A taxonomy of Gestures in Human Computer Interaction

Maria Karam and m.c. schraefel
University of Southampton

We present a classification of gesture-based computer interactions motivated by a literature review of over 40 years of gesture based interactions. This work presents a unique perspective on gesture-based interactions, categorized in terms of four key elements: gesture styles, the application domains they are applied to, input technologies and output technologies used for implementation. The classification provides a means of addressing gestures as an interaction mode across the different application domains so that researchers and designers can draw on the vast amount of research that has been addressed within the literature from an interaction perspective.

Categories and Subject Descriptors: H.5.2 [**User Interfaces**]: Input Devices and Strategies

General Terms: Gesture Interactions, Taxonomy, Classification

Additional Key Words and Phrases: HCI, Gestures, Taxonomy

1. INTRODUCTION

In this paper we present a taxonomy of gestures as a human computer interaction technique. For over 40 years, almost every form of human gesturing that is possible can be seen in the literature as a means of providing natural and intuitive ways to interact with computers across most, if not all computer application domains. In addition, almost all input and output technology has been used to enable gesture based interactions.

Though there is such a diversity within the field of gesture based computer interactions, this paper presents the foundation of a taxonomy that will provide a unified perspective of gestures within the field of computer science that spans the application domains as well as the technologies used in their enablement.

In an attempt to encompass the entire field of gesture based interactions into a single field of study, it is necssary to consider all of the different froms in which gestures are used. We propose the following approach: Gestures exist in different forms within different application domains. Within the domains, we consider the various I/O devices. Simply put, there are four categories within the proposed taxonomy: Gesture style, application domain, enabling technology (input) and system responses (output).

Our work is based on reviewing gesture based interactions from over 40 years of computing literature. Although this work does not claim to provide an exhaustive study of all the literature that considers gesture as an interaction mode, we do hope it will provide a concise overview of the space. With this approach, we hope that interaction researchers and developers of gesture based interactions can begin to build on the knowledge gained across the different domains to move gesture based interactions out of the research labs and into everyday computing applications.

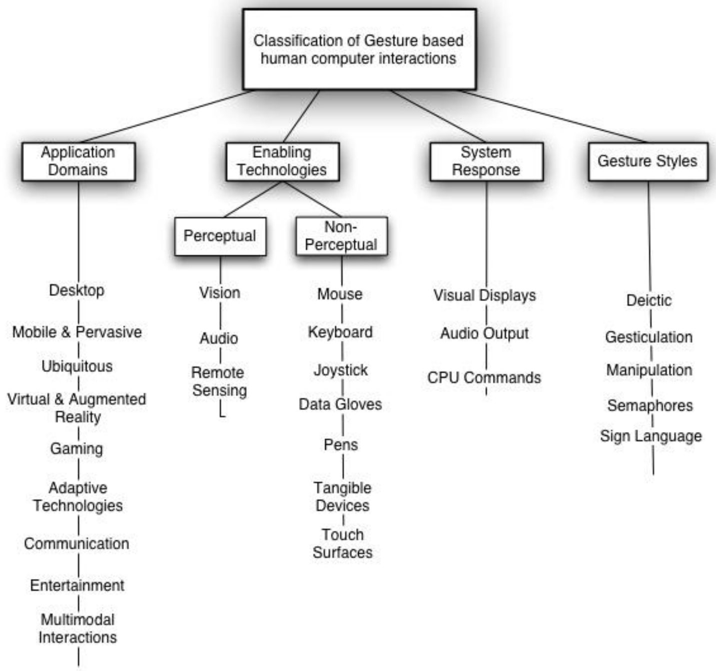

Fig. 1. Visualizing the taxonomy: The diagram shows the organization of the research on gestures based on the four categories used in the taxonomy. One thing to note with this taxonomy is that it can be rearranged based on the perspective of the research or researcher. For example if we are coming from the perspective of doing interaction research within a particular domain, then our input and output and gesture choices for the interaction are restricted by what is appropriate or available within the constraints of that domain. The diagram represents the classification of the gesture based human computer interaction literature reviewed for the taxonomy

In addition to the taxonomy, we also provide an analysis of gesture based interactions in terms of the types of evaluations that have been conducted in different domains, the problems that gestures address when used as an interaction technique and discuss some of the areas within the field that have yet to be addressed in the literature and that may provide a view into possible future directions of gestures as an interaction technique.

2. HUMAN GESTURING FOR COMPUTER INTERACTIONS

The first category of this taxonomy addresses gesture styles that are used in computer interactions. Most of the theory behind gestures as a computer interaction technique originates in the multidisciplinary research field of human gesturing, drawing on concepts from linguistics, anthropology, cognitive science and psychology for example. However, for this taxonomy, we focus on a high level classification of human gesturing styles provided by several researchers in computer science including Francis Quek [Quek et al. 2002] and Alan Wexelblat [Wexelblat 1998]. For a more detailed persepctive on the study of human gesturing, we refer readers to two external sources including The International Society for Gesture Studies and a Technical Report from Simon Fraser University [ges a; b]. In the next section, we present an overview of some commonly used forms of human gesturing as computer interaction techniques.

2.1 Study of Human Gesturing

As reviewed in the literature, several researchers in the field of human gesturing are most commonly referenced when describing the types of gestures used for computer interactions. These researchers include Ekman, Mespoulos and Lecours however it is the work of Kendon who specialized in gesture studies - and McNeil from a linguistics perspective - that have been most commonly referenced when describing the physical mechanics of a gesture [Quek et al. 2002; Eisenstein and Davis 2004; Kettebekov 2004; Wexelblat 1998]. An overview of the primary classifications referred to in some computing literature is provided by Alan Wexelblat [Wexelblat 1998].

2.2 Terms used to describe gestures for computer interactions

Within the computing literature, there are several researchers who attempt to provide a classification of the physical characteristics of gestures, including Alan Wexelblat [Wexelblat 1998], Dr. Sanshzar Kettebekov [Kettebekov 2004], Margot Brereton [Brereton et al. 2003], and Vladimar Pavlovic [Pavlovic et al. 1997]. Based on our review of the literature, we felt that by extending the the work of Francis Quek et al.[Quek et al. 2002] to include the major forms of gesturing discussed in the literature for this paper we could form the basis of this category of the taxonomy. Quek et al have proposed a framework for classifying gestures for computer interactions [Quek et al. 2002], defining three approaches to gestures, that they claim are the primary focus of existing gesture based computer interactions: manipulation, semaphores and gesture-speech approaches. Quek et al's work is based in gesticulation, or gesture and speech interfaces where gestures accompany speech for a more 'natural' interaction using bare hands. The extension to his framework proposes to include deictic or pointing gestures and language gestures for our taxonomy.

2.2.1 *Standardizing gesture terms.* Along with the terms referred to by Quek and our extension of his work, the computing literature presents many different terms to represent similar gestures. For example, gesticulations are also referred to as coverbal, pantomimes or natural gestures [Kettebekov 2004; Quek et al. 2002; Wexelblat 1995], while natural gestures are also used to refer bare handed or free handed gestures for example [von Hardenberg and Berard 2001; Baudel and Beaudouin-Lafon 1993; Eisenstein and Davis 2004]. In addition, symbolic gestures are also called iconic or stroke gestures depending on the author [Kopp et al. 2004; Koons and Sparrell 1994]. It is for this reason that we chose to build on Quek's framework which addresses many of these forms and provides clarification for the styles based on a review of past literature.

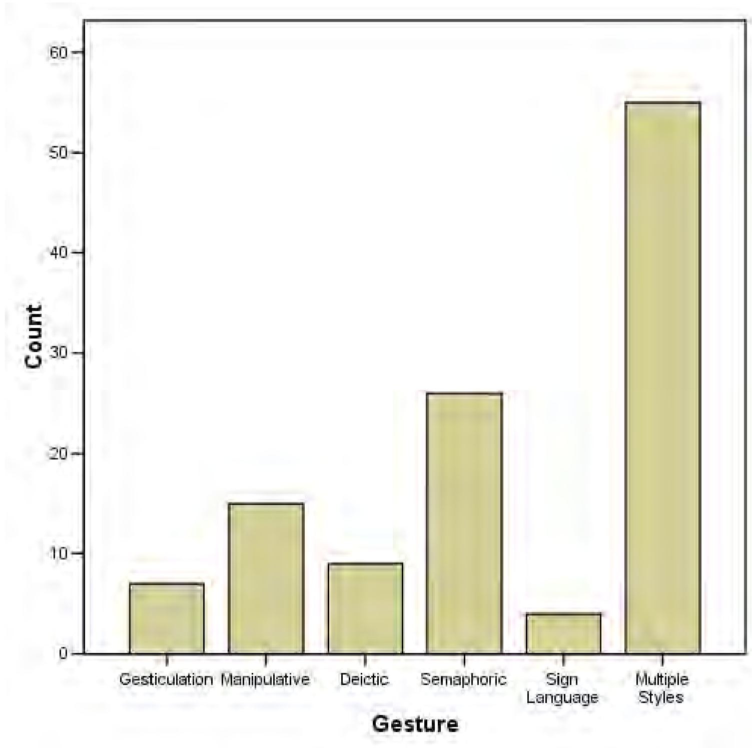

Fig. 2. The diagram represents the distribution of the different gesture styles that appear in the literature reviewed for this paper.

3. CATEGORIZATION BY GESTURE STYLES

Based on Quek's work and our review of the existing literature, the five categories of gesture styles we refer to in this category of our taxonomy are gesticulation, manipulations, semaphores, deictic and language gestures. The following section describes each of the terms based on their presence in the literature.

3.1 Deictic Gestures

Deictic gestures involve pointing to establish the identity or spatial location of an object within the context of the application domain. The application domain can include desktop computer, virtual reality applications or mobile devices for example. Although deictic gestures can be considered implicit in other forms of gestures, such as when pointing to identify an object to manipulate for example [Wellner 1991; Rubine 1992; Ward et al. 2000; Ou et al. 2003], many interactions use exclusively deictic gestures.

The first example is Bolt's "Put that there" work from 1980 [Bolt 1980]. Deictic gestures are used in conjunction with speech input in an interaction that allows the user to point at a location on a large screen display in order to locate and move objects. In addition to Bolt's work, deictic gestures are used to identify objects in virtual reality applications [Zimmerman et al. 1995], identifying objects to others in CSCW applications [Kuzuoka et al. 1994], for targeting applicances in ubiquitous computing [Swindells et al. 2002; Nickel and Stiefelhagen 2003], for desktop applications [Wellner 1991] and communication applications [Kobsa et al. 1986].

3.2 Manipulative Gestures

There are many different forms of manipulative gestures used for computer interactions however we refer to the following definition provided by Quek et al [Quek et al. 2002] who states that a manipulative gesture is one

> whose intended purpose is to control some entity by applying a tight relationship between the actual movements of the gesturing hand/arm with the entity being manipulated.

Manipulations can occur both on the desktop in a 2-dimensional interaction using a direct manipulation device such as a mouse or stylus, as a 3-dimensional interaction involving empty handed movements to mimic manipulations of physical objects as in virtual reality interfaces or by manipulating actual physical objects that map onto a virtual object in tangible interfaces. We now turn to a discussion of some different styles of manipulative gestures.

3.2.1 *Gesturing in two degrees of freedom for two-Dimensional interactions.* 2-dimensional objects in the computing literature refers to objects that are displayed on a 2-dimensional (2D) display. In some cases, the entity being manipulated is a cursor, a window or some other on screen object. The traditional interaction method for such displays normally involves the mouse, stylus or other 2 degrees of freedom (DOF) direct input devices for direct manipulation of objects in a graphical user interface (GUI). Direct manipulations are interactions frequently associated desktop applications and consist of actions such as dragging, moving and clicking objects and are not considered gestures. It is not until the user performs some gesture that must be translated as a command that we have a gesture.

One of the characteristics of a manipulative gesture is to provide parameters to the system that indicate the intent of the users request to move, relocate or physically alter the digital entity [Rubine 1992]. So that while clicking on or dragging an object is not considered a gesture, pointing at an object and then pointing at a different location to move the object is [Bolt 1980].

3.2.2 *Gesturing in multiple degrees of freedom for two-dimensional interactions.* 3-dimensional (3D) style manipulations of 2D objects involve such gestures as physically picking up and dropping a data file from one device and virtually moving it to another device located within a smart room environment [Rekimoto 1997]. Another technique for 3D style gestures involves pressure sensors that can be used to enhance direct manipulations with 2D data and displays as in a finger painting application that translates pressure into line thickness for example [Minsky 1984]. While the user must perform 2D direction movements, the additional dimensions of pressure and speed can also communicate additional information. In recent work, table top surfaces are fitted with electronic material designed to sense pressure and touch as well as movement and multiple points of finger and hand contact. Several systems including Rekimoto's smart skin enable manipulative gestures that are drawn from actual table top interactions such as sweeping and isolating groups of objects with the hands as one would do with physical objects on a table for example [Wu and Balakrishnan 2003; Rekimoto 2002].

3.2.3 *Gesturing with tangible objects for three-dimensional interactions.* Manipulative gestures styles also apply to interactions with physical objects used to represent digital objects. For example, tangible computing interfaces require the manipulation of physical objects in the real world that are mapped on to the digital objects they represent. Hinckley et al's work [Hinckley et al. 1998] uses a tangible doll's head to physically manipulate its graphical on-screen representation of a human brain. This type of interaction involves two levels of gesture manipulations: The physical manipulation of the doll's head and the resulting manipulation of it's on screen representation. For this classification, we refer to the manipulation of the physical doll as the gesture, and the transformations on the screen as the result of the output of the gesture.

3.2.4 *Gestures for real-world physical object interactions.* Additional manipulative gestures are used to enable the user to control the movements of physical objects such as a robot arm [Goza et al. 2004; Fisher et al. 1987] and a wheel chair [Segen and Kumar 1998b] for example. The literature also provides examples of manipulative gestures that involve physically manipulating computers through bumping gestures for initalizing data transfers [Hinckley 2003] or by shaking a mobile phone in order to verify the devices identity as a security measure in a public space [Patel et al. 2004].

3.3 Semaphoric Gestures

Quek et al. [Quek et al. 2002] define a semaphoric gesture as follows as

> Semaphores are systems of signaling using flags, lights or arms [Brittanica.com]. By extension, we define semaphoric gestures to be any gesturing system that employs a stylized dictionary of static or dynamic hand or arm gestures...Semaphoric approaches may be referred to as "communicative" in that gestures serve as a universe of symbols to be communicated to the machine.

Semaphoric gestures are also one of the most widely applied styles as seen in the literature even though the concept of using signs or signals to communicate information has bee a minuscule part of human interactions [Quek et al. 2002], and not natural, providing little functional utility [Wexelblat 1998]. However, with the move towards more ubiquitous computing paradigms, the use of semaphoric gestures is seen as a practical method of providing distance computing in smart rooms and intelligent environments [Bolt 1980; Baudel and Beaudouin-Lafon 1993; Cao and Balakrishnan 2003; Lenman et al. 2002b; Wilson and Shafer 2003; Streitz et al. 1999] and as a means of reducing distraction to a primary task when performing secondary task interactions [Karam and m. c. schraefel 2005]. There are several different forms of gestures that fall into the category of semaphores discussed in the literature which will be described next.

3.3.1 *Static vs dynamic gestures.* Semaphoric gestures can involve static poses or dynamic movements unlike manipulative gestures which are mainly dynamic. For example, when the thumb and forefinger are joined to represent the "ok" symbol, this is a static pose while moving the hand in a waving motion is a dynamic semaphoric gesture. These types of gestures can be performed using a hand [Alpern and Minardo 2003; Baudel and Beaudouin-Lafon 1993; Rekimoto 2002; Lee et al. 1998], fingers [Grossman et al. 2004; Rekimoto et al. 2003], arms [Nickel and Stiefelhagen 2003; Bolt 1980], the head [Schmandt et al. 2002; Davis and Vaks 2001], feet[Paradiso et al. 2000] or other objects such as passive or electronic devices such as a wand or a mouse [Wilson and Shafer 2003; Baudel and Beaudouin-Lafon 1993; Moyle and Cockburn 2003].

3.3.2 *Stroke gestures.* Semaphoric gestures can refer to strokes or marks made with a mouse or stylus which are mapped onto various interface commands. Examples include mouse strokes for back and forward control of a web browser [Moyle and Cockburn 2003], controlling avatars in virtual reality style applications using shorthand marks [Barrientos and Canny 2002], for interacting with and issuing commands for desktop style applications [Segen and Kumar 1998a; Chatty and Lecoanet 1996; Wu and Balakrishnan 2003; Ou et al. 2003; Allan Christian Long et al. 1999; Pastel and Skalsky 2004] or for screen navigation or marking or pie menu selections [Smith and m. c. schraefel 2004; Lenman et al. 2002b; Zhao and Balakrishnan 2004]. Strokes and other similar styles of gestures are also referred to as direct input techniques and include Graffiti for the palm, Jot for windows and other handwritten character sets used for mobile computers, pen computing, touch input or mouse input based strokes [Ward et al. 2000; Forsberg et al. 1998; Pirhonen et al. 2002; Rubine 1992; Cohen et al. 1997].

3.4 Gesticulation

The act of gesticulating is regarded in the literature as one of the most natural forms of gesturing and is commonly used in combination with conversational speech interfaces [Quek et al. 2002; Wexelblat 1994; Kopp et al. 2004; Bolt and Herranz 1992; Kettebekov 2004; Silva and Arriaga 2003; Eisenstein and Davis 2004; Krum et al. 2002]. Originally referred to as 'coverbal gestures' [Bolt and Herranz 1992; Kettebekov 2004], a term credited to Nespoulous and Lecours [Bolt and Herranz 1992], this gesture style has recently gained a great deal of attention in the literature and is currently viewed as one of the most challenging areas of gesture research: Gesticulations rely on the computational analysis of hand movements within the context of the user's speech topic and are not based on pre-recorded gesture mapping as with semaphores which is a problem addressed by the various research areas. Wexelblat refers to gesticulations as idiosyncratic, not taught, empty handed gestures [Wexelblat 1995] and the interfaces that use gesticulation as directive style interfaces. These are primarily multimodal, speech and gesture interfaces which attempt to create a naturalistic, conversational style interaction without the need for electronic or passive devices to detract from the natural gesticulations people naturally perform.

Unlike semaphores which are pre-recorded or trained in the system for recognition, or manipulations that track physical movements and positions, gesticulation is combined with speech and does not require the user to perform any poses or to learn any gestures other than those that naturally accompany everyday speech. Gesticulations have also been referred to as depictive or iconic gestures that are used to clarify a verbal description of a physical shape or form through the use of gestures that depict those shapes and forms for example [Koons and Sparrell 1994; Bolt and Herranz 1992; Kopp et al. 2004].

3.5 Language Gestures

Gestures used for sign languages are often considered independent of other gesture styles since they are linguistically based and are performed using a series of individual signs or gestures that combine to form grammatical structures for conversational style interfaces. In some instances such as finger spelling, sign languages can be considered semaphoric in nature. However the the gestures in sign languages are based on their linguistic components and although they are communicative in nature, they differ from gesticulation in that the gestures correspond to symbols stored in the recognition system.

3.5.1 *Communication interfaces.* Although it is feasible to use various hand signs as a means of communicating commands as with semaphoric gestures, this is not the intention of the language gestures. Rather, because sign languages are grammatical and lexically complete, they are often compared to speech in terms of the processing required for their recognition. In addition, the applications that are intended for use with sign language are communication based rather than command based and most of the literature is focused on the difficult task of interpreting the signs as a meaningful string [Bowden et al. 2003; Braffort 1996; Fang et al. 2003; Sagawa et al. 1997].

Although sign language recognition has been part of the literature since at least 1986 [Zimmerman et al. 1987], work in this area was originally based on recognition of static gestures as with finger spelling. It has been only recently that more complex algorithms and methods have been developed to improve recognition to interpret more complex words, concepts and sentence structures [Bowden et al. 2003; Braffort 1996].

3.6 Multiple Gesture Styles

Many of the systems reviewed for this paper do not focus on a single style of gesture interaction but rather employ a variety of gestures including combining deictic and manipulative [Rekimoto 1997; Fisher et al. 1987; Sharma et al. 1996], semaphores and manipulations [Joseph J. LaViola et al. 2001; Weimer and Ganapathy 1989; Sturman et al. 1989; Nishino et al. 1997; Ou et al. 2003; Grossman et al. 2004] deictic semaphoric gestures [Baudel and Beaudouin-Lafon 1993; Konrad et al. 2003; Wilson and Shafer 2003] and semaphoric, deictic and manipulative [Cao and Balakrishnan 2003; Osawa et al. 2000; Rekimoto 2002; Wu and Balakrishnan 2003], . In such systems, for example in Buchmann's FingARTips [Buchmann et al. 2004], fingers are used in an augmented reality environment to point at and manipulate objects in the AR display as well as to issue commands through semaphores. It is difficult to consider the different gesture styles in isolation for such interfaces, typically in VR and AR domains where the hand is immersed in the environment and is used

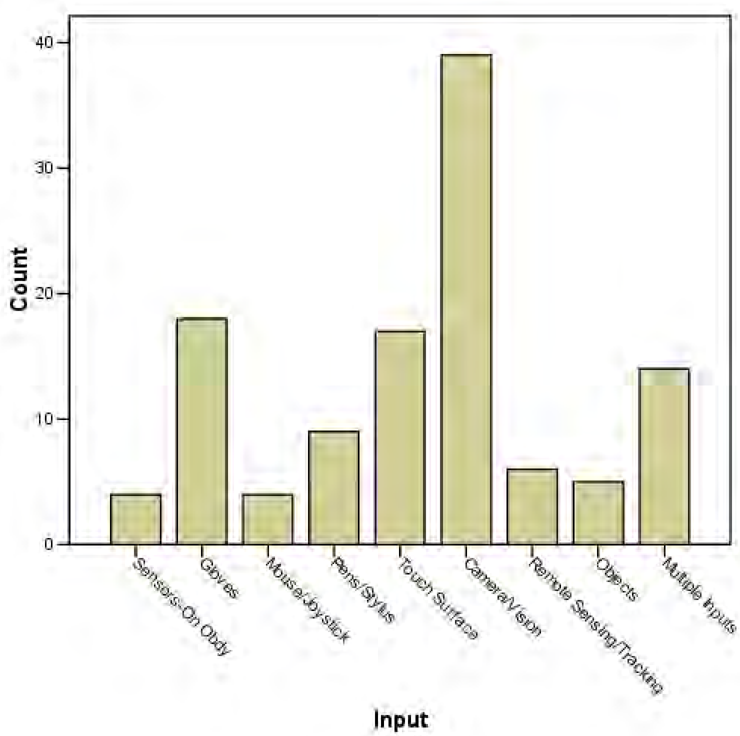

Fig. 3. The diagram represents the distribution of the input technology used for enabling the gestures from the literature.

to perform all actions including manipulations and system controls.

4. CATEGORIZATION BY GESTURE ENABLING TECHNOLOGY

In this section, we present an overview of gesture based interactions that are enabled through various input technologies. For the purpose of this work, we focus on the technology in terms of its ability to enable the gestures rather than on the specific details. Although this is not an exhaustive look at the possible technologies used for gesture based interactions, we do feel that it does provide a practical means by which to categorize the types of interactions enabled by the different styles if input that have been reviewed for this paper. We begin with a high level breakdown into perceptual and non-perceptual input and a further breakdown into the individual technologies that fall into these categories. Although one could select different hierarchical structures to present the classification of input devices, we chose this division to reflect the literature and the nature of the interactions from the perspective of the user and the different application domains that consider gestures as an input mode. The following is an overview of the input technologies discussed in this paper:

4.1 Non-Perceptual Input

We begin this discussion with non-perceptual input as a chronological approach to the technology used for gesture input over the past 40 years. Non-perceptual input involves the use of devices or objects that are used to input the gesture, and that requires physical contact to transmit location, spatial or temporal information to the computer processor.

4.1.1 *Mouse and Pen Input.* One of the first examples of a gesture based interaction system was seen as early as 1963 with Sutherland's SketchPad [Myers 1998; Sun-Microsystems ; Sutherland 2003] which used a light pen, a predecessor to the mouse, to indicate the location of a screen object. Gestures using the mouse provides an alternative to the direct manipulation or point and click method of interacting with a computer, allowing gestures or strokes of the mouse to be translated into direct commands. This type of interaction can be performed using pens [Chatty and Lecoanet 1996] or a mouse and provides a simpler and faster way to input commands to the interface, and is one of the oldest and most commonly used forms of gesture reviewed in the literature [Cohen et al. 1997; Forsberg et al. 1998; Rubine 1992; Buxton et al. 1983; Moyle and Cockburn 2003; Barrientos and Canny 2002].

4.1.2 *Touch and Pressure input.* Touch and pressure sensitive screens are another gesture enabling input technology seen in the literature since the mid 80's [Buxton et al. 1985; Pastel and Skalsky 2004]. Touch based input is similar to gesturing with direct input device however one of its key benifits is to enable a more natural style of interaction that does not require an intermediate devices like the mouse [Allan Christian Long et al. 1999; Gutwin and Penner 2002; Zeleznik and Forsberg 1999; Forsberg et al. 1998; Wolf and Rhyne 1993]. Touch sensors in particular have been widely discussed in the literature, primarily as input for gestures to mobile computing [Pirhonen et al. 2002; Brewster et al. 2003; Schmandt et al. 2002] and for tablet computer devices [Jin et al. 2004; Ou et al. 2003; Wolf and Rhyne 1993]. More recently, touch and presure sensitive material is used to enable table top computing surfaces for gesture style interactions [Rekimoto 2002; Wu and Balakrishnan 2003; Rekimoto et al. 2003; Schiphorst et al. 2002]. Touch sensors enable input for a wide variety of computer devices, ranging from desktop monitors [Minsky 1984] to small mobile screens [Brewster et al. 2003] to large interactive surfaces [Smith and m. c. schraefel 2004]. The nature of the gesture interactions they enable are similar in terms of the contact style gestures that can be performed on a surface.

4.1.3 *Electronic Sensing - Wearable or Body Mounted.* In the work presented by Bolt [Bolt 1980], electronic sensors were one of the first methods for recognizing hand and arm movements in gesture based interfaces. We will not discuss the technical details of these sensing devices, but refer to them in terms of the interactions with gestures that they enable. These interactions are characterized by the nature of the sensors in their ability to track space, position and orientation through magneto-electro sensors. These sensors, also known as polhemus sensors were and still are one of the primary devices used to directly sense body, arm or finger movements [Bolt 1980; Roy et al. 1994; Osawa et al. 2000; Joseph J. LaViola et al. 2001; Wexelblat 1995].

In 1980, Bolt describes work using position sensors or Space Sensing Cubes that are attached to the users wrists and used to track the x,y,z coordinates of their arm as it moves through space, enabling the pointing gesture to the screen in front of the room to be tracked and used as input to the system. More recently, attached sensors have been used for adaptive technology [Roy et al. 1994] and for navigating through virtual environments [Osawa et al. 2000]. They are also used in wearable devices and head tracking sets [Amento et al. 2002; Song et al. 2000; Brewster et al. 2003] however, they are somewhat inconvenient for widespread use as they are expensive and somewhat cumbersome for everyday interactions. Although there are numerous types of wireless sensors that can be attached to the body to track audio or visual data [Amento et al. 2002; Gandy et al. 2000], they remain attached to the user.

4.1.4 *Electronic Sensing: Gloves.* The next chronological stage in applying input technology to enabling gestures was the use of sensor-enabled gloves. Although there are several manufacturers of these devices, the Z-Glove and DataGlove, as discussed by Zimmerman in 1986 [Zimmerman et al. 1987] were some of the earliest referenced in the literature. The use of these gloves enabled gestures that were more detailed, involving the movement of individual fingers, wrist and hands, to allow a more flexible and accurate gesture recognition than with the polhemus type sensors. The Z-Glove consisted of a cotton glove that was fitted with sensors to measure finger bending, positioning and orientation and included a vibrating mechanism for tactile feedback. Zimmerman's system was used to demonstrate possible interactions with virtual reality applications and to manipulate computer-generated objects on a desktop computer, interpret finger-spelling, evaluate hand impairment and to interface with a visual programming language.

These interactions were usually speech accompanied, laying the groundwork for future work in multimodal speech and gesture interfaces. The gloves also enabled interactions with virtual environments and graphical objects through head mounted displays for space research [Fisher et al. 1987] and on screen and virtual object manipulations [Sturman et al. 1989; Weimer and Ganapathy 1989]. However, it wasn't until the 1990's that gesture research involving gloves gained significant attention in the literature and was more prominently used for applications such as immersed VR and autonomous agent control interfaces [Osawa et al. 2000; Song et al. 2000; Maes et al. 1997; Pierce and Pausch 2002], telematics robotics [Fisher et al. 1987; Goza et al. 2004; Silva and Arriaga 2003] and 3d graphic manipulations and navigation [Zimmerman et al. 1987]. For a more detailed description of the characteristics of gloves and their uses, we refer to Sturman's survey on glove based interactions [Sturman and Zeltzer 1994].

4.1.5 *Electronic Sensing: Sensor-Embedded Objects and tangible interfaces.* Gestures are also enabled using physical objects embedded with sensors that are manipulated by a user [Fitzmaurice et al. 1995; Hinckley et al. 1998]. This area of research is referred to as tangible or graspable interfaces [Fitzmaurice et al. 1995] so that the manipulation of these objects is called gesturing in the literature [Paiva et al. 2002; Sinclair and Martinez 2002; Patel et al. 2004]. Tangible computing gestures involve movements or manipulations of objects that translate into deictic, manipulative [Hinckley et al. 1998] and semaphoric gestures for interacting with the computer [Wilson and Shafer 2003].

4.1.6 *Electronic Sensing: Tracking Devices.* Gesture based interactions are also performed using infrared tracking devices to detect input. The WorldBeat system in 1997 demonstrated how infrared batons could be use to transmit gestures to a tracking device to control a midi player [Borchers 1997]. The infrared beam is tracked by a camera and its movements or gestures are translated into predetermined system behaviours. Further research into this approach of gesture input has seen the use of infrared tracking as a pointing device to identify computers or appliances as a means of transferring data or for remote control of devices in a smart environment [Wilson and Shafer 2003; Swindells et al. 2002]. This type of setup requires the use of infrared transmitting devices and receivers or trackers and is similar to remote controle type interactions however infra red transmitters are also being used with cameras in another form of gesture input which will be discussed in the computer vision section.

4.1.7 *Audio Input.* An alternative method of sensing gestures for large screen interactions relies on audio sensors [Paradiso 2003]. The use of audio as a means to detect the location of a knock or tap enables a selection gesture to be used for interacting with a large public display. This form of sensing is quite limited in the type of gesture detection possible however it does take advantage of the computer's capacity for audio perception and can be viewed as an alternative means of input for pointing and selection gestures. Another use for audio sensors is in a device which registers audio caused by finger and hand movements when attached to the wrist of its user [Amento et al. 2002]. Although audio is inherently a perceptual input technology, in these interactions the user is required to make physical contact with the device in order to perform the gesture which is why we place it in the non-perceptual category.

4.2 Perceptual Input

In this paper, we refer to perceptual as that which enables gestures to be recognized without requiring any physical contact with an input device or with any physical objects, allowing the user to communicate gestures without having to wear, hold or make physical contact with an intermediate device such as a glove or mouse for example. Perceptual input technology includes visual, audio or motion sensors that are capable of receiving sensory input data from the user through their actions, speech or physical location within their environment.

4.2.1 *Computer Vision.* Computer vision recognition was a major technological factor in influencing the type of gesture interactions that are explored in gesture based interactions. One of the first examples involve using video to recognize hand movements as an interaction mode as seen in Krueger et al's work from 1985 on VideoPlace [Krueger et al. 1985]. Krueger's system involved projecting a video image of the user overlaid on a projected wall display. The interaction is based on the user's image coming in contact with or pointing at objects on the display. This technique of superimposing the users's image on top of the display has been recently used in the FaceSpace system [Stotts et al. 2004], where the user is able to gesture on a desktop screen, receiving feedback through a display of their image over top of the monitor display. Computer vision is used for all forms of gesturing however poses several research problems in terms of its ability to recognize gestures while resisting changes in lighting and tracking objects for example. One work around for some of the sensitivities of computer vision has been to use LED transmitters in combination with cameras. This method allows various gestures to be performed by tracking the interrupted LED output caused by the user performing gestures [Gandy et al. 2000; Krum et al. 2002] however it does restrict the type of gestures that can be used for interactions.

4.2.2 *Remote Sensors.* In 1995, Zeleznik presented a variety of remote sensing devices that enabled the transmission of electric fields, shunted to a ground through a human body and an external electric field transmitted through the body to stationary receivers as gesture enabling technology. This form of sensing is used to detect human presence, movement and pressure, enabling full body movements for gesture based interactions. Electronic sensors have also been placed on screens for remote sensing of finger movements [Allport et al. 1995] as an alternative to mouse interactions. The iLand work mentions sensing devices and the possibility of using gesture for interactions however, the gesture interactions are not the primary focus of this work and this system presents a very broad range of sensing devices that can be used to interact within a smart room using gestures [Streitz et al. 1999].

5. CATEGORIZATION BY APPLICATION DOMAIN

We now present our classification of gesture based research focusing on the application domains they are applied to. There are many cases in which a single system is implemented for multiple application domains. For the purpose of this taxonomy, we present a view of the scope of application domains for which gestures have been considered or implemented as an interaction mode rather than focusing on the details of any specific systems. The categorization of the application domains presented in this paper is based on the literature itself, in terms of the conferences and journals the work is published and on the nature or function of the application for which the gestures were applied.

5.1 Virtual and Augmented Reality

The application domain of virtual and augmented reality represents one of the largest areas for gesture based interactions. Much of the interactions within virtual reality applications involve either semi or fully immersed displays although the physical nature of the gestures involved are relatively the same. Virtual reality applications span many domains including CSCW style interactions with the virtual environment [Sharma et al. 1996] and desktop style interactions with 3D worlds [Gandy et al. 2000]. Generally, in the domain of virtual and augmented reality, there are for main categories of gesture based interactions that we distinguish for our taxonomy that reflect the literature reviewed. Immersive interactions where the users' body is represented in the display as an avatar, navigations through virtual worlds, manipulating objects in semi-immersed virtual worlds and fully-immersed interactions to control robots or vehicles in teleoperations and telerobotic

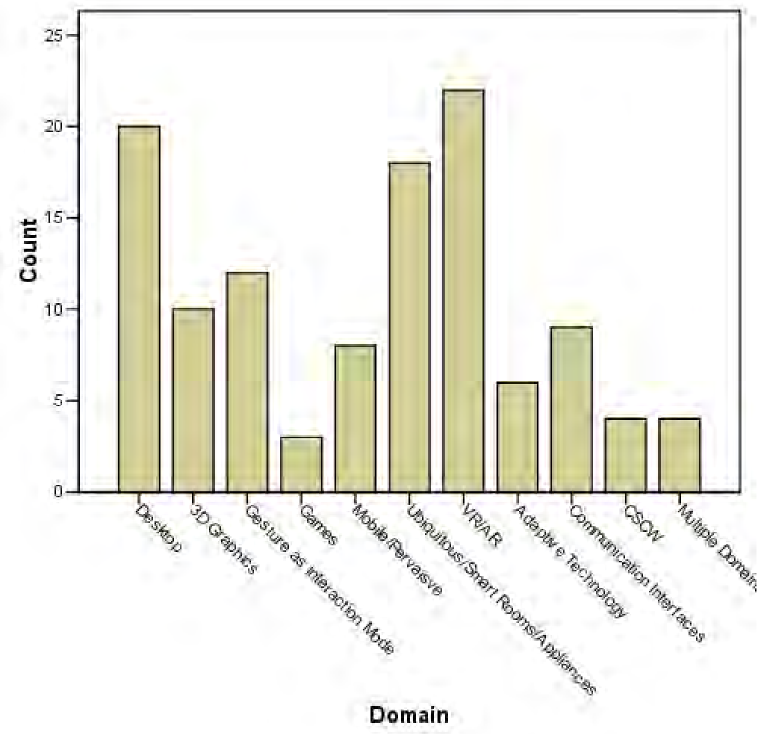

Fig. 4. The different domains for which gestures have been applied as seen in the review of the literature are shown in the bar graph.

application.

5.1.1 *Immersive Interactions and Avatars.* An early example mentioned earlier involved Krueger's Videoplace [Krueger et al. 1985]. Although this is not explicitly referred to as avatar based work, it could be considered as the precursor to the immersed avatar based interactions. Later work in controlling avatar expressions and movements involved pen gestures, where letters are drawn on a tablet input device. Each of the gestures or strokes represent different facial or body expressions that can be attributed to the avatar [Barrientos and Canny 2002]. Another gesture based interaction style for avatars in virtual worlds consist of full body movements as a means of modelling and controlling avatar movements and interactions [Thalmann 2000; Maes et al. 1997; Lee et al. 1998]. Sensors that are fitted on the user's body are used to track their body form and movements to create a virtual human on a screen. Essentially, avatars are virtual objects that are manipulated in terms of their behaviour and their movements within the virtual world.

5.1.2 *Manipulation and Navigation in Virtual Worlds.* When physical interactions are required to interact with the virtual surroundings and its objects, manipulative gestures are typically used [Fisher et al. 1987; Song et al. 2000; Pierce and

Pausch 2002; Nishino et al. 1998]. Unlike desktop style CAD or graphic displays where objects are being manipulated using abstract gestures such mapping 2D gestures to 3D graphic objects, a virtual reality interaction can involve 3D gestures based in hand or body movements. The detection of body motion and location with respect to an area in a room for example, has been implemented using sensors embedded in the floor and in the users shoes, allowing movements in the physical world to be mapped directly onto the virtual world [Joseph J. LaViola et al. 2001]. The user's location within the room is mapped to the virtual world, allowing their movements in real space to enable navigation in the virtual environment.

3D visualizations that are based on navigating around objects within in a virtual world also employ hand gestures as an interaction mode [Segen and Kumar 1998a; Nishino et al. 1998]. Gestures with the hand can enable the arranging of objects while navigating around a 3d information space such as a graph using a stereoscopic display [Osawa et al. 2000]. However object manipulations in virtual worlds through the use of gestures is a more widely discussed area in the literature.

Physical manipulation of virtual objects often involve sensor augmented gloves. During the interaction, the movements of glove are recreated within the world as a 3D object allowing a visualization of the interaction within the digital world. In 1986, Zimmerman el al. [Zimmerman et al. 1987] presented work on the development of a system that allows such interactions with graphic objects. The graphical objects are manipulated as through natural hand movements that are reproduced in the virtual world.

In 1989, Sturman et al. extend glove based interactions by investigating whole hand interactions for virtual worlds [Sturman et al. 1989]. They identify three different types of interactions for the whole hand; direct manipulations involving the user 'reaching' into the simulation to manipulate objects, abstracted graphical input such as pushing buttons or locating objects and movements referred to as a stream of tokens such as used in sign language. Gestures as an interaction technique for this type of manipulative interaction continued to grow as a research topic throughout the 90's [Weimer and Ganapathy 1989; Bolt and Herranz 1992; Koons and Sparrell 1994; Wexelblat 1995; Nishino et al. 1997] and currently remains one of primary modes for interacting with virtual objects and interactions within virutal worlds [Pierce and Pausch 2002].

5.1.2.1 *Augmented reality.* Augmented realty applications involve gesturing with tangible objects to interact with the display. Physical objects such as blocks can use a shaking gesture as a way for the user to add labels to the display [Sinclair and Martinez 2002]. Bare hands are also used in such interactions, for example as a means of selection and moving of objects in the display [Buchmann et al. 2004]. The type of interactions available in augmented reality are very similar to those in virtual reality when considering gestures and will be further discussed in the section on system response.

5.1.3 *Robotics and telepresence.* Telepresence and telerobotic applications are typically carried out for space exploration and military based research projects. The gestures used to interact with and control robots are most commonly seen as virtual reality applications as the operator is controlling a robot's actions while viewing the robot's environment through a head mounted display [Goza et al. 2004]. Gestures are used to control the robot's hand and arm movements for reaching and manipulating objects as well as the direction and speed that they are travelling.

5.2 Desktop and Tablet PC Applications

In desktop computing applications, gestures are an alternative to the mouse and keyboard interactions, enabling more natural interactions with fingers for exmaple [Iannizzotto et al. 2001]. Many gesture based desktop computing tasks involve manipulating graphics objects [Bolt and Herranz 1992; Buxton et al. 1983], annotating and editing documents [Cohen et al. 1997] , scrolling through documents [Smith and m. c. schraefel 2004], making menu selections [Lenman et al. 2002b] and navigating back and forwards web browsers [Moyle and Cockburn 2003]. In fact, if we include gestures that have been associated with direct manipulation or mouse style interactions for desktop computers, everything including inputting text [Ward et al. 2000] could be accomplished through the use of gestures when we consider the literature [Rubine 1992; Kjeldsen and Kender 1996; Cohen et al. 1997; Dannenberg and Amon 1989; Henry et al. 1990]. More novel style interactions that utilize natural gestures such as nodding have been explored as a means of interacting with responsive dialog boxes on a desktop computer [Davis and Vaks 2001].

Gesture based interfaces have also been explored for tablet computers within specialized domains such as air traffic controls [Chatty and Lecoanet 1996], adaptive technology [Ward et al. 2000], collaborative group work [Stotts et al. 2004] and musical score editing [Forsberg et al. 1998].

5.2.1 *Graphics and Drawing Applications.* One of the first application domains for which gestures were developed involved graphic style interfaces and was presented as early as 1964 [Sutherland 2003; Teitelman 1964; Coleman 1969]. The pen and tablet style interfaces of the 80's were initially based on tablets or desktop computer screens and typically used a mouse, stylus or puck for the non-touch screens. The gestures used for this style of interaction involve manipulating the input device to draw strokes, lines, circles or to make movements in different directions for drawing, controlling the functionality provided by the application, switching modes and issuing commands [Buxton et al. 1983; Rhyne 1987]. Similar applications were also implemented on touch screens or tablets using fingers or pens for gesturing [Buxton et al. 1985].

Pressure sensors are also used as a novel means of gesturing to manipulate and create graphical objects in a finger painting application [Minsky 1984]. Touch sensors were used to track the movement of the finger or input device on the screen to create the drawings, and the amount of pressure applied to the screen was used to determine the thickness of the lines being drawn for example.

Graphic applications also use gestures for navigational and view purposes as well as to control the various functions associated with the application. 2D mouse based gestures have been used to control the view provided by a cameral in a 3D graphics interface using a mouse or a pen for input [Zeleznik and Forsberg 1999]. The gestures are stroke and direction based and have been shown to be effective in controlling the virtual camera's more than 6 degrees of freedom. Graphic interfaces that deal in 3 dimensional graphic applications are a more common domain for gesture-based interactions however this area tends to lean more towards virtual reality applications.

5.2.2 *CSCW.* Gestures within the domain of CSCW systems are used within a variety of computing domains including desktop computers or table top displays [Wu and Balakrishnan 2003; Rekimoto 2002] and large screen displays [Cao and Balakrishnan 2003; von Hardenberg and Berard 2001]. Interactions such as sharing notes and annotations amoung groups of people are most common in the literature using stroke style gestures both within a single location or across a network for remote teams is explored in the literature [Wolf and Rhyne 1993; Gutwin and Penner 2002; Stotts et al. 2004]. Annotations can be transmitted on live video streams for collaborations between an instructor and a student working on tasks involving physical objects [Kuzuoka et al. 1994; Ou et al. 2003]. Images from one party are transmitted to the second party, who in turn gestures to create annotations overlaid on the image which is in turn transmitted back to the original party.

5.2.3 *Multimodal Interfaces.* There is also a large part of the literature that is focused around gestures for multimodal interactions typically through the use of gesture in combination with speech interfaces [Hauptmann 1989; Cohen et al. 1997; Gandy et al. 2000; Schapira and Sharma 2001] although the term multimodal has seen two major uses, the first specifically for speech and gesture interactions [Cohen et al. 1997] and generally for multiple input devices used in combination [Nickel and Stiefelhagen 2003]

5.3 Three Dimensional Displays

Gestures in current literature is regarded as a natural means of interacting in three-dimensions such as with a volumetric display [Grossman et al. 2004] and a 3D projection system for microbiologists [Sharma et al. 1996]. Although these displays are currently cutting edge technology and are not readily available for extensive use with gesture interactions. Krum et al. [Krum et al. 2002] use directional gestures tracked by the gesture pendant, a camera worn on a chain around one's neck, to track finger movements in different directions to navigate around a 3D visualization of the earth.

5.4 Ubiquitous Computing and Smart Environments

Gesture based interactions designed to provide users with controls over devices and displays at a distance or within smart rooms could be considered one of the earliest forms of gesture based research involving hand or arm based gestures [Bolt 1980; Krueger et al. 1985]. However, it was not until several years after Weiser described his vision of ubiquitous computing in 1991 [Weiser 1995] that gestures gain popularity as an interaction mode in this domain. By 1994, gestures were being explored for the remote control of televisions [Freeman and Weissman 1995], for sensing human presence and position within a room to enable non-contact interactions with computer devices [Zimmerman et al. 1987], to control and interact with music interfaces at a distance [Borchers 1997] and to transfer digital data between differen devices within a smart environment [Rekimoto 1997; Swindells et al. 2002].

As smart room technologies became more sophisticated, so did the use of gestures as a more natural mode for interacting with computers that were distributed throughout ones environment [Streitz et al. 1999] and as a key mode of interaction with perceptual interfaces [Crowley et al. 2000]. There are a variety of gesture styles to control or identify a variety of devices such as lights, entertainment units or appliances within sensing-enabled environments in the literature [Wilson and Shafer 2003; Fails and Jr. 2002; Lenman et al. 2002b; Nickel and Stiefelhagen 2003].

Gestures for smart room or perceptual style interactions are also used for interacting with large screen displays that can be distributed around a room or building [Paradiso et al. 2000; von Hardenberg and Berard 2001].

5.4.1 *Tangible Computing.* The idea of physically manipulating or moving tangible objects is another form of gesturing referred to in the literature. One of the first systems to introduce this form of interaction with graspable user interfaces objects was Fitzmaurice, Ishii and Buxton's Bricks [Fitzmaurice et al. 1995]. Physical objects called bricks could be moved and rotated as a method of controlling their corresponding digital computer objects. However, as early as 1991, tangible interactions with physical objects to control digital input were enabled through the DigitalDesk [Wellner 1991]. A pointing gesture was used to indicate numbers on a physical sheet of paper for input to a computer using computer vision.

In an alternative interaction mode, actual computer devices can be manipulated using gestures that involve bumping, tilting, squeezing or holding the device for interactions such as navigation within a book or document and for detecting handedness in users of such devices [Harrison et al. 1998] or for transferring data and communication between table pc's for example [Hinckley 2003].

Other forms of tangible interactions involve a shaking gesture to sprinkle labels on augmented reality display objects [Sinclair and Martinez 2002], manipulating sensor embedded dolls to communicate gestures of emotion for 3D virtual games [Paiva et al. 2002] or for interacting with smart-fabric lined objects to cause reactions in the digital creatures in artificial life applications [Schiphorst et al. 2002]. Tangible interfaces have also been used for controlling 3D visual displays for neurosurgeons using two handed interactions involving a passive doll's head and various tools that can be used to manipulate the on screen view [Hinckley et al. 1998].

5.5 Games

Finally, we look at gestures as used as an input mode for interacting with computer games. Freeman began exploring gestures that could use a player's hand or body positions to control movement and orientation of interactive game objects such as cars [Freeman et al. 1996] or for tracking spatial for navigating around a game environment [Segen and Kumar 1998a].

The control of avatar navigational tasks using pointing gestures is also used as game input [Konrad et al. 2003] and movements in virtual reality game environments [Paiva et al. 2002]. More recently, Apple has released the iToy, a game that uses a web camera to enable gesture interactions with the game and the Sony Eyetoy.

5.6 Pervasive and Mobile Interfaces

In pervasive computing and mobile computing devices, gestures have been used to allow eyes-free style interactions, enabling the user to focus more attention on their mobility rather than on having to visually interact with the devices directly [Schmandt et al. 2002; Lumsden and Brewster 2003; Brewster et al. 2003; Pastel and Skalsky 2004]. This style of gesture typically involves PDA's with touch sensitive screens that accept finger gestures or strokes as input, providing audio output to the user and enabling eyes-free interactions with mobile devices.

5.6.1 *Wearable computing.* Wearable devices that allow the user to interact within a smart room environment can create a less restrictive means of enabling gesture interactions with devices even though such devices must be carried or worn by the user. By having the recognition device worn by the user, more flexible style interactions are enabled within the space, increasing the distances with which the user is able to interact with each of the devices in the environment [Gandy et al. 2000; Amento et al. 2002] however requiring the user to wear the device.

5.7 Telematics

Although the use of computer technology in the automotive industry is a growing field that can currently be found in all new automobiles, the use of gestures for telematics applications has not received a great deal of attention in the literature. Alpern et al. explored the use of gestures within an automotive context to control various secondary tasks in the car while maintaining focus on the primary task of driving [Alpern and Minardo 2003] and in a wide range of other distraction-minimizing tasks within a car [Pickering 2005].

5.8 Adaptive Technology

Adaptive interfaces that use gesture based interactions have not been one of the primary domains investigated in the literature. Much of the research on gestures for adaptive interfaces extends existing systems to the adaptive application as seen with Zimmerman's use of gloves to measure hand impairments [Zimmerman et al. 1987] and Segen [Segen and Kumar 1998b] who presents a system for wheelchair navigation using gestures. Also, the gesture pendant has been used for adaptive applications in which gestures are performed in front of a camera that is worn around the neck like a pendant for home emergency services and to provide control of devices for home patients with vision or physical impairments [Gandy et al. 2000].

There are also systems that focus on investigating gestures specifically for adaptive technology. Keates presented work that replaced the keyboard and mouse with gestures for users with motion impairments [Keates and Robinson 1998]. Ward's Dasher system also replaces the keyboard with gestures in a novel approach to selecting and building text input based on pointing gestures used to build words [Ward et al. 2000]. Randy Pausch developed the Tailor system that uses gestures to assist with users with speech impairments [Pausch and Williams 1990]. Additional work on gestures for adaptive interfaces explores the use of face and hand or arm gestures to control the mouse movements and clicks for desktop applications [Reilly 1998]

Sign language interfaces are also considered adaptive technology for users with hearing impairment however the current systems are not considered robust enough in both recognition and processing capabilities to be used in everyday computing interactions. This form of communication interface is still in the research stage, working out problems on dealing with large vocabulary sets and limitations in image processing for example [Fang et al. 2003; Bowden et al. 2003].

5.9 Communication Interfaces

Communication interfaces are those that seek to enable a more human-human style of human-computer interactions [Wexelblat 1994; Quek et al. 2002]. Besides language gestures that are based on sign languages, gestures for communication interfaces have been considered one of the most challenging for gesture research [Eisenstein and Davis 2004]. A large body of research in communication interfaces was conducted long before the 1990's, investigating multimodal interactions that employ both speech and gestures to create more natural ways to control, identify and interact with graphic objects on desktop or virtual reality displays [Bolt 1980; Kobsa et al. 1986; Hauptmann 1989; Weimer and Ganapathy 1989]. Throughout the 90's, additional work on communicating interactions with graphical objects continued [Bolt and Herranz 1992; Koons and Sparrell 1994; Wexelblat 1995; Quek 1994] extending into the domain of speech and gesture for interacting with more general desktop style applications [Cohen et al. 1997].

However, a more generic approach to speech and gesture began to be explored as a means of enabling natural communication between humans and computers that is not as application specific as it is based on understanding how the two modes can be used to assist in better interpreting meaning [Kettebekov 2004; Robbe 1998; Quek et al. 2002; Schapira and Sharma 2001].

5.10 Gesture Toolkits

Several gesture toolkits have also been presented that are designed to support research in gesture based interactions that can enable stroke style gestures for touch screens [Dannenberg and Amon 1989; Henry et al. 1990; Rubine 1991], gestures for touch sensitive fabrics [Schiphorst et al. 2002] or for gestures through multiple inputs [Westeyn et al. 2003]. With these tools, a variety of system responses can be invoked using gestures, allowing researchers to investigate a variety of interactions within different domains. There are many toolkits available currently however much of the research in gestures is specific and few if any have been found that build on existing tookits, but instead, implement their own systems.

6. CATEGORIZATION BY SYSTEM RESPONSE

One of the key features that distinguishes the different interactions possible using gestures is the system response or output that a gesture interaction leads to. In our taxonomy, the mode of output is the end result of a gesture interaction and is a major part of the user experience. Based on our review of the existing literature, we have separated the different system responses into the following categories: Audio, visual (2D and 3D) and CPU command responses.

We now present some of the major system responses discussed in the literature and a brief overview of the different types of interactions they enable.

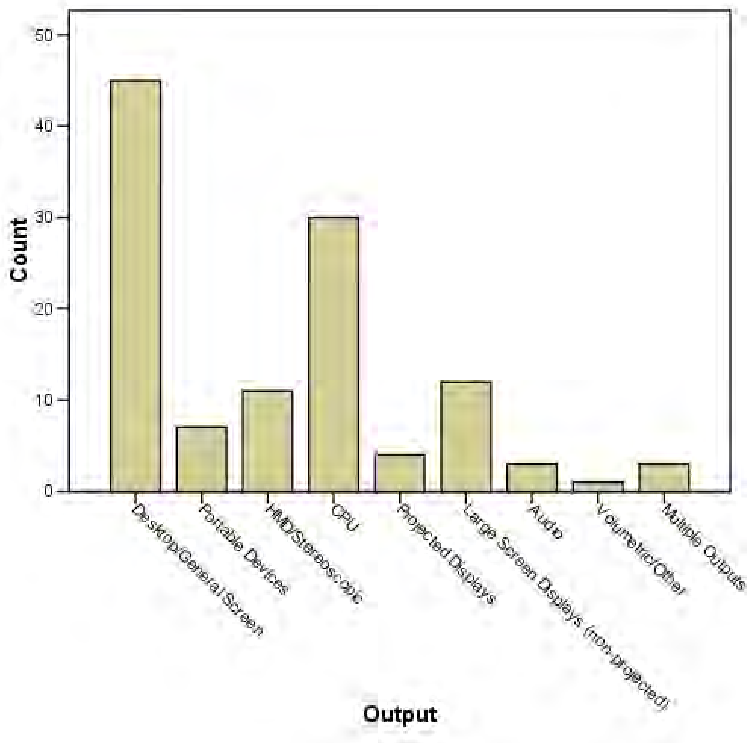

Fig. 5. The bar graphs shows the distribution of system responses through output technology used in the gesture research reviewed in this paper.

6.1 Visual Output

Most of the responses that are explored using gestures is based in visual displays. The nature of the gesture that is appropriate for the interaction varies with the type of response that is required. For example, interactions involving manipulations of on screen objects can be viewed on a screen in 2D or as 3D using a head mounted display. In addition, some gesture interactions can simply invoke a behaviour that does not require visual feedback but leads to audio or CPU commands. We will not provide an exhaustive look at gesture based system response, but will rather provide a brief summary of the different output technologies and responses that are reviewed in the literature.

6.1.1 *2-Dimensional Visual Output.* Visual output in 2 dimensions involve desktop screens, large screens, projected displays and portable or mobile devices including phones and PDAs. The decision to include a particular display when designing gesture based interactions relies on many factors, however is not normally based on the type of gesture interaction the designer is trying to create. There is the exception of creating eyes free interactions, in which case a display may not be required if audio is the preferred output [Schmandt et al. 2002; Pirhonen et al. 2002] while the input device may include an output display by default as with the PDA. In cases where the gesture simply invokes a CPU based command on some other device or computer, displays may not even be necessary citevisionWand.

6.1.2 *3-Dimensional Visual Output.* When interacting with 3D displays, including head mounted displays or stereoscopic displays, for example, there are both 2D and 3D style gestures used for interactions. However, when considering interactions with 3D displays, we turn to virtual reality applications. Within virtual reality applications, the display can be either immersive, semi-immersive or non-immersive. Each of the different types of displays has its own unique properties that enable specific gesture interactions. For example, a fully immersive head mounted display typically uses 3D gestures through glove input [Nishino et al. 1997; Song et al. 2000] as do other forms of 3D outputs including 3D projected displays [Sharma et al. 1996]. 3D graphic objects can also be presented on a 2D display or output device such as a with large screen or projected image [Wexelblat 1994; Nishino et al. 1997] or a desktop monitor display [Barrientos and Canny 2002].

Semi-immersive outputs are those that use stereoscopic glasses to enhance a screen output of 3D images [Osawa et al. 2000; Nishino et al. 1998] and non-immersive outputs are those that are presented on a 2D display [Maes et al. 1997].

6.2 Audio Output

Research into the use of audio as a output mechanism for gesture based interaction claims that this can type of output can free up primary attention for more critical tasks such as waking, driving, or other activities that are not conducive to causing divided attention [Schmandt et al. 2002; Pirhonen et al. 2002; Brewster et al. 2003]. Much of the literature that addresses the use of audio output in gesture interfaces is based in mobile and pervasive style computing. Although the use of non-speech audio output has been explored in adaptive technologies as a method of expressing graphical and spatial information to visually impaired users [Alty and Rigas 1998] the primary use for gesture based interactions remains in the pervasive and mobile computing domain.

6.3 CPU or Command Directed Output

In many of the systems reviewed for our taxonomy, the system response of a recognized gesture often is not translated onto any specific output device but rather is used as a control input command to other devices or applications. We refer to this output as CPU, where the gesture is stored for different uses. Essentially any gesture that is recognized by a system can be mapped onto a variety of output devices in a variety of formats. For example, pointing gestures that identify devices in a smart room are used to redirect the focus of the system to the specific devices indicated [Wilson and Shafer 2003] enabling the gestures to control a variety functions for the targeted device. In addition, a number of systems implemented or proposed are designed to handle a variety of gestures from different input technologies to be mapped onto different output behaviours as determined by the system [Pausch and Williams 1990; Roy et al. 1994; Keates and Robinson 1998; Reilly 1998], for communication interfaces [Quek 1994; Kettebekov 2004] or as control interfaces for interacting with smart room or ubiquitous style environments [Gandy et al. 2000; Fails and Jr. 2002; Wilson and Shafer 2003; Nickel and Stiefelhagen 2003].

7. EVALUATING GESTURE INTERACTIONS

In this section, we present a discussion on evaluations that have been conducted on individual systems or on gestures in general. Although this is not a included in the taxonomy presented in this paper, it is part of our analysis on gestures as an interaction mode as reviewed in the literature. As shown in the pie graph below, almost half of the systems, interactions, applications and devices that have been reviewed do not include any form of evaluation while about the same amount do. We will refer to this graph later in the section.

7.1 System evaluation styles and motives

Many of the systems reviewed for this paper can be regarded as point designs or examples of systems that either implement or propose novel interaction techniques and uses for gestures [Hinckley et al. 1998]. Given the nature of these point designs, it is difficult if not impossible possible to extensively evaluate such systems for usability or accuracy if they are not extended beyond the point design stage. Examples of such systems that implement gesture interactions for novel applications include a security system mobile computing [Patel et al. 2004], transferring files between computers within a networked environment using infrared pointers [Swindells et al. 2002], pick and drop gestures [Rekimoto 1997] or bumping devices together [Hinckley 2003].

Gestures have also been used in point designs that attempt to create novel or improved interactions for appliance control and home entertainment systems [Freeman and Weissman 1995; Lenman et al. 2002a],

Other point designs include the exploration of novel methods of gesture recognition such as Minsky's pressure sensing monitor [Minsky 1984], gloves [Zimmerman et al. 1987], computer vision technologies [Krueger et al. 1985] or wearable devices for detecting gestures through audio input [Amento et al. 2002].

Systems that have been evaluated typically focus on accuracy of the recognition system [Gandy et al. 2000; Wilson and Shafer 2003; Gandy et al. 2000; Zimmerman et al. 1995], the efficiency of the system in handling large sets of gestures [Sagawa et al. 1997; Kettebekov 2004; Fang et al. 2003] or the general usability of a system [Moyle and Cockburn 2003; Ou et al. 2003].

Only a few of the papers reviewed either provided an extensive evaluation of several aspects of their systems including accuracy and usability [Schapira and Sharma 2001; Gutwin and Penner 2002; Zeleznik and Forsberg 1999]. Other papers reviewed presented studies or experiments with the intent of providing empirical results and information about specific aspects of gestures such as evaluating a new method of disambiguating the division of individual gestures [Kettebekov 2004], techniques for gesturing in or designing gestures for different domains [Wolf and Rhyne 1993; Sturman and Zeltzer 1993], specific tasks [Karam and m. c. schraefel 2005] or reviews of different gesturing techniques [Kessler et al. 1995; Sturman and Zeltzer 1994; Pavlovic et al. 1997].

Many of the systems reviewed for this survey include only short user trials to test either the accuracy of the system or to report user feedback [Forsberg et al. 1998; Brewster et al. 2003; Wu and Balakrishnan 2003; Nishino et al. 1998; Cao and Balakrishnan 2003; Harrison et al. 1998].

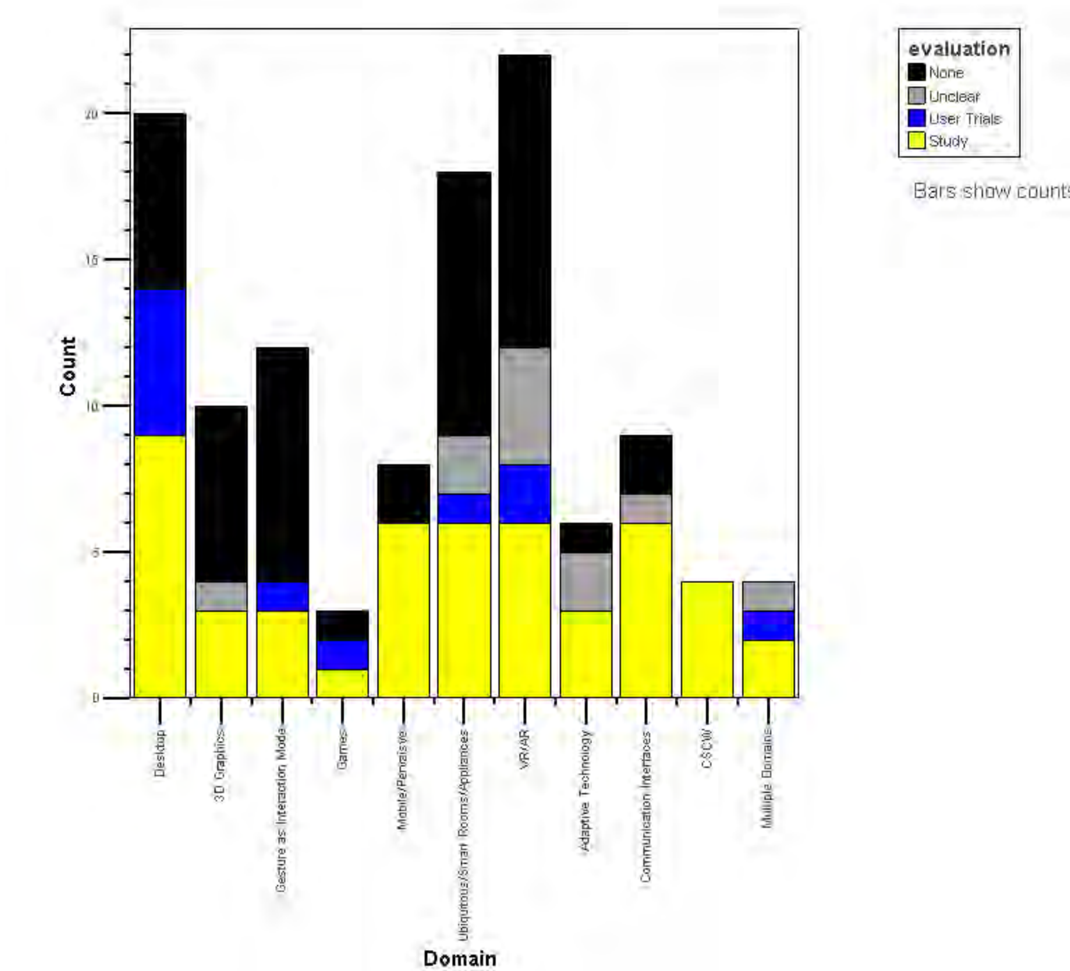

Fig. 6. Evaluations and Application Domains:The bar graph provides a general overview of the types of evaluations that were or were not conducted in the different domains for gesture interactions as reviewed in this paper. The graph shows that studies on gesture based interactions exist for each domain included in this paper. We also note that for certain domains, including 3D graphics, ubiquitous computing and for the study of gestures alone that much of the work does not present any evaluations or studies on their findings. This may be partially due to the fact that many of the research presents what Hinckley calls point designs. Another issue that is raised by this graph is that there are also some areas where actual user trials are not conducted. In the domains that do conduct laboratory studies such as with CSCW and multiple domain research, the lack of user trials may indicate that the work is at an early stage or that there is a need to explore the work using real scenarios outside of a laboratory setting for this area to progress. This will be discussed in this section.

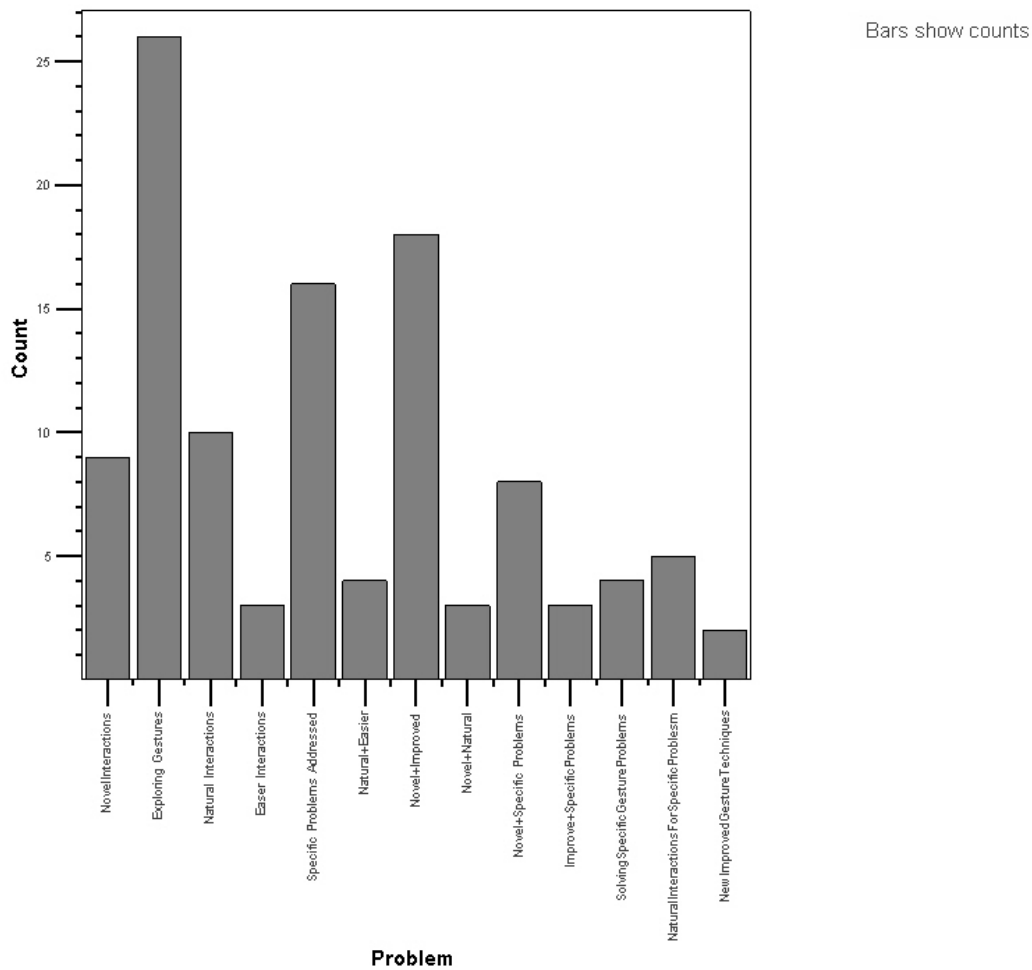

Fig. 7. Motivations and problems addressed through gesture based interactions: The graph shows that along with looking at solving specific problems and generally exploring gestures, the main motivations for looking at gestures is to create more natural, novel and improved interactions. However after 40 years of literature reviews, one would imagine that gestures at this point would be consider for some other characteristics than their novelty. This simply illustrates that future research could focus more on specific problems.

8. GESTURES: SOLVING AND CREATING PROBLEMS

In this section, we examine some of the problems that have motivated the use of gestures as an interaction technique and consider some of the problems that have been introduced when using gestures an interaction technique.

8.1 Natural Interactions

Much of the research on gesture based interactions claim that gestures can provide a more natural form of interacting with computers. Gesticulation style gestures currently considered one of the most natural solution to some of the problems encountered with multimodal interactions of speech and gesture [Wexelblat 1994; Quek et al. 2002; Eisenstein and Davis 2004; Kettebekov 2004; Koons and Sparrell 1994]. Gestures are also referred to as a means to enable more natural interactions within virtual reality applications since using our hands to manipulate objects is a natural approach to such tasks [Fisher et al. 1987; Bolt and Herranz 1992; Zimmerman et al. 1987]. However, the gloves that are required for much of these interactions can be unnatural and cumbersome for the user [Segen and Kumar 1998a]. Computer vision based input attempts to address the unnaturalness of using wired gloves when manipulating virtual objects such as in a 3D projected display for example [Sharma et al. 1996]. But computer vision does not allow the same level of accuracy or granularity of gestures that can be achieved with the glove based interaction. Gloves however remains one of the most commonly used gesture input device for virtual reality and telerobotics applications [Goza et al. 2004].

Manipulations in augmented reality applications have explored the use of markers that when placed on the users fingertips that can be tracked and used to move objects around and to issue commands in the application [Buchmann et al. 2004]. The markers can provide increased tracking power over bare hands and are not as cumbersome as the gloves are however are not as natural or flexible as bare hand interactions could be.

Natural interactions in domains other than virtual or augmented reality applications have been explored for specific applications such as editing musical scores on a screen or tablet pc using stroke gestures [Forsberg et al. 1998], an air traffic controller application [Chatty and Lecoanet 1996] and a system used for collaborative web browsing [Stotts et al. 2004].

8.2 Simplifying Interactions

Another problem addressed thorough the use of gestures in the literature include developing simpler gesture sets and interactions [Song et al. 2000; Swindells et al. 2002; Fails and Jr. 2002]. The need for simpler and intuitive interactions with computers through coverbal or multimodal speech and gesture interfaces has dominated the literature since at least the 80's, combining speech and gestures as a means of making interactions with computers in a more human to human approach [Bolt 1980; Kobsa et al. 1986; Hauptmann 1989; Weimer and Ganapathy 1989; Bolt and Herranz 1992; Koons and Sparrell 1994; Quek 1994; Wexelblat 1994; Cohen et al. 1997; Gandy et al. 2000; Quek et al. 2002; Eisenstein and Davis 2004; Kopp et al. 2004]. For example, recent work by Kettebekov explores linking the beat-like structure of gestures to concurrent speech as a method of disambiguating the individual gestures when interpreting gesticulations [Kettebekov 2004].

One of the most common problems addressed in the gesture literature involves making improvements to current interaction methods by exploring gestures as a means of simplifying existing interactions [Moyle and Cockburn 2003] or creating novel interactions that solve specific problems of existing interactions such as creating more affective interactions for avatars [Paiva et al. 2002] and solving problems with distance manipulations in virtual reality applications [Pierce and Pausch 2002].

8.3 General Improvements for Interactions

A general problem that is addressed using gestures is a general improvement in interactions with a variety of domains including desktop computing [Buxton et al. 1983; Rubine 1992; Cohen et al. 1997], 3D graphics [Segen and Kumar 1998a; Zeleznik and Forsberg 1999], ubiquitous computing [Streitz et al. 1999; Gandy et al. 2000; Rekimoto 1997] adaptive interfaces [Pausch and Williams 1990; Keates and Robinson 1998; Ward et al. 2000], pervasive and mobile computing [Amento et al. 2002] and gaming applications [Freeman et al. 1996].

Although there are many problems that gestures have been claimed to solve or at least offer a possible solution to, additional problems in gesture based interactions have also been created. We now turn to a discussion of some of the problems introduced through the application of gesture based interactions.

8.4 Problems Created

Because gestures have primarily been considered as an alternative interaction mode to traditional methods and to date are still not considered as a mode for everyday computer interactions, several examples of work dedicated to the advancing the study of gestures have attempted to address some of the issues and problems that gestures create rather than solve when applied to different domains [Wexelblat 1998; Dannenberg and Amon 1989; Allport et al. 1995; Kettebekov 2004; Allan Christian Long et al. 1999; Kjeldsen and Kender 1996; Sturman and Zeltzer 1993; Kessler et al. 1995; Eisenstein and Davis 2004].

One of the most prominent issues that arise for gesture interactions is how a computer processor can distinguish between individual gestures when they are performed concurrently. This problem is primarily an issue for open handed or bare hand gestures and involves determining when a gesture starts, when it stops and when to interpret the gestures as a sequence rather than individual gestures. Besides the gestures themselves, various input technologies used to detect gestures, computer vision being one of the most difficult inputs, are also responsible for several problems such as detecting skin tone, feature detection and noise filtering. These problems are within the domain of computer vision research and will not be discussed further in this paper.

In addition, the gesture enabling technology often restricts the type of interaction that is possible, forcing a work around in some form to address the limitations of the technology. For example, computer vision often restricts the style of gestures that are possible and often requires the user to wear coloured gloves or markers to allow recognition [Fang et al. 2003]. Although passive markers or objects create only minimum discomfort tot he user, it still restricts the user from using bare hands to gesture with. An alternative method to using the coloured gloves involves tracking the users' head and face as a means of detecting individual hands [Nickel and Stiefelhagen 2003] but in this scenario, the user is required to be in a specific location in order for the detection to be accurate and for the calibration to remain consistent.

Much of the focus of the gesture research reviewed in this paper attempts to solve or improve old problems caused in previous gesture based systems and interactions. Further problems created through the use of gesture based interactions may include gestures that are complex or unnatural for some users, fatigue caused from prolonged use, privacy issues when gesturing in public places, handling multiple users, memorizing large gesture sets or systems being costly or complex system to actually set up and use in a setting outside a research lab [Wexelblat 1998].

9. RESEARCH TRENDS

There are several trends that we have noted from the current literature reviewed in this paper. We discuss the research trends next in terms of the categories proposed in our taxonomy and provide an overview of some of the gaps in the interaction research that could be used to ensure that gesture based interactions are designed to take advantage of their affordances as an interaction technique across the many domains for future research.

9.1 Gesture Style

One of the fastest growing yet least used styles gesture interactions is gesticulation. The primary domain is in communication interfaces [Kettebekov 2004; Eisenstein and Davis 2004; Quek et al. 2002; Kopp et al. 2004] with some historical uses in VR and 3D graphics applications [Bolt 1980]. This area of research poses a great many challenges for researchers, while promising to deliver a more natural interface to solve many of the problems inherent in speech interfaces, primarily disambiguating speech input.

Along similar lines, language gestures are being explored in the literature in terms of developing more robust recognition capabilities and dealing with more complex signs through a larger vocabulary, as well as incorporating speech recognition approaches to solve many of the problems. The other domains do not typically consider gesticulations or language gestures as an interaction mode however there may be opportunities, for example in desktop computing or with games, to incorporate these natural style gesture interactions. Since most of the work in this area is primarily exploratory, it may be some time before gesticulations are ready for application to existing interaction domains.

Semaphoric gestures are the largest class of gesture seen in the literature and although they are not considered as natural as gesticulation, they make an appearance in almost all of the domains considered in this paper except for communication interfaces which try to avoid pre-recording and learning of any specific gestures. There is still a growing trend towards applying semaphoric gesture to ubiquitous and mobile computing domains. As computers move off of the desktop and into the environment, the ability to wave ones hand or gesture towards a camera to interact with a device at the other end of a room offers promising interaction benefits in certain scenarios. However it is still uncertain what those scenarios are and more research must be conducted to provide a better understanding of the affordances of

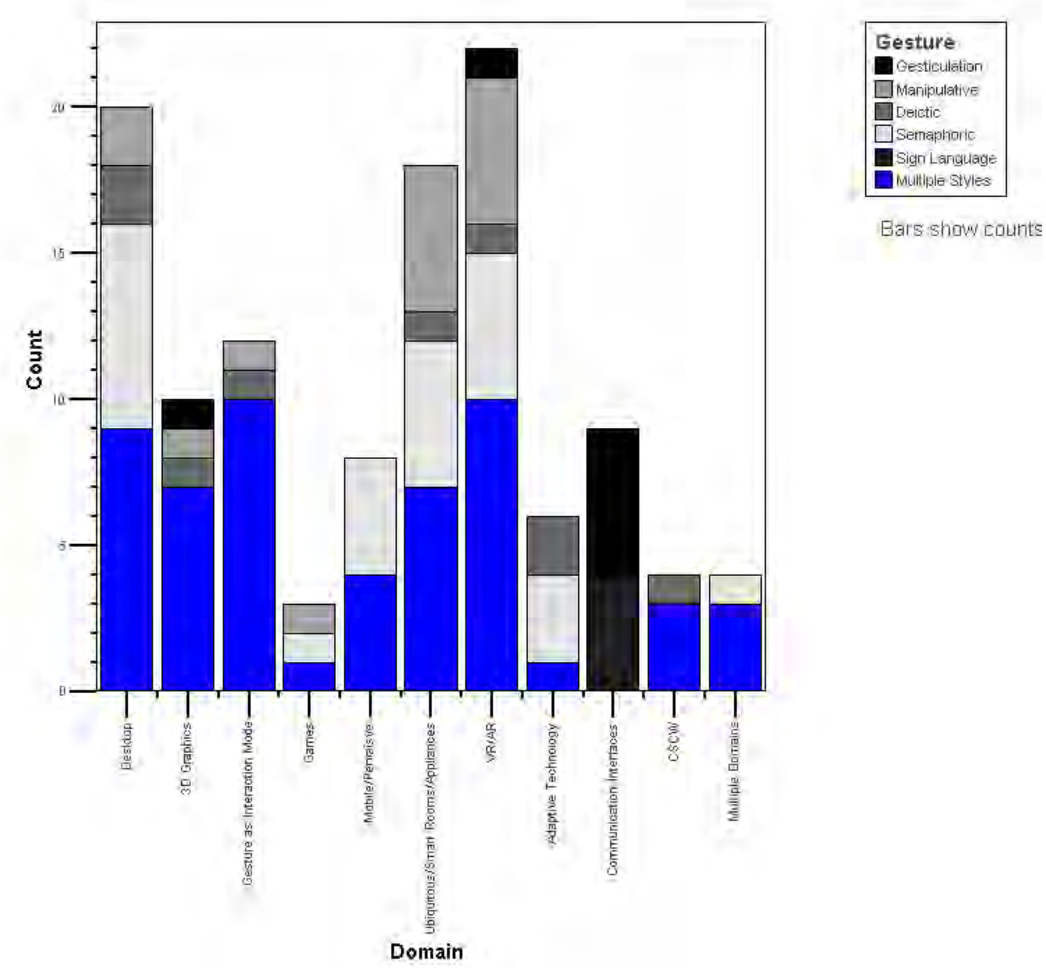

Fig. 8. Gestures in Domains: The graph shows that for all but communication style interfaces, all application domains have explored multiple gesture styles for individual contributions. If we further explore the graph, we can note certain domains that have not explored various gestures. The reasons for this may be because of inappropriateness of the particular gesture style for the domain, or simply because these have not yet been considered. One example is with gesticulation style gestures. This has not received attention in the desktop or gaming domains, however it would seem that these natural style gestures would have a valuable place in such domains. But again, this may be due to the lack of actual working cases of such gestures.

semaphoric gestures and how they can benefit the user [Karam and m. c. schraefel 2005].

Manipulative gestures is almost as widely used as semaphores except in the domain of mobile and pervasive computing where semaphores are more appropriate for the eyes free style of interactions that are required for mobile devices. However manipulative gestures do dominate the literature within the desktop computing domain.

Regarding deictic style gestures, we see this form being used in most of the domains as an interaction technique however most of the recent work tends to incorporate multiple styles of gestures including deictic in combination with semaphoric and manipulative gestures as well as their being a natural gesture that is implicitly a part of the class of gesticulations.

Multiple gesture styles are most commonly used in the literature, which may be due to the move towards more multimodal style interactions. Over half of the systems reviewed in this paper consider multiple styles of gesturing for interactions in all but the communication domains. This is likely due to the volume of work with gestures that is conducted in three main domains: desktop, ubicomp and virtual and augmented reality. As these domains are seeing some of the most promising results with the use of gestures, we will likely see more of this trend towards incorporating multiple gesture styles in single systems and interfaces.

9.2 Enabling Technology

The most commonly used class of gesture enabling technology as seen in the literature recognition is computer vision. One of the least commonly used technologies are sensors that are attached to the body. Data gloves are the next largest input used, which may be due to their application to VR and AR for gesture interactions. If we look at the desktop computing domain, and combine the mouse, pens and touch surfaces, we begin to see an equally large amount of research that focuses on gestures through direct input devices as with the vision input. Sensors, both remote and those attached to objects or the users are used least of all in the literature. In general, the trend would appear to be moving towards more perceptual style interactions which seems to be currently possible with computer vision. We would expect that as the quality of the vision improves, so will its use as a method of input for gesture based interactions.

For gestures that use direct input technologies, we already see mouse gestures and pen gestures becoming ubiquitous in everyday computing. This may be due to the level of understanding that is currently available on these devices, enabling researchers to focus on interactions rather than on the constantly changing technology that is seen in the perceptual style input such as audio and vision.

Besides the obvious observation of the use of input technologies, it is the fundamental issues that have not yet been addressed in the literature. These include understanding appropriate usage scenarios, tasks, application domains or the gestures themselves that are most appropriate for the different interactions.

For the less common input technology such as touch fabrics or electronic sensors, gestures appear to offer a more natural and intuitive mode of interaction. These input devices tend to afford specific styles of gestures which make it less confusing when making decisions about which gesture style would be most suitable for a

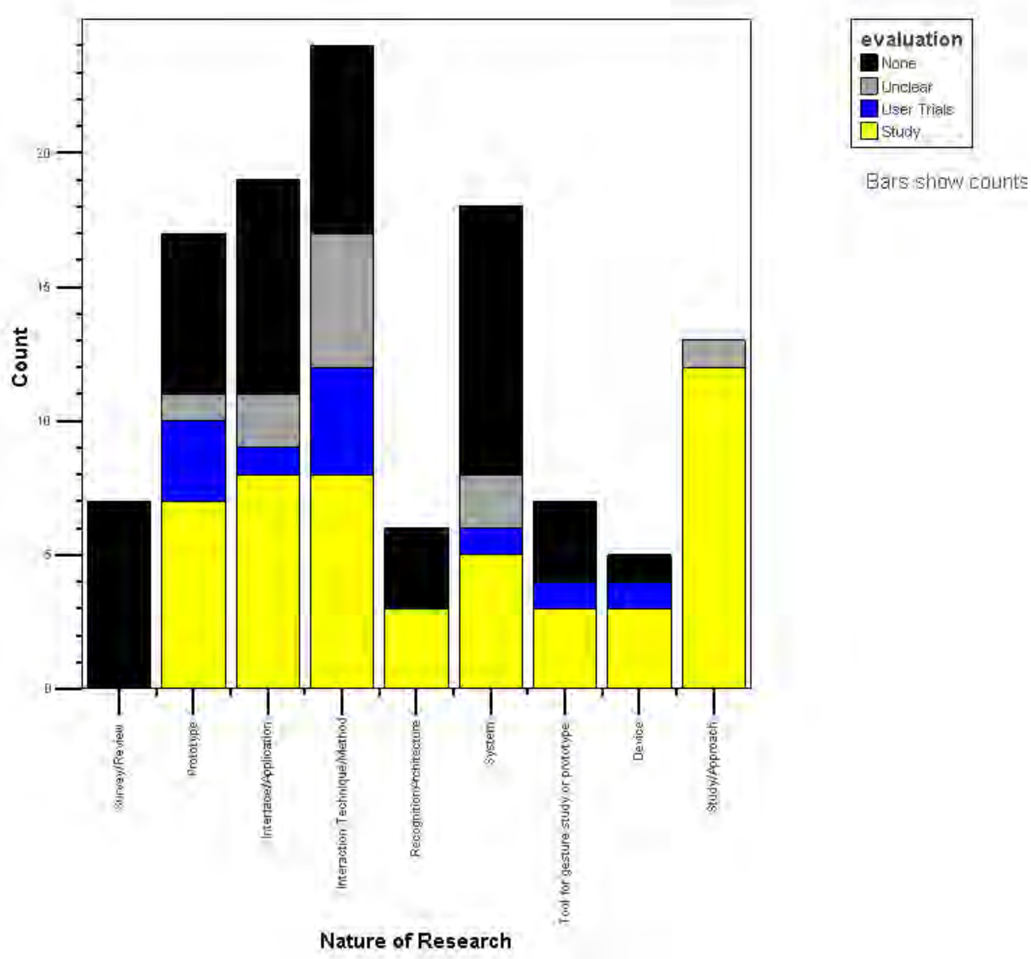

Fig. 9. Evaluations according to the nature of the literature reviewed: The bar graph represents the distribution of the different types of research that the literature presents and the evaluations conducted in each. We note that for a survey, evaluations are not relevant, whereas the study bar consists mainly of evaluations, and are just included for completeness. We do note that for all of the other categories, which focus on actual prototypes and systems, there are significant contributions that have not conducted any form of study or evaluation on their work. Again, we refer to the idea of a point design as discussed by Hinckley and note that it would be interesting to learn which of the many systems, applications, prototypes and devices have actually gone past the point design stage.

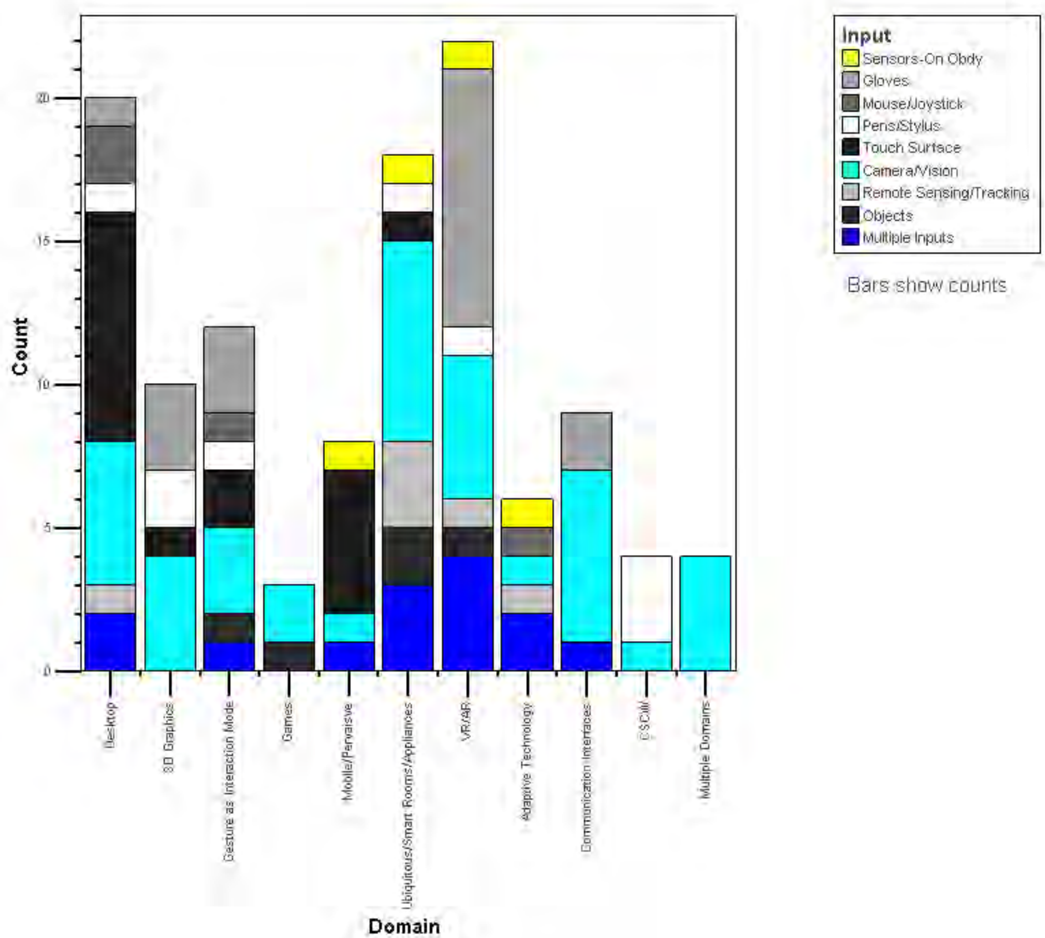

Fig. 10. Enabling technology for different domains: The bar graph shows the distribution of input technology used in the different domains in the literature to enable gestures. We note that there are several domains that have explored gestures which have yet to consider alternative enabling technologies. Although there may be limitations to the type of input possible for a given domain, it may be possible to consider new inputs for the areas that have to date explored only a few inputs such as games, CSCW and multiple domain categories.

touch table for example. However when considering the more perceptual systems that involve computer vision, there is a wider range of gestures that could apply, making it more important to consider the other elements such as application domain and output mode when designing interactions.

9.3 Application Domains

The largest application domain for gesture interaction remains virtual and augmented reality. However, with the increased attention in the relatively newer domains such as pervasive and ubiquitous computing, gestures are making a prominent contribution to interactions and may well become more prevalent in the near future as research increased within these fields.

For ubiquitous computing interfaces, in which the goal of interactions is to be invisible or at least intuitive for the user, gestures continue to be considered for both explicit interactions such as pointing or controlling devices as well as implicit interactions such as sensing users through speech and gesture interfaces.

For pervasive, mobile and wearable computing domains, gestures continue to provide a viable method of eyes-free interactions than can enable the user to concentrate on primary tasks such as driving or walking. Pervasive, mobile and wearable computing is also showing benefits from gesture interactions, primarily through touch screen stroke gestures that enable the user to control their PDA through touch and audio, feeing up the visual channel for other tasks involving mobility [Lumsden and Brewster 2003; Pirhonen et al. 2002; Brewster et al. 2003; Pastel and Skalsky 2004].

Trends in teleoperations, telematics and telerobotics are showing that gestures offer greater benifits in terms of naturalness and intuitive interactions over keyboard, joystick or other interactions that do not map as naturally to the actual task or that may distract a user from driving or concentrating on a more important task.

9.4 System Responses

The primary response for gesture interactions remains directed to the desktop style display or monitors. Possibly since it is one of the oldest forms of output for gesture based interactions, beginning in the 60's when Sutherland used pen based gestures to interact with graphics applications. However we see projected and large screen displays closely following in volume of research attention. Within each of the domains, there is a dominant system response or output, as with head mounted displays in virtual reality for example while audio is the leading output mode for mobile and pervasive research.

One of the key trends to note with system responses and the application of gestures as an interaction mode is the application to novel, cutting edge devices. However, with this novel approach to gestures, it may be difficult to progress past the point design stage until the technology becomes more common, as accessing such devices to conduct studies is difficult due to costs and availability.

9.5 General Trends and Evaluations Observed

If we consider gesture based interactions in terms of the criteria presented in this paper, we can address some of the fundamental elements of gestures that is essential

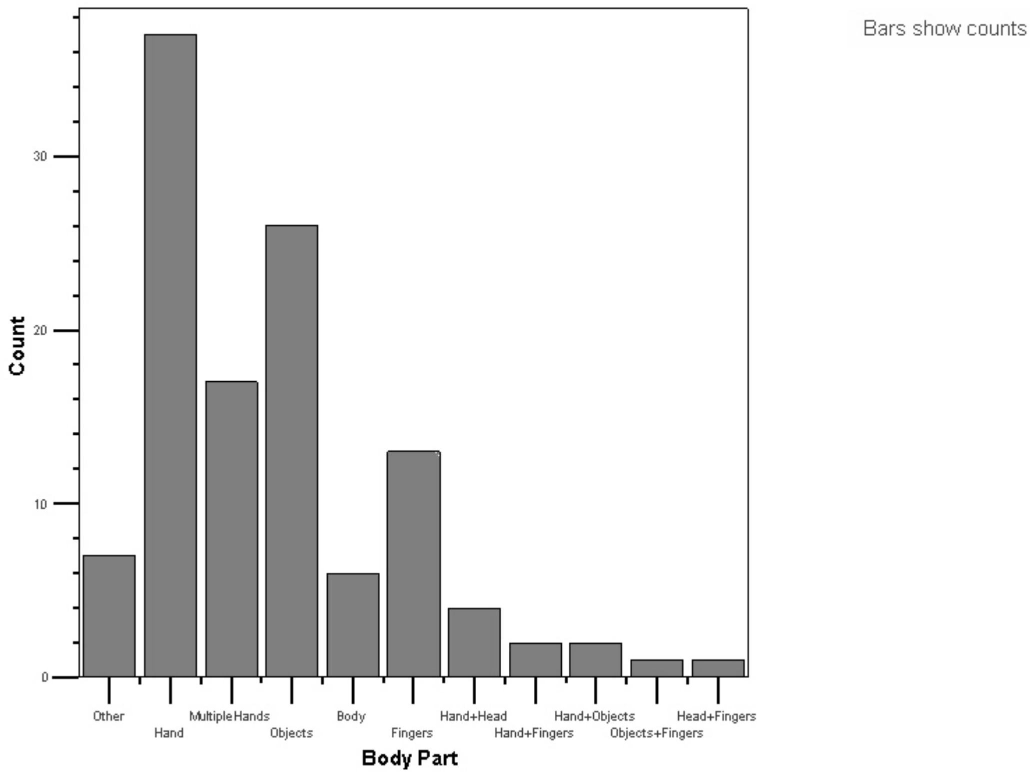

Fig. 11. Gesturing and body parts: The bar graph shows the different body parts or objects that are used to gesture with. We note that to date, hands and fingers (objects by default are controlled using the hands) are the primary gesturing body parts, which may seem trivial however, the use of fingers to perform gestures seems to be a growing area although it is one of the most used in combination with other gestures such as head and hands.

to most computing interactions: Enabling technology, system response, application domain and gesture style. These characteristics can be considered as a standard method of providing designers and practitioners with specific details about gesture interaction systems. By maintaining a standard approach to the field, consistent terminology can assist in creating a common discourse within the community. However, these are the fundamental elements of any gesture based interaction system and if we look to the future, there are several additional aspects of the interaction that should be addressed.

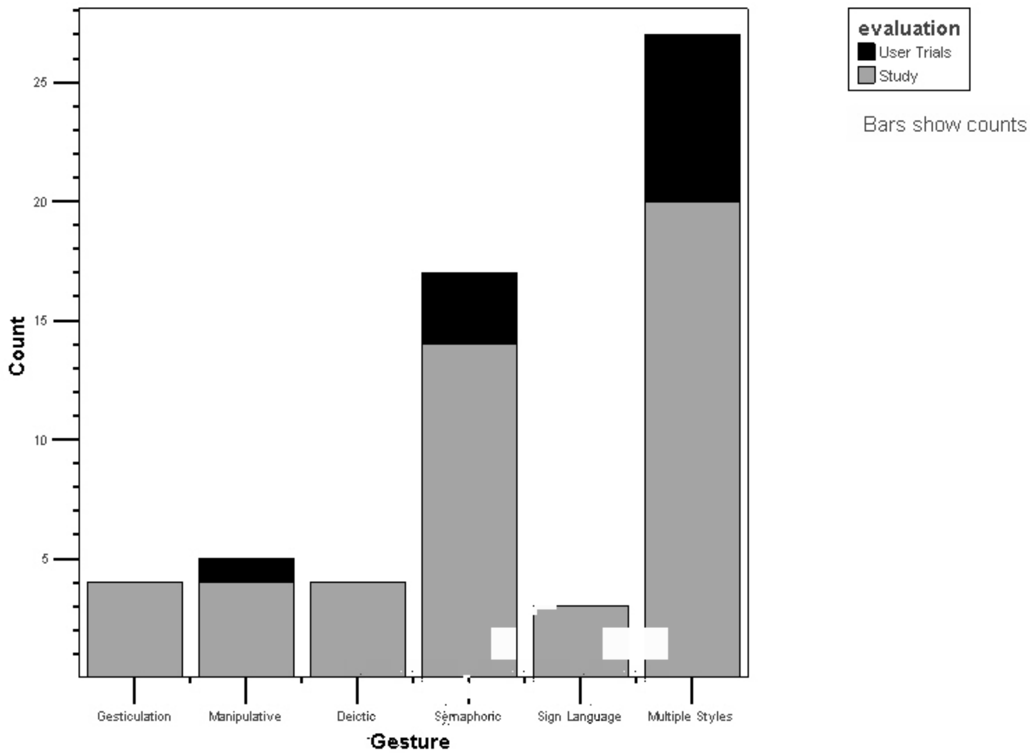

Fig. 12. Evaluations in gestures: The bar graph shows the evaluations that have been conducted within the literature in terms of the gesture styles they describe. We note that due to the lack of user trials conducted in most of the gesture styles that this may indicate why we still do not see gestures being used as an interaction mode for everyday computing scenarios. Although there are studies on each of the styles, we can deduce that either there are still findings that need to be explored before we can consider gestures for everyday computing or that the styles have just not been applied to those scenarios where there are users on which to conduct evaluation with. Again, this is simply provided as a motivation for researchers to look into conducting real user trials with some of the gesture styles in various systems.

9.6 Evaluations in the literature

As reviewed in this paper, we noted that there are several, highly relevant and necessary issues that could be addressed in future work on gesture based interactions. We discuss some of them in the following sections.

9.6.1 *Task analysis.* Although task analysis is a common approach to designing systems in HCI, there is still limited understanding of the types of tasks that gestures are most suited for. For example, there are many gesture systems that

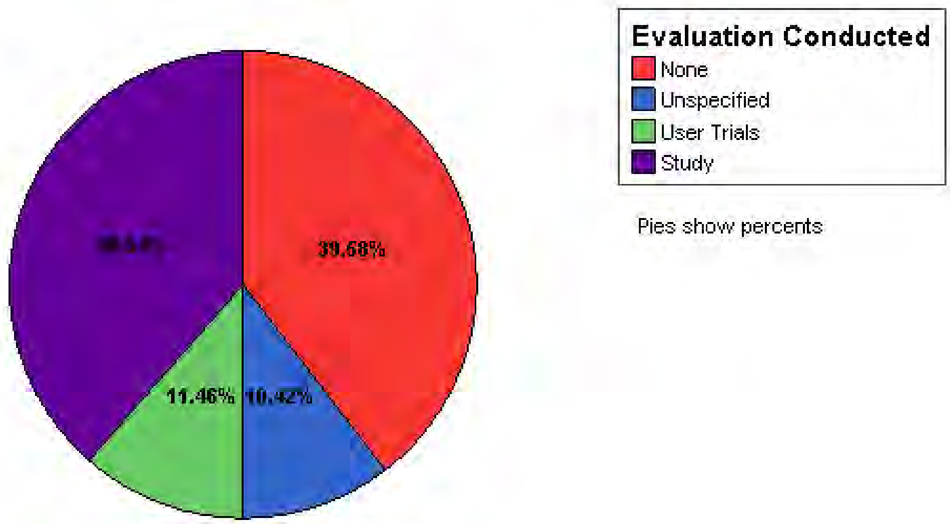

Fig. 13. Evaluations on enabling technologies: The bar graph shows the distribution of gesture enabling technology used in the literature for different application domains. The graph shows almost half of the research reviewed in this paper has not presented any form of evaluation or study that is relevant to their systems or applications. Since all of the literature included in this pie are based on systems, applications or interfaces and other tangible contributions to the field, it is surprising that most of them have not performed any form of evaluation or provided any data on the usability, accuracy or other feature of their system. Although many of the research presents novel work, it would seem that there should be some form of study to back up the legitimacy or contribution that the research gives to the field.

are designed to allow users to control their music players or appliances, however there are no user studies that we have found that address which functions users actually want to control or which types of appliances are suitable for gesture interactions. There needs to be more of a focus on the interaction aspects of gestures before designers and practitioners can realistically consider their use in computing scenarios.

9.6.2 *Computing Scenarios.* Before task analysis can occur, it would be useful to have a better understanding of the application domains and specific scenarios that would be best suited for gesture based interactions. Although gestures have been explored as an interaction mode for numerous scenarios in multiple domains, there is still a great deal of knowledge to be gained regarding the appropriateness, usefulness and practicality of the given scenarios. Currently, we see most interaction scenarios

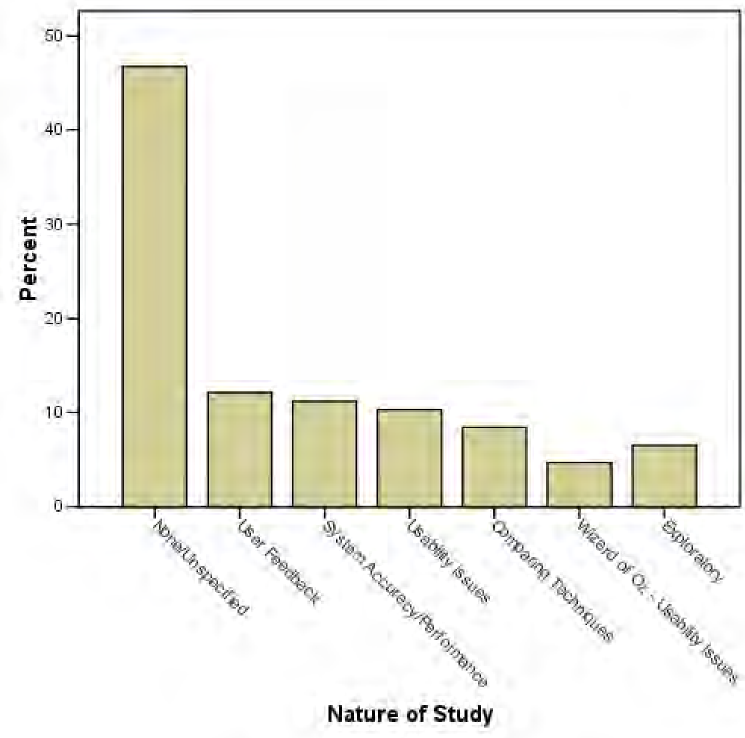

Fig. 14. The nature of the studies: In this graph, we show a generalization of the nature of the studies conducted in the literature reviewed. We do not include work such as surveys where a study or evaluation is not appropriate. We note from this graph that the different topics of study are relatively equally addressed in the research. Although the headings of the bars are a generalization, they do reflect the nature of the studies that are mentioned in the literature.

designed around the technology. In order to ensure that gesture interactions are used to enhance computing for the user, more studies are needed to provide more useful data to practitioners and designers about scenarios that are de based on more than a few user studies.

9.6.3 *System performance and user tolerance.* Again, there are few if any studies that investigate the actual performance that is required from a gesture interaction system or on the tolerance that users have for error from the system. For example, while the quality and power of cameras available for a visual gesture recognition system ranges from basic web cameras to expensive systems for high end research, there is currently no available data on what an actual recognition system would require. The same goes for algorithms and supporting technologies for gesture interactions. Unless we gain some understanding about the user's requirements for system performance and their tolerances for error, it will be difficult to make practical decisions on what technology to use for which interaction system.

As seen in the previous sections, there are several trends that have persistently motivated researchers to experiment with the various gesture based interactions including creating natural, simple, intuitive, human to human style interfaces and interactions. Also, there is the goal to develop novel interactions for new technology as well as for improving existing interactions and addressing the needs of new computing domains [Cao and Balakrishnan 2003; Paradiso 2003; Wilson and Shafer 2003; Lenman et al. 2002a; Fails and Jr. 2002; Swindells et al. 2002].

9.6.4 *As The Future Approaches.* Andrew Dillon refers to three stages of HCI research. In stage one HCI, a system is developed and then tested for usability issues. In stage two HCI, a system is developed and altered based on HCI recommendations. However in stage three approach to HCI, which Dillon refers to as the enhancement stage, HCI research informs technological developments so that systems can be built after the usability testing has been completed. In the past, gesture research has proceeded in the stage one and stage two manner. Since we have seen gestures used for all the latest cutting edge input and output technologies, the ability to test and explore gestures is limited to short studies with small user groups. By taking the third stage approach to HCI, it may be possible to predict usage scenarios and user performance with gestures. This could be accomplished in several ways including creating scenarios in which the gestures could be used and designing the interaction based on the user requirements and gesture affordances. Techniques such as the Wizard of Oz approach to evaluations, reviewing past literature and research, or working with mockups and existing prototypes could enable researchers to perform empirical experiments to demonstrate that gestures are indeed a functional, usable and appropriate mode of interaction for the given scenario.

10. CONCLUSIONS

In this paper, we have presented a classification of gesture based interaction research from over the past 40 years in terms of four main elements that apply to any gesture based interactions: the gesture style, the enabling technology, the system response and the application domain. Within these dimensions, we provided a further breakdown of the existing research into categories in order to contextualize work to date so that future research in the field can gain a better understanding of the vast field that has been considered under the term of gesture based interactions.

Rather than focus on a specific style of gesture or technology, we wanted to show the many different contexts in which the term gesture has been used in computing literature as an interaction technique. Although one could consider stroke gestures performed using a mouse as a completely different interaction mode to manipulation gestures performed with a glove, we have attempted to show that gestures are interchangeable in terms of the four dimensions used to categorize the research. That is, stroke gestures can be used to alter 3d graphic objects just as hand poses can be used to control a web browser.

The past 40 years of computer research that includes gesture as an interaction technique has demonstrated that gestures are a natural, novel and improved mode of interacting with existing and novel interfaces. But gestures remain a topic for the research lab and have not yet become a standard feature of Microsoft Windows or the Mac operating system as speech has. This was a problem that was addressed as early as 1983 by William Buxton who noted "a perceived discrepancy between the apparent power of the approach and its extremely low utilization in current practice" [Buxton et al. 1983]. This is a similarly relevant problem in today's gesture research where, as we have shown in this paper, so much has been done in theory, but so little has been applied in practice.

By presenting this overview of the vast range of gesture based research, we hope to provide a better perspective on the field as a whole and to give future researchers in gesture based literature a complete foundation on which to build their gesture based systems, experiments and interactions.

REFERENCES

Web site: International society for gesture studies.

Web site: University of chicago: Mcneill lab for gesture and speech research.

Allan Christian Long, J., Landay, J. A., and Rowe, L. A. 1999. Implications for a gesture design tool. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press, 40–47.

Allport, D., Rennison, E., and Strausfeld, L. 1995. Issues of gestural navigation in abstract information spaces. In *Conference companion on Human factors in computing systems*. ACM Press, 206–207.

Alpern, M. and Minardo, K. 2003. Developing a car gesture interface for use as a secondary task. In *CHI '03 extended abstracts on Human factors in computing systems*. ACM Press, 932–933.

Alty, J. L. and Rigas, D. I. 1998. Communicating graphical information to blind users using music: the role of context. In *CHI '98: Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 574–581.

Amento, B., Hill, W., and Terveen, L. 2002. The sound of one hand: a wrist-mounted bio-acoustic fingertip gesture interface. In *CHI '02 extended abstracts on Human factors in computing systems*. ACM Press, 724–725.

Barrientos, F. A. and Canny, J. F. 2002. Cursive:: controlling expressive avatar gesture using pen gesture. In *Proceedings of the 4th international conference on Collaborative virtual environments*. ACM Press, 113–119.

Baudel, T. and Beaudouin-Lafon, M. 1993. Charade: remote control of objects using free-hand gestures. *Commun. ACM 36,* 7, 28–35.

Bolt, R. A. 1980. Put-that-there: Voice and gesture at the graphics interface. In *Proceedings of the 7th annual conference on Computer graphics and interactive techniques*. ACM Press, 262–270.

Bolt, R. A. and Herranz, E. 1992. Two-handed gesture in multi-modal natural dialog. In *Proceedings of the 5th annual ACM symposium on User interface software and technology*. ACM Press, 7–14.

Borchers, J. O. 1997. Worldbeat: designing a baton-based interface for an interactive music exhibit. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press, 131–138.

Bowden, R., Zisserman, A., Kadir, T., and Brady, M. 2003. Vision based interpretation of natural sign languages. In *Exhibition at ICVS03: The 3rd International Conference on Computer Vision Systems*. ACM Press.

Braffort, A. 1996. A gesture recognition architecture for sign language. In *Proceedings of the second annual ACM conference on Assistive technologies*. ACM Press, 102–109.

Brereton, M., Bidwell, N., Donovan, J., Campbell, B., and Buur, J. 2003. Work at hand: an exploration of gesture in the context of work and everyday life to inform the design of gestural input devices. In *Proceedings of the Fourth Australian user interface conference on User interfaces 2003*. Australian Computer Society, Inc., 1 – 10.

Brewster, S., Lumsden, J., Bell, M., Hall, M., and Tasker, S. 2003. Multimodal 'eyes-free' interaction techniques for wearable devices. In *Proceedings of the conference on Human factors in computing systems*. ACM Press, 473–480.

Buchmann, V., Violich, S., Billinghurst, M., and Cockburn, A. 2004. Fingartips: gesture based direct manipulation in augmented reality. In *Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and Southe East Asia*. ACM Press, 212–221.

Buxton, W., Fiume, E., Hill, R., Lee, A., and Woo, C. 1983. Continuous hand-gesture driven input. In *Proceedings of Graphics Interface '83, 9th Conference of the Canadian Man-Computer Communications Society*. 191–195.

Buxton, W., Hill, R., and Rowley, P. 1985. Issues and techniques in touch-sensitive tablet input. In *Proceedings of the 12th annual conference on Computer graphics and interactive techniques*. ACM Press, 215–224.

Cao, X. and Balakrishnan, R. 2003. Visionwand: interaction techniques for large displays using a passive wand tracked in 3d. In *Proceedings of the 16th annual ACM symposium on User interface software and technology*. ACM Press, 173–182.

Chatty, S. and Lecoanet, P. 1996. Pen computing for air traffic control. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press, 87–94.

Cohen, P. R., Johnston, M., McGee, D., Oviatt, S., Pittman, J., Smith, I., Chen, L., and Clow, J. 1997. Quickset: multimodal interaction for distributed applications. In *Proceedings of the fifth ACM international conference on Multimedia*. ACM Press, 31–40.

Coleman, M. 1969. In *Proceedings of the 2nd University of Illinois Conference on Computer Graphics*.

Crowley, J. L., Coutaz, J., and Bérard, F. 2000. Perceptual user interfaces: things that see. *Commun. ACM 43,* 3, 54–ff.

Dannenberg, R. B. and Amon, D. 1989. A gesture based user interface prototyping system. In *Proceedings of the 2nd annual ACM SIGGRAPH symposium on User interface software and technology*. ACM Press, 127–132.

Davis, J. W. and Vaks, S. 2001. A perceptual user interface for recognizing head gesture acknowledgements. In *Proceedings of the 2001 workshop on Percetive user interfaces*. ACM Press, 1–7.

Eisenstein, J. and Davis, R. 2004. Visual and linguistic information in gesture classification. In *Proceedings of the 6th international conference on Multimodal interfaces*. ACM Press, 113–120.

Fails, J. A. and Jr., D. O. 2002. Light widgets: interacting in every-day spaces. In *Proceedings of the 7th international conference on Intelligent user interfaces*. ACM Press, 63–69.

Fang, G., Gao, W., and Zhao, D. 2003. Large vocabulary sign language recognition based on hierarchical decision trees. In *Proceedings of the 5th international conference on Multimodal interfaces*. ACM Press, 125–131.

Fisher, S. S., McGreevy, M., Humphries, J., and Robinett, W. 1987. Virtual environment display system. In *Proceedings of the 1986 workshop on Interactive 3D graphics*. ACM Press, 77–87.

Fitzmaurice, G. W., Ishii, H., and Buxton, W. A. S. 1995. Bricks: laying the foundations for graspable user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co., 442–449.

Forsberg, A., Dieterich, M., and Zeleznik, R. 1998. The music notepad. In *Proceedings of the 11th annual ACM symposium on User interface software and technology*. ACM Press, 203–210.

Freeman, W., Tanaka, K., Ohta, J., and Kyuma, K. 1996. Computer vision for computer games. Tech. rep., nd International Conference on Automatic Face and Gesture Recognition. October.

Freeman, W. and Weissman, C. D. 1995. Television control by hand gestures. Tech. rep., IEEE Intl. Wkshp. on Automatic Face and Gesture Recognition. June.

Gandy, M., Starner, T., Auxier, J., and Ashbrook, D. 2000. The gesture pendant: A self-illuminating, wearable, infrared computer vision system for home automation control and medical monitoring. In *Proceedings of the 4th IEEE International Symposium on Wearable Computers*. IEEE Computer Society, 87.

Goza, S. M., Ambrose, R. O., Diftler, M. A., and Spain, I. M. 2004. Telepresence control of the nasa/darpa robonaut on a mobility platform. In *Proceedings of the 2004 conference on Human factors in computing systems*. ACM Press, 623–629.

Grossman, T., Wigdor, D., and Balakrishnan, R. 2004. Multi-finger gestural interaction with 3d volumetric displays. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*. ACM Press, 61–70.

Gutwin, C. and Penner, R. 2002. Improving interpretation of remote gestures with telepointer traces. In *Proceedings of the 2002 ACM conference on Computer supported cooperative work*. ACM Press, 49–57.

Harrison, B. L., Fishkin, K. P., Gujar, A., Mochon, C., and Want, R. 1998. Squeeze me, hold me, tilt me! an exploration of manipulative user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co., 17–24.

Hauptmann, A. G. 1989. Speech and gestures for graphic image manipulation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press, 241–245.

Henry, T. R., Hudson, S. E., and Newell, G. L. 1990. Integrating gesture and snapping into a user interface toolkit. In *Proceedings of the 3rd annual ACM SIGGRAPH symposium on User interface software and technology*. ACM Press, 112–122.

Hinckley, K. 2003. Synchronous gestures for multiple persons and computers. In *Proceedings of the 16th annual ACM symposium on User interface software and technology*. ACM Press, 149–158.

Hinckley, K., Pausch, R., Proffitt, D., and Kassell, N. F. 1998. Two-handed virtual manipulation. *ACM Trans. Comput.-Hum. Interact. 5,* 3, 260–302.

Iannizzotto, G., Villari, M., and Vita, L. 2001. Hand tracking for human-computer interaction with graylevel visualglove: Turning back to the simple way. In *Workshop on Perceptive User Interfaces*. ACM Digital Library. ISBN 1-58113-448-7.

Jin, Y. K., Choi, S., Chung, A., Myung, I., Lee, J., Kim, M. C., and Woo, J. 2004. Gia: design of a gesture-based interaction photo album. *Personal Ubiquitous Comput. 8,* 3-4, 227–233.

Joseph J. LaViola, J., Feliz, D. A., Keefe, D. F., and Zeleznik, R. C. 2001. Hands-free multi-scale navigation in virtual environments. In *Proceedings of the 2001 symposium on Interactive 3D graphics*. ACM Press, 9–15.

Karam, M. and m. c. schraefel. 2005. A study on the use of semaphoric gestures to support secondary task interactions. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems*. ACM Press, New York, NY, USA, 1961–1964.

Keates, S. and Robinson, P. 1998. The use of gestures in multimodal input. In *Proceedings of the third international ACM conference on Assistive technologies*. ACM Press, 35–42.

Kessler, G. D., Hodges, L. F., and Walker, N. 1995. Evaluation of the cyberglove as a whole-hand input device. *ACM Trans. Comput.-Hum. Interact. 2,* 4, 263–283.

Kettebekov, S. 2004. Exploiting prosodic structuring of coverbal gesticulation. In *Proceedings of the 6th international conference on Multimodal interfaces*. ACM Press, 105–112.

Kjeldsen, R. and Kender, J. 1996. Toward the use of gesture in traditional user interfaces. In *Proceedings of the 2nd International Conference on Automatic Face and Gesture Recognition (FG '96)*. IEEE Computer Society, 151.

Kobsa, A., Allgayer, J., Reddig, C., Reithinger, N., Schmauks, D., Harbusch, K., and Wahlster, W. 1986. Combining deictic gestures and natural language for referent identification. In *Proceedings of the 11th coference on Computational linguistics*. Association for Computational Linguistics, 356–361.

Konrad, T., Demirdjian, D., and Darrell, T. 2003. Gesture + play: full-body interaction for virtual environments. In *CHI '03 extended abstracts on Human factors in computing systems*. ACM Press, 620–621.

Koons, D. B. and Sparrell, C. J. 1994. Iconic: speech and depictive gestures at the human-machine interface. In *Conference companion on Human factors in computing systems*. ACM Press, 453–454.

Kopp, S., Tepper, P., and Cassell, J. 2004. Towards integrated microplanning of language and iconic gesture for multimodal output. In *Proceedings of the 6th international conference on Multimodal interfaces*. ACM Press, 97–104.

Krueger, M. W., Gionfriddo, T., and Hinrichsen, K. 1985. Videoplace an artificial reality. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press, 35–40.

Krum, D. M., Omoteso, O., Ribarsky, W., Starner, T., and Hodges, L. F. 2002. Speech and gesture multimodal control of a whole earth 3d visualization environment. In *Proceedings of the symposium on Data Visualisation 2002*. Eurographics Association, 195–200.

Kuzuoka, H., Kosuge, T., and Tanaka, M. 1994. Gesturecam: a video communication system for sympathetic remote collaboration. In *Proceedings of the 1994 ACM conference on Computer supported cooperative work*. ACM Press, 35–43.

Lee, C., Ghyme, S., Park, C., and Wohn, K. 1998. The control of avatar motion using hand gesture. In *Proceedings of the ACM symposium on Virtual reality software and technology*. ACM Press, 59–65.

Lenman, S., Bretzner, L., and Thuresson, B. 2002a. Computer vision based recognition of hand gestures for human-computer interaction. Tech. rep., Center for User Oriented ID Design. June.

Lenman, S., Bretzner, L., and Thuresson, B. 2002b. Using marking menus to develop command sets for computer vision based hand gesture interfaces. In *Proceedings of the second Nordic conference on Human-computer interaction*. ACM Press, 239–242.

Lumsden, J. and Brewster, S. 2003. A paradigm shift: alternative interaction techniques for use with mobile & wearable devices. In *Proceedings of the 2003 conference of the Centre for Advanced Studies conference on Collaborative research*. IBM Press, 197–210.

Maes, P., Darrell, T., Blumberg, B., and Pentland, A. 1997. The alive system: wireless, full-body interaction with autonomous agents. *Multimedia Syst. 5,* 2, 105–112.

Minsky, M. R. 1984. Manipulating simulated objects with real-world gestures using a force and position sensitive screen. In *Proceedings of the 11th annual conference on Computer graphics and interactive techniques*. ACM Press, 195–203.

Moyle, M. and Cockburn, A. 2003. The design and evaluation of a flick gesture for 'back' and 'forward' in web browsers. In *Proceedings of the Fourth Australian user interface conference on User interfaces 2003*. Australian Computer Society, Inc., 39–46.

Myers, B. A. 1998. A brief history of human-computer interaction technology. *interactions 5,* 2, 44–54.

Nickel, K. and Stiefelhagen, R. 2003. Pointing gesture recognition based on 3d-tracking of face, hands and head orientation. In *Proceedings of the 5th international conference on Multimodal interfaces*. ACM Press, 140–146.

Nishino, H., Utsumiya, K., and Korida, K. 1998. 3d object modeling using spatial and pictographic gestures. In *Proceedings of the ACM symposium on Virtual reality software and technology*. ACM Press, 51–58.

Nishino, H., Utsumiya, K., Kuraoka, D., Yoshioka, K., and Korida, K. 1997. Interactive two-handed gesture interface in 3d virtual environments. In *Proceedings of the ACM symposium on Virtual reality software and technology*. ACM Press, 1–8.

Osawa, N., Asai, K., and Sugimoto, Y. Y. 2000. Immersive graph navigation using direct manipulation and gestures. In *Proceedings of the ACM symposium on Virtual reality software and technology*. ACM Press, 147–152.

Ou, J., Fussell, S. R., Chen, X., Setlock, L. D., and Yang, J. 2003. Gestural communication over video stream: supporting multimodal interaction for remote collaborative physical tasks. In *Proceedings of the 5th international conference on Multimodal interfaces*. ACM Press, 242–249.

Paiva, A., Andersson, G., Höök, K., Mourão, D., Costa, M., and Martinho, C. 2002. Sentoy in fantasya: Designing an affective sympathetic interface to a computer game. *Personal Ubiquitous Comput. 6,* 5-6, 378–389.

Paradiso, J. A. 2003. Tracking contact and free gesture across large interactive surfaces. *Commun. ACM 46,* 7, 62–69.

Paradiso, J. A., Hsiao, K., Strickon, J., Lifton, J., and Adler, A. 2000. Sensor systems for interactive surfaces. *IBM Syst. J. 39,* 3-4, 892–914.

Paradiso, J. A., Hsiao, K. Y., and Benbasat, A. 2000. Interfacing to the foot: apparatus and applications. In *CHI '00: CHI '00 extended abstracts on Human factors in computing systems*. ACM Press, New York, NY, USA, 175–176.

Pastel, R. and Skalsky, N. 2004. Demonstrating information in simple gestures. In *Proceedings of the 9th international conference on Intelligent user interface*. ACM Press, 360–361.

Patel, S. N., Pierce, J. S., and Abowd, G. D. 2004. A gesture-based authentication scheme for untrusted public terminals. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*. ACM Press, 157–160.

Pausch, R. and Williams, R. D. 1990. Tailor: creating custom user interfaces based on gesture. In *Proceedings of the 3rd annual ACM SIGGRAPH symposium on User interface software and technology*. ACM Press, 123–134.

Pavlovic, V. I., Sharma, R., and Huang, T. S. 1997. Visual interpretation of hand gestures for human-computer interaction: A review. *IEEE Trans. Pattern Anal. Mach. Intell. 19,* 7, 677–695.

Pickering, C. A. 2005. Gesture recognition driver controls. *IEE Journal of Computing and Control Engineering 16,* 1, 27–40.

Pierce, J. S. and Pausch, R. 2002. Comparing voodoo dolls and homer: exploring the importance of feedback in virtual environments. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press, 105–112.

Pirhonen, A., Brewster, S., and Holguin, C. 2002. Gestural and audio metaphors as a means of control for mobile devices. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press, 291–298.

Quek, F., McNeill, D., Bryll, R., Duncan, S., Ma, X.-F., Kirbas, C., McCullough, K. E., and Ansari, R. 2002. Multimodal human discourse: gesture and speech. *ACM Trans. Comput.-Hum. Interact. 9,* 3, 171–193.

Quek, F. K. H. 1994. Toward a vision-based hand gesture interface. In *Proceedings of the conference on Virtual reality software and technology*. World Scientific Publishing Co., Inc., 17–31.

Reilly, R. B. 1998. Applications of face and gesture recognition for human-computer interaction. In *Proceedings of the sixth ACM international conference on Multimedia*. ACM Press, 20–27.

Rekimoto, J. 1997. Pick-and-drop: a direct manipulation technique for multiple computer environments. In *Proceedings of the 10th annual ACM symposium on User interface software and technology*. ACM Press, 31–39.

Rekimoto, J. 2002. Smartskin: an infrastructure for freehand manipulation on interactive surfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press, 113–120.

Rekimoto, J., Ishizawa, T., Schwesig, C., and Oba, H. 2003. Presense: interaction techniques for finger sensing input devices. In *Proceedings of the 16th annual ACM symposium on User interface software and technology*. ACM Press, 203–212.

Rhyne, J. 1987. Dialogue management for gestural interfaces. *SIGGRAPH Comput. Graph. 21,* 2, 137–142.

Robbe, S. 1998. An empirical study of speech and gesture interaction: toward the definition of ergonomic design guidelines. In *CHI 98 conference summary on Human factors in computing systems*. ACM Press, 349–350.

Roy, D. M., Panayi, M., Erenshteyn, R., Foulds, R., and Fawcus, R. 1994. Gestural human-machine interaction for people with severe speech and motor impairment due to cerebral palsy. In *Conference companion on Human factors in computing systems*. ACM Press, 313–314.

RUBINE, D. 1991. Specifying gestures by example. In *SIGGRAPH '91: Proceedings of the 18th annual conference on Computer graphics and interactive techniques*. ACM Press, New York, NY, USA, 329–337.

RUBINE, D. 1992. Combining gestures and direct manipulation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press, 659–660.

SAGAWA, H., TAKEUCHI, M., AND OHKI, M. 1997. Description and recognition methods for sign language based on gesture components. In *Proceedings of the 2nd international conference on Intelligent user interfaces*. ACM Press, 97–104.

SCHAPIRA, E. AND SHARMA, R. 2001. Experimental evaluation of vision and speech based multimodal interfaces. In *Proceedings of the 2001 workshop on Percetive user interfaces*. ACM Press, 1–9.

SCHIPHORST, T., LOVELL, R., AND JAFFE, N. 2002. Using a gestural interface toolkit for tactile input to a dynamic virtual space. In *CHI '02 extended abstracts on Human factors in computing systems*. ACM Press, 754–755.

SCHMANDT, C., KIM, J., LEE, K., VALLEJO, G., AND ACKERMAN, M. 2002. Mediated voice communication via mobile ip. In *Proceedings of the 15th annual ACM symposium on User interface software and technology*. ACM Press, 141–150.

SEGEN, J. AND KUMAR, S. 1998a. Gesture vr: vision-based 3d hand interace for spatial interaction. In *Proceedings of the sixth ACM international conference on Multimedia*. ACM Press, 455–464.

SEGEN, J. AND KUMAR, S. 1998b. Video acquired gesture interfaces for the handicapped. In *Proceedings of the sixth ACM international conference on Multimedia*. ACM Press, 45–48.

SHARMA, R., HUANG, T. S., PAVOVIC, V. I., ZHAO, Y., LO, Z., CHU, S., SCHULTEN, K., DALKE, A., PHILLIPS, J., ZELLER, M., AND HUMPHREY, W. 1996. Speech/gesture interface to a visual computing environment for molecular biologists. In *Proceedings of the International Conference on Pattern Recognition (ICPR '96) Volume III-Volume 7276*. IEEE Computer Society, 964.

SILVA, V. H. Z. AND ARRIAGA, H. H. A. 2003. Evaluation of a visual interface for gesticulation recognition. In *Proceedings of the Latin American conference on Human-computer interaction*. ACM Press, 159–165.

SINCLAIR, P. A. S. AND MARTINEZ, K. 2002. Salt and pepper: Tangible hypermedia using the artoolkit. In *Poster in: Proceedings of the First IEEE International Augmented Reality Toolkit Workshop*. ACM Press.

SMITH, G. M. AND M. C. SCHRAEFEL. 2004. The radial scroll tool: scrolling support for stylus- or touch-based document navigation. In *Proceedings of the 17th annual ACM symposium on User interface software and technology*. ACM Press, 53–56.

SONG, C. G., KWAK, N. J., AND JEONG, D. H. 2000. Developing an efficient technique of selection and manipulation in immersive v.e. In *Proceedings of the ACM symposium on Virtual reality software and technology*. ACM Press, 142–146.

STOTTS, D., SMITH, J. M., AND GYLLSTROM, K. 2004. Facespace: endo- and exo-spatial hypermedia in the transparent video facetop. In *Proceedings of the fifteenth ACM conference on Hypertext & hypermedia*. ACM Press, 48–57.

STREITZ, N. A., GEIßLER, J., HOLMER, T., KONOMI, S., MüLLER-TOMFELDE, C., REISCHL, W., REXROTH, P., SEITZ, P., AND STEINMETZ, R. 1999. i-land: an interactive landscape for creativity and innovation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press, 120–127.

STURMAN, D. J. AND ZELTZER, D. 1993. A design method for whole hand human computer interaction. *ACM Trans. Inf. Syst. 11,* 3, 219–238.

STURMAN, D. J. AND ZELTZER, D. 1994. A survey of glove-based input. *IEEE Comput. Graph. Appl. 14,* 1, 30–39.

STURMAN, D. J., ZELTZER, D., AND PIEPER, S. 1989. Hands-on interaction with virtual environments. In *Proceedings of the 2nd annual ACM SIGGRAPH symposium on User interface software and technology*. ACM Press, 19–24.

SUN-MICROSYSTEMS. Sketchpad: The first interactive computer graphics ph.d. thesis, 1963 – mass. institute of technology.

Sutherland, I. E. 2003. Sketchpad: A man-machine graphical communication system. Tech. rep., Computer Laboratory. September.

Swindells, C., Inkpen, K. M., Dill, J. C., and Tory, M. 2002. That one there! pointing to establish device identity. In *Proceedings of the 15th annual ACM symposium on User interface software and technology*. ACM Press, 151–160.

Teitelman, W. 1964. Real time recognition of hand-drawn characters. In *AFIPS Proceedings of Fall Joint Computer Conference*. Spartan Books, 559–575.

Thalmann, D. 2000. The virtual human as a multimodal interface. In *Proceedings of the working conference on Advanced visual interfaces*. ACM Press, 14–20.

von Hardenberg, C. and Berard, F. 2001. Bare-hand human-computer interaction. In *Proceedings of the ACM Workshop on Perceptive User Interfaces*. ACM Press, 113–120.

Ward, D. J., Blackwell, A. F., and MacKay, D. J. C. 2000. Dasher – a data entry interface using continuous gestures and language models. In *Proceedings of the 13th annual ACM symposium on User interface software and technology*. ACM Press, 129–137.

Weimer, D. and Ganapathy, S. K. 1989. A synthetic visual environment with hand gesturing and voice input. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press, 235–240.

Weiser, M. 1995. The computer for the 21st century. 933–940.

Wellner, P. 1991. The digitaldesk calculator: tangible manipulation on a desk top display. In *Proceedings of the 4th annual ACM symposium on User interface software and technology*. ACM Press, 27–33.

Westeyn, T., Brashear, H., Atrash, A., and Starner, T. 2003. Georgia tech gesture toolkit: supporting experiments in gesture recognition. In *Proceedings of the 5th international conference on Multimodal interfaces*. ACM Press, 85–92.

Wexelblat, A. 1994. Natural gesture in virtual environments. In *Proceedings of the conference on Virtual reality software and technology*. World Scientific Publishing Co., Inc., 5–16.

Wexelblat, A. 1995. An approach to natural gesture in virtual environments. *ACM Trans. Comput.-Hum. Interact. 2,* 3, 179–200.

Wexelblat, A. 1998. Research challenges in gesture: Open issues and unsolved problems. In *Proceedings of the International Gesture Workshop on Gesture and Sign Language in Human-Computer Interaction*. Springer-Verlag, 1–11.

Wilson, A. and Shafer, S. 2003. Xwand: Ui for intelligent spaces. In *Proceedings of the conference on Human factors in computing systems*. ACM Press, 545–552.

Wolf, C. G. and Rhyne, J. R. 1993. Gesturing with shared drawing tools. In *INTERACT '93 and CHI '93 conference companion on Human factors in computing systems*. ACM Press, 137–138.

Wu, M. and Balakrishnan, R. 2003. Multi-finger and whole hand gestural interaction techniques for multi-user tabletop displays. In *Proceedings of the 16th annual ACM symposium on User interface software and technology*. ACM Press, 193–202.

Zeleznik, R. and Forsberg, A. 1999. Unicam - 2d gestural camera controls for 3d environments. In *Proceedings of the 1999 symposium on Interactive 3D graphics*. ACM Press, 169–173.

Zhao, S. and Balakrishnan, R. 2004. Simple vs. compound mark hierarchical marking menus. In *UIST '04: Proceedings of the 17th annual ACM symposium on User interface software and technology*. ACM Press, New York, NY, USA, 33–42.

Zimmerman, T. G., Lanier, J., Blanchard, C., Bryson, S., and Harvill, Y. 1987. A hand gesture interface device. In *Proceedings of the SIGCHI/GI conference on Human factors in computing systems and graphics interface*. ACM Press, 189–192.

Zimmerman, T. G., Smith, J. R., Paradiso, J. A., Allport, D., and Gershenfeld, N. 1995. Applying electric field sensing to human-computer interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co., 280–287.