# Electric Field Sensing and the "Flying Fish"

David Allport, Thomas G. Zimmerman, Joseph 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, showing the different ways in which multiple EF sensors can be used to track the position of human bodies, limbs or hands, depending upon the size, shape and placement of the sensors. We give a detailed description of the implementation of the Flying Fish, which provides an interface for navigating in 3D virtual spaces just by using hand gestures, without the need for wearing or holding anything. We also discuss ways in which this type of device can be made to feel natural for 3D navigation, introducing the idea of "invisible metaphors" to describe the different ways that we can experience navigational interactions in 3D virtual worlds when the user interdace contains no graphical display of an analogue of the control device.

## Keywords

user interface, input device, non-contact sensing, electric field, gestural navigation, interface metaphor, 3D virtual worlds, natural interaction.

## 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 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 give an indication of the breadth of applications of EF sensing, some of which we have mentioned in previous work [ZSP+95], [ARS95]. Section 3 here gives a much more 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 presence of a user's hands in front of the screen and uses this to control 3 dimensions of movement simultaneously.

Section 4 here examines another novel aspect of our work; the design issues involved in making the use of free-hand gestures for 3D navigation feel natural. 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]). We analyse more closely what this means. In particular, we introduce the notion of "invisible metaphors", and discuss their role in creating an intuitive conceptual model of gestural navigation in virtual space, since there is no unique "natural" interpretation for a given gesture.

# 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.

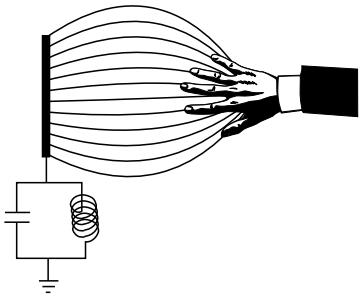

Figure 1: In the Theremin, each inductor/capacitor resonant tank circuit emits an electric field which is intercepted by the hand. The resonant frequency of the circuit changes as the hand moves in and out of the field.

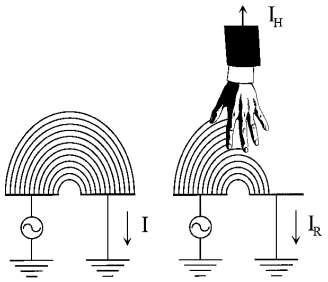


Figure 2: In what we call "Human Shunt Mode", an electric dipole field created between an oscillating transmit electrode and virtual ground receiver electrode is intercepted by a hand. Displacement current to ground $\mathbf{I_R}$ decreases as the hand moves further into the dipole field.

## 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 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).

## 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

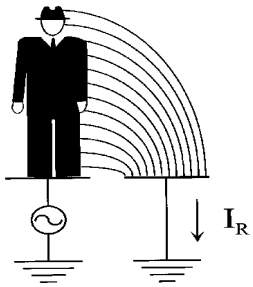


Figure 3: In what we call "Human Transmitter Mode", energy from an oscillator is coupled into a person near or in contact with the transmit electrode, making the person an electric field emitter. As the person moves any body part closer to the grounded receive electrode, the displacement current into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

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 2.6) 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. 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 4) 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 4. 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.

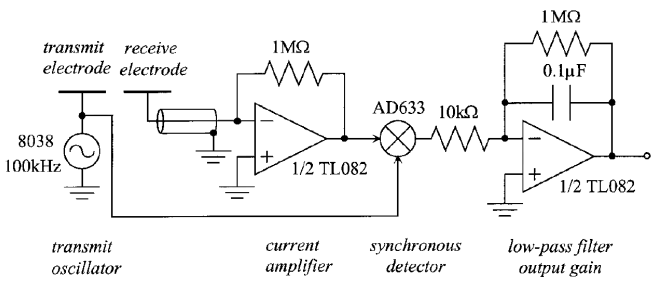


Figure 4: Synchronous detection circuitry

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, 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 of 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 [FF78]. 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 EF Sensing Compared with Other Approaches

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 a relative economy of data;

only three channels are required to locate an object of fixed shape, size and orientation in 3D (approximately 30K bits per second). 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.

The only problem in developing accurate EF sensing devices for measuring the positions of human body parts is that hands, arms etc. are not of fixed shape, size and orientation. 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.

We can easily derive 3D change measurements from just four sensor channels, telling us that a person is nearer or further than before, more to the left or right, etc., and the physics of EF fields allows us to detect hand movements as small as a millimeter within a volume of several cubic feet. However, the absolute values from the sensor channels must be calibrated for a particular sensor configuration in order to derive absolute positional information. Even given auto-calibration techniques, in practice it may require at least as many as ten channels of EF information to derive position, shape and orientation information with millimeter accuracy for the hands of any user, and the full theory of how to combine information from multiple channels to provide EF imaging is the subject of ongoing research [Smi95].

## 2.9 Electric Field Sensing Applications

Fortunately, not all applications require complete body imaging with absolute position measurement, and four sensor channels are enough to be able to build a wide variety of simple, cheap and useful input devices. 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 EF sensing.

### 2.9.1 The Finger Mouse (transmit mode)

We have implemented a two-dimensional finger-pointing mouse on a laptop computer (figure 5). 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.

The pointing finger does not need to be in contact with, or even close to the screen, thereby

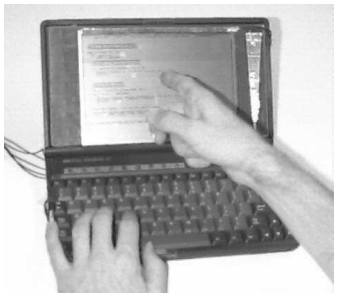

Figure 5: Two-dimensional finger-pointing mouse

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.

#### 2.9.2 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.

#### 2.9.3 Magic Chair (transmit mode)

We built a chair for the magicians Penn & Teller with a transmitter in the seat. This 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]

#### 2.9.4 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".

### 2.9.5 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.

### 2.9.6 "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.

### 2.9.7 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 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].

### 2.9.8 Future EF Sensing Applications

We have outlined a number of other potential EF sensing applications elsewhere [GZA95], [ZSP+95]. 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 a way that such devices can communicate with each other and adjust their transmission frequencies autonomously when new devices are introduced into the sensing space.

# 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 6 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 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. By adding an analog logarithmic amplifier before the analog-to-digital stage on the Fish board, as described in [PG95], we obtain the equivalent of a 12-bit reading on each sensor channel that varies in linear proportion to the distance of the hand(s) from the receiver electrode (the logarithmic amplifier will be unnecessary with the "Smart Fish" described in 2.6).

By comparing the values received at the four electrodes we can derive a 3D vector representing a direction and amount of displacement of one or two hands from a central position. 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. Note that it is not necessary for us to derive absolute positions for the hands: we simply have a very sensitive mechanism that detects small or large displacements of the hands from a "virtual center", whose absolute position will vary slightly from user to user. However, given the symmetry of human bodies, the place where the absolute values of all sensor readings are the same will have the greatest variation between users only in the forward/back direction, and it is very easy for any user to adjust to this simply by moving slightly closer to or further away from the screen.

## 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 2.9.1 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.

It was not obvious that controlling a virtual 2D mouse with a 3D motion of two hands was

Figure 6: 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.

something that anybody would want to do. What was very 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 in principle "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 (although we would require many more EF sensor channels to provide accurate enough *absolute* 3D position measurements for two hands *independently*, as required for some virtual device types).

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 Navigation

Although the Flying Fish is functionally like a 3D virtual joystick, it has no "physical affordances" other than the relationships between hand movements and functional behaviour that we design for it. In this section we examine the issues surrounding the design of mappings that feel "natural".

In principle it would be possible to design a *language* of gestures for the Flying Fish, and to have recognition algorithms that interpret (some) hand movements as valid examples of commands in our language. There would then be a question of how natural our gestural language was, and here we would need to consider issues such as the number of items in our lexicon, the ease of performing each gesture, etc. Moreover, if we were using such a gestural language to manipulate individual objects in complex worlds, there would be further questions about the ease of unambiguously mapping each recognised gesture onto a command for manipulating a unique object in the domain (see e.g. [BH92]).

However, we have not taken a gesture-recognition approach. We have used the Flying Fish purely

for 3D navigation where there is a direct correspondence between the amount of movement of the users hands (modulo use of gas pedal), and the amount of movement of *all* objects on the screen. This is unlike any previous free-hand gestural work, and it may seem prima facie that it is trivially easy to define the appropriate mapping, as Baudel et al ([BBL93]) have remarked:

> "Navigational tasks can be easily associated with gestural commands. For example, the hand should move upward for a "move up" command.

We argue in the following subsections that even in the restricted case of mapping hand positions continuously onto navigational movements, there are still a large number of design choices to be made in the mappings, and these will strongly affect how "natural" the interface feels. This is because the feeling of naturalness in an interface is constrained, if not fully determined, by many interacting factors, including 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. Essentially, our experience with the Flying Fish shows how, with different types of spatial representation, different combinations of functional mapping suggest different invisible metaphors, which in turn express different conceptual models of what the 3D navigation task is.

## 4.1 Application Task

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:

- all the objects in the world
- the position (and/or orientation) of 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 seen through the windscreen translate rightwards.
- 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 objects seen through the camera lens appear to move *up*, versus raising hand-held binoculars, where objects in the field of vision move *down*

No matter how we interpret the navigation task, the visual feedback given by the movement itself is sufficient to accomplish it successfully, *no matter what kind of input device one is using.* This is unlike the task of on-screen target-acquisition, where for some input devices (like touch-screens), it is sufficient to give feedback about when the target has been acquired, whereas for others (such as mice, joysticks etc) it is necessary to have a tracking symbol such as cursor, or some other form of feedback about the difference between one's current state and the goal.

There may be other reasons for providing additional visual cues for 3D navigation, such as a virtual compass to help prevent getting lost, or a virtual cockpit to re-inforce a flying metaphor. But these are not essential to achieve successful 3D motion, and we have chosen to explore the possibilities for navigational interfaces that do not rely upon explicit visual cues.

## 4.2 Functional Mappings

The Flying Fish implements isotonic rate control, with a control gain that is varied by the foot pedal. Results such as Zhai and Milgram's, which compare 6DOF input devices, cannot easily be applied to a device like the Flying Fish (their devices all had fixed gain, and were compared on a target-acquisition task). We have had plenty of informal experience of people visiting the Media 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, we have yet to perform quantitative performance studies of our device in action.

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 in 3D virtual worlds the part of the space that is visible to the user at any one time is often 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 in each of 3 dimensions, the *position* of the hands should control the *rate* of motion, there are still non-trivial choices to be made as to which *way* it should control that motion. These choices will constrain which of the possible physical models of 3D navigation feels like a "natural" interpretation of the interaction. There are six degrees of freedom possible in 3D navigation, and for our 3DOF Flying Fish, we can choose various combinations of translation or rotation along each axis, and in each case the mapping may be in two different directions (see figures 6-10).

## 4.3 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 the entire virtual world like an object between one's hands to manipulate it, as illustrated in figure 8. This illusion would be difficult if not impossible to achieve with a mouse, joystick, spaceball etc, although it could be done with a Polhemus or DataGlove.

When using devices like mice, joysticks etc, it is easy for us to experience our navigation as if we are in a vehicle, using the vehicle's controllers to move through the world. But with completely free-hand interfaces like the Flying Fish, it is also especially easy to feel that it is our self that is moving directly through the environment.

## 4.4 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. This can strongly suggest that one is holding it as a virtual object between one's hands.

In contrast, the Galaxy of News space [Ren94] is non-Euclidean; it has 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

Figure 7: Here the user makes a gesture which is identical to that of figure 6 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 8: 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.

intuitive gestures to indicate directions in which we want to translate (but not rotate) through the space.

Finally we have implemented a standard Euclidean 3D "virtual world" with perspective projection. Here navigation can naturally be interpreted in any of the ways listed in 4.1 above.

### 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 (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, particularly the way in which real-world analogues are suggested by graphical features. Gaver has recently given a useful analysis of how graphical computer interface metaphors such as "windows" differ from everyday metaphors; instead of highlighting similarities between existing objects they create new (graphical) objects, where the *conceptual* mapping of the interface metaphor differs from the *perceptual* mapping used to express it [Gav95].

But our experiments with the Flying Fish have revealed the possibility of an entirely different kind of interface metaphor, where there is no newly created graphical object that expresses it; the only *perceptual* expression of the metaphor is the dynamic functional mapping of hand-movements to movements of all the objects on the screen. To emphasise this absence of new graphical interface objects, we call these "invisible metaphors".

Moreover, 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 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*.

Figure 9: 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 10: 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.

We believe that the feeling of "naturalness" in using a device like the Flying Fish arises out of the correspondence between two things: (a) the conceptual metaphor for physical motion suggested by the particular functional mappings chosen, and (b) the user's conceptual model of what is happening, which may not even be conscious. For example, the hand-drawn annotations on figures 6 and 7 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. We do not need to provide any explicit visual representation of virtual handlebars to suggest the invisible metaphor of Figure 7; we only need to choose an appropriate functional mapping.

Figure 8 shows what we would call *direct manipulation* in the truest sense (because there is no graphical interface object mediating the interaction). 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 being "outside" the world in which we are navigating and grasping it like an object between our two hands. This is illusion and magic in exactly the sense called for by Kay [Kay90], since we never have an experience of navigating through the world in this way, yet it still feels natural because there is a natural correspondence between this type of interaction and everyday object manipulation.

Figures 9 and 10 also 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 8, that it may also be 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. Graphical interface metaphors can be helpful in getting users to think about a particular type of interaction in a particular way. But we believe that a large part of the fun of moving through 3D spaces is the feeling of being free to move anywhere, and since there are so many different ways of mapping a 3DOF device onto 6DOF movement, it increases the fun for the user if they can also choose to move in any *way* that they choose. Since invisible metaphors can be easily changed just by changing the functional mapping and/or the way the user thinks about the interaction, they can potentially be used to easily allow users to choose whatever navigation style that feels natural to them at the time and space they happen to be in.

## 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. 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.

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 to interact with.

We have detailed the many factors that contribute to a feeling of "naturalness" in an interaction, and have noted that individuals differ greatly 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 the possibilities for using invisible metaphors that take advantage of proprioception more than 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, the team of Teresa Marrin, Pete Rice, Edward Hammond, John Crouch, Ryan Christensen, and Alexander Sherstinsky for their work on the person-sensing room, and Lisa Strausfeld and Earl Rennison for their discussions on scale and movement in 3D representations. 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.

[BH92] Richard A Bolt and Edward Herranz. Two-handed gesture in multi-modal natural dialog. In *Proceedings of UIST 92*, pages 7–14. ACM Press, November 1992.

[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.

[FF78] R Franklin and F Fuller. Electronic wall stud sensor. US Patent No 4,099,118, July 4, 1978.

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

[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.

[Gav95] William W Gaver. Oh what a tangled web we weave: Metaphor and mapping in graphical interfaces. In *CHI 95 Human Factors in Computing Systems*, pages 270–271, Denver, Co, 1995. ACM Press.

[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.

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

[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.

[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 tomography. Master's thesis, MIT Media Lab, August 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.