1 Introduction

Avatar worlds promise new modes of rich interpersonal interaction and communication partly because the avatars themselves provide the ability to communicate nonverbally using body movement. These worlds are popular even in the non-immersive desktop environment. Current interfaces for producing avatar-based nonverbal communication (NVC) exist, but permit only a small range of NVC behavior. Most of these interfaces map facial or postural displays of emotion to buttons or menu items in the user interface. These discrete controls are inadequate for producing more spontaneous and continuous forms of NVC such as coverbal hand and arm gesture. This dissertation describes an interaction technique employing pen gesture as a means for controlling avatar gesture. This pen-based technique supports a range of expressive behavior not possible previously.

1.1 The avatar

The word *avatar* is used to describe many kinds of virtual world entities. Here we describe the way that we use the word avatar. We go on to describe how avatars might ideally be used in a virtual world and some of the potential problems with their usage.

1.1.1 Representation and embodiment

An avatar is an object that represents some real person visiting a virtual world. It functions as that person's virtual body. A visitor drives the avatar around the world in order to "get around." When people meet virtually, their mutual presence is partially indicated through their avatars' presence. When visitors interact, their avatars enact their body movements and gestures.

In a graphical virtual environment, the avatar has a visual aspect. Usually it looks like a human though not necessarily like the human it represents. Upon seeing an avatar in the environment, the observer immediately "sees" that avatar as another person. The observer behaves towards the other avatar in a way analogous to the way they would behave towards another person's physical presence. Visitors move their avatars together when they are speaking with one another, and separate them when the conversation ends. They change the avatar's facial expression to indicate emotion. They manipulate the avatar's limbs in order to gesture.

1.1.2 Value of avatar worlds

Avatar worlds are a manifestation of the desire of humans to create in the digital environment a world as rich we have in real life. The design of virtual worlds points to an intent to recreate the feel of face-to-face interaction. Virtual worlds are made up of recognizable elements like buildings with floors and walls, towns with homes and streets, or landscapes with mountains, boulders, meadows and blue sky backdrops. Many scenes recreate real life public meeting places like bars, art galleries, parks and plazas. And, despite the fact that computer graphics can create fantastical creatures of all kinds, avatars often look like humans.

In these worlds humans create new kinds of places in which to interact. These worlds have advantages over the physical world and even over other kinds of communication media.

Like other media, they allow people who are spread about in physical space to meet in a virtual space. The environment can be designed to create the proper feel for a meeting—a formal conference room for a corporate meeting or a coffee house for a casual meeting. With avatars, people can play with their appearance, mask aspects of their identity and preserve confidentiality. Communication can be presented and recorded seamlessly among multiple channels, such as audio and voice, text, diagrams and even gesture.

As elements in the virtual world, the design of avatars—including their appearance, behaviors and controls—condition the expectations and behaviors of the virtual world visitors. In three-dimensional graphical virtual worlds, visitors can maneuver the avatar around the scene. The activity is not unlike a child playing with a doll in a doll house. Although the child is sitting outside the doll house, they imagine themselves into the house through the doll. Similarly, the virtual world visitor projects themselves into the avatar. Technology aids this projection. As the user looks through a browser into the virtual world, they are looking out the eyes of their avatar. Wherever they navigate their avatar in the virtual world is where they *are* in the virtual world. And when they see another avatar, their expectation is that some human is driving it.

1.1.3 Limitations to expression through avatars

There is a catch to this virtuality. The avatar user cannot express themselves through their avatar body in the same way they do through their physical body. Two levels of disruption lie between the avatar user's intention to express themselves through movement and their ability to do so. The first is the difference between the capabilities of the avatar body and a real human body. In construction, the avatar body is simpler than the human body. The avatar's movement capabilities are limited by its geometric and kinematic design, and by the software

that governs its motions. The avatar simply cannot move like a real human. Therefore the user must map their intentions to the *possible* expressions available to the avatar.

The second level of disruption lies between the human and the computer interface to the avatar's control. In the physical world, humans produce bodily expression without conscious thought to configuring the limbs or tensing specific muscles. In the virtual world, the human effects their expression through issuing avatar control commands at the computer. A clumsy interface for controlling avatar expression draws attention away from interaction with the other person and could disturb the flow of conversation. A good interface design can make nonverbal and verbal communication seamless.

1.1.4 Designing avatar controls

The designer of an avatar communication system addresses these disruptions. Through the kinematic and behavioral features of the avatar, the designer selects what kinds of intentions can be communicated. The intentions could be of an affective, imagistic or linguistic nature. The design of the controls directs how the user will access and control the avatar's expressive behaviors.

To understand how the avatar can be used for personal expression and nonverbal communication, it is useful to consider what kinds of nonverbal communication exist in the physical world, and how these function in inter-personal interactions. This will give us a basis to understand how nonverbal behaviors are being used in avatar worlds, and how to design systems which enable them.

1.2 Nonverbal communication

Nonverbal communication encompasses the whole range of behaviors that humans (and other animals) use to signal messages to one another without the use of formal language.

(In contrast, verbal communication consists of behaviors, such as writing, speaking and signing, that transmit meaning through the use of words.) These behaviors can include the way a person dresses, their facial expression and posture. Social scientists analyze and categorize these behaviors in a number of ways. One categorization uses the sensory modality used to perceive the expression such as touch, vision or hearing. Other categorizations include the temporal or spatial extent of the signal, the parts of the body used, the function of the signal and whether or not intention plays a role in the production of the signal. Here we describe a few modalities of nonverbal communication to demonstrate the range of possible behaviors, and to introduce some of the terminology associated with the this field of study.

1.2.1 Range of nonverbal behavior

Appearance refers to the way a person dresses, cuts their hair or decorates their body. Often the way a person dresses is based on their culture or other group affiliation. It can provide cues about a person's status in an organization or society. It can be an intentional kind of display, a medium for personal expression. From it, one can discern something about a person's personality and even mood.

Proxemics, a term coined by Edward Hall [46], refers to behaviors involving the use of space around a person's body during interpersonal interactions. For instance, all cultures have norms for what is sometimes called *personal space*. While engaged in conversation, people will draw closer to each other, but a person becomes uncomfortable if the other one gets too close to them. Personal space is usually maintained subconsciously, but it can also be used intentionally in order to control or influence others [48]. Observations of police interrogations find that "spatial invasion" is a tool commonly used by police. Another study, involving a confederate

and student subjects, found that physical proximity increases the likelihood of people to agree to participate in user studies.

Any use of visually perceived body movement to communicate falls under the field of study called *kinesics*. Though these displays can be decoded, not all of them are produced in order to send a signal. Darwin, who was one of the first to study kinesic behavior, argues in his book *The Expression of the Emotions in Man and Animals* that expressions of emotion are specifically not communicative. He believed them to be vestiges of previously adaptive behaviors that often accompanied emotion in ancestral generations.

Expressions of emotion are most commonly seen and understood in *facial displays* and body *posture*. Researchers have proposed that some facial displays, corresponding to so-called basic emotions are universally understood [37]. Cues to emotion or internal state are also sometimes referred to as *affect* displays.

Gaze behavior, the movements of the eyes, are a kind of facial display. It is often treated separately because of the different functions it participates in. In particular, gaze is used in conversation regulation. The eyes gesture to let someone know when a conversation is still interesting and when it should be let go.

Body movements that are perceived through touch are sometimes referred to as *haptics*. Haptic behaviors can carry formal meanings, such as a handshake, or be a sign of affection such as an embrace.

1.2.2 Gesture

We treat gesture separately because it is the focus of our research. Gesture is a particular kind of nonverbal behavior that usually accompanies speech. Many gestures are spontaneously created in order to participate in the efficient encoding of whatever the speaker is

talking about at that moment. They are sui generis forms whose meanings are derived from their physical form and their temporal relationship with speech. Though gesture can be performed using any part of the body, it usually refers to movements of the hands and arms. Critchley [27] describes gesture this way:

> 'Gesture'—and its diminutive 'gesticulation'—refers to those bodily movements which accompany audible articulate utterance. If pantomime is silent communication, then gesture is a kind of italicized speech. Its role is to augment, elaborate, intensify, expand, modify, maximize, or in someway alter the reference-function of what is being said by the one who is talking.

Communicative, conversational gesture covers such a wide variety of behaviors, it is difficult to say exactly what it is. In most circumstances, when people see it they know it [58].

1.2.2.1 Varieties of gesture

Gestures have been analyzed a number of ways: function, especially function related to speech, form, intentionality, and cultural specificity. A number of classification schemes exist, and each is suited for certain kinds of analyses. Though the categories change, all of these systems recognize the same kinds of movements as belonging to gesture. Most of the gestures described are *coverbal gesture.* Coverbal gestures only accompany speech, and carry meaning that is shared with the spoken words or by placement in the sentence. *Speech independent* gestures, such as emblems like the "thumbs up" sign, may be used along with speech, but would also make sense without accompanying speech. A few definitions, drawn from [36][76] are presented here:

Deictics. These gestures are usually made with the hands or head, and are used for pointing to objects. Deictics may also point to abstract things. For instance, gesturers often assign

meanings to spaces around their bodies, and gesturing towards a space may mean emphasis of that idea.

Iconics. Iconics describe physical things or actions. For instance, in describing a birthday cake, one might draw a circle in the air to describe the shape of the cake.

Metaphorics. Metaphorics describe abstract ideas or thought processes. They are called metaphorics because the gestures often depict images derived from the metaphors embedded in the language [65]. For instance, the idea of a genre as an object may be manifested in a gesture that implies containment.

Beats. Subordinated to the rhythm of the speech, instead of semantic units, beats are simple movements that mark particular words as important to the discourse. They serve a similar function as putting vocal emphasis on particular words in a sentence.

Adaptors. These are very small movements, usually of the extremities, that are considered to be highly unconscious. These movements, especially those of the legs and feet are often more revealing of a person's true state of mind since they are less controllable than movements of the hands and face. These are among the chief cues used in detecting deception [38].

Emblems. Emblematic gestures are those gestures whose meanings have become so embedded in a culture that they can take the place of words in a verbal utterance [39]. A typical example is the "OK sign" used in North American culture made by forming an "o" with the thumb and forefingers and extending the third, fourth and fifth fingers up while orienting the palm away from the body. Because emblems can replace words or phrases, they are a king of *language independent* gesture.

Haptics. Haptics describes gestures in which people touch, such as shaking hands. It excludes self-touching behaviors.

Regulators. Regulators are the movement of the eyes, head and sometimes the body to regulate the flow of conversation. The are used both by the speaker and the listener. These differ from other gestures in that they do not relate to the discourse [90].

1.2.2.2 Gesture and speech

Researchers believe gesture arises from the same generative process that produces speech [59]. As such, gesture is closely related to speech. The relationship varies along a continuum first proposed by Kendon [61]. The continuum is shown schematically in Figure 1-1.

Gesticulation ⟶ Language-like Gestures ⟶ Emblems ⟶ Signed Languages

Figure 1-1. Kendon's gesture/language continuum.

Along this continuum, the gesture varies in both form and function. Gesticulation falls at one end of the continuum. It is characterized by gesture that is idiosyncratic and shares meaning with accompanying speech. Formal sign languages fall at the other end. These signs have canonical form and function exactly the way spoken language does. Emblems fall closer to the sign language part of the spectrum. Language-like gestures are not formally structured in the way emblems are, but they fill a grammatical slot in a sentence without an accompanying word or phrase. An example (drawn from [76]), is "the parents were all right but the kids were [gesture]."

Other researchers separate gesticulation completely from the rest of the continuum [96]. Gesticulation, they find, is fundamentally different in form and function from other types of gesture.

1.2.2.3 Importance of gesture

Because verbal communication can be understood in the absence of gesture, it is sometimes believed that gesture is a separate channel from verbal utterance that can be dropped with little loss of information. The invention of technologies such as writing and printing has forced humans to encode meaning into purely verbal channels. In many domains, verbal communication seems to have a kind of primacy over nonverbal communication. The combination of speech and gesture is actually a more natural way to communicate. Language itself is learned "embedded in a matrix of other channels"; that is, one learns language within a context of tonal, facial, hand and arm movement variations [112]. Developmental psychologists have found that as children acquire language skills, their gesture develops and expands at a rate similar to speech [9].

One of the functions of gesture is to encode meaning. As a person speaks they distribute the message among different channels. Spoken language presents ideas in a segmented and linear fashion. Gesture encodes differently from speech because it has both spatial and temporal aspects [60]. When it accompanies speech, it can represent concepts simultaneously and wholistically. Researchers have found that the combination of gesture with speech forms a kind of composite signal, and that this compound signal encodes a message [40][61][75].

Gesture can be instrumental in helping a speaker generate an utterance [31][41][60]. Studies have also found that gesture has a effect on the person who is gesturing. When asked to over emphasize or under-emphasize pain effects during an experiment, researchers found

that this activity actually affected the perceived pain and the autonomic response of the subject themselves [48].

Spontaneous gesture is sometimes described as unwitting gesture because speakers are only peripherally aware of what they are doing. Researchers have developed ways of decoding gestures to find insights into a speaker's thought processes. These include thoughts that speakers had not necessarily intended to reveal [38][76].

Other studies show that gestural behavior can be brought under conscious control. Whether or not a person produces gestures depends on whether or not they believe that the gestures will be available to a receiver [22][23]. When gesture and other means of communication, such as diagrams, are available to the speaker, they will combine these different modes into integrated units of composite signals [40][43]. When mothers talk to babies, they will modify both their speech and their gestures [10]. If subjects are asked to be more persuasive, they will produce more frequent gesture [77]. It has also been found that increased motivation for approval will result in increased gesticulation [89].

Listeners also attend to the information in gesture. Listeners rely on gestural cues when speech is ambiguous [106], or when white noise interferes with the transmission of speech [86]. In one study, listeners were shown narrations that were designed so that the information conveyed by gesture was distinguishable from the information conveyed by speech. Following the narration listeners were able to accurately relate the information that was conveyed to them through gesture [19]. In a study of the communicative efficacy of iconic gestures, researchers found that verbal descriptions accompanied by gesture facilitated the ability of listeners to identify objects and to recall lists of words and short stories [86].

Gesture also encodes messages that are not strictly related to spoken verbal conversations. These messages are used to smooth over social intercourse. [48] suggests that "the greatest function of nonverbal behavior is that it allows foreshadowing or partial communication of matters that might be demeaning or offensive if actually verbalized." When a speaker gestures when they are in the process of generating an utterance, listeners attend to the gesture as an indication that the speaker is thinking and does not want to be interrupted [31].

1.3 Desktop virtual reality

Desktop virtual reality is a virtual reality experience that is constrained by the technologies that one would normally use at an office desk. Graphical interaction with the virtual world takes place through a computer monitor, the representation of a three dimensional world being projected onto a two dimensional plane. The user interacts with the world using the staples of desktop computing, a keyboard and a pointing device.

Graphical virtual worlds can be two dimensional or three dimensional. In the most advanced worlds the objects are three dimensional and the avatars are highly articulated and animated. The avatars are often capable of facial expression and some kinds of user controlled animation.

1.3.1 Typical virtual world application interface

The primary form of communication in virtual worlds is verbal. In most of these worlds, verbal communication is transmitted via chat text. The chat text may appear above the avatar's head, sometimes in a balloon, and, depending on the software, may also appear in a window next to the virtual world browser. A few worlds enable verbal communication through voice. In OnLive!, visitors communicate with each other using their own voices, and the avatar faces are animated to sync with the avatar user's speech.

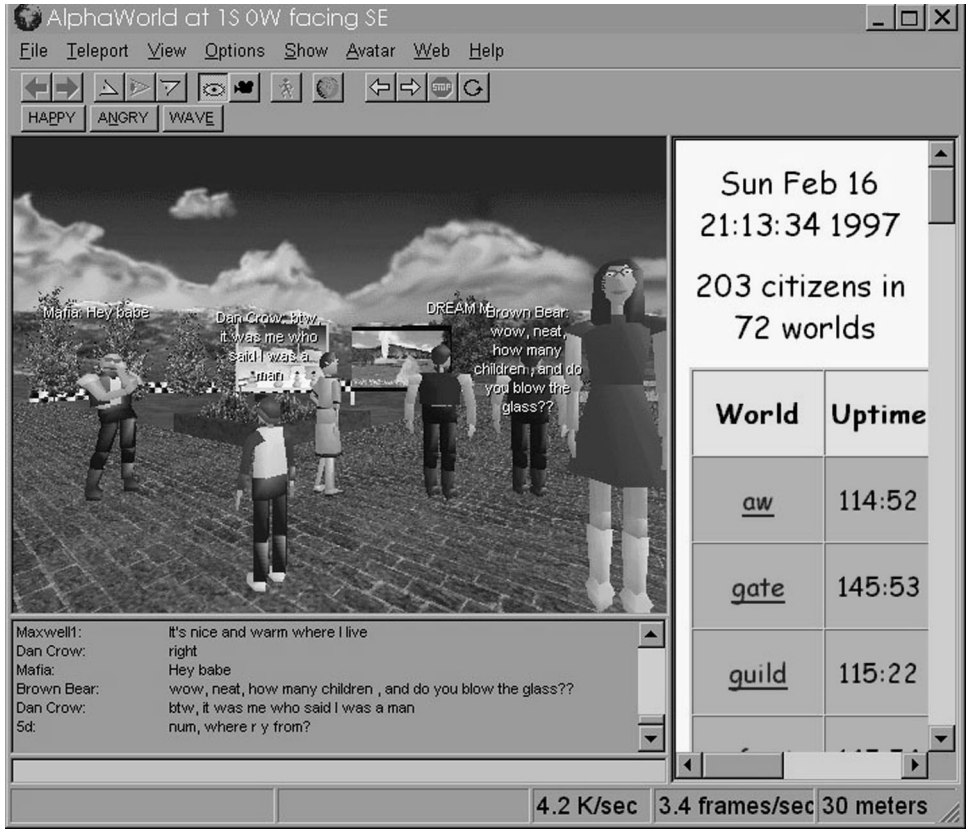

Figure 1-2. Active Worlds virtual world browser.

The main feature of the interface is a window for viewing the world. Controls consist of buttons and menu items for communication and navigation. User's can navigate around the current world or jump to other worlds. Controls for nonverbal communication are found among these interface items. An example world from ActiveWorlds is shown in Figure 1-2.

1.3.2 Nonverbal communication and personal expression

Sociologists have studied how physical environments affect the behavior of humans. In the field of architecture, theories of spatial perception have been proposed [11][51], and tested [12][50]. Architects employ this knowledge in their designs. Analogously, the design of virtual spaces shapes the expectations and interactions of people in the virtual world. If a virtual place is meant to be inviting to new visitors, it will probably mimic a bar or a public plaza.

In these places, visitors will feel encouraged to hang out and meet new people. In fact, these types of virtual environments are perceived by people as a social space [55].

Recognizable elements in these virtual worlds provide context to guide the interactions of the humans who visit them. Given the familiarity of the elements in the virtual world, one expects that interpersonal interaction will look the same as in the physical world. When avatars look humanoid, people expect their behaviors to be just like the behaviors of actual people.

Signals sent by the avatar are processed in same manner as signals sent by a physical human body. Though the human knows they are not looking at an actual person, they react as if they were [85]. The most basic signal sent by the avatar is *presence*. Visitors become aware of others in the world through moving around and visually scanning the space around them [55]. They use these visual cues despite the fact that the software may explicitly indicate the presence of others by displaying textual lists of other visitors logged in.

Another form of nonverbal expression in virtual worlds is appearance. This means a slightly different thing in virtual worlds because avatars are designed and not "born." Many virtual worlds give users a choice of premade avatars to use when they visit. In Worlds Away, users select a head and torso for their avatar from a number of basic models. Other virtual worlds allow users, for a price, to import their own avatar designs. Those who do not pay are relegated to using standard issue avatars. The cool crowd in a virtual world are those who design their own avatars. Status is accrued through having a collection of sexually appealing avatars. [111] Users can use more than one avatar during a single visit. Those who are familiar with avatar switching will sometimes rapidly switch in a spectacular visual display [28].

Proxemic behaviors are automatically enabled in all virtual worlds by virtue of the fact that avatars can move around the world. Moving one's avatar near another avatar can be an

invitation to a conversation. In combination with audio conferencing (as in the OnLive! world), avatar movement is used to form and dissolve conversation groups [30]. During conversation, interactants maintain personal space around their avatars [100]. Running one's avatar through another's (when collision detection is absent) is a way of annoying people, a kind of virtual haptics. Control over body position also permits a primitive form of dancing. In worlds where there is music that can be heard by all, people might rock their avatars to the rhythm and whirl to the melody. Behaviors in avatars, including those specific to particular virtual world communities are described further in [28].

Most virtual worlds allow users to express some emotion through their avatars. In some worlds, avatars are designed with different visual aspects representing different emotions. During a chat, users can choose to put a happy or a sad expression on their avatar's face. Body movements are also put under high level control. In WorldsAway, user's can select *wave* or *bow* as a body expression.

Some virtual worlds give their avatars continual motions that are not controlled by the user. These motions usually correspond to the bodily movements that are performed unconsciously by humans. OnLive! avatars blink and appear to breath. Though this is not really a form of expression, it does try to communicate the "aliveness" of the avatar.

1.3.3 Psychology of Avatar Communication

Telecommunications technologies attempt to replicate as faithfully as possible the communication channels available during face to face conversation. By this measure, virtual reality is a step up from telephony because it adds a visual channel. Compared to video conferencing, it at first appears less successful since video can transmit a person's true appearance.

In fact virtual reality preserves other cues that are lost during video conferencing, such as eye-contact and proximity.

Physical world fidelity is not the only measure of a medium's usefulness. In [52] Hollan et al argue that electronic communications tools should not necessarily try to recreate face to face interaction. Each medium may have affordances that better serve people's communication needs than physically proximate conversation. For instance, people may be more truthful in email because of the anonymity it affords [103]. In another study, participants reported feeling more like themselves during phone conversation and computer chat than in face to face [25].

Avatars can preserve anonymity because the user's real image is not seen. Even when complete anonymity is not desired, the ability to mask appearances proves helpful. A video phone call at an inopportune moment may require that the video camera be turned off. Avatars could provide a stand-in at these times.

Other useful affordances of avatars are yet to be discovered. Avatars provide a different way to express oneself because a person can be represented by more than one avatar. Humans are complex and we think of ourselves as having many sides. Avatar's make this more explicit.

1.3.4 Applications for drivable avatars

Entertainment offers many examples of successful applications. Avatars are particularly suited for role playing games. Chat worlds and now avatar worlds are places where people can construct and explore a new identity. Playing with gender identity and ambiguity online is a well documented activity [15][108].

Traditionally offline performing arts have explored the use of virtual worlds. Networked virtual theater is one example [84][104]. In these performances, the actors, audience and directors may only virtually be in the same theatre. The BBC has studied the feasibility of bringing actors and directors together for rehearsal in a virtual environment setting prior to television performances [99]. The rock bands Metallica and Aerosmith created online worlds to attract fans. In these virtual environments they virtually interact with their fans [107].

In the work place, virtual environment software can be combined with other tools to document and playback meetings. One example is the PAVE tool, built on top of the Palace [2]. Collaborative tools also serve during more informal encounters, fostering a sense of presence and therefore community among collaborators who are spatially separated [32][34].

Gesture exists in more formal domains. In the legal realm, legal gesture signifies a legal change or condition. One example is raising the hand before swearing an oath. In a real sense, law is embodied through gesture. In [49] the author speculates that virtual reality will "re-animate" law by bringing people together in virtual legal spaces. Legal gestures can be redesigned to be conceptually and socially appropriate to the proceedings.

Avatars and virtual reality systems have proved advantageous in psychotherapy treatments to cure various phobias. In therapies for fear of heights or fear of flying it provides a reasonable amount of reality while allowing the patient to maintain a sense of safety [66][91]. Virtual spiders that can be held and even squashed with a virtual hand have successfully treated spider phobias [18].

1.4 The problem of interactive control of avatar gesture

A central difficulty in controlling avatar body movement arises from the disparity between the simplicity of the input device and the complexity of the avatar. In the desktop

environment the avatar user relies on the coordination of their hands and eyes to control the movements of an entire avatar body—an unnatural mode of gesturing. The movement of a simple pointing device specifies only a few numbers: an x coordinate, a y coordinate and possibly a pressure level and one or two of their derivatives. Humanoid avatars, on the other hand, have many joints whose motions need to be coordinated all at once.

Though capture (e.g., using computer vision) and reproduction of an avatar user's own body motion in the avatar seems the most direct solution to controlling avatar gesture, it may not be the most desirable solution. First, the user's body is situated in a radically different environment from the avatar's. The avatar may be "walking" all around a virtual world, but the human is sitting or standing in front of a computer. The user's body movements are constrained by physics and their own physical abilities while the avatar's body can have physically impossible abilities.

Also, the context of the avatar's actions may differ greatly from the user's context. The avatar may be performing athletic demonstrations while the user is ensconced in an office with coworkers who are engaged in their own activities. Given the differences in context between the user's physical and virtual self, avatar control in which the user must physically act out body movements may be undesirable and even inappropriate.

1.5 Goals

The ultimate goal of the research described in this dissertation is an interface in which the user can communicate via an avatar using the fluid mix of speech and gesture that are used in face-to-face conversation. We assume that speech is used for verbal communication.One way that gesture encodes meaning is through its timing with speech. So one focus is the design of controls that can naturally be used while speaking. By "naturally," we mean that the user can

produce avatar gesture spontaneously, it is possible to synchronize the gesture with a spoken utterance, and controlling the gesture does not require a mental context switch. The user action should, as much as possible, feel like part of the communicative effort in the way that coverbal gesture feels like part of the speech effort.

A second goal is to make the interface *expressive*. By expressive, we mean that the user should have continuous control over the avatar gesture. Part of the communicative power of gesture is its capacity to encode meaning through variations in physical variables such as speed, size, and forcefulness. These variables change continuously. We feel that the avatar user will feel more engaged with the avatar and feel a greater sense of personal self-expression if they can control these variables. The interface must be able to extract from the user's movements the expressive features that should be passed on to the avatar.

1.6 A solution using pen gestures

Our solution is a pen user interface for avatar control. Instead of using graphical elements like buttons to select an avatar gesture, the pen gesture selects and invokes an avatar gesture. The main idea is depicted in Figure 1-3. In our case, the pen gesture set consists of letters of the alphabet and each corresponds to a particular avatar gesture. In addition, the way

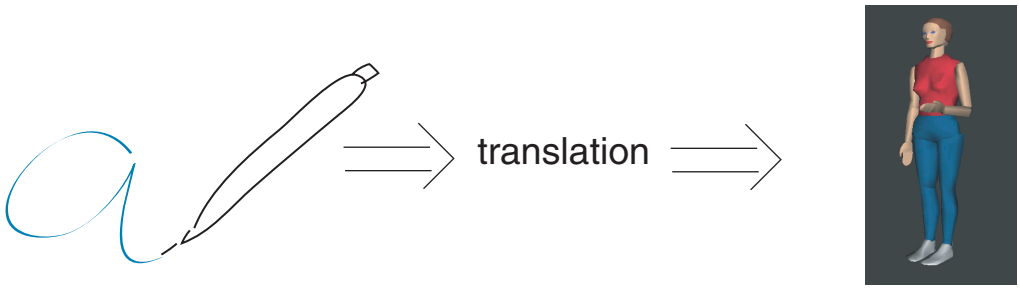

Figure 1-3. Pen gesture controls avatar gesture.

the pen gesture is written determines the way the avatar gesture will be animated. A pen ges-

ture-based user interface provides a user with seamless and spontaneous control of expressive avatar gesture.

1.7 Contributions

This interaction technique enables nonverbal communication through spontaneous avatar gesture. While previous solutions focus on speech independent gesture and affective display, this technique addresses the problem of producing and controlling coverbal gesture on avatars.

This technique applies a pen user interface to the problem of controlling the avatar unlike other solutions which use standard graphical user interfaces. In addition to using the pen gesture identity, the technique extracts continuously valued stylistic information from the way the gesture was written. We developed a framework for mapping this stylistic information to continuous variations in the avatar gesture.

The user not only selects an avatar gesture but also specifies continuously how the gesture will be performed. These variations in the avatar gesture are continuous and can vary in more than one dimension. This technique allows multiple variables to be specified with a single user action.

We also developed an application architecture that allows expressive avatars to be viewed in a multi-user virtual world without requiring viewers to use special software. With this architecture, avatar gesture and animation designers do not need to commit to a particular virtual world platform. They can write a single application and demonstrate their designs to a wide audience.

1.8 Dissertation overview

This chapter has reviewed the problem of controlling avatars for the purpose of nonverbal communication. Nonverbal communication is an important facet of face-to-face communication and has the potential to add previously unseen richness to virtual communication through avatars. In the rest of this dissertation, we will show how a pen gesture interface can provide spontaneous and expressive control over avatar gesture. We find that previous techniques for controlling avatar nonverbal communication are inadequate because they rely on traditional graphical user interfaces. Related work is described in Chapter 2, and the critique of previous techniques is found in Chapter 3. Chapter 3 goes on to describe the aspects of gesture and gestural movement that are most important to its communicative power and how our interaction technique gives the user control over these aspects. This chapter also describes the design of this interaction technique. We show that this technique is viable by describing specific algorithms and techniques for producing expressive movement from pen gestures. These techniques and algorithms are described in Chapter 4 and a speculative variation of the design is described in Chapter 5. Our implementation of the technique as a working application provides a proof of concept. Chapter 6 describes our implementation and its architecture. Avenues for future exploration of controlling avatar gesture control using pen gesture are found in Chapter 7.