Electric Field Sensing and the "Flying Fish"

David Allport,* Thomas G. Zimmerman, Joeseph A. Paradiso,
Joshua R. Smith, Neil Gershenfeld
MIT Media Laboratory - Physics and Media Group
20 Ames Street E15-487
Cambridge, Mass 02176-4307
dea,tz,joep,jrs,neilg @media.mit.edu

June 1995

Abstract

We have developed a novel input device called the "Flying Fish", which uses Electric Field (EF) sensing. We describe the basic physical principles of EF sensing, and show how it provides a simple, cheap mechanism for using hand gestures to navigate through virtual spaces, without the need for wearing or holding anything. We also discuss the implications of this type of device for interface designers, introducing the concept of "invisible metaphors", and illustrating its role in making interactions with 3D virtual worlds feel natural.

1 Introduction

When most computer interface designers think of ways of deriving input from a user, they think in terms of the human senses of touch, sight, hearing, possibly smell and (only in special circumstances) taste. Those familiar with the varieties of VR input devices may also think in terms of using intermediary objects that the user wears or holds, which can be detected by magnetic or radio sensors, or else they may think in terms of extensions of the human sense capacities like ultrasonic or infra-red.

But there is an entirely different naturally occurring sense modality which humans do not have, but which can be used to directly sense the presence of human bodies or body parts.

Weakly electric fish are extremely good at sensing electric fields [Bul82]. Some, like sharks and catfish, just passively sense the presence of electric fields, but others such as the Gymnotiformes also generate their own fields, and detect changes in these fields induced by the presence of objects in the water around them. The Gymnotiformes generate tiny electric fields in their tails and sense electric potentials along the length of their bodies, using amplitude, modulation and spectral changes to determine the size, shape, conductivity, distance, and velocity of objects around them. They also use electric fields for social communication, identifying sex, age, and dominance hierarchy. They perform jamming avoidance when they detect the phase beating

*Visiting Scientist from Hewlett-Packard Laboratories, Bristol

of their transmitted field with the field from another approaching fish: the fish with the lower transmit frequency decreases its frequency, and the fish with the higher frequency raises its frequency. Some saltwater weakly electric fish have adapted their sensing ability to detect EF gradients as low as 5nV/cm.

We speculate that apart from unconscious anthropomorphism, there are two reasons why EF sensing has hitherto not been used extensively in human-computer interfaces. First, it is only recently that inexpensive electronic components have become available to measure the small signals produced by EF sensors. Second, it has not been obvious how to control electric fields and interpret changes in them in such a way as to derive position coordinates from field measurements.

We will describe in section 2 the various ways in which electric fields can be measured, and the particular sensing hardware that we have implemented which provides fast (millisecond), high resolution (millimeter) measurements at a cost of a few dollars per channel, while using very low power detectors (milliwatts), with high immunity to noise ($> 72dB$). This sensing apparatus is not affected by clothing, surface texture or reflectivity, and can operate on length scales from microns to meters. We describe a variety ways in which we have applied EF sensing.

Section 3 gives a detailed case study of one of these applications, where we have used EF sensing to create the Flying Fish input device. This is a virtual joystick which senses the position of a user's hands in front of the screen and uses this to control 3 dimensions of movement simultaneously. We also discuss different combinations of pedals and EF sensing to produce a variety of different virtual input devices.

The developers of gestural interfaces of all kinds, whether using image processing, gloves, or other sensing mechanisms, often claim that gestural interfaces are "more natural" (see e.g. [SZ94],[BBL93]). Section 4 examines more closely what it means for a gestural interface to 3D navigation to feel "natural". In particular, we describe what we call "invisible metaphors", and show that there is no unique natural interpretation for a given gesture. We suggest that creating an intuitive conceptual model of navigation in a virtual space involves considering

- the application task
- the physical properties of the input mechanism
- the formal representation of the object space
- the functional mappings from the input device to movements in the domain
- both visible and invisible metaphors
- the user

We conclude by comparing our EF work with other research in remote sensing. We also discuss some of our ideas for future uses of EF sensing, and what we see as the key research issues in this area.

2 Varieties of Electric Field Sensing

2.1 The Theremin

The first well-known use of EF sensing for a human-machine interface was Leon Theremin's musical instrument [Gar67]. He used two omnidirectional antennas to control the pitch and amplitude of an oscillator. A hand or other body part brought near an antenna introduces additional capacitance to ground which detunes a resonant tank circuit (see figure 1). The effect of body capacitance on electric circuits was also well known to radio's pioneers, who saw the effect as an annoyance rather than an asset.

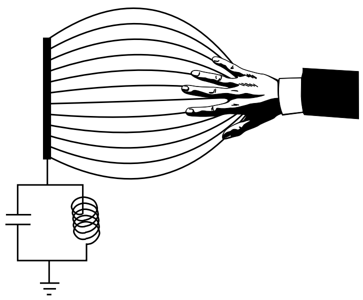

Figure 1: In the Theremin, each inductor/capacitor resonant tank circuit emits an electric field which is intercepted by the hand. The resonant frequency of the circuit changes as the hand moves in and out of the field.

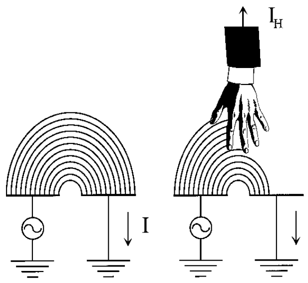

Figure 2: In what we call "Human Shunt Mode", an electric dipole field created between an oscillating transmit electrode and virtual ground receiver electrode is intercepted by a hand. Displacement current to ground $\mathbf{I_R}$ decreases as the hand moves further into the dipole field.

2.2 Human Shunt Mode

Our research on EF sensing for human-computer interfaces also grew out of a musical project. We wanted to measure the position of the bow to Yo-Yo Ma's cello in two axes with minimum impact on the instrument and its playability [Ger93]. As a result of a side-effect that occurred during this project we discovered that it was possible to create a *directed* electric field and use it to make measurements of human hand position.

Figure 2 shows the EF sensing mode that we call Human Shunt Mode. An electrical potential (voltage) is created between an oscillator electrode and a virtual ground electrode. A virtual ground is an electrical connection kept at zero potential by an operational amplifier, allowing current $\mathbf{I_R}$ to ground to be measured. The potential difference induces charge on the electrodes, creating an electric field between the electrodes. If the area of the electrodes is small relative to the spacing between them, the electrodes can be modeled as point charges producing dipole fields. The dipole field strength varies inversely with distance cubed. In practice the measurable field strength extends approximately two dipole lengths (distance between the transmitter and receiver electrodes). As the electrodes are moved farther apart, a larger electrode area is required to compensate for the decrease in signal strength.

When a hand, or other body part, is placed in an electric field the amount of displacement current $\mathbf{I_R}$ reaching the receiver decreases. This may seem counter-intuitive since the conductive and dielectric properties of the hand should increase the displacement current. However, if an object is much larger than the dipole length, the portion of the object out of the field serves as a charge reservoir, which is what we mean by "ground". The hand intercepts electric field lines, shunting them to ground, decreasing the amount of displacement current $\mathbf{I_R}$ reaching the receiver.

A major difference between this mode of sensing and the Theremin mode is that we can now use a single transmitter and multiple receivers in different locations, and place the receive electrodes to measure different dimensions, giving object position or size. This is because we get a dipole field (having a 3D ellipsoid shape) between the transmitter and each receiver. Each dipole field can be used to measure one degree of freedom, either giving object position or size. A single dipole cannot distinguish a close small object from a large distant object, since both might shunt the same number of field lines. A second dipole operating on a longer length-scale (greater electrode spacing) can be used to distinguish these two situations, or to measure two spatial coordinates of a single fixed-size object. Three dipoles can measure the 3D position of an object of fixed size, or determine the 2D position and size of an object. Four dipoles can determine the size and 3D position of an object. Five dipoles can determine the 3D position, size, and elongation of an object. We are working on the continuum limit of adding more dipoles, to perform low-resolution imaging [Smi95].

2.3 Human Transmitter Mode

During our experiments with a variety of applications we also discovered that it is possible to sense fields in another mode [GZA95]. Here a person can stand on, sit on, touch, or otherwise be near the oscillating transmit electrode. This capacitively couples low frequency energy is into the person's body, making the entire person an EF emitter (figure 3). One or more receiver electrodes can then be placed around the person. Instead of sensing the *reduction* in current $\mathbf{I_R}$ received when the human enters the field as in Human Shunt Mode, we now measure the *increase* in the current $\mathbf{I_R}$ received as the person moves closer to the receiver. At close proximity, the person and the receiver electrode behave as ideal flat plates, where displacement current varies with the reciprocal of distance. At large distances, the person and the receiver electrode behave like points, where displacement current varies with the reciprocal of distance squared.

The transmit method provides large receive signals, operates over large areas, and can distinguish between multiple persons by sequentially selecting each person as an EF emitter, or by having each person emit a different frequency. However, capacitively coupling energy into a person requires continuous close contact with the person (in contrast to shunt mode, which does not require this). We have used transmit electrodes ranging from 5 to 150 square cm, depending on proximity to the person. The transmit electrode can be incorporated into the seat of a chair, a section of a floor, the back of a palmtop computer, or a wristwatch band. Direct conductive contact with the person's skin requires a much smaller electrode area ($<$ 5 square cm).

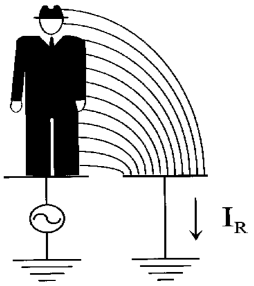

Figure 3: In what we call "Human Transmitter Mode", energy from an oscillator is coupled into a person near or in contact with the transmit electrode, making the person an electric field emitter. As the person moves any body part closer to the grounded receive electrode, the displacement current into the receiver $\mathbf{I_R}$ increases.

2.4 Mode Crossover

In practice, when we are sensing in Human Shunt Mode, some amount of Human Transmit Mode is also occurring. When a hand (or any object that is large relative to the dipole length) approaches the dipole field shown in figure 2, the displacement current $\mathbf{I_R}$ decreases. This is because the amount of field being shunted away by the hand far exceeds the amount that it is re-transmitting to the receiver. However, when the hand gets very close (much less than a dipole spacing) the displacement current $\mathbf{I_R}$ begins to increase again. This is because the amount of signal being coupled into the hand from the transmitter and re-transmitted from the hand to the (very near) receiver now exceeeds the amount being shunted away. This is the effect that led us to discover "Human Transmit Mode" as a separate form of sensing. There are simple practical ways of reducing the effect (eg making the electrode spacing large relative to the size of the hand), and our next generation of sensing hardware (see next section) will be able to detect it directly.

2.5 Sensing with the Fish

To assist researchers in exploring EF sensing, our group has produced a small microprocessor-based EF sensing unit, supporting one transmitter and four receivers (figure 4). We call it the "Fish" after the amazing sensing abilities of weakly electric fish, and also because fish comfortably move through 3D volumes, whereas mice are typically limited to 2D surfaces. The Fish board supports MIDI, RS-232, and RS-485 serial communication protocols.

A synchronous detection circuit (figure 5) is used to detect the transmitted frequency and reject all others [HH89], acting as a very narrow band-pass filter. The displacement current can be measured with approximately 12 bits accuracy (72 dB) using the components shown in figure 5. There is a trade-off between update rate (sample rate/number of samples averaged) and accuracy (signal-to-noise ratio). The signal-to-noise ratio increases as the square root of the number of samples averaged. For example, averaging 64 samples increases the signal-to-noise a factor of eight (+18 dB), with a corresponding 1/64 update rate.

The small displacement currents we are dealing with require good shielding, but the capacitance of shielded coaxial cable is orders of magnitude greater than the capacitance between the transmit and receive electrodes. Cable capacitance effectively performs a low-pass filter on the received signal, which for typical values would limit the operating frequency to 30kHz, and introduce a phase shift that needs to be compensated for in the synchronous detector (not shown). However,

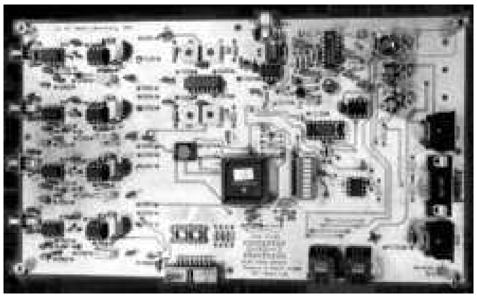

Figure 4: The Fish sensor board

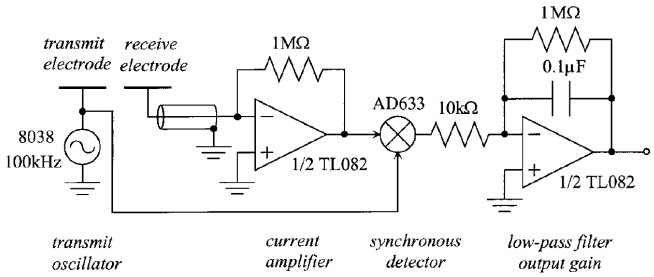

Figure 5: Synchronous detection circuitry

placing the current amplifier at the receiver electrode allows higher frequencies, limited by the amplifier's slew rate. For example, attaching the receive electrode directly to the TL082 current amplifier allows an operating frequency of 220 kHz.

2.5.1 Transmitter Power

We typically use the Fish at a frequency of around 30 to 100kHz. Electric field sensing for human-computer interfaces is generally possible in a frequency range of 10 kHz to 1MHz. Below this range, displacement currents and update rates are too small. Above this range FCC power regulations become more stringent [Com95]. At our operating frequencies, the amount of field that users are exposed to is about the equivalent that one gets from wearing stereo headphones, which is several orders of magnitude below FCC limits.

The signal wavelength at 50-80kHz is several kilometers, so the distance between our electrodes is a very small fraction that. This means that no appreciable energy is radiated. The only power consumed by the transmitter is the energy required to charge the capacitance of the transmitter electrode to the oscillating voltage. In practice the transmitter power is about a milliwatt. This allows the design of very low power systems with no radio interference.

2.5.2 Electric Field Geometry

In both transmit and shunt modes, the absolute signal strength received depends on the coupling of the person to a reference (i.e. the transmit electrode for transmit mode, or ground for shunt mode). This coupling acts as a global system gain. It is therefore the *relative* signal strength of multiple sensors that contains position information.

In Human Shunt Mode, there is the advantage that the user does not have to couple into a transmitter in order to be sensed. There is a corresponding difficulty, however, which is that the value returned by a sensor bears a more complex relationship to the position of a human hand. Each receiver measurement constrains the position of a small object (relative to dipole spacing) to an ellipsoid centered on the dipole axis. The problem of inverting the sensor readings is equivalent to the geometrical problem of finding the intersection points of these ellipsoids. Often additional constraints (prior knowledge) must be imposed to select one solution from the many symmetric cases that are consistent with the data. For example, to make a two-dimensional mouse using only two dipoles, we must impose the constraint that the hand is on one side of the dipoles. We give an overview of our approach to field geometry solving in [ZSP+95], and full details are given in [Smi95].

2.6 The Smart Fish

We are currently implementing the Smart Fish, which is a second generation EF sensing board using a digital signal processor for signal modulation and demodulation. This gives us 16 bits of signal resolution, and also allows automatic calibration and the exploration of more complex detection strategies for example using frequency-modulation chirps (as used in radar), frequency hopping, and code modulation (e.g. spread spectrum).

The Smart Fish also measures the power loading of the transmitter, which enables it to disambiguate the mode crossover described in section 2.4. The current drawn from the transmit electrode always increases as an object approaches the transmit electrode.

2.7 Other modes of Electric Field Sensing

Mathews [Mat90] developed an electronic drum that detects the 3D location of a hand-held transmitting baton relative to a planar array of antennas by using near-field signal-strength measurements. This system requires a shielded cable to the baton, and cannot detect human body position directly.

Lee, Buxton, and Smith [LBS85] measure the charging time of a resistor-capacitor network, in order to detect multiple contacts on a touch-sensitive tablet. This system only operates over the surface of the tablet.

Other types of capacitive sensors can measure proximity without contact. To assist robots to navigate and avoid injuring humans, NASA has developed a capacitive reflector sensor [VM92] that can detects objects up to 30 cm away. Their sensor uses a driven shield to push EF lines away from grounding surfaces and towards the object. Wall stud finders use differential capacitance measurement to locate wood behind plaster boards by sensing dielectric changes [Fra78]. Linear capacitive reactance sensors are used in industry to measure the proximity of grounded objects with an accuracy of 5 microns [Fos89].

In the medical domain, electrical impedance tomography places electrode arrays on the body to form images of tissue and organs based on internal electric conductivity [Web90].

2.8 Electric Field Sensing Applications

Prototyping interfaces with the Fish is basically an "arts and crafts" project, consisting of cutting out electrodes (typically aluminum foil and copper tape), taping them down, and wiring them up to the Fish board. This section notes some of the diverse applications we have developed using Electric Field Sensing.

2.8.1 The Finger Mouse (transmit mode)

We have implemented a two-dimensional finger-pointing mouse on a laptop computer (figure 6). The input device is activated by touching a small transmitter electrode with the fourth (little) finger of the left hand. Energy is coupled into the person, and the EF emitted from the pointing finger is sensed at two receiving electrodes. A thin uniform copper strip running across the top of the screen senses Y position, and a tapered strip along the side of the screen senses X position. The taper renders the electrode more sensitive to the EF emitted by the pointing finger and less sensitive to the field emitted by the arm. The shaped electrode physically implements an analog spatially varying signal gain. A third small receiving electrode, placed below the spacebar, allows the thumb of the left hand to generate click signals.

Figure 6: Two-dimensional finger-pointing mouse

The pointing finger does not need to be in contact with, or even close to the screen, thereby avoiding screen smudges and occlusion of the cursor by the pointing finger. Position sensing is easily disabled by lifting the fourth finger off the transmitting electrode, the equivalent of lifting and putting down a mouse, facilitating relative position control.

2.8.2 Person-Sensing Room (transmit mode)

In an installation piece at the MIT Media Lab, we covered the entire floor of a room with a single transmitter electrode. A person standing in the room became an EF emitter, and four receiver electrodes on the walls measured relative signal strength, indicating the location of the person. A computer program, controlling a multitude of synthesizers and sound sources, created a complex sonic terrain based on the location of the person, allowing navigation of a sonic environment.

2.8.3 Magic Chair (transmit mode)

We built a chair for the magicians Penn & Teller with a transmitter in the seat. This makes the body into an EF emitter enables them to perform music by waving their arms near four receivers mounted on poles in front of the chair. Hand position controls various sound parameters produced by computer-controlled sound synthesizers. [PG95]

2.8.4 A Simple Pressure Sensor (shunt mode)

Since electric fields pass through non-conductors, passive materials that apply force and viscous friction may be incorporated into EF sensing devices, providing haptic feedback. We have constructed a pressure pad and a viscous 3D workspace based on this principle.

We placed an ordinary foam pad on top of a transmit/receive dipole pair. Pressing on the foam produces force feedback. But since force is proportional to position (Hooke's law), and we can measure finger position by EF sensing, we can therefore measure finger force. A passive piece of foam on an EF sensor then becomes a "pressure sensor".

2.8.5 Haptic Feedback in 3D Space (shunt mode)

We fitted a plastic box with electrodes on three sides to measure hand position in 3D. The box was filled with bird-seed (millet) to provide a viscous medium for haptic feedback. The seed allows users to rest their hand in space, reducing fatigue, and provides something to grab. Slight compression of the seed increases viscosity. Perhaps a computer-controlled piston, bearing on a movable wall of the box, could provide a simple way to simulate an environment with variable viscosity.

2.8.6 "Smart Furniture" (shunt mode)

We have experimented with a variety of smart furniture. One example is a smart table which has a co-linear dipole pair (i.e. receiver-transmitter-receiver) placed underneath a wooden table to measure hand gestures. A computer screen displays an electronic newspaper whose pages can be flipped forward and backward by sweeps of a hand across the table (X-axis). Placing the hand down on the table (Z-axis) advances to the next section, lifting the hand up displays the previous section. Gestures are detected by applying a threshold to the X and Z velocities. Position in the X-axis is approximated by differencing the two receiver signals; position in the Z-axis is approximated by the sum of the receiver signals.

2.8.7 Personal Area Networks

Finally, in an application which does not use the Fish hardware, but employs a form of wearable EF field sensor, we have demonstrated that we can code information into a modulated EF

Figure 7: Using the Flying Fish to navigate in 3D. The "Wings" around the monitor contain four receiver electrodes for the four quadrants, which are shaped to enhance their sensitivity to gestures at the corners of the screen. This gives good channel separation. Here the user is making a gesture that he thinks of as having a natural interpretation of "move to the left". He will expect objects on the screen to move right.

transmitter signal. This information can then be communicated to another person's EF sensor device by shaking hands or other form of close contact. This work is described fully in [Zim95].

3 The Flying Fish

The previous section has given an overview of the large range of applications in which EF sensing can be applied. In this section we give details of one particular application, in which we have developed the Flying Fish 3D navigational controller. We describe how we have experimented with a variety of hardware and software configurations, and then in section 4 we will discuss the higher-level design issues involved in making the use of such a device feel "natural".

3.1 Adding "Wings" to a Fish

Figure 7 shows a pair of Wings mounted on a workstation monitor. These are made from simple cardboard, with four separate segments of metal foil mounted behind the upper left, upper right, lower left and lower right corners. These act as receiver electrodes and their sensitivity is completely unaffected by the cardboard and paint in front of them, which are there for reasons of robustness and aesthetics. The shape of the Wings is not, however, purely aesthetic. Each electrode has a large area at the corner, tapering to small areas near the centers of the sides. Since the value measured on each channel represents a net amount of field received over the entire area of the electrode, their shape gives them a high sensitivity at the corners and lower sensitivity near the centers of the sides.

We use Human Transmitter Mode here, and the user is capacitively coupled to the Fish's transmitter electrode by sitting on a thin cushion which contains a sheet of thin, flexible metal foil connected by a shielded wire to the Fish. As we noted in section 2.3, the amount of field we can transmit from a hand to a receiver falls off with the square of the distance as as we move away from the receiver electrodes. In practice, this begins to happen as soon as the hand is more than a few inches away. In order to compensate for this effect, we have added an analog logarithmic amplifier before the analog-to-digital stage on the Fish board, as described in [PG95]. The net result is that the 8-bit reading we get for the sensor value on each channel varies in linear proportion to the distance of the hand(s) from the receiver electrode, and also has the equivalent of 12 bits of dynamic range.

By comparing the values received at the four electrodes we can derive a 3D vector representing a direction and amount of motion corresponding to an arbitrary position of one or two hands. To a first approximation, our method is simply to take the sum of the values on all channels and map these onto the forward/back component of the vector, take the difference between the two upper and two lower sensor values for the up/down component, and take the difference between the left and right pairs for the left/right component.

3.2 Virtual 3D Input Devices

It is possible to implement a number of different virtual input devices using the Flying Fish set-up just described. We first tried implementing a virtual 2D mouse in a similar style to the finger-mouse described in section 2.8.1 above. The difficulty with this was that the workstation screen functions as a large ground, and the receivers on the sides are relatively far from a finger pointing near the center of the screen. So with only 8 bits of signal resolution on each channel, getting good tracking of one hand/finger was hard. A large part of the signal being received at the electrodes was coming from the user's arm, which from the user's point of view was *not* what should be controlling the movement. However, we noticed that it was much easier to control 2D mouse movement with two hands (the sensor signals were much stronger since the hands were most comfortable at the sides of the screen). This felt as if one was holding a pointer threaded on a virtual string held between the hands, with which one could pull the mouse from side to side, or up and down.

But it was not clear to us that controlling a virtual 2D mouse with a 3D motion of two hands was something that many people would want to do. What was also striking, however, was (a) how different it *felt* to interact with the computer in this way and (b) how correspondingly frustrating it can *feel* to try to control a 3D motion with a 2D mouse.

As we discuss in the following section, we have found that the most natural use for the Flying Fish is as a controller for 3D navigation. We note here, however, that the device is "extensible" in the sense discussed in Brown et al [BBM90], since it could implement a number of different virtual devices like a virtual 3D tablet or virtual 3D mouse by using different functional mappings from sensor values to control values.

The Flying Fish uses hand gestures in combination with foot pedals, and we have experimented with a variety of configurations. Initially we used two simple footswitches, with which we could select four different navigation modes. The current set-up uses a single pedal, which contains a simple variable potentiometer. We use this like a car's gas pedal. When the pedal is not depressed, no motion occurs, and as the pedal is depressed more and more, the gain of the input device increases. Thus in combination with the Wings we have a proportional 3D gestural input device that controls motion with variable granularity.

4 "Natural" Gestural Interactions

It is common for designers of systems that use gestural input to motivate their work by the claim that gestural interfaces are "natural" [SZ94]. It is often assumed that the logical interpretation of gestural inputs is straightforward, and the key research issue is the technology for gestural recognition. For example, Baudel et al ([BBL93]) remark that:

> "Navigational tasks can be easily associated with gestural commands. For example, the hand should move upward for a "move up" command. ... Drawing or editing tasks also have natural gestures associated with them (e.g., select, draw a circle, move this here)"

We argue here that the what constitutes a "natural" gestural interaction is itself a complex issue worthy of detailed research. Furthermore, as Buxton has pointed out [Bux90], the mere fact of a language being "natural" does not mean that it is therefore universally accessible: non-native speakers of Chinese will have to spend considerable time learning that language, despite its naturalness for native speakers.

We suggest in the following subsections that the feeling of naturalness that a user has in interacting with a computer is constrained by, if not fully determined by, the application task; the physical properties of the input mechanism being used; the formal representation of the object space; the functional mappings from the input device to movements in the domain; metaphors both visible and invisible; and (last but not least) the user. In each case we discuss how this applies to our experience with the Flying Fish.

4.1 Application Task

We do not believe that most users would find it natural to perform a C++ programming task using (only) a 3D gestural command language. The task we have chosen to address is that of navigation through 3D virtual spaces. Here there is a mapping from continuous 3D hand positions to continuous motions that is relatively straightforward (but see 4.5 below).

When we are navigating through a through a large multi-dimensional space, the important issue for the user is not how to manipulate objects quickly and accurately. The challenge for interface design this kind of space is to find ways of helping the user not to get lost; to convey a sense of place, and of how the small part of the space displayed on the screen relates to the huge part of the space which is not displayed.

4.2 Physical Properties of the Input Mechanism

The fact that the Wings are mounted on the sides of the screen mean that there is a feeling of directness in the interaction; the position of one's hands can feel part of the virtual space itself, without the hands obscuring the user's view of the screen.

It also means that it is possible to feel as if one is holding a virtual object between one's hands to manipulate it, as illustrated in figure 9. This illusion would be difficult if not impossible to achieve with a mouse, joystick, spaceball etc.

4.3 Formal Representation of the Object Space

As reported in [ARS95], we have used the Flying Fish to move through several different types of virtual space. The Financial Viewpoints space [Str95] is formally a Euclidean 3D space, but it contains a single virtual structure that either rotates around its own center or else moves closer or father away in response to input controls. It often feels as if one is holding this virtual object between one's hands. In contrast, the Galaxy of News space [Ren94] is non-Euclidean; it has

Figure 8: Here the user makes a gesture which is identical to that of figure 7 from the point of view of the physics of the input device, but here he is thinking of moving the handlebars of a virtual bicycle to the right, and will expect objects on the screen to move left.

Figure 9: Here is a third "natural" interpretation of the same gesture; if the user sees a single virtual object on the screen which rotates around its own center, it will feel natural to think of this gesture as grabbing the virtual object and rotating it clockwise around the vertical axis.

depth, but with an orthogonal rather than perspective projection, and it dynamically reconfigures itself as the user moves through it. We never experience a physical space that is like this, but we can still use intuitive gestures to indicate directions in which we want to move through the space. Finally we have implemented a Euclidean 3D "virtual world" with perspective projection. Here it feels natural to imagine oneself as moving within the space, either in one's own body or else in some kind of virtual vehicle. It can also feel natural to imagine oneself as viewing the world through a virtual camera which one is manipulating with one's hands.

4.4 Functional Mappings

Two functional types of input sensing are common: (i) "isotonic" or "moving control" (as in a Dataglove), and (ii) "isometric" or "pressure control" (as in a Spaceball). Zhai and Milgram [ZM93] have demonstrated clearly that one cannot compare "isometric" sensing with "isotonic" sensing modes without also considering interactions with the mapping functions. Both sensing modes can be combined with either (a) position mapping, where the input sensor controls the position of a manipulated object, or (b) rate mapping, where the sensor controls the velocity of the manipulated object's motion. Zhai and Milgram's experiments used a fixed control gain which they had found to be optimum for practiced users in their experimental task. They also noted that there was a large effect of control gain on performance at the task.

The Flying Fish implements isotonic rate control, with a control gain that is varied by the foot pedal. It is therefore unclear whether Zhai and Milgram's results would apply to our device. As we noted in section 3.2, the physics of the Flying Fish allow us to implement a number of different functional mappings. We chose to implement a virtual 3D joystick (i.e. rate mapping) for navigation, because we assume that the part of the space that is visible to the user at any one time may be small compared to the size of the whole virtual world. A virtual 3D mouse would therefore be tedious, because of the constant "nulling problem" [Bux86].

But even once we have decided that, for example, the left/right *position* derived from sensor values should control *rate* of left/right motion, it is still a non-trivial question to decide which *way* it should control that motion, as we shall show next.

4.5 Visible and Invisible Metaphors

Some years ago, Alan Kay remarked that:

> "One of the most compelling snares is the use of the term *metaphor* to describe a correspondence between what the users see on the screen and how they should think about what they are manipulating." [Kay90]

Although Kay's view is acknowledged in theoretical discussions of metaphor (see e.g. [Mad94]), it is still the case that most practitioners of metaphor design for computer interfaces concern themselves primarily with our perception of the interface, and the way in which real-world analogues are suggested by visually perceptible features.

However, since the Flying Fish is used with relatively large-scale hand gestures, our sense of *proprioception* also plays a large role in our experience of the human-computer interaction. Everyone has a strong internal sense of what it feels like to make certain movements of our limbs, and we are accustomed to associating these proprioceptions with the motion of our bodies through the world. The Flying Fish interface can take advantage of this fact to suggest what we call "invisible metaphors", which arise out of the experience of the interaction itself, and are a result of the way that certain kinds of gestural navigation *feel*.

In a navigational interaction, where we do not see one object moving with respect to a background of other fixed objects, but rather we see all objects moving, there are several possible physical models of what could be moving:

- the object(s)
- the self
- a vehicle mediating between self and objects - here there may be an additional level of indirection, as when I *rotate* a car's steering wheel anti-clockwise, the car *translates* towards the left, and objects in the world translate rightwards. Different types of vehicle implement different mappings between their controllers and motion in the world
- a viewer mediating between self and objects - again there may be an additional level of indirection, as when we move the end of a camera or telescope mounted on a tripod up, and the objects in the world appear to move *up*, versus raising hand-held binoculars, where objects in the field of vision move *down*

Invisible metaphors are implemented primarily in the mind of the user, and may not even be conscious. The hand-drawn annotations on Figures 7 and 8 illustrate two different interpretations of the same "natural" gesture, but the feeling of naturalness of the movement of objects on the screen is governed by which model of motion the user is implicitly assuming. There is of course a visible difference between the "natural invisible metaphors" of figs 7 and 8, in that the objects

Figure 10: Here the user is thinking of the upward gesture of his hands as meaning that he wants to move upwards; he will expect objects on the screen to move down.

Figure 11: Here the user is imagining that he is viewing the world through a virtual camera mounted on a tripod. When he moves the camera handles up, he will expect objects on the screen to move up.

move in opposite directions, but we do not need provide any explicit visual representation of virtual handlebars, for example, in order to suggest the metaphor of a virtual vehicle.

Figure 9 shows what we believe is *direct manipulation* in the truest sense. In sharp contrast to the "desktop metaphor" sense of direct manipulation, here there is no visible icon acting as a metaphorical representation of an object, which a user must manipulate in order to manipulate the object itself. All the user sees is the object in on the screen, and his movements directly control the motion of the object. However, as we have seen above, the physical properties of the Flying Fish, the formal representation of the object space (in this case Euclidean with polar coordinates) and the functional mapping used (left/right sensing to clockwise/anticlockwise rotation) combine to give a very strong impression of directly grasping the object between the two hands. This is illusion and magic in exactly the sense called for by Kay [Kay90].

Again contrary to the straightforward view of Baudel et al quoted above (4), figures 10 and 11 show that there is more than one way of thinking about the meaning of an upward gesture. We have not added another figure, but the reader will be easily able to see by analogy with figure 9, that it may also be very natural to think of this gesture as rotating an object up and forward around a horizontal axis parallel to the screen.

4.6 The User

If a particular user is accustomed to flying an aeroplane, or doing a lot of scuba-diving, or riding a bicycle every day, they may have a strong sense of what feels like a "natural" way to move in 3D space that is not shared by the majority of users. It is our belief that this individual diversity should be supported in interfaces: if there is no functional need for standardisation, none should be imposed. In our experimental applications, we have allowed the mappings of gestures into navigational movements to be customisable by the user, so that movement in each dimension can be reversed if the user feels that it is "backwards".

The differences between different user's conceptual models of navigational models may not be explicit, and since our navigational metaphors may be implemented in the subconcious mind of the user, a single user may entertain multiple different possible models for the same hardware and software settings. For example, some users may experience the difference between the functional mappings illustrated in figs 7 and 8 as the difference between *translating* one's own body left and *rotating* one's own body right. Others may think of the rotation in fig 8 as moving a vehicle controller, as we have shown. But a distinction between thinking of translating or rotating ones own body makes no difference in the vertical dimension: upward translation or rotation of the body both lead one to expect objects on the screen to move down.

However, we do not wish to suggest that there are no generalisations to be made about what users experience as being natural gestural navigation. All the factors we have outlined above have an influence, and there are some behaviours in the virtual world that strongly discourage the user from experiencing particular navigational metaphors. For example, if in response to input gestures a user sees a single object move in the virtual world with respect to a fixed background, it will be very hard to experience this as a movement of the self or a vehicle, without in some metaphorical way *identifying* with the object that is moving, and it would be difficult if not impossible to experience this as a movement of a virtual viewer.

5 Comparison with Other Approaches

We compared our work with other EF sensing research in section 2.7. There is a large body of research on other forms of remote sensing. Mechanisms used include acoustic, optoelectronic, microwave, ultrasonic, video, laser, and triboelectric (detecting static electric charge) [Fra93]. Many of these mechanisms have been adapted to measure the location of body parts in three dimensions, motivated by military cockpit and virtual reality (VR) applications [MA92]. However, they all have their problems.

Acoustic methods are line-of-sight and are affected by echoes, multi-paths, air currents, temperature, and humidity. Systems based on reflection are affected by surface texture, reflectivity, and incidence angle of the detected object. Microwaves pose potential health and regulation problems. Simple pyroelectric systems have very slow response times (> 100 msec) and can only respond to changing signals. Lasers must be scanned, can cause eye damage, and are line-of-sight. Triboelectric sensing requires the detected object to be electrically charged. Infrared systems require significant power to cover large areas.

Optical systems are also line-of-sight, require controlled lighting, are saturated by bright lights, and can be confused by shadows. Video has a slow update rate (e.g. 60 Hz) and produces copious amounts of data (on the order of 75 megabits per second) that must be acquired, stored, and processed.

Electric field sensors do not have these drawbacks. The effect of EF sensing does not depend on object surface texture and reflectivity. The data from EF sensors is continuous with resolution limited only by transmission strength and noise rejection. There is an economy of data; only three channels are required to locate a hand of known size in 3D. An EF system operating at 100 kHz can average 100 samples, provide a 1 kHz update rate, with 1 millisecond lag time.

Electric field systems can be extremely small, light-weight and low power, as required by the ever shrinking real estate and energy capacity of lap, palm, and watch based computers. Since electric fields penetrate non-conductors, sensors can be hidden, providing protection from weather and wear, as well as adding an element of magic to the interface.

However, EF sensors detect a bulk effect, integrating the amount flux intercepted by (in shunt mode) or transmitted from (in transmit mode) the entire body. Therefore care is needed in the design of input devices which are intended to measure isolated parts of the body, such as individual hands or fingers, so that the receivers will be maximally responsive to movements of these, and the effect of the rest of the body can be offset by cross-referencing between sensor channels.

6 Future Work

6.1 The Flying Fish

We noted in section 2.6 that work is already underway in implementing the second generation EF sensing board called the "Smart Fish". This will have 8 channels of sensing with improved (16-bit) resolution. We plan to use this to do away with the need for the user to sit on a transmitter cushion. The capacitive coupling of the transmitter signal into the user will be done through the foot-pedal, and we will be able to compensate for the greater variation in coupling to the user (due to different shoes etc) by using the additional channels for cross-referencing. The DSP and greater dynamic range of the analogue-to-digital conversion components will also do away with the need for the logarithmic amplifier on the front end.

We have had plenty of informal experience of people visiting the lab and trying the Flying Fish, and this has been very instructive in showing the variety of ways in which users can experience 3D navigation in virtual spaces as being natural. However, it will be interesting to see some more formal studies of the device in action, particularly making quantitive measures of the performance of this type of isotonic rate controller with variable gain for navigation tasks.

We have so far integrated the Flying Fish with information spaces and an abstract virtual world. However, the interface is eminently suitable for computer games that involve 3D motion, and we believe that there is also much to be learned from using the Flying Fish in this context (it's also fun).

6.2 Other Applications of EF Sensing

We noted some of our existing EF sensing applications in section 2.8 above. Others are described in [GZA95] and [ZSP+95]. Several of these are clearly rich areas for further development, such as the portable information communications devices of PAN [Zim95]. Before concluding, let us also note a few others.

The Tailor project [P+92] allows disabled individuals to run computer applications by mapping the unique anatomical movement ability of each individual to control signals. Combining EF sensing with such mapping techniques could provide a person in a wheelchair with individually tailored, unobtrusive, invisible, low-power, and low-cost computer and machine interfaces.

Hermetically sealed EF sensors in a palm top could determine when the case is open, when the unit is being held, and could create a large control space around the small device. Foam EF buttons could provide force and tactile feedback, detect finger approach and finger pressure, and distinguish between slow and fast presses.

Multiple transmitters and receivers, multiplexed in time, frequency, or by coding sequence, could be placed under a carpet to determine the number and location of people in a room. When an electrode under a person is activated, that person becomes the EF source. Smart floors can be used for multi-participant VR simulations without the burden of wires or the complexities of video cameras.

Researchers are currently exploring direct manipulation of instrumented real objects to facilitate 3D orientation and manipulation [H+94, MN94]. Electric field sensors could be incorporated into objects to measure object deformation, position, and orientation.

Attempts have been made to instrument whiteboards using video cameras [IK92] and optoelectronics [E+92]. Both systems require rear imaging to record stylus movement. A conventional plastic whiteboard could be fitted with an array of EF sensing electrodes to measure the location of a metal-cased marker in the hand of a shunting or transmitting person.

An array of dipoles can turn a table into a multidimensional digitizing and gesture input device. Such an EF sensing matrix may substitute for or augment a video camera for video desk applications [NW92]. Perhaps visual ambiguities and occlusions could be arbitrated by EF sensing, indicating hand location to the video analysis system.

As EF sensing devices begin to proliferate, there may be a multitude of small EF sensing devices scattered about in a single room. The jamming avoidance mechanism of weakly electric fish [Bul82] suggests that such devices can communicate with each other and adjust their transmission frequencies autonomously when new devices are introduced into the sensing space.

A key research issue remains how to best isolate the position of an individual body part given the bulk nature of EF sensing and the non-uniform nature of dipole fields. This is a currently more of a pragmatic craft than a science, but we have found a number of specific solutions, and we are actively pursuing both theoretical and empirical approaches to finding a general solution to this problem [Smi95].

Conclusions

The abilities of weakly electric fish to perform object detection, communicating, and jamming avoidance demonstrate what is possible with EF sensing. We have described the basic mechanisms of EF sensing, which allow the same detection scheme to be scaled from microns to meters simply by scaling electrodes and their spacing. EF sensing provides high resolution proximity information. The authors know of no other sensing mechanism or system that can deliver non-contact sensing with millimeter resolution at kilohertz sample rates and millisecond lag times for a few dollars a channel.

We have also described a number of applications of EF sensing using our Fish hardware, and given an in-depth discussion of one case; using the Flying Fish to navigate in 3D virtual spaces. We believe that navigation in large multidimensional spaces will be increasingly important, especially now that 3D browsers for the World-Wide Web are becoming available. Our experience with the Flying Fish shows that EF sensing can provide a very intuitive and natural gestural interface that is fun and magical to interact with.

We have argued against the notion that there should be a set of “standard user interface guidelines” for all users and applications, and have detailed the many factors that contribute to a feeling of “naturalness” in an interaction. We have found that individuals differ strongly as to which interpretations of gestural inputs they prefer. But we do not wish to suggest that "naturalness is in the eye of the beholder"; we would rather emphasize invisible metaphors and proprioception over perception, and say that for gestural interactions, "naturalness is in the body of the beholder".

Acknowledgments

The authors would like to thank Henry Chong and Barrett Elwood for their programming and hardware assistance, and the team of Teresa Marrin, Pete Rice, Edward Hammond, John Crouch, Ryan Christensen, and Alexander Sherstinsky for their work on the person-sensing room. We would also like to thank the Hewlett-Packard Corporation, Festo Didactic Corporation and the "News in the Future" consortium for their support.

References

[ARS95] David Allport, Earl Rennison, and Lisa Strausfeld. Issues of gestural navigation in abstract information spaces. In *CHI 95 Human Factors in Computing Systems*, pages 206–207, Denver, Co, 1995. ACM Press.

[BBL93] T Baudel and M Beadouin-Lafon. Charade: Remote control of objects using free-hand gestures. *Communications of the ACM*, 36(7):28–35, July 1993.

[BBM90] E Brown, W Buxton, and K Murtagh. Windows on tablets as a means of achieving virtual input devices. In D Diaper et al., editors, *Human-Computer Interaction - INTERACT '90*, pages 675–681, North-Holland, 1990. Elsevier Science Publishers B.V.

[Bul82] T H Bullock. Electroreception. *Ann. Rev. Neuroscience*, 5:121–170, 1982.

[Bux86] William Buxton. There's more to interaction than meets the eye: Some issues in manual input. In D A Norman and S W Draper, editors, *User-centered Design: New Perspectives on Human-Computer Interaction*. Lawrence Erlbaum Associates, Hillsdale, N.J, 1986.

[Bux90] William Buxton. The "natural" language of interaction: A perspective on nonverbal dialogues. In Brena Laurel, editor, *The Art of Human-Computer Interaction*, pages 405–416. Addison-Wesley, Reading, Ma, 1990.

[Com95] Federal Communications Commission. Part 15 radio frequency devices. Technical report, FCC, Washington, DC, 1995.

[E+92] S Elrod et al. Liveboard: A large interactive display supporting group meetings, presentations and remote collaboration. In *CHI 92 Human Factors in Computing Systems*, pages 599–607, Monterey, May 1992. ACM Press.

[Fos89] R Foster. Linear capacitive reactance sensors for industrial applications. In *40th Annual Earthmoving Industry Conference, Peoria, Ill*, April 1989. SAE Technical Paper 890974.

[Fra78] F Franklin, R Fuller. Electronic wall stud sensor. US Patent No 4,099,118, July 4, 1978.

[Fra93] J Fraden. *AIP Handbook of Modern Sensors*. American Institute of Physics, New York, 1993. pp 312-346.

[Gar67] L Garner. For that different sound, music a la theremin. *Popular Electronics*, pages 29–33,102–103, November 1967.

[Ger93] Neil Gershenfeld. Method and apparatus for electromagnetic non-contact position measurement with respect to one or more axes. US Patent No 5,247,261, Sept 21, 1993.

[GZA95] Neil Gershenfeld, Thomas G Zimmerman, and David Allport. Non-contact system for sensing and signalling by externally induced intra-body currents. US Patent Application, May 8, 1995.

[H+94] K Hinckley et al. Passive real-world interface props for visualization. In *CHI 94 Human Factors in Computing Systems*, pages 452–458, Boston Ma, April 1994. ACM Press.

[HH89] P Horowitz and W Hill. *The Art of Electronics*. Cambridge University Press, New York, 1989. page 889.

[IK92] H Ishii and M Kobayashi. Clearboard: A seamless medium for shared drawing and conversation with eye contact. In *CHI 92 Human Factors in Computing Systems*, pages 525–540, Monterey, May 1992. ACM Press.

[Kay90] Alan Kay. User interface: A personal view. In Brena Laurel, editor, *The Art of Human-Computer Interaction*, pages 191–207. Addison-Wesley, Reading, Ma, 1990.

[LBS85] S Lee, W Buxton, and K Smith. A multi-touch three dimensional touch-sensitive tablet. In *CHI 85 Human Factors in Computing Systems*, pages 21–24. ACM Press, 1985.

[MA92] K Meyer and H Applewhite. A survey of position trackers. *Presence*, 1(2):173–200, Spring 1992.

[Mad94] Kim Halskov Madsen. A guide to metaphorical design. *Communications of the ACM*, 37(12):57–62, December 1994.

[Mat90] M Mathews. Three dimensional baton and gesture sensor. U.S. Patent No. 4,980,519, Dec 25, 1990.

[MN94] T Murakami and N Nakajima. Direct and intuitive input device for 3-d shape deformation. In *CHI 94 Human Factors in Computing Systems*, pages 465–470, Boston Ma, April 1994. ACM Press.

[NW92] W Newman and P Wellner. A desk supporting computer-based interaction with paper documents. In *CHI 92 Human Factors in Computing Systems*, pages 587–598, Monterey, May 1992. ACM Press.

[P+92] P Pausch et al. One-dimensional motion tailoring for the disabled: A user study. In *CHI 92 Human Factors in Computing Systems*, pages 405–411, Monterey, May 1992. ACM Press.

[PG95] Joseph A Paradiso and Neil Gershenfeld. Musical applications of electric field sensing. article submitted to Computer Music Journal, June 1995.

[Ren94] Earl Rennison. Galaxy of news: An approach to visualizing and understanding expansive news landscapes. In *Proceedings of UIST 94*, November 1994.

[Smi95] Joshua R Smith. Toward electric field imaging. Master's thesis, MIT Media Lab, June 1995.

[Str95] Lisa Strausfeld. Financial viewpoints: Using point-of-view to enable understanding of information. In *CHI 95 Human Factors in Computing Systems*, pages 208–209, Denver, Co, 1995. ACM Press.

[SZ94] David J Sturman and David Zeltzer. A survey of glove-based input. *IEEE Computer Graphics & Applications*, pages 30–39, January 1994.

[VM92] J Vranish and R McConnell. Driven shielding capacitive proximity sensor. U.S. Patent No.5,166,679, Nov. 24, 1992.

[Web90] J G Webster. *Electrical Impedance Tomography.* Adam Hilger, Bristol, 1990.

[Zim95] Thomas G Zimmerman. Personal Area Networks (PAN): Near-Field Intra-Body Communication. Master's thesis, MIT Media Lab, June 1995.

[ZM93] Shumin Zhai and Paul Milgram. Human performance evaluation of manipulation schemes in virtual environments. In *Proceedings of the IEEE Virtual Reality Annual Symposium*, Seattle, Wa, Sept 1993. IEEE.

[ZSP+95] Thomas G Zimmerman, Joshua R Smith, Joseph A Paradiso, David Allport, and Neil Gershenfeld. Applying electric field sensing to human-computer interfaces. In *CHI 95 Human Factors in Computing Systems*, pages 280–287, Denver, Co, 1995. ACM Press.