multi-channel form of communication. Many factors combine to form the meaning and affect each other simultaneously. Digital environments allow access to pure geometric space in a way that is not always feasibly represented in reality. Multimodal interfaces further extend these potentials by allowing multiple methods of simultaneous influence that are similar to real-world kinesthetic experience.

By addressing multimodality, this project was a small step toward developing a gestural interface. Aspects of this interface that begin to adopt gestural qualities include recognition of interrelationships between points of input. For example, the shape scaling that is control by the distance between tools. Many of the interactions in the Geometry Drawing interface involve contextual interpretation of the circular rotation and position of the tangible tools. I am currently considering methods of finger pattern recognition that could be incorporated into the interface as well. Finger tracking capabilities that have recently been incorporated into the reacTIVision engine have led me to consider other ways to incorporate gestures like shape and movement patterns produced by users fingertips. Beyond pointing gestures, there are several possibilities that are being developed. Grasping gestures have been considered for placing, transforming, and removing shapes, lines and points. Developing finger pattern recognition is a much more complicated task to develop, so I felt before committing to a development approach, I began to develop plans for user-centered research that would help me to determine the most worthwhile approach.

Overall, this process made clearer the complexity of the task, but did not resolve issues of usability and discoverability, which had troubled me throughout several of my previous works. At this point I continued to develop the framework for several of the more successful concepts I developed, but it became apparent that there was a need for more systematic research. In order to better understand what constitutes a usable and discoverable gesture-based interface, I first needed to reference a group of gesturers to see if there were useful patterns that could be transferred into computing experiences. I was looking for research that could tell me, for example, what is the best way to communicate actions like “connect”, “bisect” or “combine” through gestures. The issue was also complicated by the question of implementation. I realized that documentation about successes failures related to user experience and technical implementation were very important for gestural interface designers to create more effective applications. I proceeded to work towards creating a study that would clarify this approach.

2.1.4 Usability testing with the DNA Workbench

All the projects previously discussed helped to make the need for usability research related to gesture-based interactivity become more apparent. I was interested in developing a research study that would present feasible strategies for further user centered research. Currently, most literature on usability and user centered design research focuses specifically upon web and software development. I found only a few publication related to gestural interaction and usability (Nielsen, Michael, Storring, Moritz, Moeslund, Thomas B. and Granum, Erik., 2003)

I began to hypothesize that the gesture sensing technologies with the greatest potential for collaboration included touch tables, tangible user interfaces, and video-based motion tracking. These technologies allowed multiple points of input. This means multiple people can interact simultaneously. By incorporating these technologies into educational scenarios, I believed the potential for collaboration could have a positive impact on learning outcomes. To be sure, I needed to test this hypothesis.

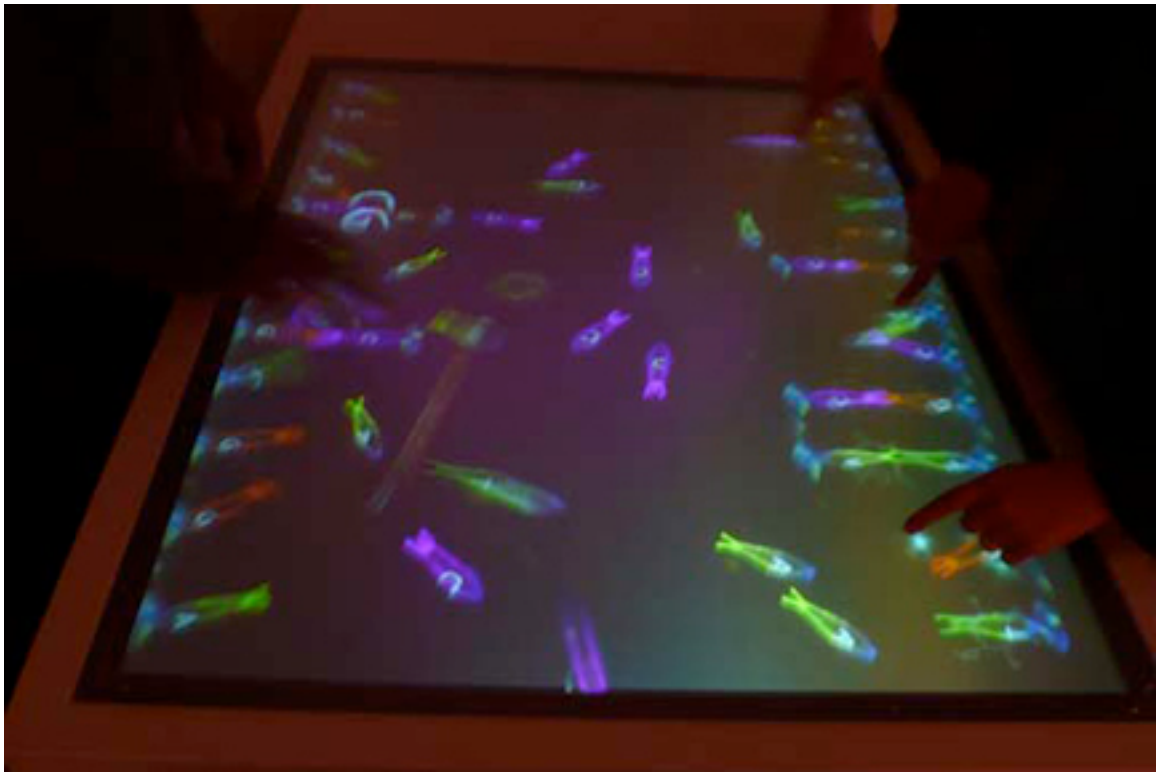


Figure 22 & 23 - The DNA Workbench

I approached Professor Alan Price who had recently exhibited a multi-user interface that was on display at COSI (Price & Berezina-Blackburn, 2008). Professor Price collaborated with several other researchers from the Ohio State University including Dr. Susan Fisher from the College of Biological Sciences, Vita Berezina-Blackburn from the Advanced Computing Center for the Arts and Design and Norah Zuniga-Shaw from the Department of Dance. They worked to develop a multi-touch, multi-user interface called the *DNA Workbench* (Ibid). The user interaction helped COSI visitors to intuitively learn the order of DNA nucleotide pairing by interacting with the system. Alan allowed me to use the *DNA Workbench to* conduct a research study that analyzed the differences in user experiences that arose when participants interacted with the DNA Workbench. The study compared the learning outcomes of individuals interacting alone vs. individuals interacting as part of a collaborative group. Several key differences between the two test groups were discovered:

1) Individual interacting as part of a group recalled more information about DNA after using the table
2) Individuals interacting as part of a group felt they had exhausted their potential for learning much quicker than individuals interacting alone.
3) Individuals in groups were able to articulate issues, problems and

questions when they arose because they had an opportunity to debate and share information related to a topic with other users. Individuals who acted alone expressed confusion about several topics but had no one to whom they could address their questions.

4) Individuals who were acting as part of a group were more confident with regards to the information they learn and are more likely to express information based upon their assumptions

Gestural interfaces have great potential in terms of collaborative computing. After conducting this study it became clear to me that we have only begun to explore multi-user scenarios. For both design and development it bring new considerations to the table. For example, interactive schemas should no longer be thought of as linear sequence of static screen with a limited set of choices to be made at any given time by a single user. There are many new technological issues to be explored, but also many issues related to collaborative communication. For example, what roles will more knowledgeable users have in facilitating interaction for first time users? During cross table communication, there is a greater opportunity for gestural communication that can be an aid to communication regardless of whether or not it is recognized by the computer. There is an opportunity to gesture in the context of digital imagery, which in itself is an augmented form of communication. Perhaps in the future technologies can become less the focus of our attention and play a greater supporting role in terms of human-to-human communication. A gestural interface does not need to be an interface with sensors that recognizes gesture, but merely an interface that encourages people to gestures one to another by providing stimulating visual feedback. Computing devices could be shaped to facilitate communication with people across a table or become places for sharing and presenting digital content. Gesture recognition could also impact collaboratively communication by allow people to conducting digital scenarios in real time. This could be helpful for simulating variations in the way complex systems work together. Interface could be controlled in a puppeteer-like fashion allow people to role play complex scenario in real time order to come to a group consensus about new directions or approaches.

For multi-touch interfaces, designers and developers need to consider relationship between the multiple hands and constantly evolving digital content. They must also consider the relationships between hands and fingers of a single user as well as the relationship between the hand motions of multiple user users. For tabletop displays, content needs to be visible from 360 degrees as well. More important are considerations related to the context in which a given device will be used. For example, will a facilitator such as a teacher or more knowledgably user be present to help guide them through the experience. If not how can functionality be discoverable when users will not likely be familiar with interactive capabilities.

For video based scenarios, designers and developers need to consider how body shape and orientation changes over time and how the body shapes and orientation relate to constantly evolving digital content. These relationships have for the most part been facilitated through recognition of iconic forms, which harkens back to static “button-push-like” interactions. The more interesting developments will come from attempts to contextualize motion through temporal context, visual-spatial context, or environmental contexts.

2.1.5 User Centered Design Research

This section is a review of a proposed methodology that may be used by designers during the early design research stages of gestural interface design. It includes a formula for a participatory research study that can be used to generate a reference library of gestures that can inspire new gestural interfaces. In this section and the next, several parameters of this study will be discussed in further detail. Each of the following concerns will be addressed:

1) Testing conditions that are conducive to gathering relevant gestures
2) Imagery capable of provoking a comprehensive range of gestures
3) Methods for sorting, comparing and analyzing results
4) Methods of applying results within new interactive schema

Through my design approach, I reached a point where I wanted to incorporate several of the new definitions of gesture I had discovered through my work and literature review, but I was overwhelmed with the complexity of the computing aspects of the endeavor. As discussed, the things that a person understands as intentionally communicative have very distinctive features that relate to the context in which they are used. Many factors can combine to create the context for a gesture. A person can refer to elements in their immediate surroundings or they can establish a reference point for understanding a gesture through prior gestures. Computationally, this provides too many scenarios, which could not all be accounted for because computers must be told exactly what to recognize and when to recognize it.

In order to incorporate gestures into interactive scenarios that fit my developing definition of gesture, I needed to address several questions that had kept me from moving forward. First, I needed to create a context for understanding gestures that related to visual experience. While I gained a lot of useful information from researchers who had studied gesture in the context of speech, I did not know if I could translate these finding into a visual context or whether gestures would still operate in the same way. After isolating specific gestures, I would need to determine the features of each gesture that were most significant to its meaning. After establishing specific features that were critical to deciphering intentionally communicative features of a gesture, then I would be able to determine the simplest way to sense these features and what the most appropriate technology might be.

It seemed that a feasible solution to gestural interaction would require that users be presented with potential gesturing strategies that could be combined and build off each other to create more complex interactions. A limited set of gestures would need to be presented so that people would not be overwhelmed. This limited set needed to be something that users could combine to create limitless variation. Much like a musical instrument, I felt that the results of interactions should allow users to combine a limited set of gestural repertoires into an infinite set of outcomes. An ideal result would allow endless exploration that meant more control could come with practice and exploration of possibilities. In order to accomplish this, I wanted to determine which sets of motions could enable large amounts of variations and allow users to articulate visual staging and animation scenarios.

Some qualities of gesture can be more easily translated into computing than others. For example, with the mouse we see the two dimensional representation of space being mapped directly to the cursor's position on a two dimensional plane. Two-dimensional mouse movements have also been mapped to three-dimensional space in many software applications like Maya. Other kinds of gestural motion are more difficult to map, and could not be as easily translated to virtual space. Some gestures are truly three-dimensional and may require a higher level of spatial or temporal recognition to be interpreted. For example, someone might gesture with both hands to describe how two object interact in 3D space, this may be more difficult to describe because it involves factors such as the shape of both hands, the distance between hands, the movement patterns of each hand, and the timing and syncopation of all of these elements.

When a computer sees a gesture it does not always need to see everything. In fact, it would not perform as well if it did see every piece of information related to one's movement. Most gestures do not require full resolution of spatial and temporal information in order to establish what someone intends to communicate. Contrarily, inappropriate resolution could cause important information related to some subtle gestures to be indistinguishable to a computer. To program gestural interaction requires awareness of specific features of gestures that are used to communicate naturally. Anything beyond the minimal information needed to recognize the gesture puts unnecessary strain on the system's storage and processing power.

The question is therefore what is the minimal amount of information that the computer needs to see in order for it to recognize the meaning a user intends to express. The minimal amount of information could vary greatly depending upon the gesture. For example, pointing is quite simple, but pantomime gesture are much more complex. Before we can know how to program an interface we first need to establish what gestures should be recognized and distinguish features of these gestures that should be recognized.

I decided to develop this study because I needed to see a relationship between gestures and visual representation. I could not just use gestures that were produced in the context of speech because these gestures would not necessarily relate to a visual interface. In order to find gestures that would be relevant within the context of a visual interface, I needed to see gestures that were produced in the context of visual cues. The role of the visual component of a gestural interface includes providing information that helps people to understand how and why the computer is responding to their gestures. In order to make sense of this relationship, a person should be able to connect the response of the visual interface to temporal and spatial components of real-world experience.

For the connection between the visual component of an interface and visualizations in the mind of users to be accurate, I needed to establish a relationship between real-world visual and spatial transformations and the gestures produced by people observing those scenarios. This could help to establish a connection between what people see and the gestures they relate to what they see.

In producing the visual stimuli that I asked people to respond to, I had an opportunity to influence the specific types of movement that participants produced. I had a chance to investigate some of my hypothesis regarding how to best represent temporal and spatial change through gesture. I was most interested in discovering which aspects of the stimuli test subjects would respond to and represent through gesture. In additional, I was looking for the commonalities that occurred between among groups of people in terms of strategies for interpretation of video clips and methods of representing each clip.

When I choose the specific video clips I did not have plans for a specific interface in mind, but I had considered the benefits of gestural interfaces for collaborative content creation and interaction. I imagined that applications that could benefit directors of stage and screen, project managers, event organizers or planners that require synchronization or participation of many individual efforts. I could also envision educational scenarios. The gestures investigated include those that people might use to express forms of spatial movement, organization, containment, interacting timelines, and several expressive qualities of motion. Because the gestures are enacted within the context of video cues, a reference point for understanding the way in which gestures can be used to express visual-spatial transformations could be established.

Before generating imagery, I began by defining some desired tasks or processes that I felt could be more easily accomplished through gestural interaction. Several visual ideas were considered that address important issues outlined in each category (for a full shot list with still frame images, please refer to appendix B) Video was shot and narrowed down to the most effective clips. Five key concepts will be addressed:

1) Spatial Movement of People, Objects, and Point of View
2) Ordering and Organization
3) Containment and Transfer of resources
4) Timing

5) Descriptions of quality motion (light, floating, abruptness)

Early tests of imagery were conducted in order to make sure the imagery was not biasing study results. Tests of the visual stimuli included surveying several people and asking them how they would use their hands or body to express a given image. The process was similar to a game of charades. The surveys produced wide variations of responses from each person, but also provided important feedback that helped to improve the stimuli used within the study.

When choosing visual stimuli it was discovered that it was best to avoid imagery that might bias users towards their preconceptions of interactive media. For example, for an early form of this study, vector-based imagery was developed. This imagery was determined to be too leading in nature after initial testing because it brought to mind familiar interactive scenarios. For example, in response to vector imagery, one participant repeatedly tried to make hand gestures on a table-top similar to mouse-pointer and tried to double click using their fingertip. I changed my approach and began using video imagery. This connected the action in the imagery to real world spatial and temporal transformation and helps to dislodge the participants' preconceptions about other interactive systems.

Through this research study, I wanted to encourage participants to produce gestures that would inform and inspire possible interactive and technological solutions that have not yet been considered. For this reason I avoided limiting the range of interactions that they are asked to produce. They were simply asked to retell what they saw visually. However, because I was interested in how visual displays might impact the gestures produced, I varied the ways in which imagery was displayed during the study. I included three variations: a large up-right projection screen, a horizontal projection on a table surface, and a small screen on a laptop (See Appendix for illustrations and imagery of test conditions). I was interested in how gestures might be altered based on the environment in which they were produced and how screen size and orientation might affect the gestures produced by participants.

In this section I have discussed the process of developing a body of visual imagery that could be the stimulus for gestural expression during a participatory research study. I have discussed the logic behind my decisions and the early testing that helped to reaffirm my approach. In addition, I have discussed the test conditions for the study. In order to provoke a range of responses I varied the presentation format and ordering of video clips shown to each participant. In the following chapter I will discuss in greater detail the methods of sorting, comparing and analyzing the results of this research study. I will also discuss potential applications of the results and new directions based on initial findings.

CHAPTER 3 ANALYSIS

3.1 Initial Methods of Categorization

As was discussed in the previous section, the imagery developed for this study was developed with some predefined categorizations in terms of the quality and range of expressions I hoped to obtain. This was not designed to represent and exhaustive set of potential gesture vocabularies, but rather was designed to represent a large range of a particular scope. The scope of gestures included within the study was designed to represent five major groups of gestures developed from a consensus of various gesture researchers (See Appendix A). While gesture researchers categorize semiotic gestures into various sets, based on their use in daily communication, they generally address a specific set of gestures that I have grouped into five major categories. These include:

1) Deictic Gestures (pointing) – typically used to define a referent
2) Spatial Movement Gestures – used to define that trajectory of motion
3) Pictographic Gestures – used to define form
4) Kinetographic Gesture – represent bodily motion through mimicry
5) Beat Gestures – denote pacing or syncopation

When developing visual stimuli for the study, my aim was to produce imagery that would provoke a range of gestures that would represent each of these groups. These forms of gestures have been defined and are commonly used by gesture research to describe their observations of gesture performed in the context of speech. While these categorizations are well established and quite effectively used by researchers who are analyzing speech-accompanying gestures, I felt that these categorizations might be inadequate for my purposes. I was interested in gestures that could be used to control a computer interfaces and I was therefore more interested in gestures that would be useful in the context of interaction with computers.

Because the gesture produced in this study would be produced in response to visual imagery, it was not clear whether or not the categorizations of gesture defined by gesture researchers would provide sufficient direction to designers of interactive media. Most research viewed gestures in terms of its role in development of modification of spoken language. Therefore, I kept these initial categorizations, but planned to add some categorizations based on common methods of expression I observed during the research study. I had expected to see gestures that could fall in one or many of these categories simultaneously. I had also expected that given the nature of my investigation, I would see correlations between the gestures that would relate more specifically to gesture-based interaction design.

Through the video clips, it was my intent to present prepositions and verbs – like in, over, through, etc. or roll, tilt, swing, align, etc. (For a full list see appendix C) . I had hoped that this would allow me to compare various forms of gestural representation that could be used to describe action and orientation with meaning similar to their linguistic counterparts. By comparing the gestures used to describe verbs and prepositions to those used to describe form or action in a single video clip, I suspected that similarities would emerge that would make commonalities among gesturers' expressions more salient. I had designed each video clip to address a specific type of action, process, or transformation in order to provoke people to represent these types of phrases in gestural form. In many cases I was successful, but in others the attention of users focused other distinct or relevant aspects of a scenario.

3.2 Analysis methods

For this study, I presented 24 video clips to 15 different participants, resulting in 300+ video clip sets, data that needed to be transformed into an analyzable form. During the study all participants were asked to view each of the roughly 30-45 second clips and then retell what they saw in each sequence using hand gestures. They were given the opportunity to speak aloud, but they were asked to only use words that reinforced what they were attempting to express using their hands. Participants were recorded from two angles, front and side.

While observing participants and reviewing video recorded during the study, I developed several insights related to patterns of representation that could be used to analyze and categorize results of the study. Developing more explicit categorizations required a top-level view of a fairly complex data set. In order to develop these categorizations and optimize my analysis techniques, I developed a digital application designed to facilitate this process. Development of this tool was, as many of the endeavors previously discussed, a highly iterative process. Several features of this tool addressed apparent needs related to analysis. Expansion of functionality would be a result of findings gathered once simpler functionality was in place.

The tool was designed as both a research tool and presentation platform. As a research tool, I needed the application to allow me to view several videos simultaneously and store annotations relative to the temporal context in which they occurred. This would simplify the review process by easing my cognitive load and allowing me to focus on developing a more effective process of analysis. Once I entered time-coded data, I needed to view multiple layers of temporal co-occurrences simultaneously so that I could get a better view of common co-occurrences and archive it in a manner that was easy to analyze and navigate. I needed the application to allow me to focus in and make notations on a clip-by-clip basis, but I also needed to zoom out and see larger correlations in occurrence across the entire video set.

In addition to the role the application played as an aid to my research, it needed to simplify the presentation of information related to annotations. Through simplified presentation information I could establish more concrete methods of analyzing and comparing clips. Because all the notations that I generated during my analysis were stored in a database, it would allow me to search, sort, and cross analyze based on features of expression such as reference clips where certain categories of gestures manifested themselves most frequently or instances where various gestures commonly co-occurred. This would simplify the process and allow me to focus on developing more consistent methods of annotation.

This initial landing screen on the interface includes a large co-occurrence map, which represented each clip and each tag permutation on a map grid that plotted and highlighted the tags present in each clip. This allows a top-down point of view of tag occurrence throughout the entire video archive. Each video was manually tagged when a tag term related to the gestural pattern occurred. Selecting an element in the co-occurrence map added the clip to a list of videos to be viewed. Clicking on a tag category would select all video clips in which a tag occurred and add them to the list. Once I was finished searching and building a video set for review, I hyperlinked to a new screen in which the selected videos could be viewed in consecutive order.

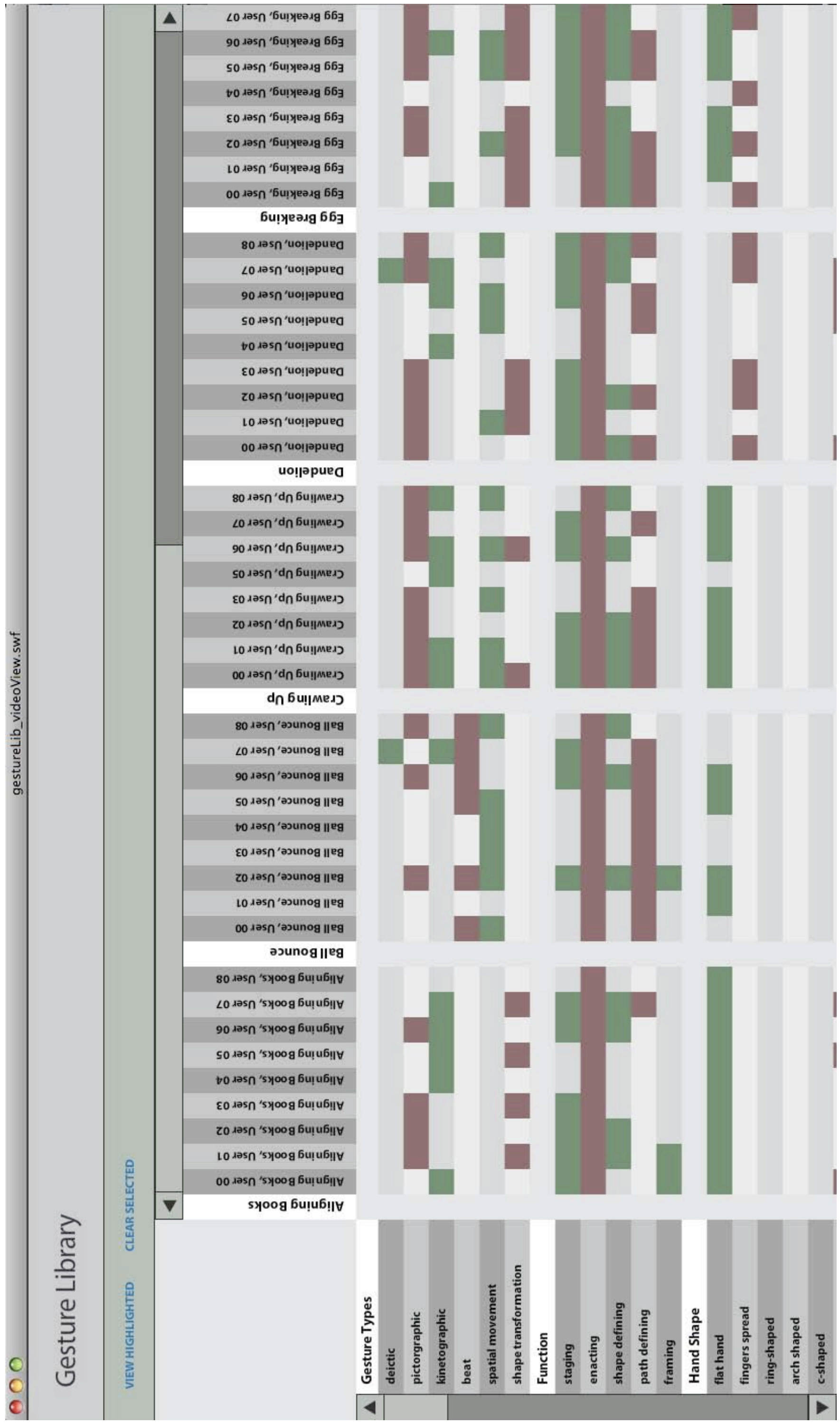


Figure 24 - Gesture Library Analysis Tool, Co-occurrence map

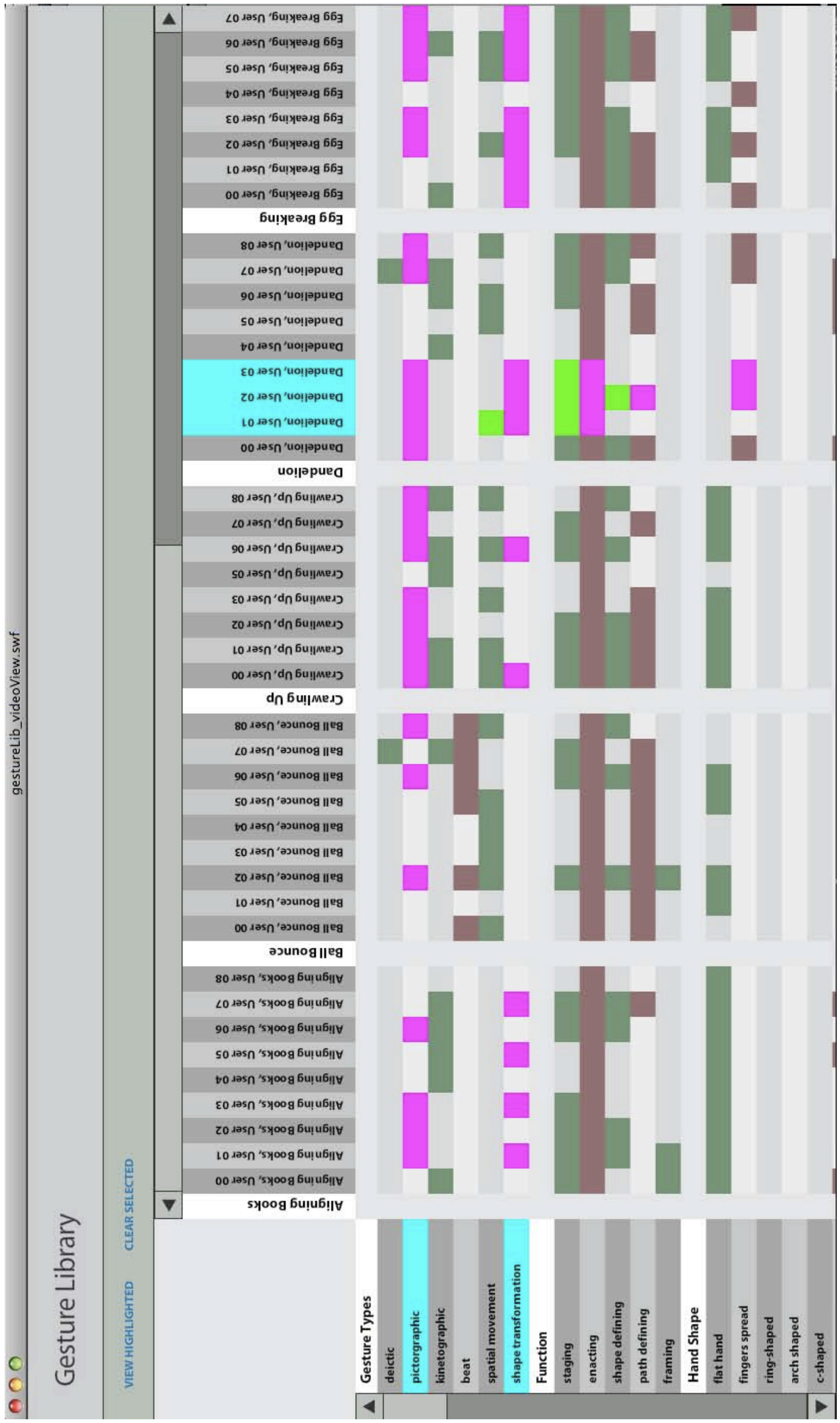


Figure 25 - Gesture Library Analysis Tool, with highlighted search elements

From the initial landing page I could gather related clips and then I could view these clips in an organized manner in the video player. I developed the video player that would play 3 views of the scenarios simultaneously, one view for each of the two camera views and one view of the video clip as seen by the participant. This would allow me to simultaneously analyze the front and side views of the participants and the video clip to which they were responding. In addition, this would allow me to navigate the large database of video clips with greater ease.

The video view player was built to include a multi-layered timeline, which would allow me to store tags and comments and view correlations between them. Tags would include the various groupings of gestures discussed previously. This would also include new tags and tag categories that would emerge from the process of reviewing and analyzing videos. Comments could be used to highlight more specific features of the gestural expressions. The interface allowed me to manually enter tags and comments as I viewed three of the video clips simultaneously. After entering tags and comments related to a video clip, they appeared on the timeline, and when clicked, would allow me too hyperlink to a specific time location within the video clip where there was a relevant gestural expression. This would allow me to compare manifestations of various forms of gestures across the larger set of videos, making it easier to see expressive commonalities emerging.

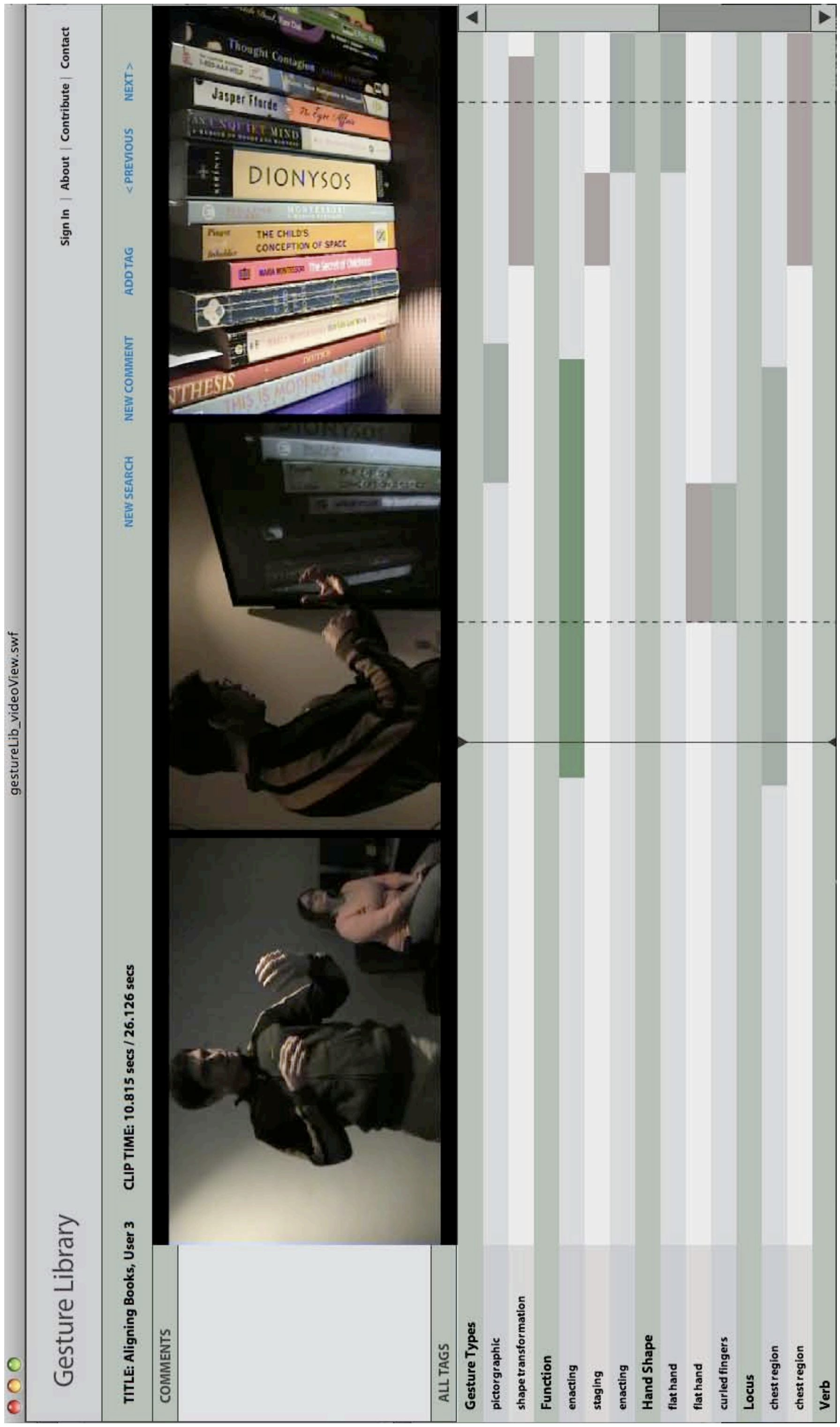


Figure 26 - Gesture Library Analysis Tool - Video Viewer

3.3 New categorizations arising from analysis

Although the gestures that were produced were far from predictable, there were reoccurring themes of usage that could be correlated to the nature of the visual presentation to which people responded. Correlations could also be made between subjects' verbal descriptions of what they were attempting to communicate and the types of gestures that they produced. Several more focused categories of inquiry arose in my mind during the process of viewing the test subjects' gestural representations in the context of visual stimuli and their verbal interpretations. I will elaborate further in the following sections.

3.3.1 Articulation within body space

In addition to the meaning of gestures expressed, I had planned to analyze the body space in which gestural expressions occurred. Gesture researchers often analyze the *locus*, or body space, in which gestures are performed by breaking down gestures based on areas or regions that are relative to the body, such as in front of the face or chest, above the head, or extending away and to the front or sides of the body (Rossini, Nicla, 2004). People often use the body as a reference point or to represent gestural subject matter. By analyzing the body space in which gestures occurred, I was looking for commonalities in spatial articulation as well as commonalities in the meaning expressed by the form or spatial transformation of the hands and limbs. I looked for gestural expressions in several body regions including the face region, the chest region, the waist region, above the head and below the waist. I considered locus of motion on the sides or behind the body to be occurring in the "body periphery". In addition I considered gestures that occurred while contacting hands, contacting an arm, or contacting a table (which was common in the table-top projection scenarios.

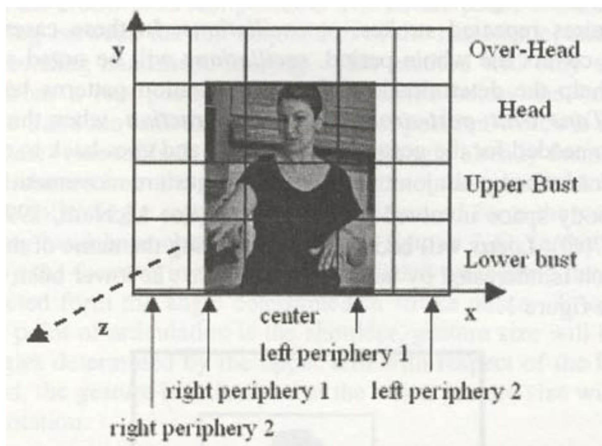


Figure 27 - Locus of gestures

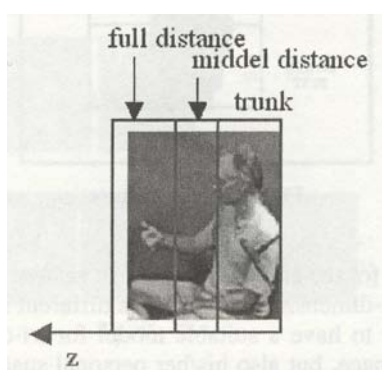


Figure 28 - Locus of Gestures, lateral

I found that there was a relationship between common body space regions of articulation and the method in which visual stimuli was presented to users. For example, when video clips were presented to participants on a large screen, gesticulation commonly occurred in the chest or waist region and hands more commonly extended forward in the direction of the screen. When imagery was presented on a horizontal tabletop screen, gesturer frequently reached over the screen or touched the screen. Participants who view video clips on the table were more likely to gesture in the waist region of the body. When participants viewed video clips on a laptop, gestures were produced in a smaller body space and were typically closer to the body.

Because both the table top and laptop scenarios provided the participants with a flat horizontal surface, these participants were more likely to contact the table surface in response to video clips like the video of a ball bouncing. Participants who were standing in front of a large projection screen did not have a flat surface in range, but would often use a flat horizontal palm or horizontal forearm to represent a flat surface (See Figure 29).

Figure 29 - Gestures describing a ball bouncing

Contacting between hands or between a hand and an arm represented either the interaction between two objects or the force impacting an object. For example, in response to the video clip of a boy playing on a slide, a diagonally oriented for arm was often used to represent the slide and the second hand moved across the forearm to represent the boy moving across the slide (See Figure 30). When describing books aligning, one person pushed a flat palm against the side of the other flat palm to represent the hand pushing the book spine to align the books (See Figure 31).

Figure 30 - Gesture representing a boy sliding down a slide

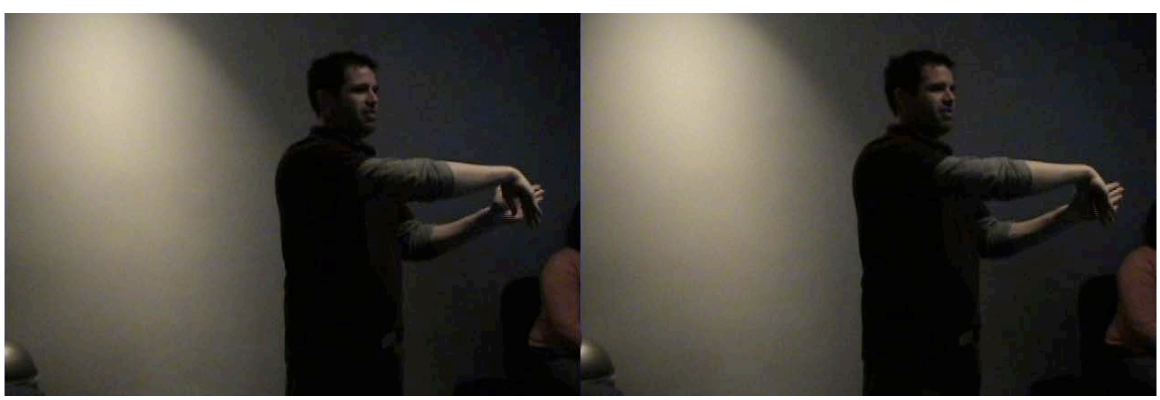


Figure 31 - Gesture representing books being aligned

Hands also contacted when a person began defining a continuous shape. When drawing a shape in the air, such as a curved, triangular or cylindrical shape, people often move their hands in bilateral symmetric movements. They often begin and end these movements by contacting their hands to indicate the point of symmetry related to the shape they are describing (See Figures 32-34).

Figure 32 – Gesture describing a triangular swing set

Figure 33 – Gesture drawing the arc of the swing

Figure 34 – Gesture defining the shape of a tube

3.3.2 Hand shape and orientation

An additional area of interest was commonalities in the use of hand shapes and orientation of the hand and limbs. The use of hand shapes during gestural expression is significant for several reasons. The shapes produced by hands can be used to represent the form of object or organism and can also be used to represent the changes in the shape of an object or organism over time. When people use their hands to represent various forms and types of shape transformation, the joint hierarchies of the body limit them to certain sets of hands shapes and hand shape transformations. As a result, specialized repertoires of motion become apparent when observing the use of hand shapes for expression in various scenarios. Similar hands shapes and hand shape transformations express information of a similar nature.

This phenomenon became apparent to me after a read some of the work of Adam Kendon (2004). In *Gesture: Visible Action as Utterance*, Kendon explored the used of specialized families of hand shapes that were used to modify the meaning of speech during speech accompanying gestures. He grouped gestures into several families of expression based on hand shape and orientation during articulation. Example included open *hand prone* (palm up*), open hand supine* (palm down), ring shaped hands, and *grappolo* gestures (precision grip or finger bunch). What he found was commonalities in the meaning of expression produced in each hand shape family. Most of his findings were relevant to the interpretation of meaning in speech accompanying gestures, but I had wondering if I could find similar patterns of expression when looking at gestures that described visual phenomenon.

As it turns out, I did find patterns of usage, although they were somewhat different that the types of expression discovered by Kendon. Gestures produced in responses to visual imagery tended to be more literal and less abstract in nature than gestures used to parse or augment speech. I found common hand shape and orientation groups of expression that were similar in form. These include: palm facing up, palm facing down, fist, curved fingers, spreading of fingertips, flat hand or shaping the hand into a ring or cylinder (See Appendix B for a detailed list).

Flat palms were used very frequently throughout this study and presented different types of meaning related to the orientation or movement of the hand. Gestures with flat palms facing down were often combined with a horizontal or vertical movement such as a flat surface (see Figure 35 – a flat palm is held vertically to represent the table that the ball bounces on) or the water level in a glass (See Figures 36 – a flat palm is moved upward to represent a rising water level). After viewing a video clip that portrayed a close-up image of books being stacked, many people enacted the scenario by stacking flat-downward-facing palms repeatedly (Figure 37).

This finding suggests that flat hand gestures with palm facing down may be well suited for defining planer surfaces, barriers or boundaries that can be acted upon. These "surface defining gestures" are often used to

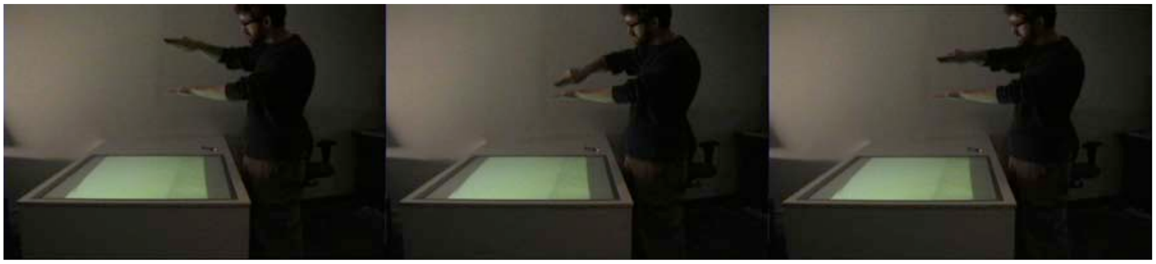


Figure 35 - Gesture defining ball bouncing on a table surface

Figure 36 - Gesture describing a glass filling with water

Figure 37 – Gesture describing a book being stacked on top another book

establish a spatial relationship between an object and a horizontal plane. When people responded to video clips of the ball bouncing or egg breaking, they often began by establishing the ground plane, which established the relationship between the ball or egg and the ground plane and the movement of shape of trajectory of these objects as they moved toward the established plane.

When two flat palms faced each other it was typically when people were attempting to describe changes in orientation or position. When one palm faced the other and the wrists were rolled or pitched in unison this was typically describing tilt or lean of an object such as aligning of books (Figure 38). When two palms faced each other and were moved in unison to the same position and then retracted, this typically represented placement of an object (Figure 39). Sometimes when palms faced each other it was in order to describe shape and/or position. When one palm faced the other and the hands moved toward the other, this typically represented things moving into alignment (Figure 40). In response to the video of a boy rolling down the hill, several people oriented two flat palms towards each other and one on top the other, then oscillated their palms horizontally in a bilaterally asymmetric fashion (Figure 41). When referring to an object like a glass, book or other object in view, several people moved two hands forward and slightly toward each other as if patting a form between two hands. When people described wrapping a box or the organization of books they might start by saying “there was a box” or “there was a book” by extending there

Figure 38 - Gesture describing books being tilted upright

Figure 39 - Gesture describing books being straightened

Figure 40- Gesture describing books being aligned

Figure 41- Gesture describing boy rolling down a hill

hands forwarding with two flat vertical palms in a single chop-like motion. This gesture is typical a gesture that denotes placement or location of an object within a scene.

Placing or positioning gestures of this nature might be used within interactive scenarios to establish a reference point for the object so that following actions are understood. Once a spatial reference to an object is established, gestures happening in relationship to the referenced shape, orientation or location of an object have an established spatial relationship to the objects. These placement gestures may be use method of interaction or ways of defining or establishing a referent to which interactive gestures should refer. Once a referent is established they might also be used for positioning and orienting objects within digital space.

When a hand was oriented forward or towards the screen, they were typical used to define vertical surfaces or forces acting upon a vertical plane. A flat palm facing forward with fingers spread was often used to draw the paths of motion related to a force, like the wind blowing through trees (See Figure 41). Forward facing palms were also frequently used to describe alignment of books (See Figure 43). Two vertical, forward facing palms were often moved forward in unison to describe a process of alignment. Alignment gestures of this nature could be useful in interactive for describing orientation of objects along the z-axis of the body. Using a bilaterally symmetric gesture with two-hands facing forward, several people

Figure 42 - Gesture representing wind blowing through trees

Figure 43 - Gesture representing alignment of books

Figure 44 - Gesture describing misaligned books

were able to describe the depth alignment of books as uneven being uneven from front to back (See Figure 44).

Precision Grip Hand Shapes were sometimes used when pantomiming gripping motions, but were more frequently used to describe form through the implied shape of the hand when gripping a form presented within the video. For example, in both the egg breaking and ball bouncing video clip, people frequently formed their hands into a shape similar to the shape their hand might make when holding the ball or egg (See Figure 45). They would hold their hand in this “gripping” shape while they moved their hand to describe the location of the object as it moves across the scene.

In the example with the water glass, there was no image of the hands within the cameras view, yet many people described the glass by shaping their hand into a shape similar the shape the hand makes when gripping a glass (See Figure 46). In the image of the dandelion many people represented the dandelion by holding their hand out in front of the body and pantomiming the hand shape similar to holding a dandelion (See Figure 47). Some people went as far as to pantomime picking the dandelion, even though the video clip presented to users showed no view of human hands. The placement of small objects within the scene was often described using pinching. In response to the video clip of a boy swinging, one person described the fulcrum point of the swing by holding his hand in a pinching shape in the location where the fulcrum point was located relative to the arching path he drew in the air (See Figure 48).

Figure 45 - Gesture describing an egg being dropped then breaking

Figure 46 - Gesture describing glass filling with water and overflowing

Figure 47 - Gesture describing picking a dandelion then blowing on it

Figure 48 - Gesture defining fulcrum point on a swing (pinch)

While precision grip gestures are typically mimic motion of the human body, but they can also be used to describe shape as they seem to grip an imaginary form. The shape of the hands during precision grip gestures can often be used to establish the shape of a referent through the negative space that the hands create. I think this is an important consideration for interactive gestures because it may be a possible mode of referring to objects within digital space. For example, if I want to grip a virtual cylinder I could make a cylindrical gripping motion, but if I want to grip and cube my gripping shape might be more rigid. If I want to pick something small like a small flower perhaps a smaller pinching gesture would be more appropriate.

3.3.3 Movement Repertoires

Examples of common hand motions include vertical or horizontal chopping, spreading hands from clenched to open, grasping motions, two hands oscillating asymmetrically or forming hands as if around an imaginary cylindrical form or planer surface. As gesturers used their hands to represent the various forms shown in the videos, commonalities emerged from analysis of the independent expressions of a larger group of people attempting to describe the same imagery. As will be discussed further in the following section (3.5 Common gestural repertoires), there were many commonalities related to co-articulation of the hands or hand-shape form and motion combinations.

3.3.4 Functional role of gestural phrase

As I approached the process of analyzing gesturers' responses during this study, I had some unique concerns in my mind related to recognition and interpretation of gestures. Unlike other gesture researchers that I had previously reviewed, the purpose of my investigation was centered on issues of pattern recognition for computing applications. I was therefore looking for ways that gestures could be a form of input for interface control. This lead me to a unique method of categorization that I define as "staging" and "enacting" phases of gestural expression. This area was not preconceived, but rather it arose from my observations of participants during the research study. It became clear that there were some strategies that were common among participants that would reoccur and present themselves in many ways throughout the study. Gestures could typically be sorted into two categories: 1) Gestures used to define the scenario in which other gestures took place and 2) Gestures used to describe the action taking place within a scenario. It was clear that certain gestural expressions had functional roles related to defining the space (staging phase) in which gestures later gestures would be enacted (enacting phase).

These two types of gestures played off of each other during gestural expressions and combined to form the larger meaning of gestural expression. Often the meaning of "enacting" gestures were tied to previous "staging" gestures because staging gestures had lingering meaning that was often referenced by later actions through co-articulated space, a relationship to a predefined path or inflection towards the areas where something was defined or expressed in the past. Both staging and enacting gestures include three general subsets of gestures related to modes of gestural interface input. These include shape defining gestures, path defining gestures, and framing gestures (gestures that define points of view). Below I have defined the five main "functional" categories of gestures that I utilized during categorization of my results :

1) Staging - gestures that define the setting or spatial relationships within in which action takes place
2) Enacting – addressing the primary action taking place through gestures that express kinesthetic or transformative features of motion
3) Shape Defining – define the shape of a place or a thing either within the static environment or objects changing over time
4) Path Defining – define that path or direction of motion
5) Framing – define the main area of action, point of view or perspective from which gestures should be understood

Shape defining, path defining and framing gestures can occur in either the staging or enacting phases of expression. Typically the staging phase of an expression will precede the enacting phase of an expression. Exclusions to this rule exist, but in most cases in is a result of attempts to reiterate the scenario presented within a video clip.

Through this study, I observed many examples of the combined use of staging and enacting gestures. For example, in response to one clip that portrayed water being pour into a clear glass, many participants would begin by defining the glass by forming two hands into a circular form, similar to the shape hands adopt when holding a glass (See Figure 49). This would often be followed by a pouring gesture that would occur above or in relationship to a hand that mimicked holding the glass. Sometimes, gestures used to define the shape of the glass would be followed by a gesture which overlaid the area where the glass-shaped hand was formed. A flat vertical hand was used to represented the water level rising within the glass (See Figure 36).

Figure 49 - Gesture defining water glass filling with water then overflowing

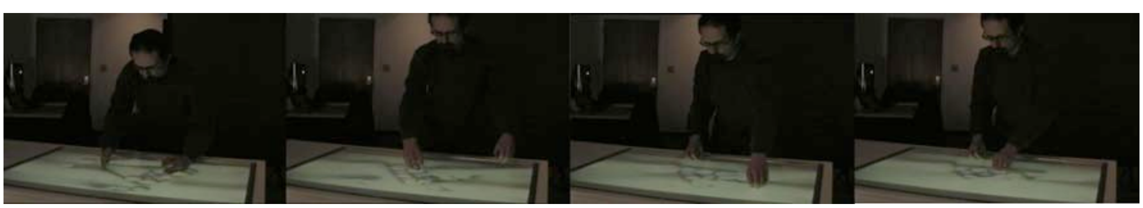


Figure 50 - Gesture defining pool ball being placed in a triangle

Finger and hand motion performed in response to staging gestures often represent spatial transformations. After viewing the movie clip that showed pool balls being racked, many people define the pool rack by making a triangle with their hands then pantomiming grasping the pool balls and placing them inside the space where the pool rack had been previously defined (See Figure 50). Also in the water glass video clips, the form of the glass would be defined and then the gesture that followed would be enacted in relationship to where the form had just been defined.

It was clear upon seeing these examples that it was possible to represent containment through a process of defining the external boundaries of a form and then enacting in relationship to established boundaries. Once a form had been established, later motions could be contextualized based on their location to mean either “inside” or “outside”. This ability to contextualized motion through sequential staging and enacting gestures has implications for gesture-based interactivity because it suggests that gestures can be interpreted based on discrete functionality. The meaning of consecutive gestures can operate based on the meaning established by previous gestures. For example, in the process of creating and manipulating visual geometry, staging gestures could be used to create a geometric form to position and orient the form within a virtual space. These gestures could be following by enacting gestures that determine how the geometry will be animated along a path or how the geometry will transform over time.

3.4 Common gestural repertoires

In this section I will discuss various manifestations of the themes discussed previously, both in greater detail and in a more general conclusive manner. I will begin with some general observations and then discuss the implications of these observations.

3.4.1 Gestures for defining form

Hand and arm shapes were frequently used to represent objects. For example, participants using hand shapes represented the pool balls, pool rack, plants, paper, and the glass. Using the shape of their hands, they established the form and orientation of a referent to which later motions could refer. Sometimes these motions were enacted while the hand remained in the shape of the referent, but sometimes it was easier to articulate a path of motion through air drawing in relationship to an area where other relevant gestures occurred.

When defining shape or form, people used several strategies. They might enact motions similar to the movements that would be created if their hands followed the external surface of the form. In response to the video with the water glass, people commonly held their hands as if contacting the out edge of a glass (See Figure 36 or 49). Other examples included a ring-shaped hand combined with a vertical motion used to represent a cylinder (See Figure 34). A child climbing through a large tube on a playground was often represented by one arched hand moving horizontally and then another hand passing between the sides of the arch shape (See Figure 51).

When someone attempted to represent the form of a stationary object they would typical form their hands into the shape of the object, hold the shape in the air momentarily and then retract their hand or enact other gestural phrases.

Some shape defining gestures involved sustained hand shapes coupled with movements that defined the surface of the form. For example, in response to a video clip of the dandelion being blown apart in the wind, one individual used one hand to represent the dandelion with a hand shape that mimicked holding a dandelion. They then wiggled the fingers of their secondary hand over the other hand to represent the seeds on top the dandelion (See Figure 52). These two forms of expression were the most common means of expressing spatial form. This suggests that recognition of form defining gestures should involve the recognition of sustained hand positions and sustained hands shapes coupled with stationary or moving hands. Shape defining gestures could have many roles within virtual environments because they are effective for both 2D and 3D representation of forms. They could be effective means of content creation for animators, gamers, engineers and many other disciplines.

3.4.2 Gestures for defining spatial transformation

Shape transformation gestures are a type of gesture that is used to describe how the shape of an object or group of objects changed. For example, the video clip of an egg breaking often provoked shape transformation gestures describing the change in the shape of the egg. Two types of representation manifest themselves in participants. In the first, the hand shape changed from closed fist to open hand with fingers spread, indicating the egg splattering on a flat surface (See Figure 53). In the second, two vertical curved hands with palms facing each other were followed by a bilateral upward wrist rotation to indicate a “cracking” egg (See Figure 45).

Shape transformation gestures typically include two consecutive hand shape, hand orientation or hand motion patterns that are contradictory. The contradictory motions typically refer to the transformational characteristics of a given gesture. For example, when describing the spatial transformation of a misaligned shelf of books becoming aligned, someone might begin by oscillating two forward-facing flat hands forward and backward in a bilaterally asymmetric fashion (See Figure 43). To indicate the transformation of books into an aligned state, someone might either push two flat palms forward in a bilaterally symmetric fashion (See Figure 42) or move a hand along an imaginary vertical or horizontal plane (See Figure 55). Sometimes alignment was also represented

Figure 51 - Gesture defining a boy crawling through a tube on the playground

Figure 52 - Gesture defining the shape of a dandelion’s top

Figure 53 - Gesture describing an egg falling and breaking

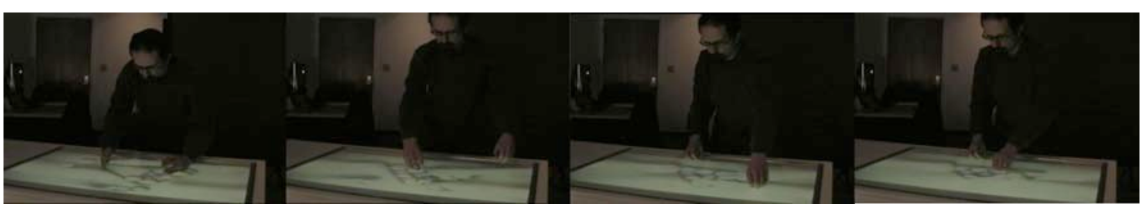


Figure 54 - Gesture defining pool ball being placed in a triangle

with hands transitioning from a bilaterally asymmetric to a bilaterally symmetric movement pattern indicates a transformation in alignment (See Figure 56).

Several clips portrayed plant-life in windy scenarios including a large tree with oscillating leaves, a field of corn blowing in the wind and a dandelion with seeds blowing off and scatter in the wind. In response to each of these scenarios, oscillating hand motions often portray a tree blowing (See Figure 57) in the wind or fluttering seeds blowing off of a dandelion (See Figure 58). In each of these scenarios there was fairly consistent levels of continuous motion, which is a repetitive form of back and forth transformation. Oscillating hand motions seem to represent "back and forth" patterns of transformation. It seems that the best approach to recognizing a gesture that is intended to represent a transformation of some sort is to determine in advance some positions that represent balanced or counterbalanced motion. For example, a bilaterally symmetric motion followed by a similar motion that was bilaterally asymmetric would signify that a transformation is occurring. Oscillating or "back and forth" motions are also frequently used to represent transformations in movement.

Figure 55 - Gesture describing the aligned books

Figure 56 - Gesture describing books moving into alignment

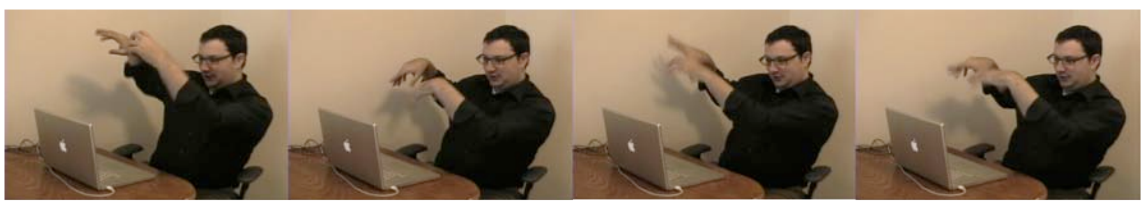


Figure 57 - Gesture defining tree leaves oscillating in the wind

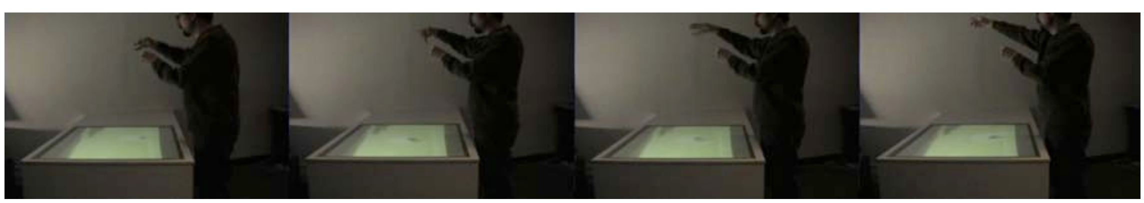


Figure 58 – Gesture defining dandelion seeds blowing away in the wind

3.4.3 Gestures for defining position

Positioning gestures typically include vertical or horizontal chopping gestures or single bilaterally symmetric gestures. When describing books being placed on a table many people performed a similar “placing” gesture that involved starting with two bilaterally symmetric clenched hands and then moving hands forward and downward (See Figure 59).

Two hands are often used to describe the spatial relationship of two objects, one to the other. Objects moving apart or together can be described with bilaterally symmetric motions with hands either moving together or apart (See Figure 60). The placement of repetitive row of objects are often described with chopping gestures which seem to be effective at describing repetition in space. When describing a row of books or a stack of books many people used vertical chopping motions (See Figures 61 & 62). Similarly when describing rows of corn, people frequently used chopping gestures to describe row.

These observations suggest that positioning and orienting gestures has several distinct characteristics that could be recognized using computer vision or other motion sensing techniques. Recognizing a placing gesture could involve recognizing bilateral motion when hands are extended away from the body and then retract. Gestures that involve bilaterally symmetric flat palms facing each other could be an effective for expressing orientation or placement of objects within virtual space.

Figure 59 - Gesture describing stacking books

Figure 60 - Gesture defining pool ball being placed in a triangle

Figure 61 - Gesture defining books in a row (chopping)

Figure 62 - Gesture defining books in a row (one chop then outward motion)

3.4.4 Gestures for defining path of motion

Spatial movement or path defining gestures occur in several forms. If the gestures are describing the movement of an object or person, the gestures can occur while the hand is in a shape meant to represent the person or object. For example, when describing a ball falling and bouncing people often formed their hand into the a fist, representing the ball, and then moved their hand along the path of the ball while their hand was still in the shape of a fist (See Figure 63). In some cases, simple air drawing could be used to define the path of motion that the referent follows. For example, to describe a boy rolling down a hill, many people drew a looping diagonal path (See Figure 64).

When drawing a path of motion people can use a single finger or many fingers simultaneously. People would often draw the path of a single object by pointing or moving a group of fingers in a "finger bunch" shape (See Figure 65). When describing the path of multiple objects people would typically separate their fingers. When describing plants blowing in the wind, many people chose to describe the path of the wind rather than the motion of the plant being blown by it. For example, many people drew wavy lines with their fingers in the air (See Figure 42).

Because the hand shape of an individual performing a path defining gesture is fairly distinct, it seems as though this might simplify the process of recognizing hand shape and triggering path defining modalities within an

Figure 63 - Gesture describing a ball bouncing

Figure 64 - Gesture describing a boy rolling downhill

Figure 65 - Gesture defining dandelion seeds coming off of the dandelion

Figure 66 - Gesture describing a boy on a swing

interactive scenario. It seems that when individuals enact path-defining motions, the fingertips define and emphasize motion. Sometimes, the hand and arm seem to be oriented toward the fingertips so that fingertips lead the hand and arm. For example, in response to video clips of the boy swing or going down the slide it was common for the path of motion to represented using a flat horizontal palm moving forward towards the finger tips (See Figure 66).

These characteristics of path defining gestures suggests that they posses distinct hand shapes such as pointing with one or multiple fingers, flat hands with spread fingers or movement patterns with fingers leading the hand (like an imaginary airplane). Many path-defining gestures can already be easily recognized with multi-touch surfaces. Perhaps an understanding of the relationship between fingers on each hand could be considered as a means of influencing drawing modalities on these surfaces. Video-based gesture recognition could also support recognition of hand shape during air drawing.

3.4.5 Gestures for defining point of view

For the most part, individuals in this study adopted a first person point of view when describing a scenario, but at times they enact gestures that either modified their point of view as described by their gestures or indicate framing such as the location at which items exit or enter the cameras view. When describing point of view one individual used two bilaterally symmetric hand motions with fingers leading flat palms horizontally and pointing downward to indicate he was “looking down on the scene” (See Figure 67).

Many of the participants within the study addressed elements entering or exiting the frame. When doing so they use a vertical palm with horizontal fingers facing forward to indicate the boundary of the cameras view and then enact an object crossing over the boundary while saying, “entering the frame” (See Figure 68). These are both examples of framing gestures, which are gestures that are used to represent point of view.

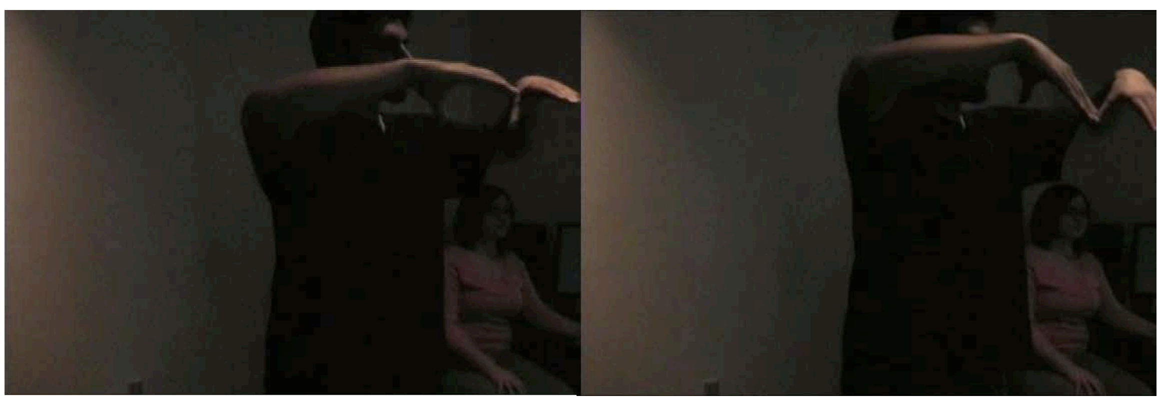


Figure 67 - Gesture defining dandelion seeds coming off of the dandelion

Figure 68 - Gesture defining dandelion seeds coming off of the dandelion

In a few cases, one individual used a single hand with vertical fingers bunch and palm oriented upward, the typical *grappolo* gesture, to begin saying “ There was a ball” or “There was a person”. These gesture were held in positions that suggested the individual’s point of view in relationship to the subject. By using a single hand shape signifier like the grappolo gesture and considering its relationship the face of the gesturer the point of view of the gesturer could be understood. If the grappolo hand is held close to the face it could represent a close-up point of view. If it is held away from the body or in the periphery one’s view, it could represent an object on the edge of one’s view.

It is possible that framing gestures could be used in virtual scenarios where the placement of a virtual cameras or changes in point of view are being determined by a user. Often times, when describing a situation, information can be more optimally described from one angle than another.

For this reason, people often use gestures to identify the angle from which a problem should be understood. Within digital interfaces this is true as well. In 3D animation software, such a Maya, it is often affective to switch between points of view like top-down, side view, or front view in order to describe the motion that is taking place within a scene. Similarly, framing gestures can be used to alter the point of view from which a gestural expression should be understood. There is a potential to utilize framing gestures to contextual the point of view from which a path of motion or shape is being defined in 3D space.

3.4.6 Gestures for defining body movement

Often when hands or a figure were present within the video clip, participants responded by creating a kinetographic or pantomime gesture that mimicked the movement created by the hands or figure. For example, in response to the video clip of a boy swinging, many people enacted grabbing the chains on the swing by raising to clenched hands in bilaterally symmetric fashion and then swinging their hands forward and backward (See Figure 69). After viewing a pair of hands rolling up a piece of paper into a tube many people raised two hands, formed them into matching ring shapes and then pitched their wrists forward and backward in a bilaterally asymmetric fashion (See Figure 70).

When describing the body movement of another human, it is common to utilize a hand shape where the index and middle finger come to

Figure 69 - Gesture describing a boy swinging

Figure 70 - Gesture describing rolling a piece of paper

Figure 71 - Gesture describing a boy walking up a slide

represent legs. In several video clips, a young boy could be seen on a playground swinging, playing on the slide, climbing through a tube or rolling down a hill. In response to these clips people frequently describe the actions of the boy by pantomiming his movement using an extended index and middle finger to represent the motion and path of his legs (See Figure 71).

In many situations, I observed people performing pantomime gestures that related to actions that would have occurred outside the cameras view in the video clips that they were responding to. In these cases, even though the hand motions were not depicted in the scene, the motions of the hands seemed to be implied to the individuals observing the video. For example, one video clip portrayed books being stacked from a close-up angle so that only the front spine of the books stacking one on top the other could be seen. Many people depicted this scene by moving two hands downward while hand shapes were moving from clenched to open, in a fashion similar to place a book on a table (See Figure 59). Another example of this phenomenon was a frequent response to videos that depicted an egg being dropped into the camera view from above. Although the video did not show any hands in view, many people began describing this scenarios by lifting one hand in a shape similar to that formed when holding an egg, then opening the hand in place with fingers spread to pantomime dropping the ball into the scene (See Figure 45).

Because certain types of visual transformation seem to have implied relationships the kinesthetic motions that they are a result of, it may be useful to utilize kinetographic gestures like pantomime and direct manipulation gestures to generate transformations within digital space. Coupling between direct manipulation or pantomime gestures and the transformations that they trigger within the digital space could allow people to logically couple their influence on digital space with their experience of manipulating the physical world. As with many types of digital interaction, this relationship need not be a one to one relationship. For example, certain pantomime gestures that would be used to spread, break or stretch a small object in the real world could be couple to more metaphorical actions like spreading ranges of numbers, breaking apart linear timelines, or stretching the size of a window or the angle of a virtual camera.

3.5 The structure of gestural representation

Participants tended to have a consistent structure of representation throughout the study. For example, some people consistently began describing a video clip by first staging the scenario, then enacting the action that took place within the scene in the area constructed by their staging gestures. A few subjects went as far as to discuss the framing of the picture plane within each shot and the location at which elements entered and exited the scene. Some people expressed only the focal point of action within the scene. Using this approach, someone might represent the motion of a ball bouncing but not define the flat surface upon which it bounced.

If an individual was detail oriented with their gestural expressions, they tended to address each scenario with a consist level of attention throughout the study. For detail oriented gesturers, gesturing typically began with setting, was followed by the subject, then moved to predicate and then described the action. This is similar to the *resilient properties of language* discussed in Section 1.1.3 . Individuals who began with the action or predicate tended to be less detailed in their descriptions of action as well. For example, when responding to the various scenarios that involved plants blowing in the wind, a detail-oriented gesturer might explain the setting, then the structure of the plant, and then describe the variations in movement in each section of the plant. In this case, the description of each scenario would vary greatly. In comparison, an action-focused gesturer might have very similar gestures for each windy scenario, each generalizing the overall force of the wind. A field of wheat could be represented with the same gesture as a tree blowing in the wind. Each scenario might only describe the overall waving motion of the plant. This suggests that top-level elements that relate to generalizations of actions taking place within a scene tend be represented more abstractly than smaller more intricate elements of the scenario.

This notion has implications in terms of the scope of influence that an interactive gesture should possess. The scope of influence or the impact that people intend to have upon an interactive scenario might vary greatly. Understanding the scope that is intended is a complex problem because it not only involved contextual interpretation based on the status and visual-spatial characteristics of the virtual scene, but it is also a factor related to the personal approach of a given user. While this research can provide no further direction for addressing this issue, it does raise new questions that may need to be address in future studies.

There is a hierarchy of complexity involved in the level of detail that the gesturer addresses. Communicating broader views of the scenarios involves more abstract representations while describing smaller more specific elements involves more concrete relationships between the subject and predicate gestures. For example, in the video clip that involved breaking apart pool balls, a gesturer could either describe the movement of the entire set of balls with a single gesture that involved the fingers spreading in outward in all directions, or they could separate the movements of each ball into discrete motions of each ball on at a time. Pool balls are all the same shape, so it's easy to generalize about their movement. However, if someone was attempting to describe distinctively different shapes they might need to establish a pictorial reference for each. Once you staged or established references, then you can refer to them through pointing gestures or encircling gestures that refer to the objects. Once you have established a reference point, the action of the gestural phrase is described by the following gestures.

One issue related to interpreting intent based on hand shape is that the same hand shapes were often used to represent different things. For example, a flat hand was used to represent a piece of paper and a table surface. A chopping motion was used to describe books as well as consecutive rows of corn. Although they were used in various contexts, the use of the gestures had similar functions within the gestural phrase. For example, a flat vertical palm is effective for representing both paper and a tabletop because they are both flat. Chopping motions were useful for describing objects in linear rows because it speaks to consecutive ordering within space. When similar hand shapes were used to represent different objects, the visual representations that they signify typically share common features.

Understanding what a hand shape or pattern of motion is meant to represent requires an understanding of the context in which the gesture is produced. For this reason, interactive gestures need to be contextualized by the visual scenarios that are being presented within the interface. Designers need to address the visual interface with consideration of the role it plays in terms of situating or contextualizing motion. Within the information hierarchy of gestural interfaces, there may need to exist some kind of framework for interpreting the meaning of a gestures based on co-occurring visual imagery.

Many features of gestural expression produced during this study affirm the notion that, although gesturers are spontaneously generated, there is still a rational involved in the production of gestures. When people began to describe the scene, they paused when they realized that they had not defined necessary information through staging or framing. In this case, they might either start over or communicate the needed information before beginning again. I found this interesting because it suggests that they are not operating off of a clear plan. They are adopting new techniques as they go and they had the ability to quickly change their gesturing technique when the information required it or their previous method would not work. As in speech, communication methods operate more efficiently when they involve a plan or a well thought out strategy. Sometimes it seems as though a plan can be developed quickly and on the fly, but when the plans do not seem to work, there is a point at which learning occurs and a new strategy can be developed.

As with other interactive systems, gestural systems of interaction will need to provide feedback to users that help to clarify methods of behavior modification that will lead to more desirable outcomes. This is a much more complex issue for gestural modes of interaction than for traditional modalities which have more linear sequential methods of address users status or progression. One discussion of the pen interface used for the Palm Graffiti interface made clear to me importance of this issue. The following relevant story is paraphrased from Bill Moggridge's book *Designing Interactions* (2007).

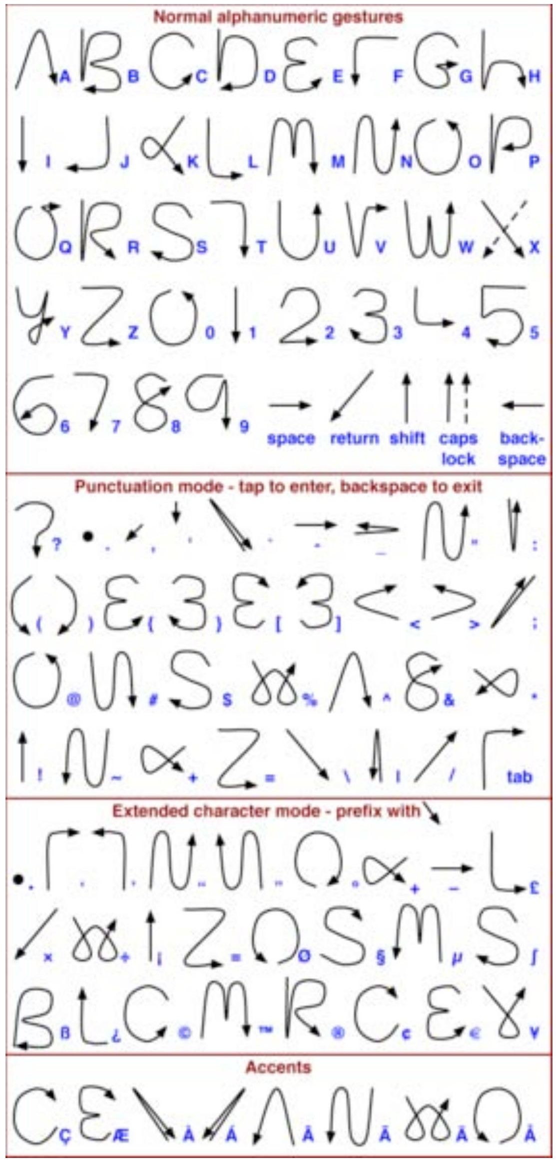


Figure 72 - Gestures used by original Palm OS handheld computers

When Jeff Hawkins began developing the handwriting recognition software for Palm Pilots he was troubled in his attempts to recognize various handwriting styles. Early prototypes seemed to only be efficient at recognizing styles of handwriting that were similar to *his* writing style. His moment of epiphany came from his discoveries during his study of neuroscience. He became familiar with the fact that human brains want a consistent model. People want to be able to predict what the results of their action are going to be. They also want to know how to modify their behavior in order to achieve more predictable results. He decided that asking people to write in a specific way created more simplified and reliable system than trying to recognize the differences between writing styles.

In a similar fashion, the most effective models of gestural interaction will not likely be systems that recognize variations in styles of gesticulation. Instead, the most effective models will provide users with simple systems of representation that they can easily adapt to. Optimal systems will provide feedback that helps users understand how they can modify their behavior to get more desirable outcomes. This is not to say that gestures should be arbitrarily defined so that we can conform our gestures to those expected by the computer. On the contrary, it means finding the most consistent model that represents a commonly agreed upon method of representation, such an alphabet.

Similar to handwriting style, we all possess unique styles of gesture. Yet as with handwriting, there are elements of a typified form. As with any language, conformity to a *type form* allows more articulate expression because there is a standard of communication in which to express one's self to others in a common language. Developing gesture libraries involves defining "type forms" of expression. *Gestural Type forms* are gestural phrases that include hybrids combination of several variables of gestural expression and can include several co-occurring gestural phrases. Unlike other, more codified languages, gestures have a recognizable visual or formal relationship to the things that they represent. Type forms of expression often correspond to the visual-spatial information that they represent. For this reason, their interpretation is closely tied to the context in which they are used. Through a process of contextualization, interaction design can create a visual-spatial framework for understanding the meaning of a gesture and the potential influences it can have on an interactive scenario. Similar patterns of hand shape, movement, and orientation often have fairly consistent meaning when produced in response to the same visual stimuli, which suggests that there is a potential for development of standardized gestural repertoires for gestural interface interaction design.

3.6 Findings

As a result of this study, much of the potential for integration of gesture into existing interactive schemas has become clearer to me. First, there is the implication that the duality of gestural staging and enacting has upon the design of interactive schemas. By isolating these two distinct representational spaces, it will be easier to contextualize gestures into interactive schemas based on those intended to be constructive within interactive space and those intended to influence the created space.

Second, I see opportunities for greater contextualization of traditional and gestural interaction through the use of gestural signifiers. Through observation of gestures it can be seen that certain gestures have roles as modifiers or augmenters of other co-occurring gestural phrases. Similarly, gestures used within future interfaces could contextualize interaction or influence the modality in which other more traditional modes of interaction occur. As we look more deeply into the spatial and temporal relationships in which gestures involve, we will see many potential to remap these relationships into methods that provide users with new ways of directing their interactive experience.

Lastly, there are several potentials for gestural vocabularies that will be more easily adopted than others because they are coupled with common methods of gestural representation. Moving forward from elements of our common gestural repertoires means greater consistency and unity in development approach, but it also means greater potential for fluid articulation on the part of users. By using gestures that align to common gesture usage, users may feel as though the interface is oriented toward their method of communication, rather than feeling as though they must conform their actions to those expected by the interface. Greater universality will be necessary for successful integration and greater public acceptance of gestural technologies.

In the previous section several gestural repertoires were discussed which may inspire or informed the process of developing new gestural libraries for gesture-driven interactive prototypes. The results of this approach may produce valuable information that can be used to inform the design process through iterative development, prototyping and testing. Collective results may lead to more universal and usable gesture libraries and methods of applying gestural expression. This method could potentially lead designers closer to the development of standards for gestural interface design. Results from this approach can vary widely, which is why many iterations of this study could be conducted without fully exhausting the potential for new discoveries. This is just one possible approach to expanding our understanding of gesture in a way that produces actionable information for designers. This study may be conducted with different gestural-interaction-related subjects of inquiry and still produce relevant information. As this study is reproduced, various gestural interaction schemas can be gathered to create a broader record and more articulate definitions of the design potentials for gestural interaction.

As I developed and conducted the test plan for my gesture archive, several factors seemed to require further attention including: 1) the impact of personality type of users 2) the influence of the test plan and 3) test monitor script and the influence of the presentation of visual stimuli. Variations of these elements could produce a wide range of gestural expressions. Exploring these territories could lead to many yet to be discovered factors related to gestural interface design.

Re-mapping features of representation to different kinds of gestures has several benefits. First, several characteristics of gesture can be represented simultaneously, which begins to address interrelationships between several variables and points of input. Second, by simultaneously mapping patterns like 3D planes of motion, locus of motion, hand shape and orientation or other patterns of motion these motions become contextualized one to another. It is not sufficient to include static hand positions or single insular strokes. It is easier to synchronize various components symphonically through input devices that capture a larger array of spatial movement and temporal syncopation. Through this research it was never my intent to define sets of static poses or sequences of hand poses with definitive meaning. Rather, I hope my readers can begin to consider how hand gestures could be used in a many similar to the way they are used by a musician in relationship to his instrument. Musicians use a limited set of gestural repertoires, yet through careful practice they develop skills that allow them to articulate limitless variations in expression. As with the design of any quality instrument, it is crucial that gestural interaction designers define appropriate gestural repertoires that allow people to express themselves in a fluid and articulate manner.

In terms of computer use and the communication of ones' intent, there are many types of gestural representation that offer simpler and more effective modalities of digital interaction than those yet seen. Perhaps the best integration of gestural representation will be as a communication layer that exists amidst other forms of representation such as written language, sounds, and visual imagery. Defining a path toward successful integration will require a combined understanding of both existing computing infrastructures and modalities of gestural communication that enable fluid and articulate expression.

Results gathered from this study will allow gesture-based interaction designers to move forward with greater confidence because it provides a large amount of evidence that they can respond to and build upon. During this study, many of the visual scenarios that were presented to subjects provoked similar or even identical forms of gestural representation. This suggests that there is a potential for implementing standards within gesture-based interactive scenarios.

Evidence from this study suggests a rationale arising within gestural expressions that is linked to common visual and kinetic experiences. The gesture archive created for this research study provides specific visual and spatial reference points from which the relationship between visual-spatial imagery and gestures can be decoded. This process of decoding is a necessary simplification of gestural expression that will be required if gestural expression will be more fully integrated into computing experiences.

CHAPTER 4 CONCLUSIONS

From this final research study, I hope that readers have gained a new perspective on the potentials of gesture-based interaction. It is my intent to reach out to other designers and help them to develop new ways of thinking about integration of gestures into interactive design. After reviewing many of the scenarios describing the gestural expressions produced by research subjects, I hope to reaffirm the value of approaching gesture-based interaction design through user-centered research. By including early design research processes and discussing developments that lead to and help define my goals for the final research study, I hope to have adequately framed some of the issues that will be relevant to designers as they attempt to design gesture-based interactive experiences. This research study proposes new ways of framing design issues related to gestural interface design. Emphasizing the importance of designers and there role in the development of interactive media will be an important step that prevents technological issues from overshadowing designers' visionary role.

For me, this body of work represents a transformational journey in which I learned a great deal about human expression and human experience. I see many new potentials and a clearer path towards integration than I could have imagined prior to undertaking this research endeavor. As I create new works, I will attempt to implement the knowledge I have discovered, but I go forward with the assumption that the more I know the more I will realize there is to know. As a result of this research, I know future work will involve application of some of the unique discoveries I made. In the future I will work to develop interactive scenarios that demonstrate paths toward greater feasibility. This will continue to be a process of experimentation and iterative design.

In addition, the need for new presentation methods has become more apparent to me. Improvements in the documentation and communication methods related to gesture based interaction design will be necessary, because representing gestural interaction means representing multiple layers of co-occurrence related to hand shape, motion and orientation. This will involve developing new information hierarchies, development processes and approaches to digital content management. Unlike static wireframes, which are traditionally used by user experience designers, gestural interface design will require presentation of multiple layers of information simultaneous. It may also need to be understood in relationship to the changing shapes of the hands and their orientation to the body. I have already begun to experiment with methods of video compositing using graphical elements overlaying staged motions that pantomime a users method of interaction. Methods such as this will be needed in order for designers to communicate the experiences they desire to create. Similar design methodologies will also be a means of testing approaches prior to more expensive, time consuming and deterministic technological implementation. Proper presentation of an interface will makes for a much more convincing argument when seeking stakeholders, investors or collaborators.

It is important to consider the fact that gesture-driven media appeals to our kinesthetic intelligences, a form of human intelligence that is frequently under-represented by current technology. If technology is, as McLuhan suggests, an extension of our abilities and senses, then many of our sensory abilities related to touch and kinesthetic action are under represented by our technologies. Perhaps the reason we have focused less on extending our tactile and kinesthetic selves is because these are sense and abilities we cannot understand as easily. McLuhan's ideas present the notion that media is an extension of ourselves that forever changes our makeup and shifts the balance of our conscious experience. We can see evidence of his statement and feel the imbalance of which he writes when observing technologies effect on modern society. In human history you are unlikely to find a more sedentary, disembodied and detached technological cultures. Following McLuhans's logic only a greater awareness of the deficit that our media has created within our culture might lead to new, more balanced solutions. As McLuhan suggests that, ".. it is the tactile sense that demands the greatest interplay of all the senses"

REFERENCES

3M (2009) "3M Touch Systems" Retrieved March 30, 2009 at <http://www.3M.com/touch>

Alibali, Martha W. (2005). Gesture in Spatial Cognition: Expressing, Communicating, and Thinking About Spatial Information. SPATIAL COGNITION AND COMPUTATION, *5*(4), 307–331. Lawrence Erlbaum Associates, Inc.

Arduino (2009) "Arduino" Retrieved on January 30, 2009 at <http://www.arduino.cc/>

Berda, Tomer .(2004) "DirectCommunication Xtra™ Version 2.0 Documentation". DirectXtras. Retrieved January 26, 2009 from <http://www.directxtras.com/DComm_doc.asp?UUID=2277260>

Bodanis, David. Electric universe : the shocking true story of electricity. New York : Crown Publishers, c2005

Buxton, W. (1990). The Natural Language of Interaction: A Perspective on Non-Verbal Dialogues. In Laurel, B. (Ed.). *The Art of Human-Computer Interface Design*, Reading, MA: Addison-Wesley. 405-416.

Buxton, Bill. (2007) Sketching User Experiences: getting the design right and the right design. Elsevier, Inc.

Brown, Stuart. (2008) "Hands-On Computing: How Multi-Touch Screens Could Change the Way We Interact with Computers and Each Other " Scientific American, June, 2008. Retrieved on January 30, 2009 at <http://www.sciam.com/article.cfm?id=hands-on-computing>

Camurri, Antonio, Mazzarino, Barbara, Volpe, Gualtiero (2003) "Analysis of Expressive Gesture: The EyesWeb Expressive Gesture Processing Library". Gesture-Based Communication in Human Computer Interaction: 5th International Gesture Workshop, GW 2003 Genova, Italy. Springer.

Cassel, R. and Collet, C. (2004) "Tracking of Real Time Acrobatic Movements by Image Processing". Gesture-based Communication in

Human-Computer Interaction: 5th International Gesture Workshop. Genova Italy.

Chun, Ben (2006) “flosc: Flash OpenSound Control”. Retrieved on September 15, 2007 at <http://www.benchun.net/flosc/>

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. Cognition, 23, 43-71.

Cravotta, Robert. (2007) “Recognizing gestures: Interface design beyond point and click”. EDN. Retrieved January 20, 2009 from <http://www.edn.com/index.asp?layout=article&articleid=CA6466206>

Dejean, V. (2009) “Autism, Tomatis and Sensory Integration” Retrieved January 30, 2009 from <http://dejean.net/beforewords.html>

Donald, Merlin. (1991) Origins of the Modern Mind. Three stages in the evolution of culture and cognition. Harvar University Press.

Ehrlich, Stacy B. , Levine, Susan C. and Goldin-Meadow, Susan . (2006) The Importance of Gesture in Children’s Spatial Reasoning. The American Psychological Association, Developmental Psychology, Volume 42, Number 6, p. 1259-1268.

Emory, Karen. (2002) Language, Cognition and the Brain: Insights from sign language research. Lawrence Erlbaum Associates.

Endo, Ayaka and Kahura, Yasuo. (2008) “Rhythmic Instruments Ensemble Simulator Generating Animation Movies Using Bluetooth Game Controller” *Proceedings of the 2005 International Conference on New Interfaces for Musical Expression 2008 (*NIME08) June 4-8, 2008, Genova, Italy

Faconti, G. P. (1996) “Reasoning on Gestural Interfaces through Syndetic Modelling” SIGCHI Bulletin Volume 28, Number 3, July 1996, p. 71-77

Fitzmaurice G, Ishii H, Buxton W. (1995) “Bricks: Laying the Foundations for Graspable User Interfaces”. In Proceedings Of the CHI ´95, 1995, pp. 442-449.

Forbus, Ken. “CogSketch: Software for studying and enhancing spatial reasoning” Spatial Intelligence and Learning Center. Retrieved January 20, 2009, from <http://www.silccenter.org/projects/cogsketch_index.html>

Gale, Bruce. (2007) “Conducting”. The Concert Band. Retrieved March 6, 2007,
from < http://www.theconcertband.com/conducting.htm>.

Gleicher, M. and Ferrier, N. (2002) "Evaluating Video-Based Motion Capture" Proceedings of Computer Animation 2002.

Goldin-Meadow, S. (2000). Beyond words: The importance of gesture to researchers and learners. Child Development , 71, 231-239. Links

Goldin-Meadow, S. (2003). Hearing gestures: How our hands help us think. Cambridge, MA: Harvard University Press.

Goldin-Meadow, S. (2003). The Resilience of Language: What Gesture Creation in Deaf Children Can Tell Us About How All Children Learn Language. New York: Psychology Press.

Goldin-Meadow, S. (2004). Gesture's role in the learning process. Theory into Practice 43, 314-321.

Goldin-Meadow, S., Kim, S., & Singer, M. (1999). What the teacher's hands tell the student's mind about math. Journal of Educational Psychology , 91, 720-730.

Han, Jefferson (2005) "Low Cost Multi-touch sensing through frustrated total internal reflection" ACM. *UIST'05*, October 23–27, 2005, Seattle, Washington, USA.

Handspeak: American Sign Language Online Dictionary, Baby Sign, International Sign, and more. (2006). Jolanta A. Lapiak. <www.handspeak.com>

Inition, inc. (2007) "Monkey 2" Retrieved January 20, 2007 at <http://www.inition.co.uk>

Infomus Lab (2007) "InfoMus Lab: EyesWeb Tutorials". Retrieved on May 15, 2007 at <http://www.infomus.dist.unige.it/EyesWeb/EywTutorials.html>

Kaltenbrunner, Martin , Bencina, Ross. (2007) "ReacTIVision: a computer-vision framework for table-based tangible interaction", Proceedings of the 1st international conference on Tangible and embedded interaction, February 15-17, 2007, Baton Rouge, Louisiana

KCP Technologies (2008) "Geometer's Sketchpad Resource Center" Retrieved March 30, 2009 at <http://www.dynamicgeometry.com/>

Kendon, Adam. (2004) Gesture: Visible action as Utterance. Boston: MIT Press.

Kuhlman, Lane (2007) “Building Virtual Environment”. Retrieved January, 20, 2009, http://accad.osu.edu/~lkuhlman/building_ve.html

Kuhlman, Lane and Price, Alan (2008) “Multi-User Interface Research”. Retrieved May 20, 2009 at <http://accad.osu.edu/~lkuhlman/multi_user_interface.html>

Lee, Johnny Chung (2008) “Johnny Chung Lee > Projects > Wii” Retrieved January 30, 2009, <http://www.cs.cmu.edu/~johnny/projects/wii/>

Make (2009). “Make: technology on your time” Retrived on January 30, 2009 at <http://makezine.com/>

McNeill, David (2000) Language and Gesture. Cambridge University Press.

McNeill, David (2005) Gesture and Thought. University of Chicago Press.

Microsoft (2009) “Microsoft Surface”. Retrieved January, 20, 2009, <http://www.microsoft.com/surface/>

Michigan State University Communication Technology Laboratory. (1997) “American Sign Language Browser”. Retrieved January 20, 2007 at <http://commtechlab.msu.edu/Sites/aslweb/index.html>

Mitchel Resnick , Fred Martin , Robert Berg , Rick Borovoy , Vanessa Colella , Kwin Kramer , Brian Silverman. (1998) “Digital manipulatives: new toys to think with”. Proceedings of the SIGCHI conference on Human factors in computing systems, p.281-287, April 18-23, Los Angeles, California, United States

Moggridge, Bill. (2007) Designing Interactions, Boston: MIT Press.

Molich, R., and Nielsen, J. (1990). Improving a human-computer dialogue, Communications of the ACM 33, 3 (March), 338-348.

Montessori, Maria. (1967) The Discovery of the Child. Fides Publishers, Inc.

Music Technology Group (2009) “Reactivision 1.4” Reactable. Pompeu Fabra University, Barcelona, Spain. Retrieved on January 30, 2009 at <http://mtg.upf.es/reactable/?software>

Nielsen, Michael, Storring, Moritz, Moeslund, Thomas B. and Granum, Erik. (2003) A procedure for Developing Intuitive and Ergonomic Gesture Interfaces for HCI. Gesture-Based Communication in Human Computer Interaction. 5th International Gesture Workshop, GW 2003. Springer-Verlag Berlin, Heidelberg.

NUI Group (2009) “NUI Group” Retrieved on January 30, 2009 at <http://nuigroup.com/>
O'Malley, C. and Stanton-Fraser, D. (2004). Literature review in learning with tangible technologies. Nesta FutureLab Series, report 12.

OpenCV (2008) “Open CV Wiki” Retrieved on January 30, 2009 at <http://opencv.willowgarage.com/wiki/>

O’Sullivan, Dan & Igoe, Tom. (2004) Physical Computing. Boston: Thomas Course Technology, 2004.

Perceptive Pixel (2009) “Perceptive Pixel” retrieved on January 30, 2009 at <http://www.perceptivepixel.com>

Parallax (2009) “Parallax, Inc.: Basic Stamp Information” Retrieved on January 30, 2009 at <http://www.parallax.com/tabid/295/Default.aspx>

Pergams, O. R. W. and P. A. Zaradic. 2006. Is love of nature in the US becoming love of electronic media? 16-year downtrend in national park visits explained by watching movies, playing video games, internet use, and oil prices. Journal of Environmental Management 80:387-393.

Phiffer, D. and Zer-Aviv, M. (2006) “Atlas Gloves: A DIY hand gesture interface for Google Earth”. Retrieved January 30, 2009 at <http://atlasgloves.org/diy>

Price, Alan (2007) “Virtools Plugin”. Retrieved January 20, 2007 at http://accad.osu.edu/~aprice/works/plugins

Processsing (2009) “Processing 1.0” Retrived on January 30, 2009 at <http://processing.org/>

Raffle, Hayes Solos, Parkes, Amada J., Ishii, Hiroshio, (2004) Conference on Human Factors in Computing Systems,Volume 6, Issue 1, 2004: Topobo: A Constructive Assembly System with Kinetic Memory New York: Association for Computing Machinery, Inc.

Reactable (2009) “Reactable” Retrieved April 20, 2009 at <http://accad.osu.edu/~aprice/works/plugins>

Rhyne, Jim (1987) Dialogue Management for Gestural Interfaces. Computer Graphics, Volume 21, Number 2, April 1987. Pages 137-143.

Rizzolatti, Giacomo and Arbib, Michael A. (1998) Language within our grasp. Trends Neurosci. 21, 188–194

Rohr, K. (1997) Human movement analysis based on explicit motion models, Kluwer Academic Publisher, Dordrecht/Boston.

Rossini, Nicla (2004) The Analysis of Gesture: Establishing a Set of Parameters. In Camurri, A.-Volpe, G. (Eds) Gesture-Based Communication in Human-Computer Interaction. 5th International Gesture Workshop, GW 2003, Genova, Italy, April 2003. Selected Revised Papers: 124-131. Springer-Verlag: Berlin Heidelberg New York.

Saffer, Dan (2007) "Interactive Gestures" Interactive Gestures Pattern Library. Retrieved January 20, 2009 from <http://www.interactivegestures.com>

Saffer, Dan (2008) Designing Gestural Interfaces. O'Reilly Media, Inc.

Sanders, Liz (1992) "Converging Perspective: Product Development Research for the 1990s" Design Management Journal, Volume 3, Number 4. Design Management Institute.

Sanders M, McCormick E (1993) Human Factors in Engineering and Design. McGraw Hill.

Shen, C., Forlines, C., Wigdor, D. Vernier, F. (2007) "Letter to OS Designers from the Tabletop Research Community" retrieved on March 31, 2009 at <http://www.diamondspace.merl.com/>

Smith, J. White, T. Dodge, C. Allport, D. Paradiso, J. Gershenfeld, N. (1998) "Electric Field Sensing for Graphical Interfaces" IEEE Computer Graphics and Applications, Volume 18, Number 3, pp.16-22

Standing, E.M. (1962). Maria Montessori: Her Life and Work. Fresno, California: Academy Guild Press.

Suominen, J., and Keune, A. (2007) "Gimme Sugar" Paja: Technology for music, art and design. Retreived on January 30, 2009 at <http://mlab.taik.fi/paja/?p=232>

Valli, Alessandro (2007) "Natural Interaction White Paper" retrieved on March 31, 2009 at <http://www.naturalinteraction.org/images/whitepaper.pdf>

Vicon Motion Systems, 2009. http://www.vicon.com.

Yinlin, L., Groenegress, C., Wofgange, S. and Fleischmann, M. (2004) Gesture-based Communication in Human-Computer Interaction: 5th International Gesture Workshop. Genova Italy.

Zuckerman, O. , Arida, S. , and Resnick M., (2007) Proceedings of CHI 2005: Extending Tangible Interfaces for Education: Digital Montessori-inspired Manipulatives New York: Association for Computing Machinery, Inc.

APPENDIX

A. Notes on Gesture Classification

The area of cognitive science known as gesture studies deeply investigates the use of gesture within daily human communications. Various researchers investigating gesture categorize gestures in different ways. Many different interpretations of gesture apply to interface in various ways. A few notable categorizations relate to the focus of this investigation. Cadoz (1994) groups gestures into three types based on their functions.

• semiotic: those used to communicate meaningful information.

• ergotic: those used to manipulate the physical world and create artifacts

• epistemic: those used to learn from the environment through tactile or haptic exploration

This research is primarily concerned with semiotic gestures because these gestures may be used to communicate information to a computer. McNeil (1992) divides this categorization down into 4 groups.

• Iconic Gesture - used to convey information about the size, shape or orientation of a personal, place or thing

• Metaphoric Gestures – gestures that represent information through descriptive, temporal-spatial metaphors. They represent abstract ideas rather than concrete objects.

• Deictic Gestures – use to indicate people, places or things that are real or imagined

• Beat Gestures – Beat musical time or used for emphasis

Ekmans and Friesen break down what McNeil's "iconic gestures" into 3 smaller yet more discrete categories.

• Kinetographic Gestures- Gestures used to depict bodily action

• Spatial Movement Gestures – Movement of people, places or things through space,

• Pictographic Gestures – Gestures used to depict the shape of people, place or things

To recap Kendon's definition, a communicative gestural phrase is comprised of 3 phases of movement :

1) preperation – when stroke leaves the resting posito and moves to the part where the meaningful gesture will be perfoomred.

2) nucleus – the meaningful part of the gesture in which formal, spatial and temporal characteristics of the hands and limbs express meaning

3) retraction/reposition – post-nucleus phase where gesturer returns to a static postion

These aspects have been further expanded upon to include:

1) size – distance between the beginning and end of the stroke

2) gesture timing - length of time between the beginning and end of the stroke

3) point of articulation – main joint involved in the gestural movment

4) Locus – body space involved by the gesture included head, bust, torso and periphery areas

5) x,y, and z axis – location of gesture within an imposed imaginary spatial plane.

B. Gesture Archive Classifications

1. Function
    a. Staging
    b. Enacting
    c. Shape Defining
    d. Path Defining
    e. Framing
2. Gesture Types
    a. Diectic Gesture
    b. Pictographic Gesture
    c. Kinetographic Gesture
    d. Spatial Movement Gesture
    e. Shape Transformation
    f. Beat Gesture
3. Hand Shapes
    a. Flat hand
    b. Ring Shaped
    c. Arch shaped
    d. Pinch
    e. Precision Grip
    f. Fingers Spread
    g. Fist
    h. Pointing
    i. Cupped Hand
    j. Index & middle finger extended
    k. Two-handed Triangle
    l. Curled fingers
    m. Straight fingers
    n. Fingers bunched
    o. Finger sides together
4. Hand Orientation
    a. Palm down
    b. Palm up
    c. Palm forward
    d. Palm toward body
    e. Palm facing palm
    f. Palm outward
    g. Palm inward
    h. Horizontal Palm
    i. Vertical Palm
    j. Vertical Fingers

k. Horizontal Fingers
l. Fingers Leading Hand
m. Elbow above hand

5. Movement Patterns
    a. Bi-Lateral Symmetry
    b. Bi-Lateral Asymmetry
    c. Hand Waving
    d. Vertical oscillating
    e. Horizontal oscillating
    f. Wiggling Fingers
    g. Curling Fingers
    h. Circular
    i. Vertical Planer
    j. Horizontal Planer
    k. Cylindrical
    l. Planer Rebound
    m. Vertical Chopping
    n. Horizontal Chopping
    o. Stacking flat palms
    p. Pitching wrist
    q. Rolling wrist
    r. Touching Palms
    s. Hand over hand
    t. Finger(s) drawing
    u. Clenched to open hand
    v. Open to clenched hand
    w. Interwoven fingers
    x. Along Arm
    y. Pantomime
6. Prepositions
    a. On
    b. Over
    c. Through
    d. Under
    e. Above
    f. Between
    g. Below
    h. Inside
    i. Outside
    j. Against
    k. Along
    l. Beside
7. Verbs
    a. Sink
    b. Float
    c. Pour

d. Collide
e. Organize
f. Scatter
g. Stack
h. Climb
i. Slide
j. Swing
k. Wave
l. Oscillate
m. Wrap
n. Contain
o. Tilt
p. Bounce
q. Break
r. Align
s. Straighten
t. Roll
u. Flow
v. Push
w. Pull
x. Transfer
y. Separate
z. Combine

8. Locus
   a. Chest Region
   b. Face Region
   c. Waist Region
   d. Below-Waist Region
   e. Above Head
   f. Body Periphery
   g. Vertical Plane
   h. Horizontal Plane
   i. Contacting Arm
   j. Contacting Hands
   k. Contacting Table

C. Video clips viewed by participants

Aligning books on a shelf

Tilting a shelf of books upright

Stacking books

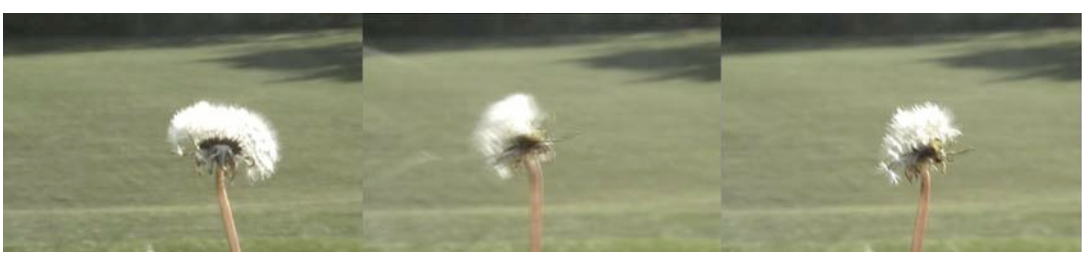
Ball Bouncing

Egg Breaking

Windy Cornfield

Tree leaves blowing in the wind

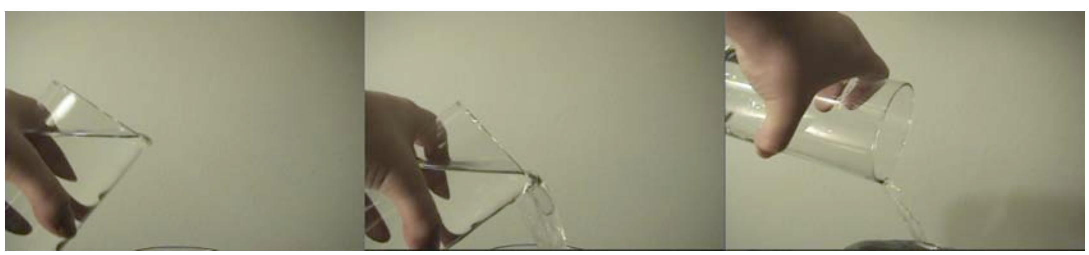
Dandelion seeds scattering in the wind



Glass being filled with water

Water overflowing from a glass

Water being poured

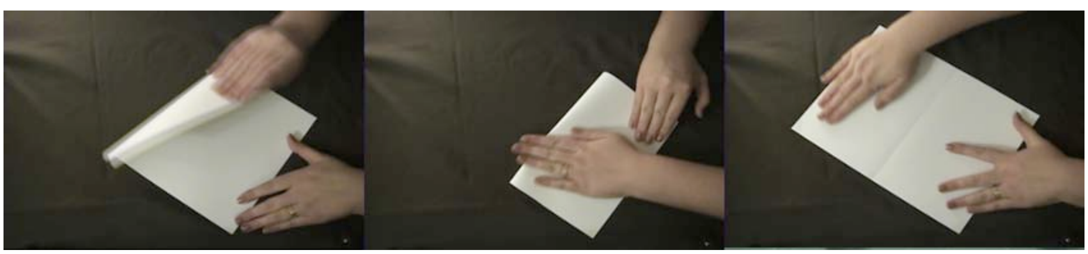


Rolling downhill

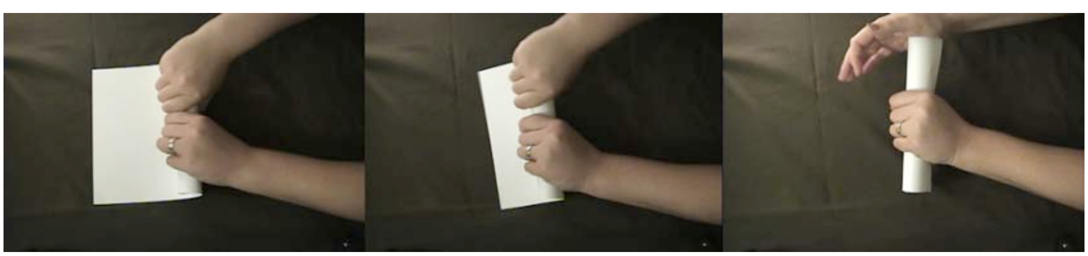
Sliding down a slide

Crawling through a tube

Folding a piece of paper



Rolling a piece of paper



Pool Ball being racked

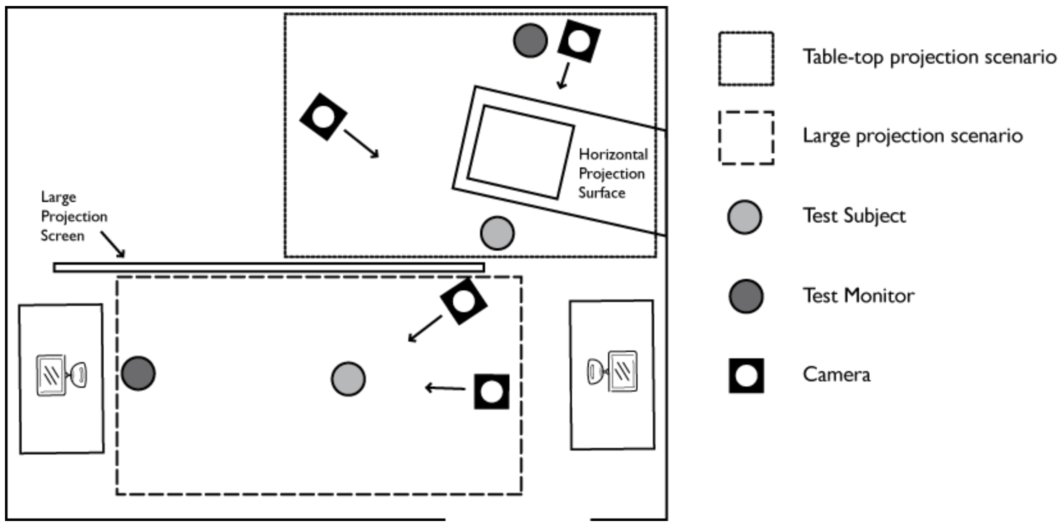
Pool Ball breaking apart

D. Test Monitor Script

"For this study you will be recorded by several video cameras. It is important that you sign a release for allowing the analysis and review of this video. Myself and other investigators may review the video. The video may also be made public to other interested researchers through presentations at conferences. Your verbal responses and video clips of your hand motions that enact during this study, which may include images of you face, may be posted online. At no point during the presentation of results or in the context of videos online will your name or any of your personal information be shared. If you accept these terms please sign the forms and we may proceed. Also please note that you are under no obligation to proceed through this study after beginning. You may withdrawal your consent to participate at any time after we begin. Your consent is voluntary.

During this study you will be presented with a series of pictures. After being presented with a sequence of images you will be asked to act out a gesture with your hands that you feel expresses the action or representation described by the image sequence. For each sequence feel free to speak the information that you are attempting to communicate out loud. You may also be asked to elaborate verbally upon the meaning of your hand gestures. The study should take less than 30 minutes to complete. It should also be noted that there are no wrong answers so don't be afraid to express yourself. If you are at all confused during the study please do not be afraid to ask for clarification. Thank you for participating in this study. If you are ready we can proceed."

E. Test Environment



NOTE: Participants who responded to laptops met in various settings at their convenience