A Low-cost Transparent Electric Field Sensor for 3D Interaction on Mobile Devices

Mathieu Le Goc, Stuart Taylor, Shahram Izadi, Cem Keskin

A Low-cost Transparent Electric Field Sensor for 3D Interaction on Mobile Devices

Mathieu Le Goc[2], Stuart Taylor[1], Shahram Izadi[1], Cem Keskin[1]
[1]Microsoft Research [2]INRIA

ABSTRACT
We contribute a thin, transparent, and low-cost design for electric field sensing, allowing for 3D finger and hand tracking, as well as in-air gestures on mobile devices. Our approach requires no direct instrumentation of the hand or body, and is non-optical, allowing for a compact form-factor that is resilient to ambient illumination. Our simple driver electronics are based on an off-the-shelf chip that removes the need for building custom analog electronics. We describe the design of our transparent electrode array, and present a machine learning algorithm for mapping from signal measurements at the receivers to 3D positions. We demonstrate non-contact motion gestures, and precise 3D hand and finger localization. We conclude by discussing limitations and future work.

Author Keywords
NUI, 3D sensing, electric-field sensing, mobile, non-optical

INTRODUCTION
Natural user interfaces (NUI) have become synonymous with contact-less 3D interaction using whole body, hands and fingers, exemplified by technologies such as Kinect and Leap Motion. At the core of many technologies in this domain is the *camera*. Examples include stereo, structured light, and time-of-flight cameras (see [16] for a review).

Whilst the camera offers a rich signal and great flexibility, it has limitations. These include: 1) the need to keep a minimal distance between the sensor and the scene, leading to line-of-sight and form-factor constraints. 2) the computational overheads of processing depth or RGB images to infer user input. 3) sensitivity to lighting changes and outdoor ambient light. And 4) the high power requirements of cameras, in particular depth cameras with active illumination. These limitations are a barrier particularly for mobile interaction, where form-factor, power, outdoor use, and computational costs are significant considerations.

These challenges of camera-based systems have led to other forms of 3D contact-less sensing in the context of mobile interaction. *Optical proximity sensors* have been used in a variety of wearable devices [19, 14, 18]. Mobile phones such as the Samsung Galaxy S4 use a single front-facing proximity sensor for coarse in-air gestures. Researchers have also

CHI 2014, April 26 - May 01 2014, Toronto, ON, Canada
Copyright 2014 ACM 978-1-4503-2473-1/14/04...$15.00.
http://dx.doi.org/10.1145/2556288.2557331/

Figure 1: Our transparent electric field sensing design can be placed on top of a mobile device to enable 3D input and motion gestures.

demonstrated proximity sensors around the periphery of a mobile phone to enable input beyond the screen [15, 2]. These systems can suffer from very coarse sensing resolution, lack of transparency, form-factor and line-of-sight considerations, and issues regarding ambient lighting. *Magnetic field sensing* has demonstrated relative 1D, 2D and even 3D tracking by pairing permanent magnets with magnetometers either by wearing both [3, 10] or leveraging the inertial measurement unit (IMU) available on most mobile devices today [11]. However, the sensing resolution can be coarse and limited to specific degrees-of-freedom (DoFs). These systems also require instrumentation of the user's hand, which acts as a barrier to unencumbered interactions.

This paper addresses some of these challenges, and focuses on another type of contact-less 3D sensing technology for mobile interaction based on *electric-field sensing* (EFS). Specifically, we contribute a thin, transparent, and low-cost design for EFS, allowing for 3D fingertip and hand tracking, as well as in-air gestures on mobile devices. Our sensor design requires no direct instrumentation of the hand or body. Further, it is non-optical, allowing for a compact form-factor that is resilient to ambient illumination. Our simple driver electronics is based on an off-the-shelf chip that removes the need to build complex front-end analog electronics. We cover considerations in our new antenna design, and present a machine learning algorithm for mapping from signal measurements at the receivers to 3D finger positions. We demonstrate non-contact 3D motion gestures, and precise hand and finger localization. Finally, we conclude by discussing limitations and future work.

ELECTRIC FIELD SENSING
The term electric field sensing (EFS) and *capacitive sensing* broadly describe proximity and touch sensing technologies which aim to sense variation in a generated electric field to

localize surrounding objects, such as the user's hands and fingers. Three main sensing configurations exist [29, 23, 24].

Loading-mode is the most typical configuration used in capacitive touch-screens and proximity sensors (such as the Theremin [8] or more recently by [27]). A periodic electric signal is applied to a single electrode, forming an oscillating electric field. As the user approaches the electrode, conductive properties of the human body form a weak capacitive link with the electrode, causing a signal change. By sensing the degree of this change, touch or proximity is detected. Typically the input frequency is kept constant, although sweeping though a range of input frequencies has been shown to reveal additional user information [22].

Shunt-mode is an alternative approach to the "traditional" mechanism for capacitive sensing above, and was explored in seminal work in the 90s [29, 23, 21, 24], and more recently [9]. Again, an electric field is generated by applying a periodic signal to a transmit electrode. This field is then captured by one or more receive electrodes in proximity. Any grounded object placed in the field leads to disturbances. The conducting properties of the human body absorbs some of radiated field and *shunts* it to ground, occasioning variations in signals at the receive electrodes. Systems such as [24, 17, 9] combine electrodes in a novel layout to support 3D sensing, hand posture detection, and even geometry reconstruction.

This sensing configuration is typically known as *active* EFS. *Passive* EFS measures existent electric fields, a bi-product of noise from power lines and appliances, which can lead to novel sensing of coarse whole body gestures, activity sensing and other interactive scenarios [5, 6, 4, 20]. Passive EFS requires a dedicated sensor placed on the body or fixed in the environment, and lacks the precision for true 3D tracking.

Transmit-mode is a final variation, where the input signal is coupled with a person's body. This coupling turns the user into a transmitter whose signal can be picked up by one or more receivers. This technique has been applied for user identification on multi-touch tabletops [7], as well as personal area networks [28], and user localization [26].

As shown EFS has been explored for 3D sensing and human computer interaction, but there is one area that the potential for EFS has been under explored, and that is *mobile sensing*. In our work we present a transparent and low-cost EFS sensor design, which can be leveraged alongside a new off-the-shelf chip, to readily create new mobile EFS sensing scenarios. Our approach makes mobile EFS sensing extremely simple and cheap to experiment with, and will hopefully lead to new application possibilities.

EFS SENSING CHIP

To simplify the design of our prototype, we adopted the use of a new low cost, off-the-shelf EFS sensing chip (MGC3130) from Microchip Inc. [12]. The device comprises of the die-level hardware necessary to perform EFS sensing (analog front-end and an I^2C digital interface), combined with a custom firmware library enabling control over the sensing process. When coupled with an external electrode array, the device senses 3D hand and finger positions within a defined region dictated by the electrode configuration.

The recommended electrode arrangement places four receive electrodes along the boundary of the rectangular sensing region, with a fifth receive electrode in the center. Below this layer sits the transmit electrode which extends to cover the entire sensing region. This arrangement is most easily constructed using a multi-layer PCB. The sensor is both low-power and high frame rate, and streams sensor readings at 200Hz in normal operation, with a typical power draw of 20mA; this drops to around $70\mu A$ in 'approach' mode, and $9\mu A$ in sleep mode. We have extended this design to create a transparent electrode array, and an algorithm with greatly improved 3D localization.

TRANSPARENT LOW-COST MOBILE EFS

Based on the Microchip sensor, our current prototype shown in Fig. 1 employs an NXP LCP1768 microcontroller, which provides a bridge between the I^2C interface and either USB or (via a separate module) Bluetooth for streaming of sensor readings. We have designed a transparent single-layer electrode array measuring 11x7cm, fabricated from a thin sheet of polyester film coated with indium tin oxide (ITO) (with intrinsic resistance of $50\Omega/sq$, and optical transmittance of $\sim$80% in the visible band). The single-layer design minimizes complexity and significantly reduces prototyping and production time. Our early multi-layer designs, although functional, were difficult to construct as they required careful alignment and inter-layer connections.

Fig. 2 shows our final electrode array design and supporting drive electronics. To create our electrode design, we use a laser cutter to score directly into the surface of the film, thereby creating separate electrically isolated regions. Using a single layer design can result in large levels of crosstalk and wide variations in the sensitivity of each of the receive electrodes, so to help mitigate these problems we have identified the following steps: (1) completely surround the receive-electrode feeder lines with the Tx electrode; although counter-intuitive, this results in a more uniform 'background' Tx signal that the sensing chip is able to filter out, and (2), minimize the width and length of the Rx feeder lines and to route them away from other Rx electrodes.

In some of our early designs, we experimented with gold based coatings (AuARE) and also a new carbon nanotube film [25]. When considering the choice of coating, the intrinsic resistance of the film becomes important: as mentioned above, we want to minimize the width of the feeder lines, however, this also increases their resistance. Using a width of 0.5mm for the feeder lines on ITO, we are able to achieve an end-to-end resistance of below $20k\Omega$. Coatings such as the carbon nanotubes exhibit values closer to $100k\Omega$, which significantly degrades the performance of the sensor.

The scoring process using the laser cutter required careful fine tuning. The goal is to create electrically isolated regions, without damaging the surrounding coating and also to minimize the optical patterning of the surface of the film. To achieve this, we opted for a multi-pass approach using a relatively low-power laser setting. This process typically required between 3-7 passes, but resulted in narrower and 'cleaner' score lines than a single higher-power pass. In our current design, the score lines are approximately 0.3mm wide.

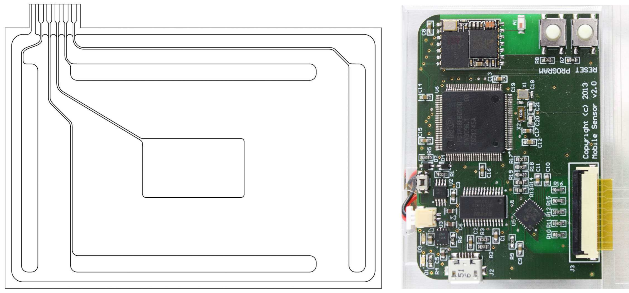

Figure 2: (a) Electrode array design, (b) drive electronics

Although the design of our sensor is achieved using a single sheet of ITO coated film, we have found in practice that when placed over a display, the display's drive electronics causes sensor interference. To overcome this and to help minimize the effects of other environmental noise, we introduce a second unetched sheet of ITO film between the sensor and the screen which effectively shields the sensor from noise. In a future implementation, this ground plane could be applied to the reverse side of the main sensor sheet, and additionally, the different electrode regions would typically be created using masking during the ITO production phase.

LEARNING-BASED 3D LOCALIZATION

The high level Microchip API that provides 3D positions is ultimately a black box and it is unclear how the low-level signal is processed. We have observed large errors when using this black-box 3D localization algorithm. Therefore we are interested in efficiently and more accurately estimating the finger or hand location by directly looking at the sensor signals (exposed through the same Microchip API).

This can be formulated as a regression problem, since we need to find a mapping from the 5D measurements into 3D world coordinates. However there are a couple of problems that make this task challenging. First of all, the hand can be in many different poses while it retains the same fingertip location. For all of these configurations, the 5D signal will be different. Thus the mapping needs to be many–to–one. The sensor noise further complicates the task, since the mapping becomes many–to–many. As a result, analytical closed-form solutions to the estimation problem will not yield reliable results. Instead, we use machine learning methods to approximate this complicated mapping. In particular, we use a random decision forest (RDF) for regression [1]. The main strength of RDFs is their speed and accuracy. Furthermore, RDFs compare favorably to a dense 5D lookup table approach, which requires a much larger memory footprint for storing samples (limiting mobile scenarios).

RDF-based Regression

Using a regression forest, we want to find a mapping from the 5D sensor measurements $\boldsymbol{x} = [x_1, x_2, x_3, x_4, x_5]$ to the 3D world coordinates $\boldsymbol{y} = [X, Y, Z]$. RDFs consist of internal split nodes and leaf nodes. Split nodes extract certain features, evaluate a test function and send the incoming input to their children, whereas leaf nodes contain statistics or a local model. For each split node, a set of test functions are learned during training. Each input sample is evaluated with all the trees independently until a leaf node is reached. The estimations from different trees are then averaged to get the final result.

In this work, we use normal distributions over the 3D world coordinates as the local models. Test functions at split nodes are of the form: $f(F) < T$. Here we use a very simple feature, namely the measurement from a randomly selected electrode (s). This reading is then compared against some threshold T:

$$f(s) = x_s, s \in 1, \ldots, 5 \tag{1}$$

This splits the data into two sets and sends each to a child node. The quality of a split is measured by how much the total variance is reduced:

$$G(s,T) = \mathrm{tr}(\Sigma_C) - \sum_{m \in \{L,R\}} \frac{|C_m|}{|C|} \mathrm{tr}(\Sigma_{C_m}) \tag{2}$$

where $\mathrm{tr}(\Sigma_K)$ is the trace of the covariance matrix for a set of observations K, and $|K|$ denotes the number of samples in the set. C denotes the set of observations prior to the split, and C_L and C_R are the sets of observations after the split. At training time, we exhaustively select all the s values, sample multiple thresholds uniformly from a large range, and select the feature which maximizes G. Once leaves are reached, the mean and variance of the 3D world coordinates are stored at each node. At run-time, $\boldsymbol{x}$ is evaluated with every tree and the means at each leaf are averaged to get the final estimation.

To apply this technique, we collected 8000 $\boldsymbol{x}$, $\boldsymbol{y}$ pairs, from four participants with different hand sizes using a high-precision stereo camera setup [13]. This yields a dataset of 32000 samples. We use half the examples for training, half for test, and optimize the forests through grid search over the depth and forest size parameters. We use six trees of depth 17. Our results for 3D accuracy are shown in Fig. 3 (left). The best average error is estimated to be 9.65mm (error is the 3D distance from current estimate to ground truth). Note this is a per-frame regression result, without temporal filtering. With Kalman filtering, our average accuracy is ~4mm. This compares favorably with our experiments using the Microchip 3D measurements (taken from our stereo rig), which yields much higher error rates $>$ 20mm. Fig. 3 (right) shows a plot of our X,Y accuracy across varying depths compared to the Microship sensor.

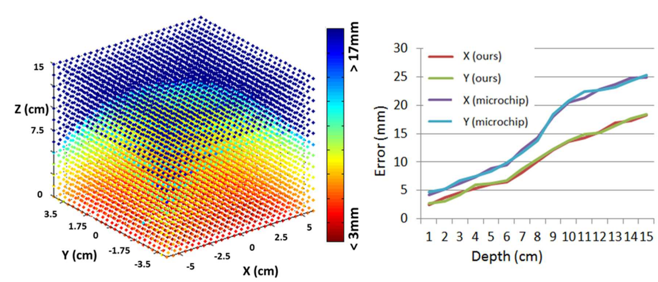

Figure 3: Left: The error (in millimeters) for our 3D localization algorithm. Right: X,Y error (in millimeters) across depths, for our method and Microchip baseline.

In realizing our EFS design and machine learning algorithms we have built a variety of demonstrations, which allow simple visualization of the 3D input sensed from the device, continuous touchless control, and simple in-air swipe gestures. Fig. 4 and the accompanying video show examples, including

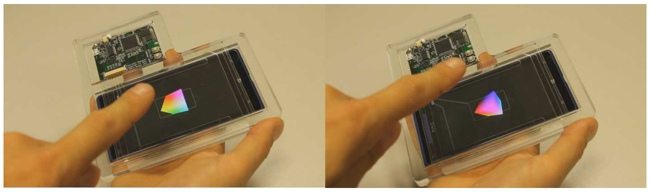

Figure 4: 3D input capabilities of our sensor in a mobile context. Please also see accompanying video.

a wireless and self contained setup which allows 3D input on top of a regular mobile phone.

LIMITATIONS AND DISCUSSION

Whilst we have built a proof-of-concept mobile device which enables precise EFS sensing, there are challenges and limitations that need to be overcome, and exciting areas of future work. As many researchers have documented, EFS can be a very challenging sensing technique to work with [29, 23, 21, 24]. Perhaps the most challenging is environmental interference, which can be an issue in mobile environments. Our initial investigations of testing the sensor whilst moving through a variety of test environments is encouraging however. This is partly as we include suitable grounding to reduce interference, but also because we are actively sensing in a limited 3D volume (above the screen). One clear area of future work which we have yet to explore is how to incorporate touch sensing either using the projective capacitive sensor available on most mobile devices in combination with EFS, or by directly leveraging our sensor.

CONCLUSIONS

In this paper we have contributed a transparent EFS sensor configuration for 3D mobile interaction, and an associated machine learning algorithm for interpreting this signal into 3D input. Given the simplicity of the antenna design, the wide availability of the Microchip sensor, and the low-cost nature of the approach, we believe that this will enable other practitioners to revisit the principle of EFS sensing in exciting new contexts, such as mobile computing.

REFERENCES

1. Breiman, L. Random Forests. *Machine Learning 45*, 1 (Oct. 2001), 5–32.
2. Butler, A., Izadi, S., and Hodges, S. Sidesight: multi-touch interaction around small devices. In *UIST'08*, ACM (2008), 201–204.
3. Chen, K., Lyons, K., White, S., Patel, S. utrack: 3d input using two magnetic sensors. In *UIST'13*, ACM (2013).
4. Cohn, G., Gupta, S., Lee, T.-J., Morris, D., Smith, J. R., Reynolds, M. S., Tan, D. S., and Patel, S. N. An ultra-low-power human body motion sensor using static electric field sensing. In *Ubicomp'12*, ACM (2012), 99–102.
5. Cohn, G., Morris, D., Patel, S., and Tan, D. Humantenna: using the body as an antenna for real-time whole-body interaction. In *CHI'12*, ACM (2012), 1901–1910.
6. Cohn, G., Morris, D., Patel, S. N., and Tan, D. S. Your noise is my command: sensing gestures using the body as an antenna. In *CHI'11*, ACM (2011), 791–800.
7. Dietz, P., and Leigh, D. Diamondtouch: a multi-user touch technology. In *UIST'01*, ACM (2001), 219–226.
8. Glinsky, A. *Theremin: ether music and espionage*. University of Illinois Press, 2000.
9. Grosse-Puppendahl, T., Braun, A., Kamieth, F., and Kuijper, A. Swiss-cheese extended: an object recognition method for ubiquitous interfaces based on capacitive proximity sensing. In *CHI'13*, ACM (2013), 1401–1410.
10. Han, X., Seki, H., and Hikizu, M. Wearable handwriting input device using magnetic field. In *SICE'07*, IEEE (2007), 365–368.
11. Harrison, C., and Hudson, S. E. Abracadabra: wireless, high-precision, and unpowered finger input for very small mobile devices. In *UIST'09*, ACM (2009), 121–124.
12. Inc, M. http://www.microchip.com/gestic, 2013.
13. Kim, D., Izadi, S., et al. Floodlight: 3d silhouette sensing for high-precision input on and above physical surfaces. In *CHI'14*, ACM (2014), Forthcoming.
14. Kim, J., He, J., Lyons, K., and Starner, T. The gesture watch: A wireless contact-free gesture based wrist interface. In *ISWC'07*, IEEE (2007), 15–22.
15. Kratz, S., and Rohs, M. Hoverflow: expanding the design space of around-device interaction. In *MobileHCI'09*, ACM (2009).
16. Lanman, D., and Taubin, G. Build your own 3d scanner: 3d photography for beginners. In *ACM SIGGRAPH 2009 Courses*, ACM (2009), 8.
17. Lee, J., Park, K. S., and Hahn, M. The 3d sensor table for bare hand tracking and posture recognition. In *Advances in Multimedia Modeling*. Springer, 2006, 138–146.
18. Lee, S. C., Li, B., and Starner, T. Airtouch: Synchronizing in-air hand gesture and on-body tactile feedback to augment mobile gesture interaction. In *ISWC'11*, IEEE (2011), 3–10.
19. Metzger, C., Anderson, M., and Starner, T. Freedigiter: A contact-free device for gesture control. In *ISWC'04*, vol. 1, 18–21.
20. Mujibiya, A., and Rekimoto, J. Mirage: exploring interaction modalities using off-body static electric field sensing. In *UIST'13*, ACM (2013), 211–220.
21. Paradiso, J. A., and Gershenfeld, N. Musical applications of electric field sensing. *Computer music journal 21*, 2 (1997), 69–89.
22. Sato, M., Poupyrev, I., and Harrison, C. Touché: enhancing touch interaction on humans, screens, liquids, and everyday objects. In *CHI'12*, ACM (2012), 483–492.
23. Smith, J. R. Field mice: Extracting hand geometry from electric field measurements. *IBM systems journal 35*, 3.4 (1996), 587–608.
24. Smith, J. R. *Electric field imaging*. PhD thesis, MIT, 1998.
25. Unidym. http://www.unidym.com/, 2013.
26. Valtonen, M., Maentausta, J., and Vanhala, J. Tiletrack: Capacitive human tracking using floor tiles. In *Percom'09*, IEEE (2009), 1–10.
27. Wimmer, R. Capacitive sensors for whole body interaction. In *Whole Body Interaction*. Springer, 2011.
28. Zimmerman, T. G. Personal area networks: near-field intrabody communication. *IBM systems Journal 35*, 3.4 (1996), 609–617.
29. Zimmerman, T. G., Smith, J. R., Paradiso, J. A., Allport, D., and Gershenfeld, N. Applying electric field sensing to human-computer interfaces. In *CHI'95* (1995), 280–287.