uTouch: Sensing Touch Gestures on Unmodified LCDs

Ke-Yu Chen[1], Gabe Cohn[1], Sidhant Gupta[2], Shwetak N. Patel[1,2]
Electrical Engineering[1], Computer Science & Engineering[2]
UbiComp Lab, DUB Group, University of Washington
Seattle, WA (USA)
{kychen, gabecohn, sidhant, shwetak}@uw.edu

ABSTRACT

Current solutions for enabling touch interaction on existing non-touch LCD screens require adding additional sensors to the interaction surface. We present *uTouch*, a system that detects and classifies touches and hovers without any modification to the display, and without adding any sensors to the user. Our approach utilizes existing signals in an LCD that are amplified when a user brings their hand near or touches the LCD's front panel. These signals are coupled onto the power lines, where they appear as electromagnetic interference (EMI) which can be sensed using a single device connected elsewhere on the power line infrastructure. We validate our approach with an 11 user, 8 LCD study, and demonstrate a real-time system.

Author Keywords
LCD; touch; capacitive sensing; EMI

ACM Classification Keywords
H.5.m. Information interfaces and presentation: Miscellaneous.

General Terms
Human Factors; Design.

INTRODUCTION AND RELATED WORK

Although touch-enabled user interfaces are rapidly becoming popular for mobile applications, the cost and complexity of large touch displays have dramatically limited their use for desktop computers. In the past, researchers have used optical sensors such as cameras [1] and infrared sensors [3, 5, 6]. Although these systems typically provide finger-level accuracy for multiple touch points, they may not scale well in all situations, as each display needs its own sensor and communication mechanism. Additionally, most of these systems can only detect when the user is touching the surface, and are unable to sense when a hand is approaching the surface or hovering above it. Specialized capacitive sensors are typically used to detect when a hand is approaching or hovering [7].

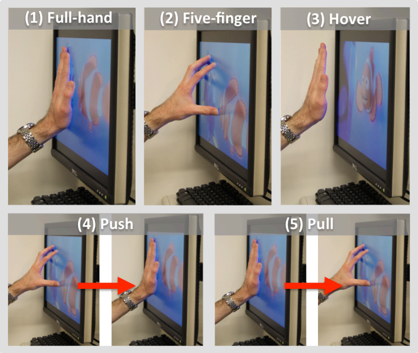


Figure 1: uTouch detects and classifies 5 gestures without any instrumentation to either the LCD or the user: (1) full-hand touch, (2) five-finger touch, (3) hover, (4) push, and (5) pull.

In this work, we present *uTouch*, a system that enables touch and hover interaction on existing non-touch LCD monitors without requiring any instrumentation to either the display or the user. Our approach utilizes existing signals in the LCD display, which are amplified when a user brings their hand near or touches the monitor's front panel. These signals are coupled onto the power lines, where they appear as electromagnetic interference (EMI) which can be sensed using a single device connected elsewhere on the power line infrastructure. Additionally, a *single* sensor can be used to sense touches on *many* displays connected to the same power line infrastructure (*i.e.*, home or office space).

This approach of indirect capacitive sensing using EMI was first shown in LightWave, which detects the capacitive coupling between a user and an uninstrumented compact fluorescent lamp (CFL) [4]. uTouch is similar in spirit to LightWave; however, the EMI noise signal being sensed and phenomenon causing the changes in that signal are quite different. In LightWave, the presence of the human body near the CFL detunes the oscillator in the switched-mode power supply, and thus changes the amplitude and frequency of the noise signals on the power line. In uTouch however, the power line EMI used for sensing touches is produced by the timing control signals in the LCD panel, and those signals are amplified due to increased power consumption of the LCD when a user touches the panel.

In this paper, we describe the theory of operation and present a feasibility study demonstrating the ability to detect and classify 5 different gestures (see Figure 1) across 11 users and 8 LCD displays (on both desktops and laptops). We describe the signal processing and machine learning needed to implement the uTouch system, and present a *real-time* demonstration (see the video figure).

CHI'13, April 27 – May 2, 2013, Paris, France.
Copyright 2013 ACM 978-1-4503-1899-0...$10.00.

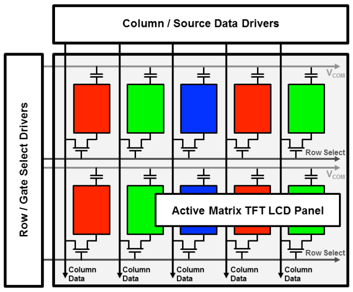


Figure 2: Simplified schematic of a small section of the LCD panel and the row and column drivers.

THEORY OF OPERATION

In LCD monitors, a backlight produces uniform white light using either a cold cathode fluorescent lamp (CCFL) or light emitting diodes (LED). The white light passes through a polarizer, liquid crystal (LC), color filter, and a second polarizer before being emitted at the front of the display. The intensity of the light is controlled by the strength of the electric field applied to the LC. Pixels are made by closely grouping red, green, and blue colored filters, which visually combine to produce the desired color.

Although the panel is made of a large array of pixels, only a single row of pixels is on at any time, and therefore small thin-film transistors (TFT) are used to enable each pixel. Figure 2 shows a small section of an LCD panel array. With the gate voltage applied only to the active row, a field is created on all electrodes in that row. Each row is selected once per frame, and enabled periodically at the refresh rate. We will refer to the rate at which the display switches active rows as the *row rate*. The row rate is dictated by the refresh rate of the display (commonly 60 Hz) and the number of rows (*i.e.*, the native resolution), and does not change when the driving resolution is changed.

As explained above, the row select lines and column data lines are changed every row, at the row rate. As a result of this, the row and column drivers consume power in bursts at the row rate. In the same way that current spikes from a digital clock couple EMI onto the power line, the current spikes from the row and column drivers result in EMI on the power line at harmonics of the row rate [2]. On some monitors, EMI is also observed at harmonics of half of the row rate. This is because some LCDs group adjacent rows in what is called line-paired inversion. In this case, if we assume that the colors of nearby pixels are typically similar, then the voltages on the column data lines only change significantly every other row. This will therefore cause EMI at half of the row rate.

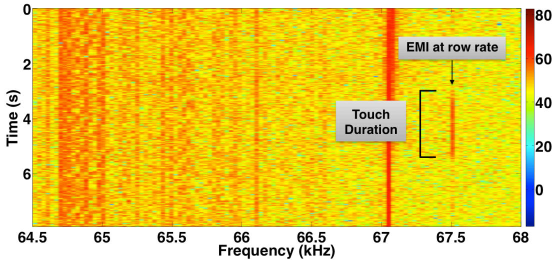


Figure 3: EMI seen at the row rate of 67.5 kHz is observed well above the noise level during the user's touch of the LCD.

We note that since the EMI is produced by multiplexing the rows of a panel, it is thus independent of the backlight technology (*i.e.*, LED vs. CCFL) and independent of the pixel-level electrode configuration (*i.e.*, TN vs. IPS). All such varieties of panels have the same type of TFT array and therefore produce EMI at the row rate in the same way.

Sensing Panel Touches

Although the row rate EMI on the power lines is typically below the noise level, when a user's hand hovers over or touches the panel, a very large capacitance to ground is added in parallel with the row select lines and column data lines. This added capacitance results in significantly higher power consumption by the row and column drivers, which causes higher levels of EMI at harmonics of the row rate (and half of the row rate in LCDs that use line-paired inversion). This EMI is both conducted onto the power line and radiated onto the power lines by the panel and user. The resulting EMI on the power line can then be seen well above the noise level as shown in Figure 3. Furthermore, the relative amplitude change of this EMI is a function of the strength of the capacitive coupling between the panel and the hand. This noise is therefore a robust signal for sensing different kinds of touches and hovers on the panel.

EXPERIMENTAL PROCEDURE AND ANALYSIS

We set up an experiment in a real home environment and conducted a user study of 11 participants (3 female) on 6 LCD monitors (M1-M6) and 2 laptops (L1-L2). Table 1 shows the variety of panel technologies used in the study.

We used 5 different touch gestures as shown in Figure 1:

ID	Model	Size / Resolution	Panel / Backlight	Interface	Refresh Rate (Hz)	Row Rate (kHz)
M1	Asus VW246H	24 in / 1920x1080	TN / CCFL	HDMI (to PC)	60.0	67.5
M2	ViewSonic VX2035wm	20.1 in / 1680x1050	TN / CCFL	DVI (to PC)	60.0	65.3
M3	Samsung 226BW	22 in / 1680x1050	TN / CCFL	DVI (to PC)	59.9	64.7
M4	Dell 2007WFP	20 in / 1680x1050	S-IPS / CCFL	DVI (to PC)	60.0	64.7
M5	Dell 1703FPs	17 in / 1280x1024	TN / CCFL	VGA (to Laptop)	60.0	64.0
M6	HP S2231	21.5 in / 1920x1080	TN / CCFL	DVI (to PC)	60.0	67.5
L1	Acer ASPIRE5736Z	15.6 in / 1366x768	TN / LED	N/A	60.0	47.1
L2	Dell INSPIRON1545	15.6 in / 1366x768	TN / CCFL	N/A	60.0	49.4

Table 1: Specifications of the six monitors (M1-M6) and two laptops (L1, L2) used in the experiment.

full-hand touch (FH), five-finger touch (FF), hover, push and pull. In a push gesture, the user was asked to first perform a five-finger touch, followed by a full-hand touch (*i.e.*, the hand is "pushed" toward the panel), and the pull was exactly the opposite. Each participant performed 6 repetitions of each touch gesture on each of 6 monitors (M1-M6), and 5 of the 11 participants also collected data on 2 laptops (L1-L2). We randomized the order of gestures to mitigate any temporal bias. For consistency, all monitors used the same background image; however, we have also demonstrated a *real-time* implementation with dynamically changing backgrounds, as shown in the video figure.

To measure the EMI on the power line, we used the same hardware used in LightWave [4]. An analog high-pass filter (HPF) with a 3 dB corner frequency of 5.3 kHz is used to reject the strong 60 Hz component. The output of the HPF is sampled at 1 MS/s using a 12-bit analog-to-digital (ADC) converter in the USRP (Universal Software Radio Peripheral) followed by transforming the signal into frequency domain using a 32,768-point fast Fourier transform (FFT), yielding a frequency resolution (or bin size) of 30.5 Hz. The signal from the USRP is then fed into the computer for data analysis. It should be noted that a USRP was used in this prototype simply for convenience. Since the required hardware is very simple (a 1st order HPF and an ADC), it can easily be integrated into a small plug-in unit, which can be installed anywhere in a home or office.

Touch Gesture Detection & Classification

The touch EMI is produced at many harmonics of the row rate; however, for our processing, we manually selected the single harmonic with the highest power for each monitor. This selection can be automated using a one-time calibration process. After selecting the EMI peak, we sum the energy of the magnitude of the FFT in the selected frequency bin with the 2 adjacent bins, and filter the result with 3 passes of a Savitzky-Golay filter with a degree of 1 and a frame length of 39 (Figure 4, top). To identify the end points of the performed touch gesture, we took the 1^{st}-order derivative (*i.e.,* sample-to-sample difference) of the filtered summed energy curve and smooth it again with 2 more passes of the Savitzky-Golay filter (Figure 4, bottom). The end points are chosen to be the positions when the derivative curve exceeds a globally defined threshold selected to be the equal error rate of 4.94% using all of the data from our user study.

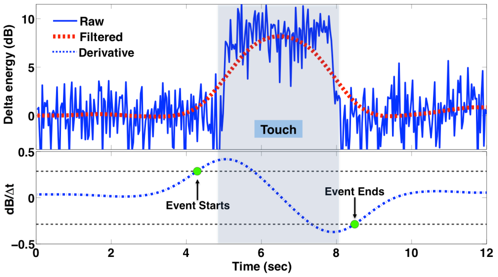


Figure 4: The summed amplitude (top, blue) of the EMI during a touch event, indicated by the gray highlight. The filtered amplitude (red) and derivative (bottom) are shown along with the threshold used to determine the start and end of the event.

After an event is detected, features are extracted from the filtered energy curve for gesture classification. From Figure 5 (top), it can be seen that the amplitude change is much greater for a full-hand touch compared to a five-finger touch. To capture these amplitude differences we compute the following three features: (1) *maximum amplitude*: maximum value during the touch duration, (2) *average amplitude*, and (3) *change in amplitude*: difference of the average amplitude during the touch duration and the average energy in the 3 seconds prior to the touch. The push and pull gestures can be distinguished using the asymmetry in the capacitive coupling (see Figure 5, bottom) using the following features: (4) *peak amplitude position*: position in time of the point of maximum amplitude relative to the segmented touch duration, and (5) *amplitude asymmetry*: difference between the average amplitude in the first and second half of the segmented touch duration.

RESULTS AND DISSCUSSION

We obtain an average detection rate of 96.4% (σ=9.5), with the rate being above 97% for all touch gestures, except for hover and push. The lower detection rate for hover is due to the hand being farther from the panel, resulting in less capacitive coupling, and thus less change in EMI. The lower detection rate for the push gesture is attributed to the way most users performed the gesture. We observed that users tended to approach the screen more slowly when performing the push gesture compared to the other gestures. Since our detection algorithm is based on the derivative of the EMI, this slow approach results in fewer detections.

To explore the feasibility of using uTouch to classify user gestures, we treat our analysis as a machine learning problem. We first trained a 5-class (*i.e.*, representing the 5 touch gestures) support vector machine (SVM) using the Weka toolkit. In order to model a realistic use case, we trained the model using *only* the first *two* examples of each touch gesture for each monitor and user, and then tested the model on all remaining examples. We obtained an average classification accuracy of 68.3% (σ=22.5%, chance=20%).

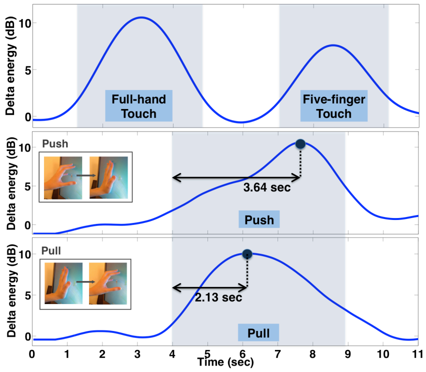


Figure 5: Summed energy curves representing full-hand and five-finger touches (top) as well as the push and pull gestures (bottom). The gray shaded area indicates the touch duration.

From the confusion matrix in Figure 6 (left), it can be seen that there is significant confusion between the five-finger touch (FF) and hover. During these actions, the palm, which represents the largest area of the hand, is at about the same distance from the LCD panel, and therefore has nearly the same capacitive coupling. These similarities cause the amplitude of the EMI to be nearly the same, which results in the observed confusion. In addition, the confusion between FH/FF and push/pull is due to the fact that the push and pull gestures are comprised of FH and FF touches. We believe that a larger training set would allow more robust classification of these 5 gestures.

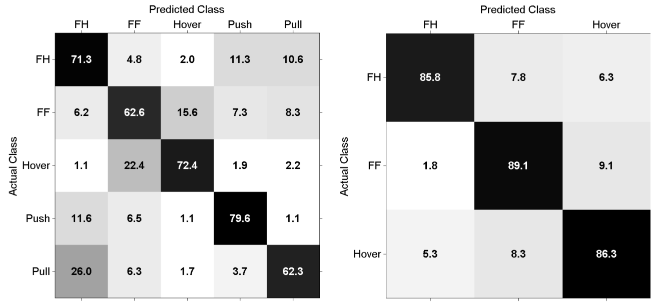
Using a reduced gesture set of 3 gestures (hover, push, and pull), the classification rate becomes 86.3% (σ=17.3%, chance=33%). Figure 6 (right) shows the confusion matrix, which shows considerably fewer misclassifications. Note that due to imperfect event detection, there are a different number of test examples in each classification run (although the training set is always the same size), and thus the accuracies reported in the confusion matrix do not average to the aggregate accuracies reported above.

In order to explore the temporal stability of the signal, two of the users in the study were asked to repeat the experiments on separate days. One user performed the experiment on 6 separate days and the other on 3. We trained a model using only the data from the first day, and then ran the classifier on the data from the remaining days. We are able to classify the 3 gestures used above with 83.8% accuracy (σ=11.8%). Since the performance was not degraded substantially over the case in which we used only data from one day, we can conclude that our signal is temporally stable over many days, even when the noise level and load on the power line changes.

It should also be noted that the high standard deviation in classification accuracy for the 5-class problem is almost entirely due to differences in the monitors. The standard deviation of the classification accuracy across all 11 users is only 5.5%, while it is 20.3% over the 8 LCD screens used. This low variation across users indicates that it may be possible to build a generic model, and thus require *no* per-user training. To test this hypothesis, we trained a 3-gesture classifier (*i.e.*, the same 3 gestures used above) on the data from 10 of our 11 users, and then ran the classifier on the data from the remaining user. We folded across all 11 users, achieving an average accuracy of 88.8% (σ=13.6%). This high accuracy suggests that it is possible to build a generic model using a large database of users, thus removing the training requirement for new users.

REAL-TIME IMPLEMENTATION

In addition to the offline analysis, it is also important to demonstrate the ability for uTouch to work in real-time. In order for uTouch to run in real-time with low latency, we simplified much of the processing and machine learning. The summed energy signal is only filtered once using a Savitzky–Golay filter with a degree of 1 and a frame length of 3. In addition, a moving average filter of length 3 was used to further smooth the signal. We used a simplified decision-tree classifier using only two of our original features (*i.e.*, maximum amplitude and peak amplitude position). This implementation has little perceivable latency and shows high event detection and classification rates when used on our best performing monitor (*i.e.*, M5). The video figure demonstrates this real-time system as well as a gaming interface and video playback controller.



	FH	FF	Hover	Push	Pull
FH	71.3	4.8	2.0	11.3	10.6
FF	6.2	62.6	15.6	7.3	8.3
Hover	1.1	22.4	72.4	1.9	2.2
Push	11.6	6.5	1.1	79.6	1.1
Pull	26.0	6.3	1.7	3.7	62.3

	FH	FF	Hover
FH	85.8	7.8	6.3
FF	1.8	89.1	9.1
Hover	5.3	8.3	86.3

Figure 6: Confusion matrices for the 5-gesture (left) and 3-gesture (right) classification.

CONCLUSION

We demonstrate an approach for passively detecting and classifying a user's touches and hand hovers over an uninstrumented LCD panel. Our approach utilizes electromagnetic interference (EMI) produced by LCD displays that is amplified when a user brings their hand near or touches the LCD's front panel. This EMI is coupled onto the power lines, which can be sensed using a single device connected elsewhere on the power line infrastructure. We show the feasibility of using this approach through an 11 user, 8 LCD study, and a real-time system. uTouch can be used in a variety of applications such as multimedia controls, gaming interfaces and home automation control.

REFERENCES

1. Abileah, A., Green, P. Optical sensors embedded within AMLCD panel: design and applications. In the *Proc. of SIGGRAPH 2007*.
2. Boer, D.W. *Active Matrix Liquid Crystal Displays*. Oxford: Elsevier, 2005.
3. Echtler, F., Pototschnig, T., and Klinker, G. An LED-based multitouch sensor for LCD screens. In *TEI 2010*.
4. Gupta, S., Chen, K., Reynolds, M., Patel, S.N. LightWave: Using Compact Fluorescent Lights as Sensors. In the *Proc. of UbiComp 2011*.
5. Hodges, S., Izadi, S., Butler, A., Rrustemi, A., Buxton, B. ThinSight: versatile multi-touch sensing for thin form-factor displays. In the *Proc. of UIST 2007*.
6. Moeller, J., & Kerne, A. ZeroTouch: an optical multi-touch and free-air interaction architecture. In *CHI 2012*.
7. Rekimoto, J. SmartSkin: an infrastructure for freehand manipulation on interactive surfaces. In *Proc CHI 2002*