Field mice: Extracting hand geometry from electric field measurements

by J. R. Smith

Several members of a family of techniques called Electric Field Sensing are described. Each sensing technique can be understood as a measurement of the value of a different component in an effective circuit diagram that summarizes all possible current pathways involving the body to be sensed and the sensing electrodes. An analytical model of the sensor response is presented, and then a probabilistic framework for inferring geometrical information from field measurements is described. The inference framework is demonstrated for the cases of a two-dimensional and a three-dimensional mouse.

Electric Field Sensing has existed in some form since a musical instrument known as the theremin was invented circa 1917.[1-3] Before that time, only aquatic animals had used electric fields to sense their environments.[4] The theremin combined analog sound synthesis and an early form of Electric Field Sensing in a single clever circuit. It is surprising how little research attention has been focused since then on the use of electric fields for measuring the human body, because electric field sensors are safe, fast (millisecond time scales), high resolution (millimeters), and inexpensive (dollars per channel), and measure robust bulk properties of the body rather than ephemeral surface properties. Nevertheless, with the exception of the work of Matthews[5] and Vranish,[6,7] little effort has been made to improve upon the early, "capacitive" form of Electric Field Sensing until recently.

There are two reasons why Electric Field Sensing has languished. The first is that at the time of its invention, very few devices were capable of being controlled electronically. The physical user interface problem of transducing voluntary human action into electronic signals had therefore not yet become important. Today, computers and other electronic devices are becoming increasingly ubiquitous, and in more and more application domains the factor that is limiting the performance of these systems is the physical human interface rather than raw processing power. Devices such as the computer mouse, the trackball, and the IBM TrackPoint* were an improvement over a simple keyboard, but are inadequate for more complex tasks, such as interacting with three-dimensional environments or animating complex articulated three-dimensional models. The DataGlove[8] is a much more expressive input device, but is a step backward from the point of view of convenience, since it must be worn and interferes with other noncomputer activity. A noncontact three-dimensional mouse, in which the hand *is* the pointing device, is an appealing compromise, since it allows richer interactions than a conventional mouse, but with lower "transaction costs" than the DataGlove, since it does not have to be put on and then removed to start and then stop using it.

In addition to the initial lack of applications for Electric Field Sensing, there was also a second, technical obstacle to its development. Since the response of the sensors is a nonlinear function of the input (for example, hand position), extracting useful information from the measurements is a difficult computational

problem and, until recently, a prohibitively difficult one. The theremin relied on the extraordinary ability of humans to learn complex mappings.[1] The performer had to devote years of practice to learn to play it, like any traditional musical instrument. A computer mouse that required years of practice to master would clearly be impractical.

In this paper, I present a noncontact three-dimensional mouse as an introduction to Electric Field Sensing and as an illustration of mathematical techniques that may be used to extract information about a matter distribution from measurements of an electric field. In the first section of the paper, I introduce Electric Field Sensing, situating the recent work of the Physics and Media Group of the MIT Media Laboratory in relation to earlier so-called capacitive-sensing techniques. In the second section, I present a quantitative forward model of the response of a sensor to a small grounded object. The third section shows how to use this forward model in conjunction with a probabilistic framework to solve the inverse problem of inferring the position of an object from sensor values. The final section describes a field mouse, a prototype noncontact three-dimensional input device whose only moving part is the user's body.

Electric Field Sensing refers to a family of noncontact measurements of the human body made with slowly varying electric fields.

Physics of Electric Field Sensing

We will use the term Electric Field Sensing to refer to a family of noncontact measurements of the human body that may be made with slowly varying (approximately 50 kHz) electric fields. Previously, several of these measurements had been lumped together under the accurate but imprecise rubric of "capacitive sensing." Close attention to the current transport pathways reveals that what has been called capacitive sensing is actually comprised of several distinct mechanisms. Furthermore, two transport pathways have been overlooked. This has led the Physics and Media Group to two new forms of Electric Field Sensing.[9-16] For reasons we will explain later, we refer to the previously known technique as a "loading mode" measurement. The new techniques involve previously unexplored "shunt mode" and "transmit mode" measurements.

In shunt mode, which is the version of Electric Field Sensing on which we will focus in this paper, a voltage oscillating at low frequency is applied to a transmit electrode, and the displacement current induced at a receive electrode is measured with a current amp. The received displacement current may be modified by the body being sensed, which need not be in contact with either electrode. The basic arrangement is shown in Figure 1.

In the figure, T is the transmit electrode, R is the receive electrode, H is the hand, and ε is the dielectric constant of the medium in which the hand is moving. R_i and C_i are the internal resistance and capacitance of the body. Each capacitor represents the total flow of displacement current from one of its terminals directly to the other. The capacitors are variable: their values depend on the geometry. In particular, C_0, C_1, and C_2 depend on the hand position.

Later in this section, we explain the physics of Electric Field Sensing in terms of an effective circuit diagram whose components correspond to current transport pathways. Both existing and new Electric Field Sensing techniques can be understood as measurements of component values in this diagram. But before explaining the effective circuit model, we describe the actual sensing circuitry developed by the Physics and Media Group.

Hardware. The Physics and Media Group has developed an evaluation board, "The Fish,"[17] for experimenting with Electric Field Sensing. A more capable, almost all-digital successor, "The Smart Fish," is currently under development.[18] The Fish consists of a transmitter that can be tuned from 20 kilohertz (kHz) to 100 kHz and four receive channels that use synchronous detection, a measurement technique described below. The transmitter consists of an oscillator connected to an op-amp. The op-amp defines the voltage on the transmit electrode, as specified by the oscillator, by putting out as much current as required to maintain the correct voltage. The amount of power that the user is exposed to is on the same order of magnitude as that received from a pair of stereo headphones, and is several orders of magnitude below the amount permitted by FCC regulations.

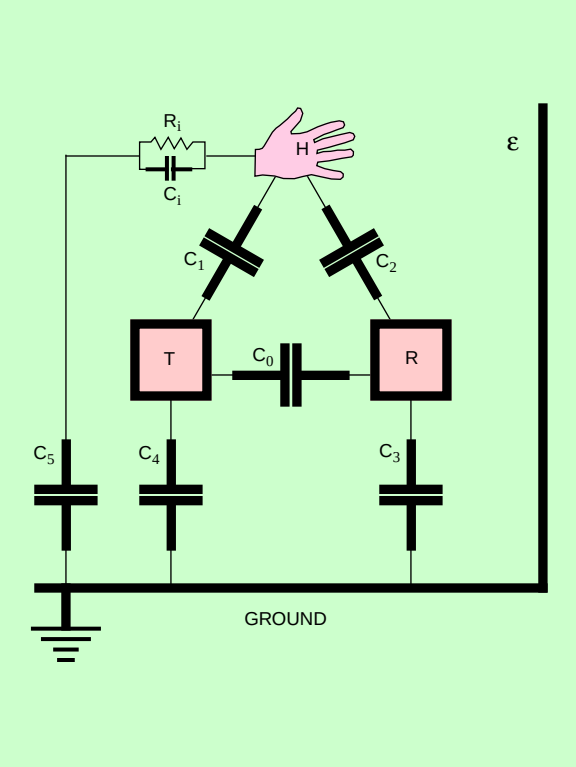
Each receive channel consists of an op-amp gain stage, a multiplier, and another op-amp used as an integrator; these components are used to implement synchronous detection. In the version of this technique used in the Fish, the received signal is multiplied by the original transmitted signal, and the resulting function is integrated over an interval of 60 milliseconds (ms). The effect of these two operations is to project out all of the Fourier components of the received signal except for the component that was transmitted. The multiplier and integrator are computing (in analog electronics) the inner product of the transmitted signal function s_t and the received signal function s_r, with a window function set by the integration time. The sense in which the multiplier and integrator project out all undesirable Fourier components is the following: because all distinct pairs of Fourier components are orthogonal, the contribution to the inner product $\langle s_t, s_r \rangle$ from all the undesirable (i.e., different from s_t and therefore orthogonal) components is zero. The input stage is therefore a very sharp filter that rejects all signals not of the proper frequency and phase. If desired, phase shifts may also be measured by performing a second, quadrature demodulation. Synchronous detection gives very high immunity to noise because of this sharp filtering.

It is also possible to describe the sensing circuitry in terms of amplitude modulation. The transmitter may be thought of as a carrier whose amplitude is modulated by the person's body. The receive multiplier mixes the carrier down to direct current (DC), and then the final low-pass filter rejects all signals other than those superimposed on the carrier.

Lumped circuit model and sensing modes. This subsection considers a lumped circuit model of a single transmit-receive pair with a single target object, in whose position we are interested. The various "modes" in which the Fish circuitry can be used have clear interpretations as current paths through the circuit diagram shown in Figure 1. For each sensing mode, we give a brief overview from a user's point of view and then explain the physics of the mode in terms of this diagram.

Figure 1 depicts the model. There are four "terminals": the transmitter, the receiver, the target object (shown as a hand), and ground. The six distinct interconductor capacitances are shown. The small resistor R_i and capacitor C_i represent the internal capacitance and resistance of the body. Capacitor C_5 is the target object's coupling to ground. If a person is being sensed, C_5 is usually dominated by the capacitance through the shoes of the person. The ground terminal may either be a groundplane in close proximity to the transmitter and receiver, or the ambient room ground.

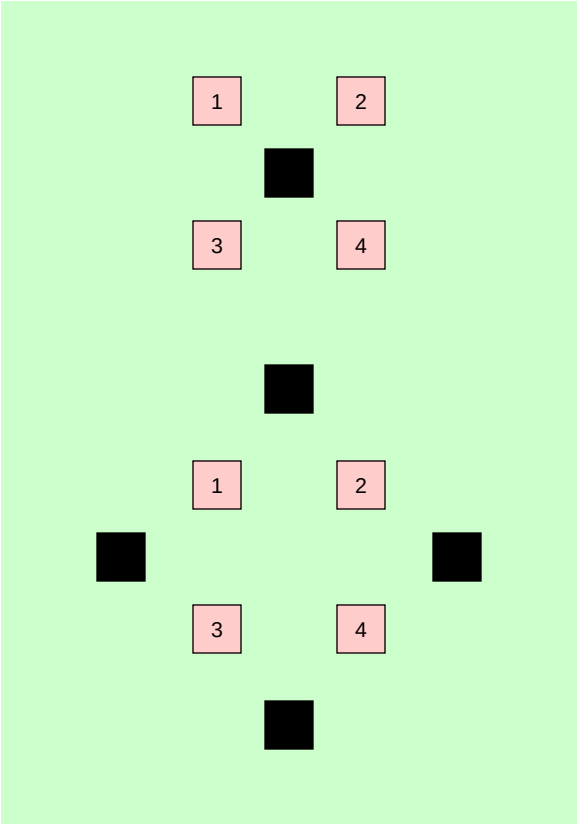
Figure 1 Lumped circuit model of Electric Field Sensing



The sensors can be used in a variety of ways, explained below, each of which modifies these capacitances differently. We measure capacitance by measuring the current arriving at the receiver, as explained in the earlier subsection on hardware.

Transmitter loading mode. This mode is the original Electric Field Sensing pathway. When a hand approaches the transmitter, the capacitance between the two conductors increases. In the theremin this new value of C_1 changes the oscillation frequency of a parallel inductor-capacitor resonant circuit, or LC tank, which is then mixed with a constant frequency to produce an audible beat. In the work of Vranish,[6,7] the value of C_1 is found by measuring the current lost through the transmitter. In loading mode, there is no receiver.

Figure 2 Two conductivity distributions that by construction yield identical loading-mode measurements and distinct shunt-mode measurements



The Smart Fish[18] has circuitry to measure the current being lost at the transmit electrode.

Transmit mode. In transmit mode,[11,16] the transmit electrode is put in contact with the user's body, which then becomes a transmitter, either because of direct electrical connection, or capacitive coupling through the clothes, which is shown as current path C_1 in the circuit diagram.

When the hand moves, the spacing to the receiver changes, which changes the value of C_2. This sensing technique is one that has been overlooked until recently.

When the spacing from the hand to the receiver is large, the received signal goes roughly as $1/r^2$, because the hand acts like a point object and the field falls off as $1/r^2$. By Gauss's law, the induced charge on the receiver also goes as $1/r^2$. Since the potentials on the electrodes are defined by the Fish circuit, we know the capacitance to be $C = Q/V$, and the received current $I_R = 2\pi fCV$, as explained in the previous subsection. When the hand is very close to the receiver, C_2 (typically) has the geometry of a parallel plate capacitor, and the signal goes as $1/r$.

Shunt mode. The remainder of this paper will be concerned with shunt mode, which is the most radical departure from previous practice, since it is a three-terminal measurement. With shunt mode it is possible to extract more geometrical information per electrode than with other modes, as we will subsequently explain.

In shunt mode, neither the transmitter nor the receiver is in contact with the user's body. When the user's body is out of the field, current flows from transmitter to receiver through the effective capacitance C_0.

When part of the user's body, such as a hand, enters the field, it functions as a third terminal, and the capacitance matrix changes, often drastically. In particular, the values of C_0, C_1, and C_2 shift. Since the voltage between the transmitter and receiver is held constant, the change in the component values between the transmitter and receiver leads to a change in the current arriving at the receiver. From the amount of current that fails to arrive at the receiver, one can infer something—what, exactly, is the question addressed in the third section of this paper—about the "amount of arm" in the vicinity of the sensor.

There is a strong sense in which shunt mode is more informative than loading mode: with shunt mode one can make n^2 measurements using only n electrodes, whereas loading mode allows only n measurements with the same number of electrodes. This allows shunt mode measurements to distinguish conductivity distributions that yield identical loading mode measurements. Figure 2 shows two distributions that yield identical loading-mode measurements and distinct shunt-mode measurements. We will explain the example in detail later, after we have discussed the effective circuit mode quantitatively.

Quantitative discussion of lumped circuit model. An expansion in a (small) time rate parameter of the field generated by the transmitter shows that 100 kHz is comfortably in the quasistatic regime for measurements on room scales (10 m) or smaller.[19] By using

the quasistatic approximation, calculating the current received at a particular electrode is straightforward.

The static charge on a conductor i is due to the static term $\mathbf{E}_0$ in the expansion of the $\mathbf{E}$ field:

$$Q_i = -\int_{S_i} \varepsilon \mathbf{n} \cdot \nabla \phi_0 da \quad (1)$$

where S_i is the surface of i, $\mathbf{n}$ is the outward normal to S_i, ϕ_0 is the potential of which $\mathbf{E}_0$ is the gradient, and the permittivity ε is a function of position, since the medium is not homogeneous. This expression relates the macroscopic charge on a measurement electrode to the microscopic permittivity field ε that we are ultimately interested in knowing. This microscopic permittivity field determines the value of the collective property of capacitance, which we will find by measuring a current.

Using the standard definition, the capacitance of conductor i due to a conductor j is the ratio between the charge on Q_i and the voltage between j and a reference. If we know the capacitance and voltages for a pair of electrodes, we can find the charge induced on one by the other. Because of the linearity of all the equations involved, the total charge on i induced by all the other conductors is the sum of the separately induced charges[20] (but note that the capacitances are not linear functions of position):

$$Q_i = \sum_j C_{ij} V_j \quad (2)$$

An element of the matrix C_{ij} represents the ratio between Q_i and V_j assuming all the other Vs are zero.

The macroscopic quantity we actually want to measure, because it encodes geometrical information, is the capacitance. But it is easier to measure current than charge, and we can extract the same information from the current. The current I_i entering receiver i is given by the time derivative of the charge on i: $I_i = dQ_i/dt$.

$$I_i = \frac{d}{dt}\sum_j C_{ij} V_j = \sum_j C_{ij} \frac{dV_j}{dt} \quad (3)$$

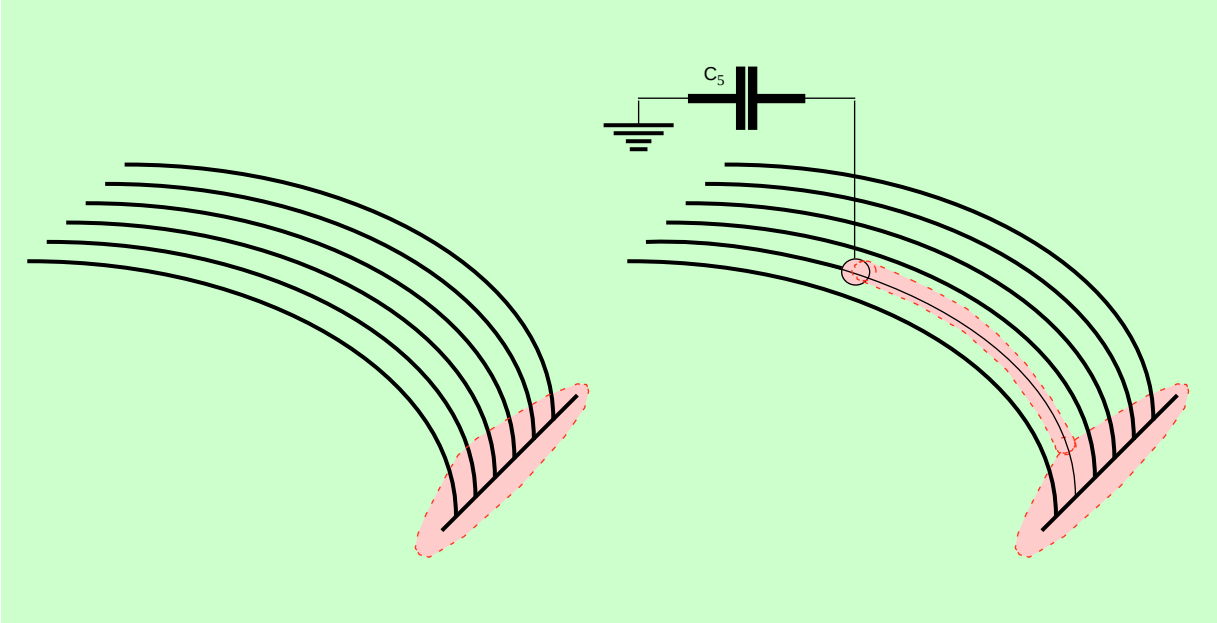
The off-diagonal terms of the capacitance matrix C_{ij} represent the ratio between Q_i and V_j for i not equal to j when all the other Vs are zero. The diagonal "self-capacitance" terms C_{ii} represent the charge on i when it is held at V_i and all the other electrodes are at zero. Thus the diagonal terms correspond to a loading-mode measurement, and the off-diagonal terms correspond to shunt-mode measurements. For pure shunt-mode measurements (no contribution from transmit mode) made with identical electrodes, sensor values are invariant under the operation of interchanging the transmitter and receiver, so the matrix is symmetrical. This is not the case in transmit mode, so it may be possible to separate the contributions from shunt and transmit mode by splitting the capacitance matrix into its symmetric and antisymmetric components.

With shunt mode it is possible to extract more geometrical information per electrode than with other modes.

Having introduced the capacitance matrix, we can now give a clearer explanation of Figure 2, the example of a pair of conductivity distributions that can be distinguished by shunt measurements but not loading measurements. For clarity we will compare the loading measurements that can be made with a set of n electrodes to the shunt-only (we assume there is no transmit mode contribution) measurements that can be made with the same electrodes. Since the shunt measurements are symmetric with respect to interchange of the transmitter and receiver, with n electrodes, $n(n-1)/2$ distinct shunt measurements may be made, whereas just n loading-mode measurements are possible. It is not until $n \geq 4$ that $n(n-1)/2 > n$, so shunt mode does not have an advantage over loading mode in terms of the number of measurements per electrode when there are fewer than four electrodes.

Figure 2 shows two distributions that, by construction, give the same loading measurements: the four small dark objects that comprise the second distribution can be moved in from infinity until the signals are the same as those from the first distribution. Thus the loading-mode measurements on all four electrodes are the same for the two distributions. So $C^1_{ii} = C^2_{ii}$, where

Figure 3 The unperturbed electric field impinging on the receiver, left, and the perturbed field, right



the superscript indicates the distribution. Even assuming that the shunt measurements around the sides of the square formed by the electrodes yield the same values for the two distributions (i.e., that $C^1_{12} = C^2_{12} = C^1_{23} = C^2_{23} = C^1_{34} = C^2_{34} = C^1_{41} = C^2_{41}$), the remaining shunt-mode measurements clearly distinguish the two cases: $C^1_{14} \neq C^2_{14}$ and $C^1_{23} \neq C^2_{23}$. In this sense shunt-mode measurements are more informative than loading-mode measurements.

Component values. Now that we have discussed the effective circuit diagram qualitatively, we will present some typical component values. The circuit diagram contains only one (small) resistor because the real impedance of free space is essentially infinite, and the real impedance of the body is nearly zero. Barber[21] gives resistivity figures on the order of 10 Ωm (ohm-meters), plus or minus an order of magnitude: cerebrospinal fluid has a resistivity of 0.65 Ωm, wet bovine bone has 166 Ωm, blood has 1.5 Ωm, and a human arm has 2.4 Ωm longitudinally and 6.75 Ωm transversely.

Zimmerman measured the capacitance between the right hand and the left foot of a living human, and found a value of 9.1 picofarads (pF).[22] A simple parallel plate model of feet in shoes with 1-centimeter (cm) thick soles gives a capacitance of 35 nanofarads (nF), using $C = \varepsilon_0 A/d$, and taking $A = 2$ feet $\times$ 20 cm $\times$ 10 cm and $d = 1$ cm. For 10-cm thick platform shoes, the value of $C = 3.5$ nF. (We have neglected the dielectric constant of the soles.)

Variations in C_5 can cause offsets of the sensor values. The C_5 of a person wearing 10-cm thick platform shoes would be one tenth that of a person wearing shoes with 1-cm thick soles. In fact, the value of C_5 can vary even more, for example, when the person is barefoot or standing on an actual platform such as a stage.[23] But C_5 can be measured for calibration purposes with a single loading-mode measurement of the current out of a transmitter that the user touches before operating the equipment.[24]

Forward problem

Having introduced the various forms of Electric Field Sensing, I now present analytical calculations of sensor readings given hand positions.

General case. The sensor values can be determined in the most general case by solving the Laplace equation

with an inhomogeneous permittivity ε

$$-\nabla \cdot \varepsilon \mathbf{E}_0 = \nabla \cdot (\varepsilon \nabla \phi_0) = \varepsilon \nabla^2 \phi_0 + \nabla \varepsilon \cdot \nabla \phi_0 = 0 \qquad (4)$$

and then using Equation 1 to find the charge induced on the receive electrodes by the field that solves Equation 4. However, for a problem as simple as the three-dimensional mouse, this model is too general. We do not need to solve Laplace's equation each time we move the mouse. In this section an approximate forward model of the response of the sensor to a single point-like grounded object will be presented, and its relevance to more general problems discussed. A more detailed discussion of the model is presented in Reference 19; for the purposes of this paper, the main point is that the model is consistent with the sensor data, as shown later in Figure 4. There are two parts to the model: the first part describes how the sensed object interacts with any electric field, and the second models the particular electric field induced by our transmitter and receiver.

Modeling the interaction: The point absorber. We want to model the effect of a small, perfect conductor h at a point $\mathbf{x}$ in space, connected to ground through a capacitance C_5 and a wire whose effect on the field is negligible. To create a simple model of the effect of h on the displacement current arriving at the receiver, we will assume that the object affects the field only in a limited way: the field lines (of the unperturbed field) are attenuated when they intersect the object, but the path through space traced out by any particular field line is unchanged.

To find the change in received displacement current using this approximation, we can start with a Gaussian surface surrounding the receiver and add a tube from the perturbation back to the surface, with the axis of the tube everywhere parallel to the field line from the perturbation to the surface, as shown in Figure 3. Since the tube is small, its surface is also parallel to the field, so there is no flux through any part of the tube except the end cap, which is adjacent to the object and perpendicular to the field at that point. The change in flux through the Gaussian surface that results from the perturbation is therefore proportional to the original field strength at the location of the perturbation diminished by an attenuation factor and multiplied by the area of the perturbation. Gauss's law then gives the charge induced on the receiver. So the received current $I_R \propto I_0 - \alpha E(\mathrm{x}) \cdot dA$, where I_0 is the unperturbed current, α is an attenuation factor between 0 and 1, and dA is a vector representing the area of the cap. If the object has no orientational dependence (i.e., if it is spherical) and its size is fixed, then the area and attenuation factors may be combined into a single constant, and the dot product replaced with the magnitude of the field strength, giving $I_R \propto I_0 - |E(\mathrm{x})|$. Having introduced a model for how an object changes the received signal that depends on the field strength at the location of the object, we now present a model of the unperturbed field itself.

Modeling the field: The dipole approximation. We will approximate the field resulting from a pair of small, identical, rectangular electrodes of dimension $b \times c$ and displaced from one another by a along the x-axis as a dipole with the same spacing. The dipole moment of a charge distribution is

$$\mathbf{p} = \int \mathrm{x}' \rho(\mathrm{x}') d^3 x' \qquad (5)$$

If the surface charge density on the electrodes had a uniform value of ν,[25] then Equation 5 yields $p_x = \nu abc$ and $p_y = 0$. The expression for p_x makes sense: νbc yields the total charge on one electrode, so we could write $p_x = Qa$. Thus the pair of rectangular electrodes displaced from one another by a and charged to $+Q$ and $-Q$ has the same dipole moment as a pair of point charges $+Q$ and $-Q$ displaced by the same amount.

To justify the dipole approximation more rigorously, we could solve for the charge distribution on the electrode surfaces and then perform a multipole expansion of the charge distribution. The dimensions at which the higher-order terms became significant would be the limits of the approximation's validity.

Modeling the sensor response. Using the point absorber model together with the dipole approximation of the field geometry, we can model the sensor response data measured using a small grounded object as a hand phantom. Figure 4 shows a plot of the function $I_0 - |E(\mathrm{x})|$, where I_0 is a constant and $E(\mathrm{x})$ is a dipole field, given by the gradient of the dipole potential $\mathbf{p} \cdot \hat{\mathbf{r}}/r^2$. The dipole moment $\mathbf{p}$ is a constant representing charge multiplied by the vector from the center of the transmitter to the center of the receiver. Measurements were made in a plane parallel to the dipole axis, using a grounded metal cube (one inch on a side) as a hand phantom. The electrodes were one inch square pieces of copper foil spaced eight inches apart. The theoretical plot is for a plane parallel to the

Figure 4 Comparison of functional form of dipole model with experimental data

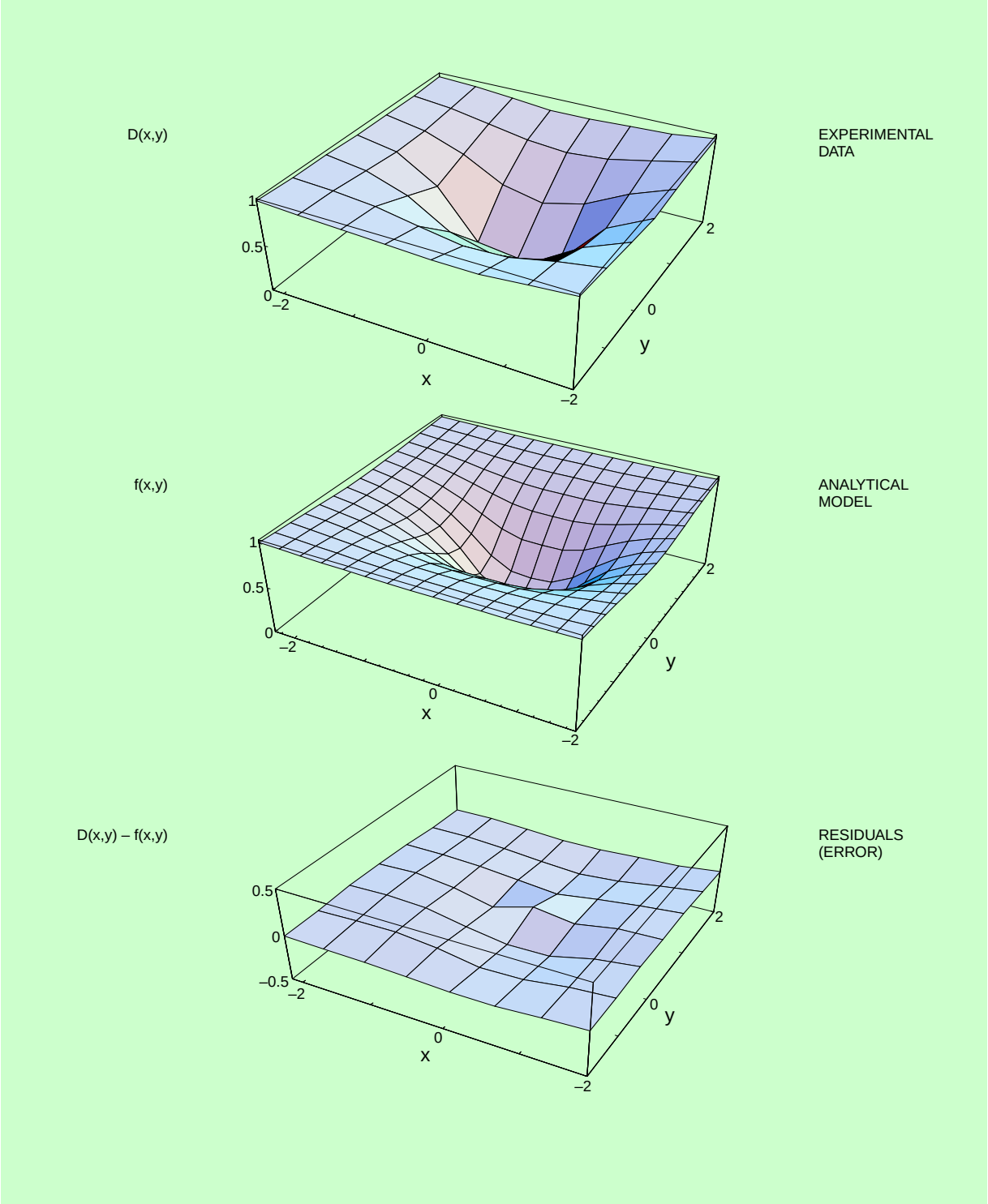
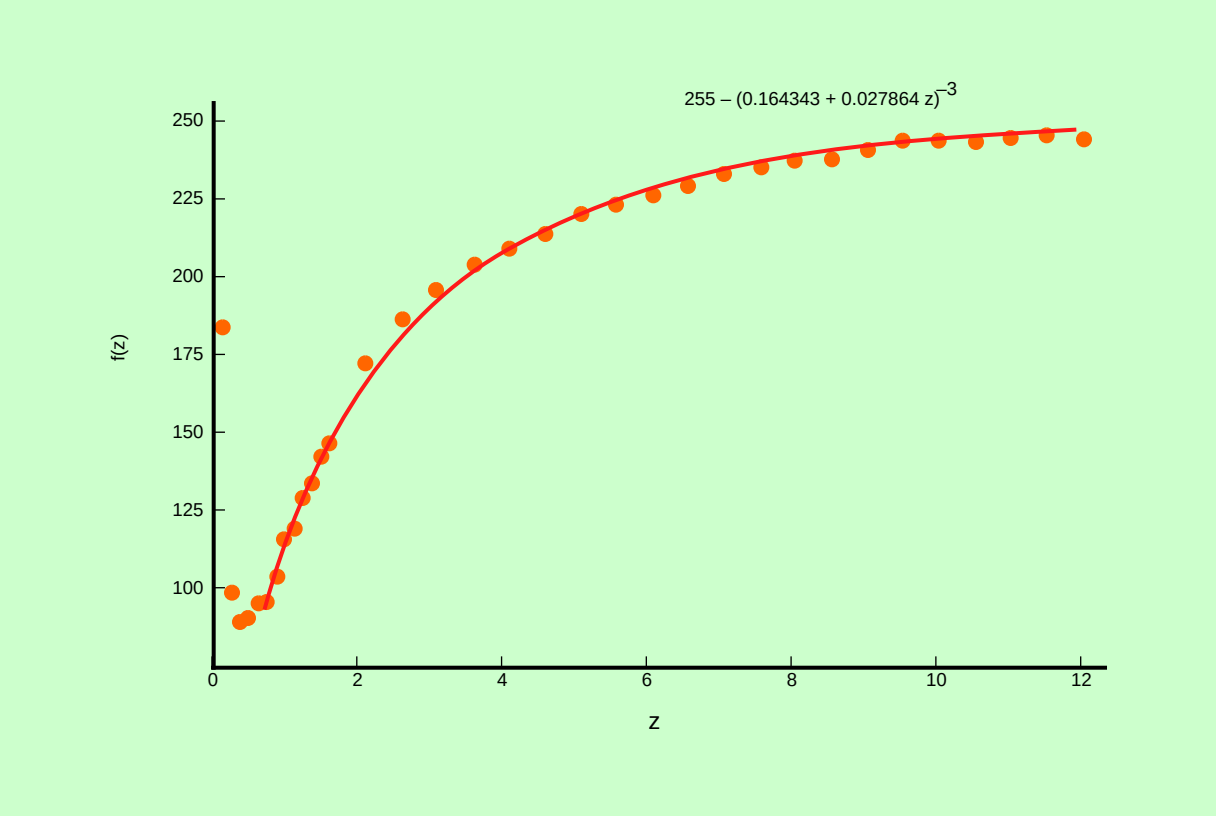


Figure 5 Comparison of functional form of dipole model with experimental data

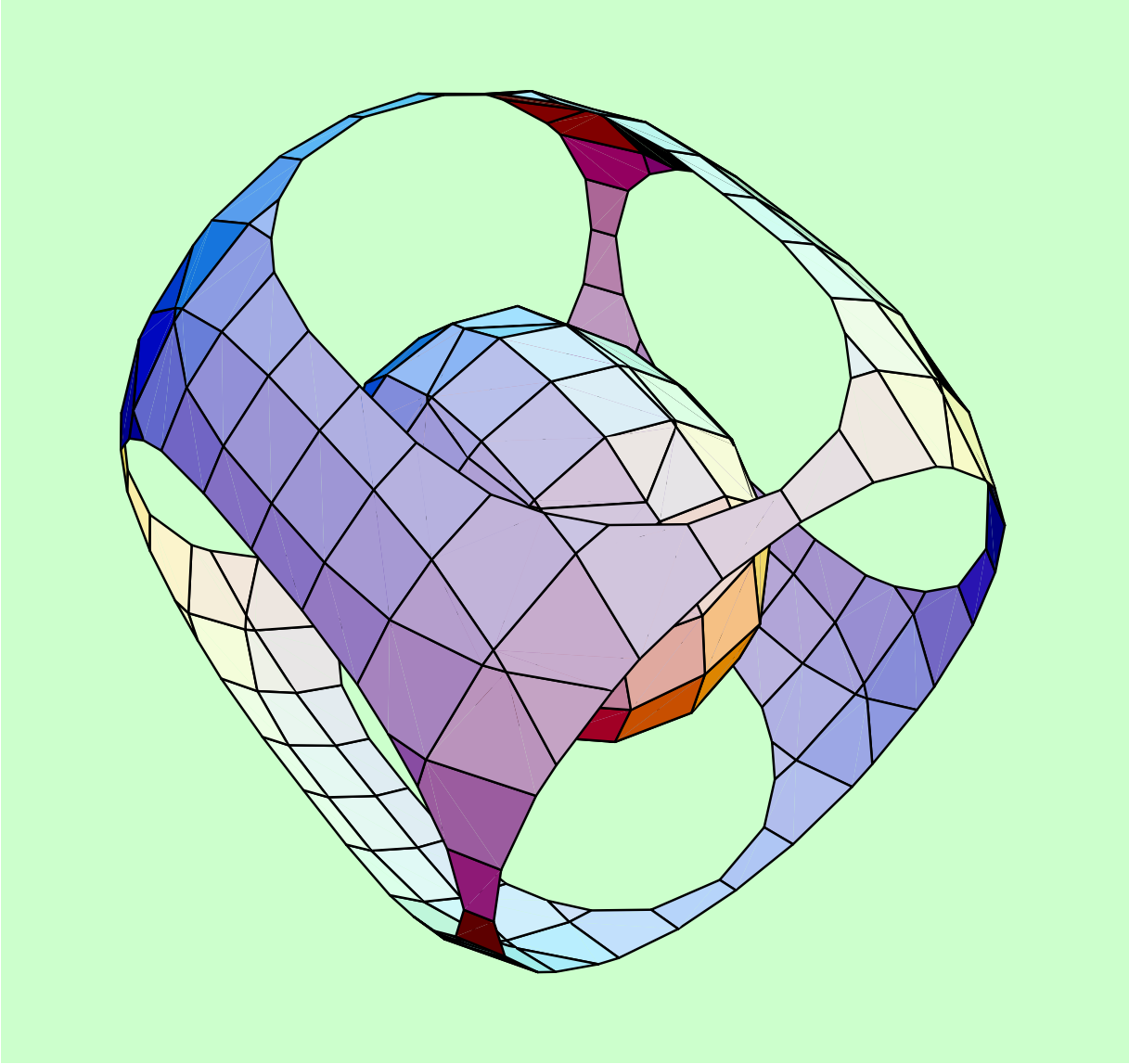


dipole axis, at a distance of .9*a* above the dipole axis, where *a* is the dipole spacing.

Figure 5 shows sensor measurements along the *z*-axis, perpendicular to the transmit-receive axis, and the dipole response model. Since *x* and *y* are zero, the dipole model simplifies to $\mathbf{E}(0,0,z) = 1/z^3$. Scale and offset parameters for the distance (abscissa) and sensor value (ordinate) have been fit to the data. The function plotted is shown at the top of the graph. Measurements were made along a line originating at the dipole origin and extending outward, perpendicular to the dipole axis. At very short distances, transmit mode starts to dominate, and the signal rises again.

Isosignal shells. Given this field model, we can plot surfaces of constant sensor readings. These plots are very helpful in shaping one's intuition about the behavior of the sensors, and the isosignal surfaces themselves are fundamental to understanding what information a measurement provides. The surfaces are ellipsoidal shells. The central axis of the ellipsoid is the dipole axis. Figure 6 shows two nested isosignal shells for two different sensor readings. The outer shell has been cut away to reveal the inner one. The dipole generating the field lies along the central axis of the ellipsoid. A shell represents the ambiguity class of a measurement, that is, the set of points in model parameter space all of which explain the data as well as possible. A point in parameter space that is not part of the shell generated by a measurement corresponds to a setting of model parameters that does not explain the data as well as the points that are part of the shell. Measurements made by additional electrodes generate additional shells in parameter space. The intersection of all the shells gives the set of points in parameter space whose members all explain the data as well as possible. In the case of a three-dimensional mouse, the parameters are spatial coordinates, so the parameter space is ordinary three-dimensional space.

Figure 6 Nested isosignal shells for two readings returned by a sensor



Later, we will also interpret these ellipsoidal shells as the ambiguity class for a single sensor reading. We will use the term ambiguity class to mean the set of points in model parameter space, all of which are maximally likely given the data readings. Our goal will be to use multiple sensor readings to reduce the ambiguity class from a two-dimensional manifold to a single point, the unique setting of model parameters that explains the data.

With the isosignal shells in mind, this is a good place to consider the generality of the point-absorber model. Can it be extended? What happens if there are two point absorbers? Though the sensor values are not linear, they do depend monotonically on distance, at least in the shunt regime. Furthermore, in the shunt regime, adding another point absorber can only decrease the received displacement current. This means that in the general case of an arbitrary number

of point absorbers, a sensor reading guarantees that the interior of the (single-absorber) shell is free of absorbers, and makes, in the worst case, no guarantees whatsoever about the exterior of the shell.

This shows that even though the capacitance is not a linear function of the absorber configuration, the point absorber still plays a crucial role in understanding the behavior of the sensors in the presence of more complex geometries. We are presently developing an electric field imaging method called the Swiss Cheese algorithm that is based on this understanding of the relationship between the single absorber response function and the response to a general distribution.

Groundplane. The dipole approximation discussed above also turns out to be a good model of the response of a sensor in the presence of a groundplane. The reasons for this are discussed in Reference 19. Here we only explain the consequences of having an analytical model for the behavior of a single transmit-receive pair in the presence of a groundplane.

Since receivers are virtual grounds, the field configuration produced by, say, one transmitter, should not be changed by replacing a patch of groundplane with a receiver. In other words, the field produced by a transmitter surrounded by groundplane is identical to that produced by a transmitter surrounded by a patchwork of receivers and ground. This in turn means that, in the presence of a groundplane, receivers do not affect one another—they are entirely independent dipoles. So we now have an approximate analytical solution to the forward problem of predicting the sensor values given a hand location, and this solution will be very helpful in solving the inverse problem.

Constructing the ambiguity class

In this section we introduce a general probabilistic framework that will allow us to solve inversion problems such as the framework needed to make the three-dimensional (3D) mouse. The framework will also suggest a means of designing maximally informative sensor geometries. This approach, as applied to imaging problems, is described by Jaynes,[26] Herman et al.,[27] Gull and Daniell,[28] and Skilling and Gull.[29]

The essence of the approach is to view hand finding, for example, as an inference problem. We define a model whose parameters—in the case of the 3D mouse, the three coordinates of the hand—we wish to know, and a probability distribution over those parameters. As more data become available (for example, as we consider additional sensors), the volume of the feasible set of model parameters decreases, and the probability distribution becomes increasingly peaked around the "true" values of the parameters.

As mentioned in the previous section, we will define the ambiguity class for a set of measurements and model parameters to be the maximally probable subset of model parameter values, given the measured values. As we remarked in the discussion of the forward model, for a single shunt-mode measurement and model (m_x, m_y, m_z), the ambiguity class is a two-dimensional manifold, an ellipsoidal isosignal (and thus isoprobability) shell. A second measurement yields a one-dimensional ambiguity class (the intersection of the two single-measurement ambiguity classes), and a third measurement yields a zero-dimensional ambiguity class, i.e., a single point or set of isolated points. Because each additional (nondegenerate) measurement reduces the dimensionality of the ambiguity class by one, each measurement allows the value of an additional model parameter to be inferred. Thus, with two measurements, we can infer the values of two position parameters or one position and one size parameter; with three measurements, we can infer two positions and one size, or two positions and one orientation, or three positions, and so on. Each additional measurement allows us to infer the value of an additional parameter characterizing the distribution.

In the presence of a groundplane, receivers do not affect one another.

Ill-posed (underdetermined) problems, in which there are more unknown parameters than measurements, can be made well-posed either by collecting additional data or by specifying additional *a priori* constraints on the feasible set. This is the Bayesian view of regularization. These constraints can be encoded in

the prior probability distribution that defines the initial feasible set.

If enough measurements have been made to yield a zero-dimensional ambiguity class, then the posterior probability distribution has isolated peaks at the maximally probable points. In this paper, we always use a prior (a prior probability distribution) to select a single one of these peaks; this corresponds to a design choice about the working region of the mouse. Given a peaked probability distribution, we can analyze the uncertainty of our estimate of the model parameter values by examining the curvature of the distribution in the vicinity of the maximum.

Assuming a Gaussian approximation to the posterior probability distribution, the inverse curvature of a peak in a particular direction gives the uncertainty of the estimate of the parameter value (or linear combination of parameter values) corresponding to that direction. The amount of information provided by a measurement can be quantified by the change in entropy of the distribution that resulted from the measurement. The problem of designing sensor arrays may be posed in terms of maximizing the expected information provided by a measurement.

Since the sensors are subject to additive Gaussian noise,[19] the probability of the data given some setting of model parameters is given by

$$p(D|m) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(D-f(m))^2}{2\sigma^2}} \tag{6}$$

where σ is the standard deviation,[30] D is a data value, $f(m)$ is the data value predicted by our analytical forward model given a model configuration (hand position) m. This distribution is normalized: if we integrate over all values of D, we get 1.

By Bayes's theorem,

$$p(m|D) = \frac{1}{\sqrt{2\pi}\sigma} \frac{e^{-\frac{(D-f(m))^2}{2\sigma^2}} p(m)}{p(D)} \tag{7}$$

For the case of a two- or three-dimensional mouse, we can choose a prior $p(m)$ that renders the inversion well-posed by, for example, restricting the possible hand positions to positive coordinate values. This restriction selects one of the two peaks in the posterior distribution. A useful prior (probability) for one of the model parameters m_x is $p(m_x) = c/(1+e^{-\beta m_x})$, defined in some finite range of m_x, where c is a normalizing constant and β is a sharpness parameter. This function is a way to approximate a step function with a closed form expression. A possible advantage of using this function over a hard step function is that numerical optimization techniques are able to follow it back into the high probability region, since it is smoothly varying. The prior for our entire model is the product of the priors for m_x, m_y, and m_z. So the posterior, with the prior for just one dimension shown, is

$$p(m|D) \propto e^{-\frac{(D-f(m))^2}{2\sigma^2}} \frac{c}{1+e^{-\beta m_x}} \tag{8}$$

Apart from the prior, which we might have chosen to be a constant over some region, the functional form of $p(m|D)$ is identical to that of $p(D|m)$. The fact that the $p(m|D)$ distribution and the $p(D|m)$ distribution, which have completely different meanings, have the same functional form is the content of Bayes's rule. However, the similarity in functional form is in some sense superficial. Consider the normalization of $p(m|D)$. Rather than performing the analytically tractable Gaussian integral over D (tractable and Gaussian because when we integrated $p(D|m)$, m and therefore $f(m)$ was fixed), we must integrate over all values of m, which means integrating our forward model composed with a Gaussian. The difficulty of performing this integration depends on the form of f. This normalization constant, which Bayesians grandly call the evidence, is not important for finding the best setting $\hat{m}$ of the model parameters, since a scaling of the dependent variable (probability) has no effect on the location of maxima. However, it does become important when making any sort of comparison between different functions f, or calculating entropies.[31]

Information collected by multiple sensors can easily be fused: simply take the product of the posterior terms $p(m|D_i)$ for each separate sensor i to find the posterior distribution given all the available data. (We are assuming that each sensor makes just one measurement; otherwise we would need separate indices for the sensors and the data values.) Thus if D now denotes the set of N measurements D_i,

$$p(m|D) \propto \prod_i^N e^{-\frac{(D_i - f_i(m))^2}{2\sigma^2}} \prod_j \frac{c}{1 + e^{-\beta m_j}} \quad (9)$$

Notice that, since log is a monotonically increasing function, if we maximize $\log p(m|D)$, we will get the same m as if we had maximized $p(m|D)$. It will be desirable in practice to work with log probabilities rather than probabilities for several reasons: we can save computation time since exponentials disappear, and multiplication and division become addition and subtraction. Furthermore, when we multiply many probabilities together, the numbers become very small, so that numerical precision can become a problem. Using log probabilities alleviates this problem and reduces computation time, since exponentials appear so often in each of our Gaussian probability distributions. Rather than maximizing the log probability, we could minimize the negative of the log probability:

$$-\log p(m|D) = \sum_i^N \frac{(D_i - f_i(m))^2}{2\sigma^2} + \sum_j \log(1 + e^{-\beta m_j}) - \log c \quad (10)$$

This quantity has the familiar interpretation of the sum of squared errors between the actual data and the data predicted by the model, with an additional error term derived from the prior.

Error bars: Local uncertainty about the maximum. Once the basic degeneracies have been broken, either by collecting sufficient data or imposing constraints via a prior, so that there is a single maximum in the log probability, the uncertainty about the best setting of model parameters may be represented by the inverse Hessian matrix A^{-1} evaluated at the maximum. To see why, we will consider the Hessian and its properties. The Hessian A gives the curvature, which is a measure of confidence or certainty. In A's eigenvector basis, in which it is diagonal, the diagonal elements (the eigenvalues) A_{ii} represent the curvature along each of the eigenvector directions (known as the principal directions). The curvatures along the principal directions are called the principal curvatures. The product of the curvatures, the Gauss curvature, which serves as a summary of the certainty at a point, is given by the determinant of A. The average curvature is given by 1/2 trace $A = (A_{11} + A_{22})/2$. Finally, the curvature in a particular direction (in two dimensions) $v = (cos\theta, sin\theta)$ is given by Euler's formula:[32]

$$\kappa = v^T A v = \kappa_1 cos^2\theta + \kappa_2 sin^2\theta \quad (11)$$

The inverse of A in the eigenvector basis is the matrix with diagonal elements $1/A_{ii}$. Thus the inverse Hessian specifies "radii of curvature" of the probability distribution, which can be used as a measure of uncertainty. The determinant and trace of the inverse Hessian are independent of coordinates, so we may use these as local measures of the "Gauss uncertainty" and mean uncertainty even when we are not in the eigenvector basis. Using Euler's formula above, we could determine the uncertainty in any desired direction. Our inverse Hessian is ordinarily known as the variance-covariance matrix in statistics, so the geometrical description is certainly not the only way to understand this quantity; however, we find the geometrical description helpful in the context of this problem.

Entropy: Global uncertainty and maximally informative sensor geometries. The most general global measure of uncertainty is the entropy. The change in entropy of the $p(m|D)$ distribution resulting from the collection of new data measures the change in uncertainty about the values of the model parameters, including uncertainty due to multiple maxima, given a set of measurements. The change in total entropy ΔH of the posterior distribution resulting from a measurement D_{n+1} is

$$\Delta H(m|D_{n+1}) = H(m|D_{n+1}) - H(m|D_n) \quad (12)$$

where

$$H(m|D_n) = -\int p(m|D_n) \log p(m|D_n) dm \quad (13)$$

The expected change in entropy when we collect a new piece of data, that is, the change in entropy averaged over possible data values, gives a basis for comparing sensor geometries. The expected value of $H(m|D)$ is

$$I = \int p(\hat{m}) H(m|D) d\hat{m} \quad (14)$$

where $\hat{m}$ is an actual object position and $D = f(\hat{m})$, with f being the forward model. Thus,

$$I = \int p(\hat{m}) H(m|f(\hat{m}))d\hat{m} \quad (15)$$

and substituting in for $H(m|f(\hat{m}))$,

$$I = \int p(\hat{m})\left[-\int p(m|f(\hat{m}))(\log p(m|f(\hat{m})))dm\right]d\hat{m} \quad (16)$$

I is a measure of the quality of a sensor geometry. By analogy with coding theory, the best measurement procedure (for single measurements) reduces the entropy as much as possible. One could therefore search for optimal sensor geometries by minimizing I.[31,33,34]

The suitability and quality of a sensor layout and prior can be checked by examining the ambiguity class.

Example: Two-dimensional field mouse. Here we use this technique to construct the ambiguity class and find the most likely model parameters given two sensor readings. We want to infer the position of the hand in two dimensions from two sensor readings. So the model consists simply of two numbers, representing the position of the object purported to explain the sensor readings. The sensor axes are oriented perpendicular to one another, and the transmit electrode is shared.

Figures 7 and 8 show the posterior distributions $p(m_x, m_y|D_1)$ and $p(m_x, m_y|D_2)$ for the two sensors, oriented perpendicular to one another. To make the figure easier to view, the noise has been exaggerated dramatically. If the actual noise levels for the sensors had been used, the features of the surfaces would be so minute that the contour plot routine would have very little to display.

Figure 9 shows the posterior probability distribution $p(m_x, m_y|D_1, D_2) = p(m_x, m_y|D_1)p(m_x, m_y|D_2)$ with a uniform prior. The arrows show the principal components of the inverse Hessian evaluated at each peak. The larger error bar on the less sharply defined peak has been scaled down by 1/3 to fit it on the page. The usable region of the mouse is the upper right quadrant of the region shown; in practice a nonuniform prior would be used to eliminate the ambiguity by removing the structure in the lower quadrants.

The "principal uncertainties," or error bars, are also shown superimposed on the maxima. The three smallest arrows have been scaled up by a factor of 10 to make them more visible. The larger arrow on the less sharply defined peak has only been scaled up by 3 1/3, so that it fits on the page.

Application: 3D field mouse

In this section we describe another application of the discussion from the previous section: a three-dimensional mouse. We choose a sensor geometry and construct its ambiguity class for an example hand position. It is possible to check the suitability and quality of a sensor layout and prior by examining the ambiguity class: if there are multiple maxima, the inversion is ill-posed, and if the peak is not sharp (if the maximum has high radii of curvature, that is, a high value of the determinant of the inverse Hessian matrix), the value is very uncertain.

Figure 10 shows the layout we selected. The electrode labeled T is the transmitter, and R1–R3 are the receivers. The surrounding square represents a groundplane.

Earlier we discussed criteria for optimal sensor design. Evaluating the entropy integrals, and averaging over all possible data values, represents a substantial practical challenge. Efficient means of doing so will probably require sophisticated numerical techniques, except in special cases.

Therefore, we will simply satisfy ourselves that this layout does not lead to ill-posed inversion problems by examining its ambiguity class. Figure 11 shows the ambiguity classes for three single sensor measurements made using this layout. The object being measured is at $(x,y,z) = (0.8, 0.6, 1.05)$, using units of the smallest transmit-receive spacing. The ambiguity shells intersect at just one point in the region of interest. The ambiguity class for the joint measurement of all three sensor values is this single intersection point. Figure 12 shows the posterior distribution for this sensor layout and the object once again at $(x,y,z) = (0.8, 0.6, 1.05)$. Each image shows a slice through the three-dimensional posterior probability distribution,

Figure 7 Posterior probability $p(x,y,0.9 \mid D_1)$ where D_1, the measurement on sensor 1, is given by $D_1 = f_1(0.9, 0.6, 0.9)$; the value of z is constrained to be 0.9.

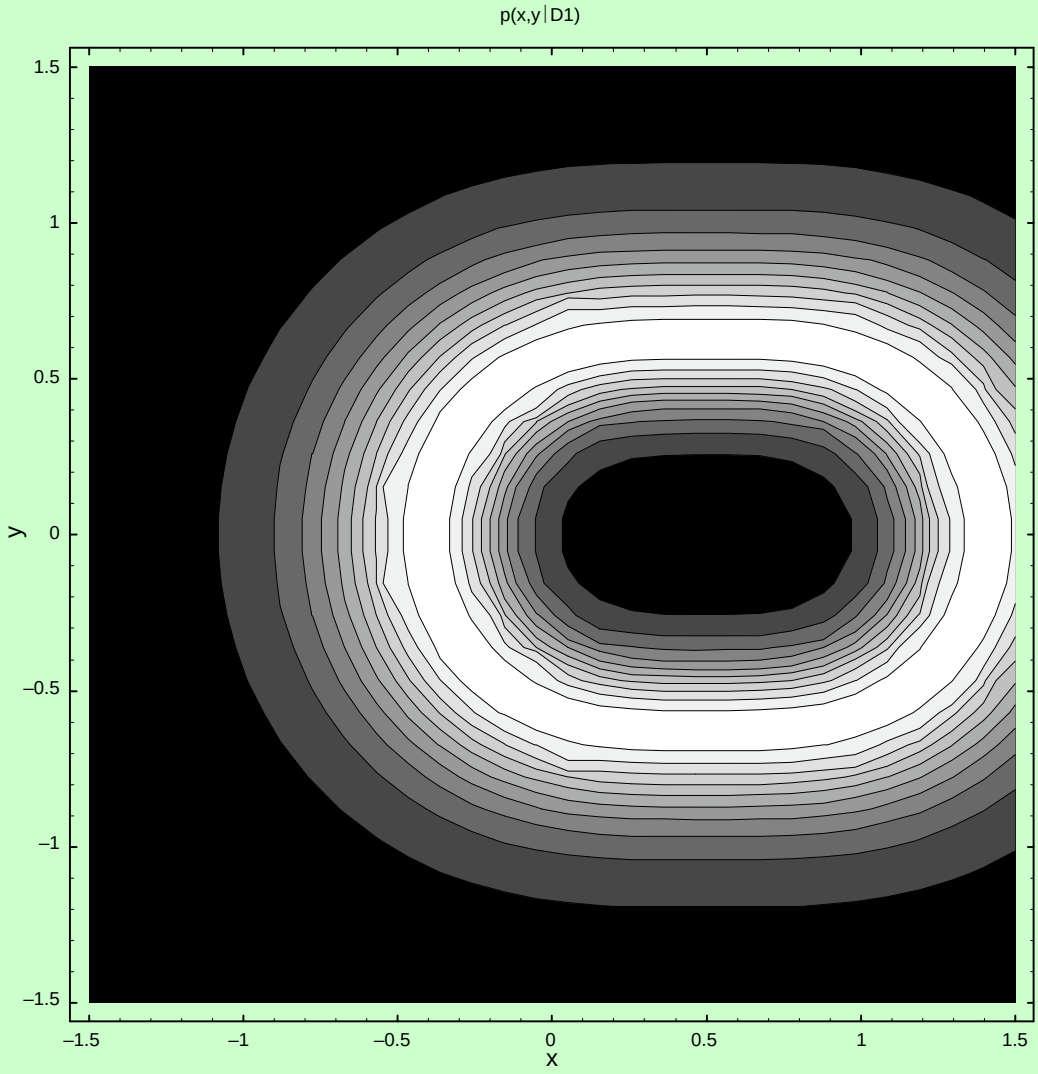


Figure 8 Posterior probability $p(x,y,0.9 \mid D_2)$ where D_2, the measurement on sensor 2, is given by $D_2 = f_2(0.9, 0.6, 0.9)$; the value of z is constrained to be 0.9.

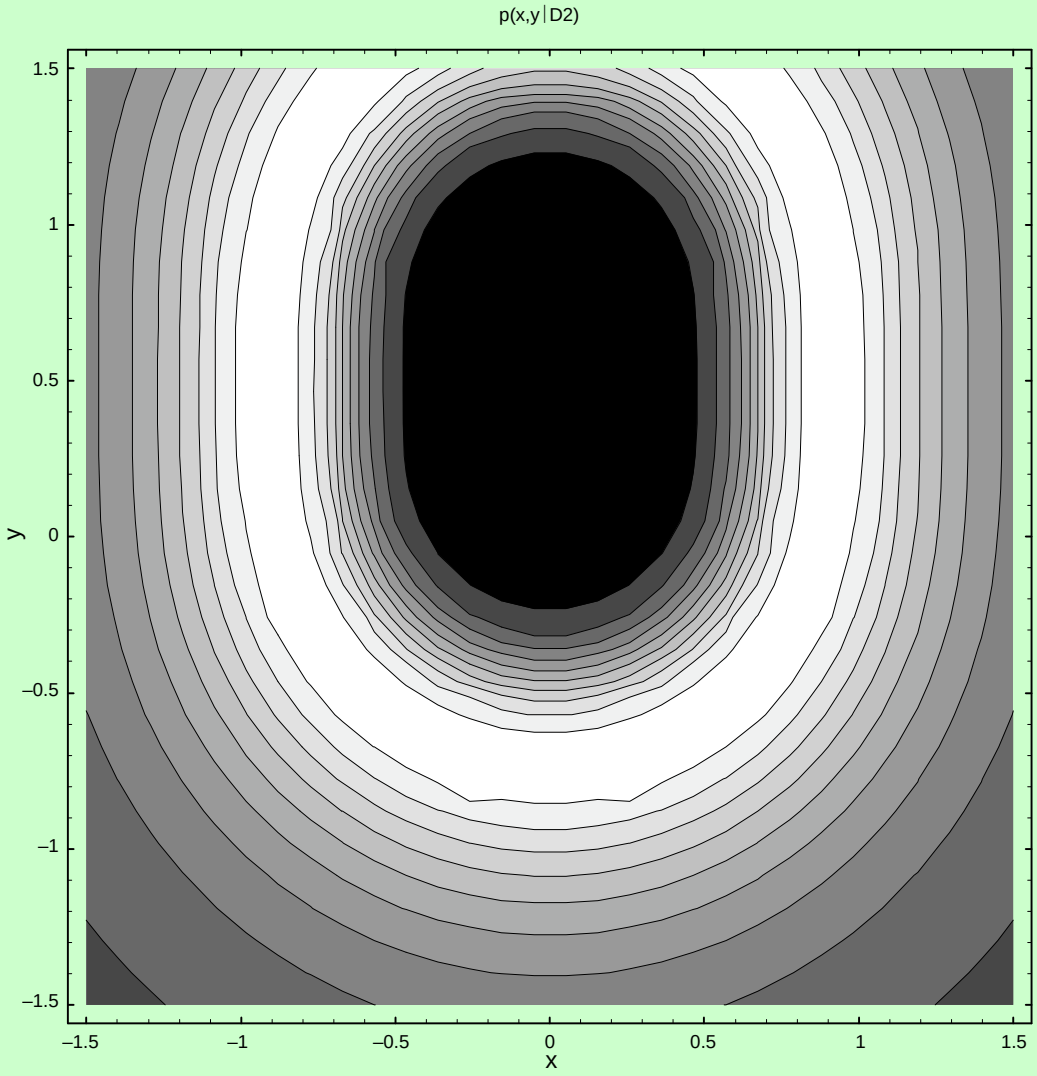


Figure 9 Posterior probability distribution with error bars: $p(x,y,0.9 \mid D_1,D_2)$ for sensors 1 and 2 given measurements $D_1 = f_1(0.9, 0.6, 0.9)$ and $D_2 = f_2(0.9, 0.6, 0.9)$; the value of z is constrained to be 0.9.

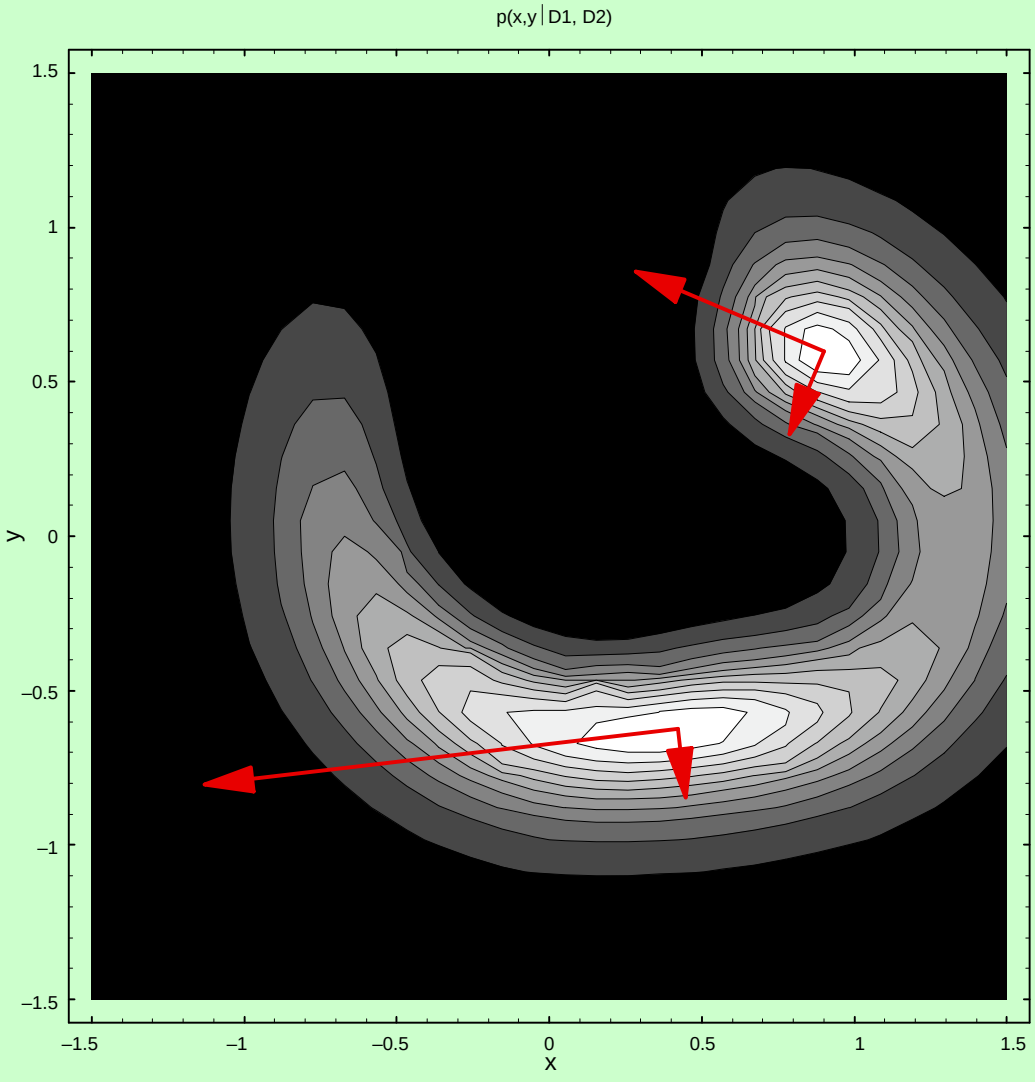
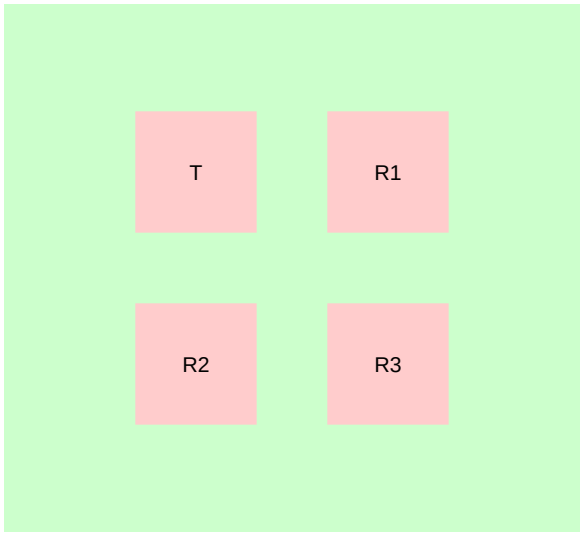


Figure 10 Sensor geometry for 3D mouse



parallel to the *X*-*Y* plane. The maximum of the posterior is at (0.8, 0.6, 1.05), as expected.

In the figure, the variance of the noise assumed on the sensors was increased 400 times over the observed noise to make the figure readable. Thus the figure illustrates the relative uncertainty of the position estimate, but not the absolute uncertainty. The figure shows that the uncertainty in the *Y* direction is least; the most uncertainty is in the joint estimate of *X* and *Z*.

The most important feature of the plot is that there is a single maximum. It gives an indication of the geometry of the uncertainty isosurfaces. We have scaled the distribution so that the maximum value is 1. The white at location (0.8, 0.6, 1.05) corresponds to 1, and the black elsewhere corresponds to values near zero.

To invert the signals we maximize the log probability, which corresponds to minimizing a prior term plus the sum of squared error between the measured value and that predicted by the current estimate of the hand position.

Figure 13 shows a screen shot of the mouse. The user's hand motion is mapped onto the motion of the hand icon. The hand can pick up the small cube shown, move it around the space, and set it back on the floor. Because we cannot yet extract hand size, we have used a "sticky hand, sticky floor" protocol for grasping and releasing the cube. The small cube starts on the floor. When the hand first touches the cube, the hand closes, and the cube "sticks" to the hand and moves with it until the hand returns to the floor, at which point the hand opens and the cube sticks to the floor, where it remains until the hand returns. More information on the 3D mouse is available on the World Wide Web.[35]

Conclusion

The main technical obstacle to the use of Electric Field Sensing, the computational burden associated with inverting the signals, appears to be tractable. We are currently working on the problem of inferring parameters beyond position (for example, size and orientation) in order to make a "Gloveless DataGlove," and also on the problem of extracting low-resolution three-dimensional images from electric field measurements, a process we call Electric Field Tomography.[19]

Electric Field Sensing could profoundly affect people's mode of interaction with machines, and their expectations about the properties of objects generally. A common first reaction to a table with embedded electric field sensors is amazement, since it appears to be magic. However, the interaction soon begins to feel transparent, natural, and ordinary, rather than magical—why should you not be able to indicate your intentions to an object by waving your hands at it? Hume, in his famous argument on the impossibility of miracles,[36] says that the scope of the ordinary or natural can always be enlarged to subsume "magical" phenomena. Our experience with Electric Field Sensing seems mainly to support his position, but we have found there is an enjoyable transient period, before the user's experience has become routinized, in which magic is possible.

Acknowledgments

The 3D graphics code used to visualize the 3D mouse was written by Barrett Comisky. I thank the other members of the Physics and Media Group who worked on the development of the Fish board: Joe Paradiso, Tom Zimmerman, and Neil Gershenfeld. For helpful comments on earlier versions of this paper, I thank Neil Gershenfeld, Seth Lloyd, Stephen A. Benton, and the anonymous reviewers. I thank David MacKay and Steve Gull for teaching me the statistical techniques used in this paper. This work

Figure 11 The ambiguity classes for three single measurements made using the sensor geometry from the previous figure

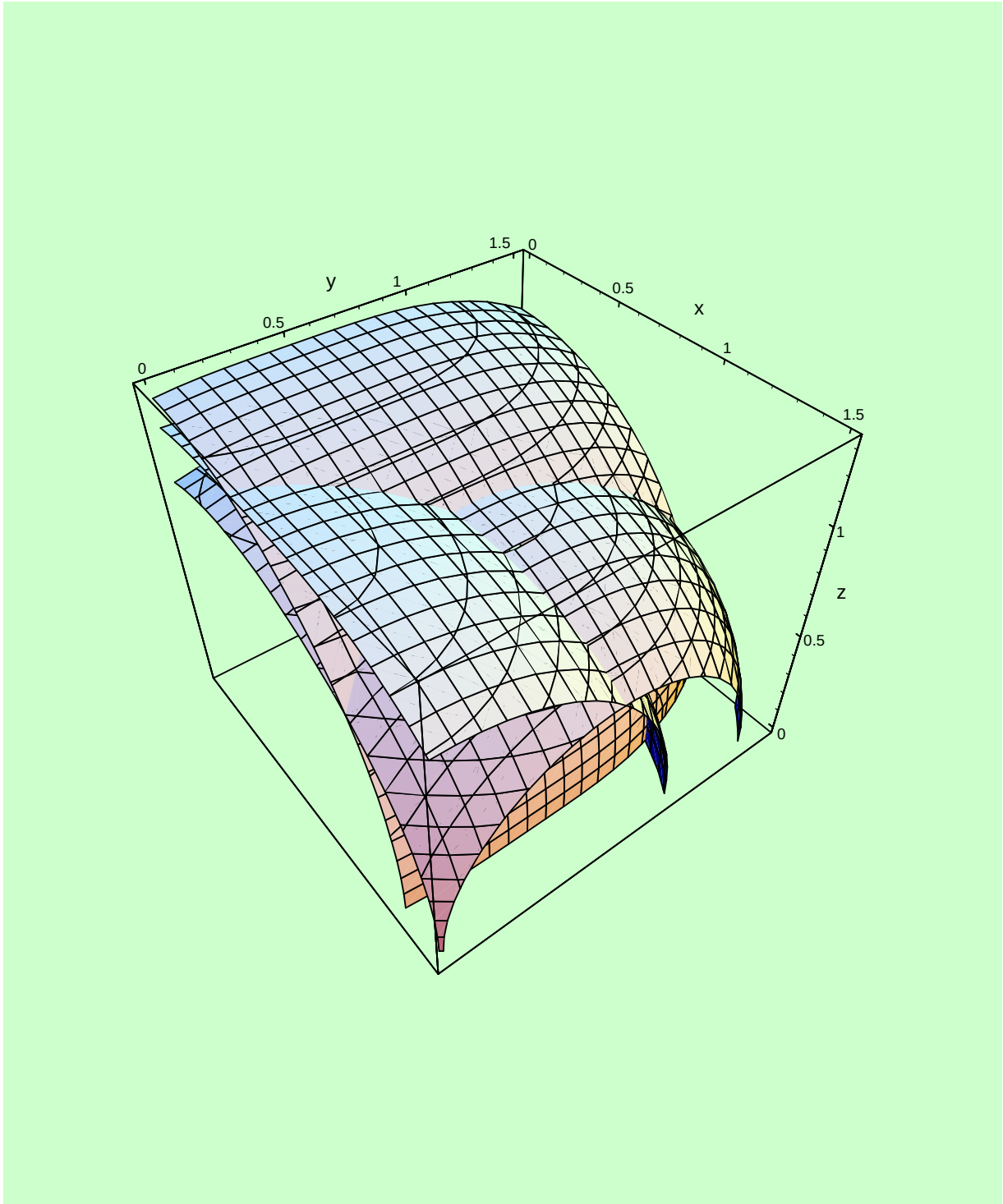


Figure 12 Posterior distribution over model parameters (x,y,z) for sensor geometry from the previous two figures: $p(x,y,z \mid D_1,D_2,D_3)$ for sensors 1, 2, and 3 given measurements $D_1 = f_1(0.8, 0.6, 1.05)$, $D_2 = f_2(0.8, 0.6, 1.05)$, and $D_3 = f_3(0.8, 0.6, 1.05)$

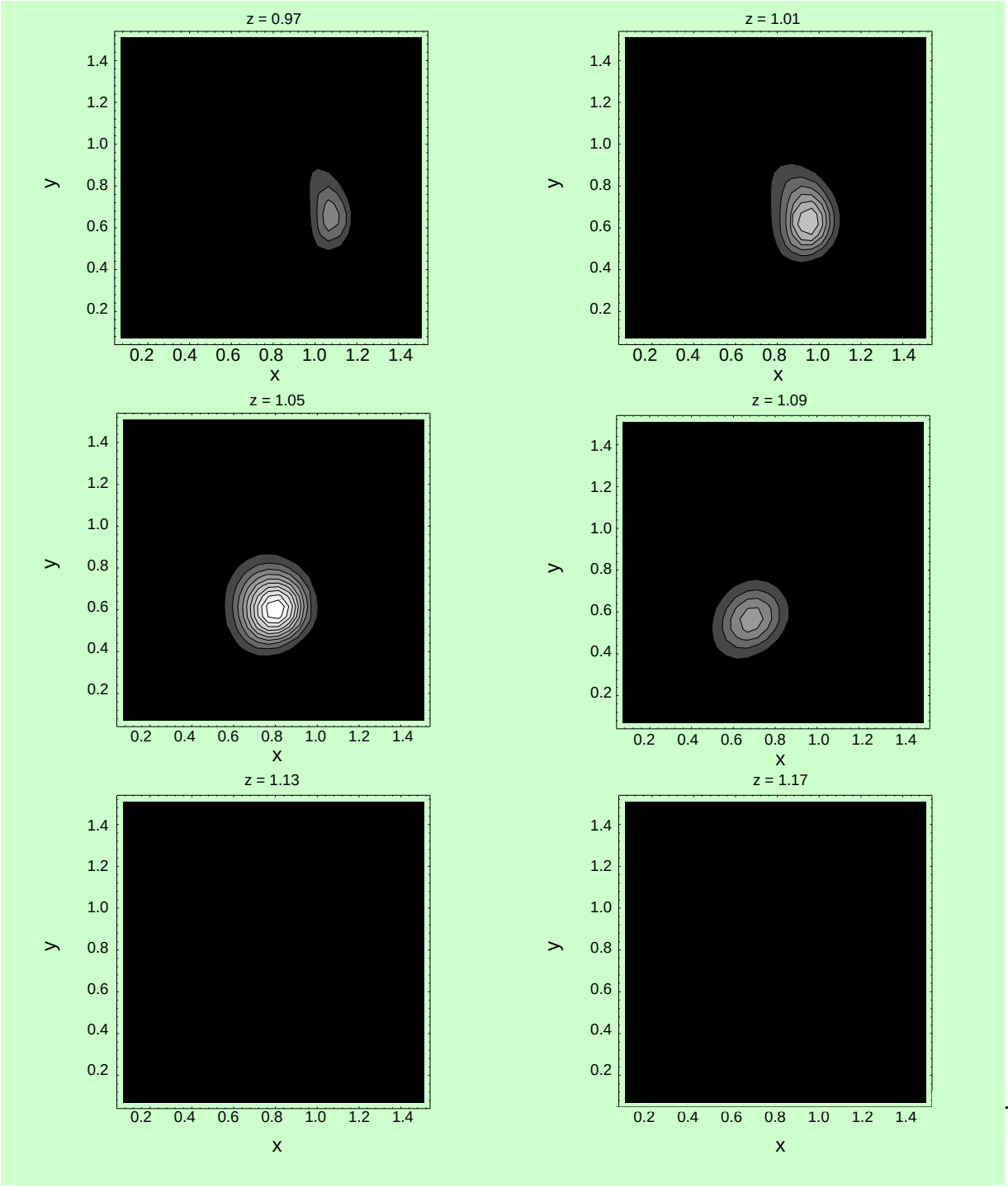
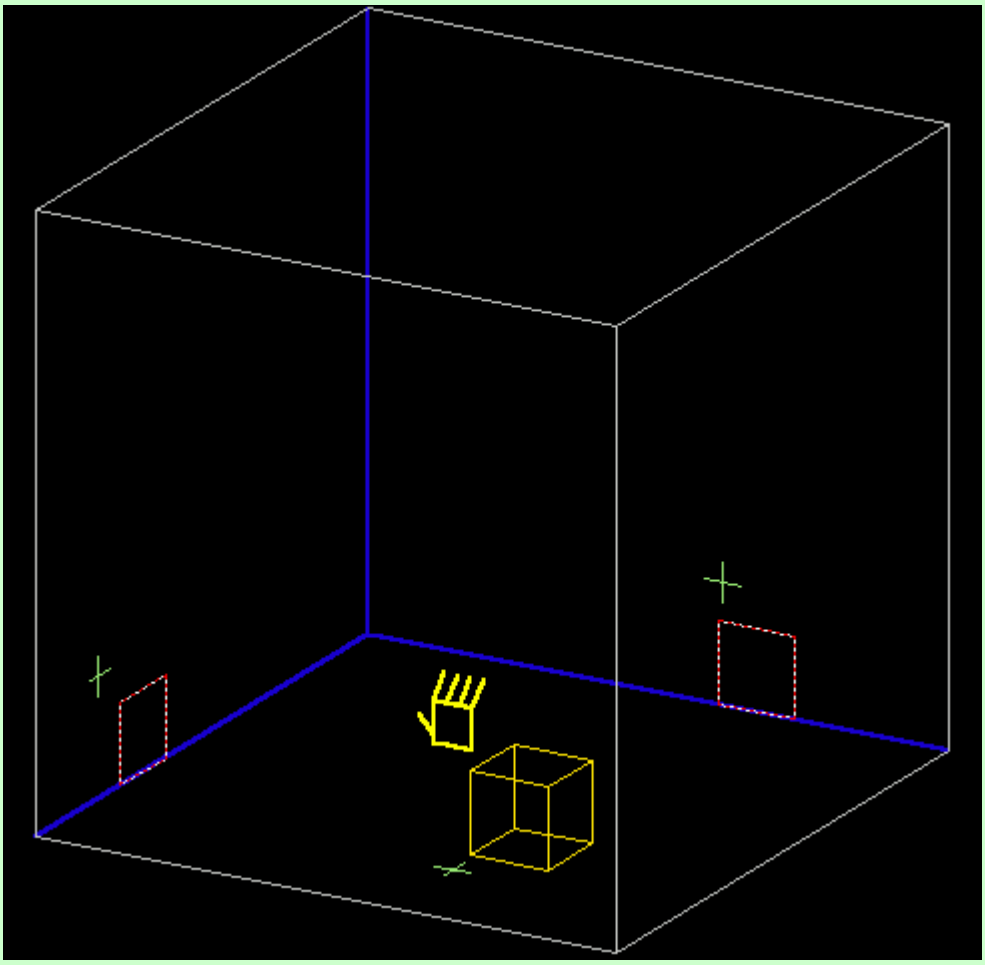


Figure 13 3D mouse



was supported in part by the Hewlett-Packard Corporation, Festo Corporation, Microsoft Corporation, Compaq Computer Corporation, the Media Lab's News in the Future and Things That Think consortia, and a Motorola fellowship.

*Trademark or registered trademark of International Business Machines Corporation.

Cited references and notes

1. S. K. Martin, *Theremin: An Electronic Odyssey*, film (1993).
2. B. M. Galeyev and L. S. Termen, "Faustus of the Twentieth Century," *Leonardo* **24**, No. 5, 573–579 (1991).
3. M. Nicholl, "Good Vibrations," *Invention and Technology*, Spring, 26–31 (1993).
4. J. Bastian, "Electrosensory Organisms," *Physics Today*, 30–37 (February 1994).
5. M. Matthews, *Three Dimensional Baton and Gesture Sensor*, U.S. Patent No. 4,980,519 (December 25, 1990).
6. Vranish et al., *Driven Shielding Capacitive Proximity Sensor*, U.S. Patent No. 5,166,679 (November 24, 1992).
7. Vranish et al., *Phase Discriminating Capacitive Array Sensor System*, U.S. Patent No. 5,214,388 (May 25, 1993).
8. T. G. Zimmerman, J. Lanier, C. Blanchard, S. Bryson, and Y. Harvill, "A Hand Gesture Interface Device," *ACM SIGCHI+GI'87 Conference on Human Factors in Computing Systems and Graphics Interface*, Denver, CO (1987), pp. 189–192.
9. N. Gershenfeld, *Method and Apparatus for Electromagnetic Non-Contact Position Measurement with Respect to One or More Axes*, U.S. Patent No. 5,247,261 (September 21, 1993).
10. N. Gershenfeld and J. R. Smith, *Displacement-Current Sensor and Method for Three-Dimensional Position, Orientation, and Mass Distribution*, U.S. Patent Application (February 3, 1994).
11. N. Gershenfeld, T. Zimmerman, and D. Allport, *Non-Contact System for Sensing and Signaling by Externally Induced Intra-Body Currents*, U.S. Patent Application (May 8, 1995).
12. N. Gershenfeld and J. R. Smith, *Displacement-Current Method and Apparatus for Resolving Presence, Orientation, and Activity in a Defined Space*, U.S. Patent Application (February 3, 1994).
13. T. G. Zimmerman, J. R. Smith, J. A. Paradiso, D. Allport, and N. Gershenfeld, "Applying Electric Field Sensing to Human-Computer Interfaces," *CHI'95 Human Factors in Computing Systems*, Denver, CO (1995), pp. 280–287.
14. J. A. Paradiso and N. Gershenfeld, "Musical Applications of Electric Field Sensing," *Computer Music Journal*, forthcoming.
15. J. A. Paradiso, "New Technologies for Monitoring the Precision Alignment of Large Detector Systems," *Nuclear Instruments and Methods in Physics Research, Section A*, submitted June, 1996.
16. D. Allport, T. G. Zimmerman, J. A. Paradiso, J. R. Smith, and N. Gershenfeld, "Electric Field Sensing and the Flying Fish," in the ACM-Springer *Multimedia Systems Journal* special issue on Multimedia and Multisensory Virtual Worlds (Spring 1996).
17. It is called "Fish" because electric fish use similar mechanisms to sense their environments, and because we hope that The Fish, which navigates in three dimensions, might be the successor input device to the mouse, which only navigates in two.
18. Physics and Media Group, *Smart Fish Technical Manual*, MIT Media Laboratory, Cambridge, MA (1995).
19. J. R. Smith, *Toward Electric Field Tomography*, master's degree thesis, MIT Media Laboratory, Cambridge, MA (1995).
20. R. M. Fano, L. J. Chu, and R. B. Adler, *Electromagnetic Fields, Energy, and Forces*, John Wiley & Sons, Inc., New York (1960).
21. D. C. Barber and B. H. Brown, "Applied Potential Tomography," *Journal of Physics E* **17**, 723–733 (1984).
22. T. G. Zimmerman, *Personal Area Networks (PAN): Near-Field Intra-Body Communication*, master's degree thesis, MIT Media Laboratory, Cambridge, MA (1995).
23. J. A. Paradiso, personal communication.
24. J. A. Paradiso, *Gesture Wall Electronics*, technical report, Brain Opera, MIT Media Laboratory, Cambridge, MA (1996).
25. The surface charge distribution is not in fact uniform, although the surface is an equipotential.
26. E. T. Jaynes, "Prior Information and Ambiguity in Inverse Problems," *SIAM AMS Proceedings* **14**, 151–166 (1983).
27. G. T. Herman, H. Hurwitz, and A. Lent, "A Bayesian Analysis of Image Reconstruction," G. T. Herman, Editor, *Image Reconstruction from Projections: Implementations and Applications*, Springer-Verlag, Berlin (1979), pp. 85–103.
28. S. F. Gull and G. J. Daniell, "Image Reconstruction from Incomplete and Noisy Data," *Nature* **272**, 686–690 (1978).
29. J. Skilling and S. F. Gull, "The Entropy of an Image," *SIAM AMS Proceedings* **14**, 167–188 (1983).
30. It will not represent conductivity in this section.
31. D. J. C. MacKay, *Bayesian Methods for Adaptive Models*, Ph.D. thesis, California Institute of Technology, Pasadena, CA (1991).
32. F. Morgan, *Riemannian Geometry: A Beginner's Guide*, Jones and Bartlett, Boston (1993).
33. S. P. Luttrell, "The Use of Transinformation in the Design of Data Sampling Schemes for Inverse Problems," *Inverse Problems* **1**, 199–218 (1985).
34. D. V. Lindley, "On a Measure of the Information Provided by an Experiment," *Annals of Mathematical Statistics* **27**, 986–1005 (1956).
35. J. R. Smith, *Field Mouse*, World Wide Web page http://jrs.www.media.mit.edu/~jrs/fieldmouse.html.
36. D. Hume, *An Enquiry Concerning Human Understanding* (1789).

Accepted for publication May 31, 1996.

Joshua R. Smith *MIT Media Laboratory, 20 Ames Street, Cambridge, Massachusetts 02139-4307 (electronic mail: jrs@media.mit.edu).* Mr. Smith is a Ph.D. candidate and research assistant in the Physics and Media Group of the MIT Media Laboratory, and is interested in the physics of information and computation. He has an M.S. from the Media Laboratory, a B.A. Hons. in natural sciences (physics and theoretical physics) from Cambridge University, and a double B.A. in computer science and philosophy from Williams College. He has had internships at the Santa Fe Institute, Los Alamos National Laboratory, the Yale University Department of Computer Science, the Williams College SMALL Geometry Research Group, and NASA's Goddard Institute for Space Studies.

Reprint Order No. G321-5626.