Tomo: Wearable, Low-Cost, Electrical Impedance Tomography for Hand Gesture Recognition

Yang Zhang  Chris Harrison
Carnegie Mellon University, Human-Computer Interaction Institute
5000 Forbes Avenue, Pittsburgh, PA 15213
{yang.zhang, chris.harrison}@cs.cmu.edu

ABSTRACT
We present Tomo, a wearable, low-cost system using Electrical Impedance Tomography (EIT) to recover the interior impedance geometry of a user’s arm. This is achieved by measuring the cross-sectional impedances between all pairs of eight electrodes resting on a user’s skin. Our approach is sufficiently compact and low-powered that we integrated the technology into a prototype wrist- and armband, which can monitor and classify gestures in real-time. We conducted a user study that evaluated two gesture sets, one focused on gross hand gestures and another using thumb-to-finger pinches. Our wrist location achieved 97% and 87% accuracies on these gesture sets respectively, while our arm location achieved 93% and 81%. We ultimately envision this technique being integrated into future smartwatches, allowing hand gestures and direct touch manipulation to work synergistically to support interactive tasks on small screens.

Author Keywords
Smartwatch; Fitness Band; Bio-impedance; Biometrics; EIT; Finger Input; Mobile Devices; Interaction Techniques.

ACM Classification Keywords
H.5.2. [Information interfaces and presentation]: User interfaces – Input devices and strategies.

INTRODUCTION
Tomography analyzes the inner structure and composition of objects by examining them with excitations such as electricity and radiation in a cross-sectional manner [4]. Electrical Impedance Tomography (EIT), proposed by Webster [13] in 1978, uses pair-wise impedance measurements from surface electrodes surrounding an object to recover the impedance distribution of the inner structure [7]. Compared with other tomographic methods such as CT scans (x-rays), PET scans (gamma rays) and magnetic resonance imaging, EIT is non-invasive and relatively cheap. It has been widely applied in the medical domain for the assessment of cardiac function, pulmonary hypertension, and lung function [7]. Since its introduction in the early 1980s, EIT instrumentation has continued to evolve in step with advances in electronics. However, even today, medical EIT systems continue to be relatively large and expensive, precluding integration into consumer electronics. This may be the chief reason why the technique has not been utilized in the Human-Computer Interaction domain, despite its great promise.

In this paper, we describe our efforts to create a small, low-powered and low-cost EIT sensor, one that could be integrated into consumer worn devices, such as smartwatches. Achieving these design properties comes at the cost of reduced precision and resolution compared to medical EIT systems. However, as we will show, our system is still able to resolve considerable detail. This ability to non-invasively look inside a user’s body (Figure 1) opens many new and interesting application possibilities. For example, muscles change their cross-sectional shape and impedance distribution when flexed. Therefore, as a proof-of-concept application domain, we use our EIT sensor for hand gesture recognition. We call this system Tomo — a sensing armband that can be worn on the wrist or arm (Figure 1).

RELATED WORK
Worn Gesture Recognition Approaches
Hand gestures extend interaction of computers beyond clicking and typing. One approach to gesture recognition is through computer vision. Pervious work has attached cameras to users’ wrists, pointing towards the hand, in order to recover a 3D hand pose model [15,24]. However, these systems have steep computational requirements and are fairly bulky, as they need line of sight to the fingers.

Figure 1. Tomo worn on the arm (left) and wrist (right) with reconstructed images of the interior shown in the background.
Another approach is to attach accelerometers and gyroscopes to a user’s wrist and fingers [17]. Accuracy largely depends on the position of the sensors, since the efficiency of the data to reveal hand gestures decays quickly as sensors move away from the fingers and towards the wrist or forearm. For example, Xu et al. [25] investigated hand gesture recognition using data collected from the wrist; the system supports dynamic hand motions (e.g., wiggling of the fingers), but not static hand poses (e.g., finger pinches).

Other researchers have capitalized on the change of arm or wrist contour when hand gestures are performed. Pressure sensors, often integrated into a band, can rest on a user’s skin and measure the change of contour [8,14]. Fukui et al. [11] and SkinWatch [16] built wristbands with IR proximity sensors to recover wrist contour. From our investigations, we have found contour-based methods are easily affected by changes in band tightness, clothing and arm movements.

There is also a growing body of research that uses electrical signals generated by muscle activation (i.e., Electromyography) for gesture recognition [20,21]. In general, this requires “wet” electrodes and/or placement on larger muscle groups (e.g., upper forearm) for there to be a sufficient signal-to-noise ratio (SNR).

Bio-Impedance & Bio-Capacitive Sensing
More similar to Tomo are approaches that use bio-impedance or bio-capacitive sensing. In the consumer domain, this sensing has been used for e.g., digital scales with body fat sensing capabilities [10,22]. In the research domain, Touché [19] recognizes complex configurations of human hands and objects based on the change of impedance by electrically exciting objects with a swept-frequency signal. Of note, these systems use a single transmitter and receiver. In contrast, Tomo uses many transmitter/receiver pairs, but only a single frequency.

Most similar to our work are methods that use many electrodes. Similar to our setup, Cornelius et al. [6] measures cross-sectional impedances between all pairs of eight electrodes resting on a user’s forearm. This is used to recognize the wearer for authentication purposes. Cheng et al. [5] created a neckband to measure the capacitance of a users’ neck. Neck-related events, such as head motions, speaking, coughing, and swallowing can lead to significant and distinguishable changes in capacitance distribution, which can then be used for activity recognition.

Lastly, GestureWrist [18] detects changes in wrist contour by measuring the capacitance between a series of electrodes integrated into a wristband. Importantly, this only measures the contact condition of an electrode to the skin (roughly proportional to pressure), and not the inner impedance. Furthermore, GestureWrist does not use a multiplexed, pairwise sensing scheme, instead relying on a fixed transmitter and seven receivers. Two hand gestures are described, but accuracy is not evaluated.

ELECTRICAL IMPEDANCE TOMOGRAPHY
Electrode Scheme
Electrical Impedance Tomography (EIT) recovers the inner impedance distribution of objects using pair-wised measurements from surface electrodes surrounding an object [7]. Four-terminal and two-terminal schemes are the most prevalent measurement strategies. We used a two-terminal scheme in this work to reduce system complexity.

Image Reconstruction
To generate the reconstructed interior image, we used Linear Back Projection (LBP), one of the most popular reconstruction algorithms for industrial applications. Compared to other algorithms, it is less computationally expensive, and thus well suited for real-time image reconstruction. Many other image reconstruction techniques exist; see [7] for an extended discussion.

IMPLEMENTATION
Electrode Band
As seen in Figures 1 and 2, we fabricated an electrode band that could be worn on the forearm and wrist. We made the band with eight evenly spaced copper electrodes. The elasticity of the band could be adjusted to provide a firm contact between the electrodes and users’ skin.

Sensor Board
We designed our bio-impedance sensing board (Figure 2) around an AD5933 Impedance Analyzer [3]. This chip includes a frequency generator and on-board ADC that allows impedance measurement at a specific frequency between 0Hz and 100KHz with a resolution of 0.1Hz. We used an excitation signal of 40KHz, which we found revealed the most distinguishable features of gestures during piloting. Our sensing board also included two 8-to-1 multiplexers (ADG1608), which allowed us to select any two of the electrodes as the transmitter and the receiver.

Our sensing board cost roughly $40 to build in individual quantities. The most expensive part was the AD5933 chip (~$20). Undoubtedly, a mass produced version could be made more compact and at lower cost. We also measured our sensor board’s power consumption when worn and fully functioning. The result indicated a power consumption of approximately 50 mW, not including Bluetooth power consumption (most if our technique was integrated directly into a smartwatch). Overall, this suggests EIT could be integrated into future mobile electronics (for reference, the Apple Watch 2015 contains a 780 mWh battery).

Figure 2. Left: Anatomy of the human forearm. Center: Our prototype electrode band. Right: Tomo sensing board.
Bio-impedance Acquisition

Internally, the AD5933 computes the Discrete Fourier Transform (DFT) of 1024 samples and returns real $R$ and an imaginary $I$ values. This process takes around 3ms. The magnitude of the DFT is given by $\text{Magnitude} = \sqrt{R^2 + I^2}$. Following the calculation process described in the AD5933 manual [3], we computed the impedance as:

$$\text{Impedance} = \frac{1}{\text{Gain Factor} \times \text{Magnitude}}$$

where the Gain Factor was computed by calibrating each electrode pair with a 10kΩ resistor.

An Arduino Pro Mini was interfaced with our sensing board, and reported the calculated impedance to a laptop over Bluetooth. The impedances between all electrodes pairs were measured sequentially without repetition, which resulted in 28 independent values (see Figure 4). The system achieved a sample rate of 10 frames per second (with each frame containing 28 sensor-pair values), which was sufficiently fast for real-time interaction.

Machine Learning

Our 28 values of cross-sectional impedance were used as features for classification. As illustrated in Figures 3 and 5, even simple gestures result in significant changes in the inner impedance distribution. We also computed the differences of all pairs (without repetition) of the 28 impedance measurements, producing 378 additional features. We found these relative measurements made our classifier more robust to global changes in bio-impedance over time. In total, this yielded 406 features. We used a support vector machine (SVM) implementation provided by the Weka Machine Learning toolkit for classification (SMO; polynomial kernel with default parameters) [12].

USER STUDY

We recruited 10 participants (3 female), all right handed, with a mean age of 24. The study had two phases, and took approximately one hour to complete; participants were paid $10 for their involvement. As all of our participants were right handed, the system was worn on the left forearm, the conventional location for wearing a watch.

Gesture Sets

In order to compare the accuracy of our approach to previous research, we adopted a widely used [8,15,24] thumb-to-finger pinch gesture set: Index Pinch, Middle Pinch, Ring Pinch, and Little Pinch (illustrated in Figure 3). We also created a hand gesture set, designed around coarse motions of the hand. This gesture set included Fist, Stretch, Right, Left, Thumbs Up, Spider-Man and Index Pinch (Figure 3). We also included a Relax gesture in both sets as the neutral state, bringing the total number of gestures in the pinch and hand sets to five and eight respectively.

Procedure

Our study had two phases of data collection, serving different experimental purposes. Data from phase one was used for evaluating within-user accuracy, the volume of training data needed, and cross-user accuracy (i.e., universality). Data from phase two was used for evaluating reproducibility. Within each phase, data was collected first from the arm and then from the wrist location; the two gesture sets were collected simultaneously. After a brief introduction, phase one began. Participants were fitted with our wrist- and arm-bands simultaneously. Of note, unlike most electromyography (EMG) systems, our approach does not require any conductive gel or other special preparation of the skin. Once comfortable, participants were asked to perform one gesture at a time. Gestures were requested visually on a laptop screen in a random order.
After the initial step of collecting data, participants were then given a ten-minute break, during which the bands were removed. Following the break, the bands were replaced on the arm and wrist (as close to the original locations as possible), and phase two began. Participants followed the same process as described above, except that gestures were only requested once (i.e., one round) for each sensor position, gathering an additional 2,200 data points.

RESULTS

The central questions we wished to answer in our evaluation were: 1) What is the accuracy of our system? 2) How much training data is needed to initialize our classifiers? 3) How stable is the data collected by our approach, both across users and time?

Within-User Accuracy

Because users have different bodies, most bio-sensing systems require per-user classifiers (e.g., electromyography [20,21,23], bio-acoustics [1,9]). Using data collected during phrase one, from a single user at a time, we trained our classifier on nine rounds of data, testing on a tenth. This procedure ensured that data points adjacent in time (which will naturally tend to be more similar) were either in the test set or train set, but never both. We evaluated all train/test combinations and averaged the results per user.

For the hand gesture set, the wrist location achieved a mean accuracy of 96.6% (SD=2.8%). The arm location achieved a mean accuracy of 93.1% (SD=4.7%). A major source of error was the confusion between Fist and Thumbs Up, contributing 28% of the misclassifications. Figures 6 and 8 provide confusion matrices and an overview of the accuracy results.

For the pinch gesture set, the accuracy of the wrist location was 86.5% (SD=10.1%). The most confused gestures were Ring Pinch and Little Pinch, contributing 27% of the misclassifications. The arm location achieved an accuracy of 89.9% (SD=10.2%). In this case, Index Pinch and Ring Pinch confusion contributed 14% of the misclassified instances. Figures 7 and 9 provide confusion matrices and an overview of the accuracy results.

Our accuracy results compare favorably to other bio-sensing gesture recognition systems with respect to accuracies and number of gestures supported. The EMG-based system created by Saponas et al. [21] supported the pinch gesture set at 77.0% accuracy (and required conductive gel to provide sufficient SNR). WristFlex [8], which used pressure sensors, demonstrated 69.0% accuracy (when providing no feedback during the study, same as our procedure). For reference, Tomo achieved accuracies of 86.5% and 89.9% for this gesture set when worn on the wrist and arm respec-
assess this stability, we do not need to be re-trained each time the device is worn. To understand how performance was affected over time, we trained each time the device is worn.

**Volume of Training Data**

To better understand how much training data is needed before classifiers are sufficiently robust, we ran an experiment varying the size of the training set (phase one data only). Specifically, we first trained the SVM on round 1 data, testing on round 2. We then trained on rounds 1 and 2, and tested on round 3, and so on up to rounds 1 though 9 being used for training and testing on round 10. The results of this analysis are shown in Figure 10. In general, accuracy increased quickly, and largely plateaued by around 50 training instances per gesture.

**Reproducibility**

The ideal bio-sensing system is stable overtime, and does not need to be re-trained each time the device is worn. To assess this stability, we trained a SVM on all phase one data from a single participant and tested on that participant’s phase two data. We evaluated all participants and combined the results (Figure 8 and 9).

Across all conditions, absolute accuracy dropped by 21.6%. The drop was more precipitous for wrist placements (29.1%) compared to arm placements (14.0%). It appeared certain gestures were more impacted than others. For example, the Right gesture dropped 5.2% in accuracy, while Ring Pinch dropped 42.3%. This suggests there may be gesture sets that are more robust overtime.

**Universality**

We also wished to test the universality of our system – that is, whether a classifier could be created that works across users, without having to train on the wearer first. To test this, we combined data from nine participants into a single training set, and used a tenth participant for testing. We evaluated all train/test combinations and averaged the results (Figure 8 and 9). The results clearly showed infeasibility for a universal classifier; at least at the scale of training data we collected. Accuracy across sensor position and gesture set condition was poor: 47.2%.

Looking at the confusion matrices, we found five gestures (Relax, Fist, Stretch, Right and Left) that retained accuracies above 70% (hand gesture set, sensor on arm). This may suggest that these five gestures are more universal. In response, we trained another universal (cross-user) classifier with only these five gestures supported: Accuracy rebounded to 82.6% (SD=16.2%).

**Sensor Location**

Gestures in the pinch set employed similar muscle movements, and as a result, they were more easily confused with each other compared to hand gestures. We also found that the wrist location beat the arm, in both gesture sets, with respect to within-user accuracy. Conversely, the arm location yielded better reproducibility and universality. However, the differences are small and merit further investigation.

**SMARTWATCH INTEGRATION**

As an interactive demo, we instrumented the wrist strap of a Samsung Galaxy Gear 2013 Smartwatch with eight electrodes (Figure 11). Our sensor board was attached to the underside of the watch, though obviously everything would be tightly integrated in a consumer device. Our sensor
board transmitted data to a laptop over Bluetooth, which performed classification and controlled the smartwatch interface over a persistent web socket. We created a simple interaction sequence, utilizing four hand gestures (see Video Figure). In this example, a user can navigate through a series of messages with Left and Right gestures. If desired, a message can be opened with a Fist gesture and return to the list with a Stretch gesture. If a phone call is received, it can be dismissed with a Left gesture, or expanded with a Fist gesture.

LIMITATIONS
Our results suggest that a general, cross-user classifier is currently infeasible, and thus, per-user training would be required. Although an annoyance from a user experience perspective, we believe the biggest limitation is reproducibility. The simple fact is, the human body changes and is in constant motion, which makes sensing challenging. Further, bio-sensing methods tend to be placement sensitive, and EIT is no exception. Moreover, since we do not use gel or any preparation of the skin, contour changes in the wrist and arm can affect the contact condition between electrodes and users’ skin. Environmental interference can also affect the measured impedance. For example, we found persistent electromagnetic interference (EMI) at several frequencies, most notably around 50KHz, caused by fluorescent light ballasts. Partly for this reason, we chose an excitation signal of 40KHz, which was reasonably free of ambient noise.

FUTURE WORK
In our immediate future work, we hope to find approaches that mitigate issues surrounding reproducibility and universality. A major contributing factor to these limitations is varying placement of the electrodes, which change slightly each time the device is worn, and certainly across users. As shown in recent work by Amma et al. [2], higher resolution electrode arrays can be used to compensate for differences in placement, essentially by digitally rotating the signal to a uniform position, from which machine learning features can then be generated. We also believe superior band ergonomics and electrode material would yield a more stable signal, as would a four-terminal EIT sensing configuration. Finally, we hope to improve the frame rate of our work by replacing the AD5933 chip with a separate DDS and ADC. This should allow Tomo to support continuous gesture sets and overall, more dynamic applications.

CONCLUSION
We have presented Tomo, a wearable, low-cost and low-powered Electrical Impedance Tomography system for hand gesture recognition. It measures cross-sectional bioimpedance using eight electrodes on the wearers skin. Using 28 all-pairs measurements, our software can recover the interior impedance distribution, which is then fed to a hand gesture classifier. We evaluated two gesture sets (hand and pinch sets) and two body placements (wrist and arm). User study results show that our approach can offer high accuracy hand gesture recognition when the system is trained on the wearer. However, like most other bio-sensing systems, results degrade when the system is re-worn at a later time, or worn by other users.

ACKNOWLEDGEMENTS
This research was generously supported by Qualcomm, Google, and the David and Lucile Packard Foundation. We also thank Gierad Laput and Chen Gong for their help with demos and figures respectively.

REFERENCES


22. Skulpt, Inc. http://www.skulpt.me

