A Simple Wristband Based on Capacitive Sensors for Recognition of Complex Hand Motions

Jingyuan Cheng ¡¿
Embedded Systems Lab.,
Uni Passau, Germany
jingyuan.cheng@uni-passau.de,
http://esl.fim.uni-passau.de

Gernot Bahle, Paul Lukowicz
Embedded Intelligence Lab,
DFKI Kaiserslautern, Germany
{gernot.bahle,paul.lukowicz}@dfki.de

***Abstract*—We demonstrate that an unobtrusive wristband with embedded textile capacitive sensors can provide reliable information about complex hand motions. We describe the design considerations resulting from FEM field simulations and the actual hardware design. We then show example signals from different motions and present a quantitative evaluation on the recognition of 10 digits and 26 letters that users "write in the air".**

I. Introduction

Designing and evaluating novel, unobtrusive activity sensing and input methods is at the core of ubiquitous computing. Hand motions especially can provide very detailed information about a user's actions. In addition, it is natural for us to control our environment through hand movements [10], [11]. As a consequence, there have been many proposals to capture these motions, often utilizing inertial motion sensors. However, most of them rely on heavy augmentation of the hand or wrist, in many cases requiring the user to wear gloves or other sensing devices directly on the back of his hand [3]. At the same time, it is well known that people dislike wearing devices on this location [8].

As an alternative approach, this paper shows that an unobtrusive textile wristband can be used to detect a variety of complex hand motions through capacitive sensing. The idea is that a capacitive sensor reveals subtle changes inside the wrist which in turn are correlated with the muscles moving the hand and the fingers. The wristband itself requires no special fixation (shown by the test person taking it off and putting it back on between experimental runs). While capacitive sensing is an established principle, adapting it to the specific purpose of sensing subtle motions and changes inside the wrist involves a number of problems. In this paper we describe how we solved this problems and present the evaluation of our sensor on a complex hand gesture recognition task.

A. Related Work

Capacitive sensors are a well-known sensing modality in industry and wearable computing. Rekimoto presents capacitive sensors embedded in the wrist and in a jacket and uses them to detect arm gestures [13]. Compared to our setup, the gestures recognized are very crude and the setup requires the user to wear several capacitive textiles distributed on the torso and arm. Oum et. al. uses metal plates integrated in PCB as pads to get breath and pulse from the wrist [12]. Cheng et. al. uses conductive metal textile as pads so that the sensor becomes soft and twistable. It is used to observe inner changes like breathing and heartbeat and a broad range of activities including swallowing, head position, walking pace etc. [4]. Amft et. al show how to use electrodes directly attached to the muscles of the lower arm to sense muscle activity and fatigue [2]. Amft et. al. use motion sensors embedded in a wrist worn device as input for a questionnaire [1]. However, most gestures in the related work are quite crude compared to the ones used in this paper [9], [14]. There are also quite a number of data gloves and wearable keyboards. Unfortunately, they all require heavy augmentation of the user's wrist/hand [6], [1], [10], [11]. Additionally, all of these devices use a combination of motion sensors, cameras/light sensors and microphones to perform input or context recognition. Therefore, they can be seen as complimentary to textile capacitive sensing.

II. Description of the Device

We can form a capacitive sensor by putting two conductive pads near the human body, utilizing the body as the insulating material. The capacitance changes with the pads' shape and their relative position to each other and to the human body. We embed four capacitive sensors into our elastic wrist band, located at the top, bottom, left and right side around the wrist. Each sensor is composed of two $3cm \times 0.6cm$ metal textile pads parallel to each other (center to center distance $1cm$). The sensors are isolated from the human body by a thin layer of soft tissue. Hand and wrist movements change the sensor shape and the skin/muscle alignment underneath the band, which results in capacitance change.

A. Sensor Design Through E-field Simulation

To understand how overall capacitance and its change is influenced by sensor design (physically), we choose SEMCAD X, a program deducing E-M field distribution using a Finite-Difference Time-Domain method to get the capacitance of an object whose mechanical structure is fixed. A typical system is shown in Fig. 2), simulating a small change underneath

978-1-4577-1767-3/12/$26.00 ©2012 IEEE

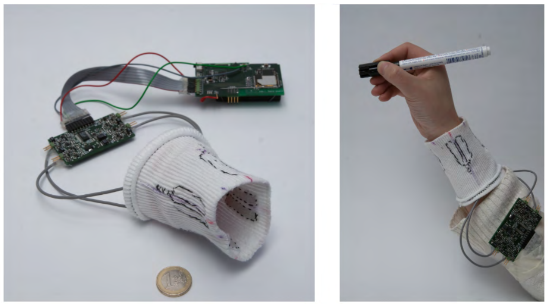

Fig. 1. wristband prototype with 4 sensors (marked with black threads)

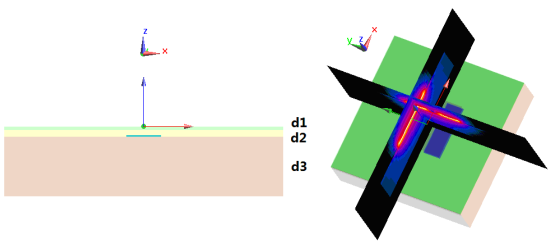

Fig. 2. simplified model of one sensor and the human body (side view and E-density distribution), Layers along axis z from top to bottom: sensor pads(blue,directly on the top of isolating layer), isolating layer(green, thickness d_1), skin(yellow), small change at the bottom of skin(aqua), muscle(red).

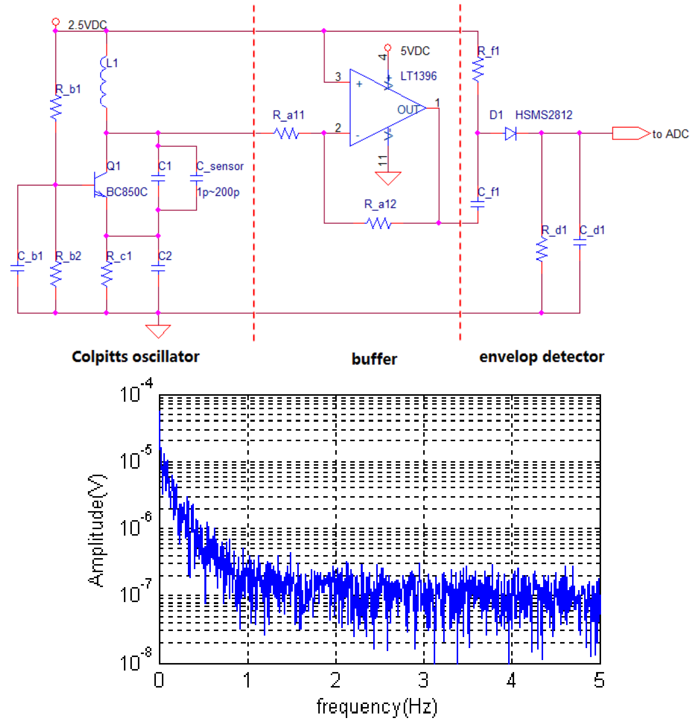

Fig. 3. analog design and noise level (when no C_{sensor} is attached)

skin covered by an isolating layer. Because the same tissue has different permittivity and conductivity when exposed to a sinusoidal E-field of different frequencies [7], frequency is also a factor to be considered. When sensor pads are driven by two sinusoidal signals with the same amplitude(1V) and frequency(f) and 180° phase delay, the overall capacitance is then: $C = \Sigma \frac{dU}{dV}(x, y, z)\Delta x \Delta y \Delta z$, where $\frac{dU}{dV}(x, y, z)$ is the energy density given as simulation result. The main conclusions from simulation are:

1) Overall capacitance: can be estimated as two capacitors in parallel, where the human body is viewed as an ideal conductor and the distance between one pad and the human body is the isolating layer's thickness (d_1 in Fig. 2) with a frequency range of 10KHz-100MHz. The pads' material doesn't matter as long as it is highly conductive.

2) Measuring frequency: According to the simulation, the frequency of a sinusoidal signal applied to the sensor pads should be as high as possible in the range of 10-50MHz. This is especially true for detecting inner changes (e.g. pulse). The highest frequency is limited by the distance from sensor to measuring circuit. A rule of thumb is to keep two-way propagation delay to less than 10% of one period. To minimize the extra capacitance and the propagation delay, the wire connecting sensor pads to the analog circuit should be as short as possible.

3) Most sensitive area: lies directly under the sensor. Sensor pads should cover the change but no more.

4) Isolating layer: should be as thin as possible. Sensitivity is inversely proportional to its thickness.

5) Shielding: is not practical with frequencies in the MHz range. Low-frequency systems (760KHz in [12], 60-240KHz in [5]) can use a voltage follower to generate the second signal which is exactly the same as on the sensor pad and is applied to another pad right behind the sensor pads. Because sensor pad and shielding pad have the same voltage, the E-field between them is zero no matter what is behind the shielding pad, i.e. the sensor pad is sensitive on one side and isolated on the other side. As the simulation showed, in our high-frequency system, propagation delay in wires and through chips and signal distortion caused by the amplifier make the second signal differ too much from the original. Another layer of isolation material with high permittivity doesn't help either.

B. Analog Design

To detect minor changes (pF level or less) from already quite a small overall capacitance (100pF level), high sensitivity and precision is required. To meet both requirements, we modify the circuit design in paper [12]. The analog circuit is composed of three stages:

1) Colpitts oscillator: generates a neo-sinusoidal E-field, whose amplitude and frequency changes with the capacitance between two sensor pads (shown as C_{sensor} in Fig. 3).

2) Buffer: isolates the very sensitive oscillator from the rest of the circuit and provides enough driving capability.

3) Envelop detector: picks up the voltage change which shows the capacitance change.

The main differences are:

1) Frequency: is pushed to around 20MHz according to E-field simulation.

2) Frequency change to amplitude change: Both frequency and amplitude change when C_{sensor} changes. According to real measurement, amplitude change is higher with our setting (frequency 20MHz). The low pass filter is changed into a high pass filter(R_{f1} and C_{f1}), which helps to remove pink noise from transistor Q1.

3) Wearablity: is increased by simplifying the circuit. We use 24 bits $\Delta - \Sigma$ ADC and remove secondary amplification. With a 1700mAH battery, 4 channels, a sample rate of 20Hz and wireless data transfer, our wearable prototype runs continuously for 6.5 hours.

Special care is given to crosstalk and noise level. Each channel is tuned differently so that their central oscillation frequencies differ around 2MHz from each other. The 2.5V DC power supply and diode D1 are the main noise source. The noise level we have achieved after carefully picking proper chips is shown in Fig. 3, which is far more than enough for hand writing, as seen in Fig. 5, where effective signal(mV level) fully exceeds noise (pink noise maximum 10μV lower than 1Hz and $0.2\mu V$ above 1Hz).

III. Evaluation

For this article, our goal was to demonstrate that a large amount of useful information about hand gestures can be provided by a very unobtrusive, easy to put on sensor. To emphasize this point, we selected a natural, large and difficult to accurately recognize set of gestures of 26 letters and 10 digits. The dataset consists of five sets of recordings, performed by five different subjects, one of them female. Each recording is composed of ten runs. The wristband has been taken off and on between the individual runs to create placement variance. Each run incorporates three repetitions of the 36 letters and digits; this amounts to 30 instances of every class per recording. The division into 10 runs provided a natural split for training / testing purposes. Each run was used as a test set once, with the remaining 9 runs used for training. Each person specific analysis thus consists of ten distinct evaluations. As we asked our subjects to just write as they usually do, letter shapes vary a lot between test persons. For this reason, we have refrained from any person-independent classification. To evaluate our results, we used three measures: recall and precision averaged over all classes (both given with mean and standard deviation over all ten runs) as well as a "best k of 36" measure (meaning, the correct answer is among the k most likely candidates for a given test sample). The latter makes sense for applications like word recognition where the possible combinations of letters can be severely limited by a dictionary of a given language. All of our results can be found in Fig. 6.

A. Evaluation Strategy and Results

To evaluate the data thus collected, we use the following strategy: first, initial classification results are obtained by two different methods.These are then fused into a final output using set operations and ranking.

1) Dynamic Time Warping distances: The first technique utilizes Dynamic Time Warping (DTW) to measure the distance between training instances and test samples on all four channels. Let the training samples be divided into sets C_1 to C_{36} corresponding to their class. For each test sample, calculate the distance to all elements of C_1, then take their mean. Repeat this for all sets C_i of training data. This yields one vector of average distances to all classes per test sample. Finally, sorting by distance gives a list of most likely to least likely class for each sample.

	Results (Averages of 5 subjects)							
	Recall		Precision		Best 2		Best 3	
	mean	std	mean	std	mean	std	mean	std
DTW Avg.	**45%**	**26%**	**61%**	**28%**	**60%**	**12%**	**68%**	**12%**
LDA Avg.	**47%**	**14%**	**50%**	**16%**	**60%**	**10%**	**69%**	**9%**
Fusion Avg.	**58%**	**14%**	**65%**	**21%**	**72%**	**10%**	**77%**	**9%**

Fig. 6. Recall, Precision, Best 2 and Best 3 values (mean and standard deviation) for approaches 1, 2 and the fusion of both

2) LDA on frequency based feature space: Our second approach employs frequency based features like the frequency centroid and cepstrum coefficients calculated on the original four signal channels, as well as on their first and second derivative and also on the pairwise ratios between channels (ch_i/ch_j, for $i, j = 1..4, i \neq j$). The resulting feature space is then transformed using linear discriminant analysis (LDA). Afterwards, actual classification is done using the WEKA machine learning toolbox, specifically the Naive Bayes classifier. For each test sample, this classifier outputs the probabilities of the sample belonging to class 1 to 36. Once again, by sorting we obtain a list of most likely to least likely class.

At this point, for each test sample, there exist two vectors v_1, v_2 of class affinities, sorted from most likely to least likely. To fuse them, we first intersect the first 10 elements of each set. This results in two new, shorter vectors containing only the classes deemed likely by both techniques. Note that, while the classes themselves are identical for both vectors, their order does not have to be. Thus, in a final step, we calculate a new, fused order by averaging each classes' position in v_1 and v_2. In case of a collision, the placement in the lda-based vector takes precedence.

One thing becomes obvious immediately: both approaches perform far above random chance (which, for a 36 class problem, would be about a 2.8% chance to roll the right class for a given test sample). In 77% of all cases, the correct letter can be found within the top 3 candidates. Fusing DTW and LDA results increases recall by an average of 13 points compared to the DTW and 11 points compared to the LDA. It also manages to improve upon the precision of the DTW approach by 4 points, while significantly lowering its standard deviation (by 7 points). The best 2 and best 3 scores are similarly boosted. In 77% of all cases, the correct letter can be found within the top 3 candidates. This is a very promising foundation for further work towards real life applications (e.g. text recognition).

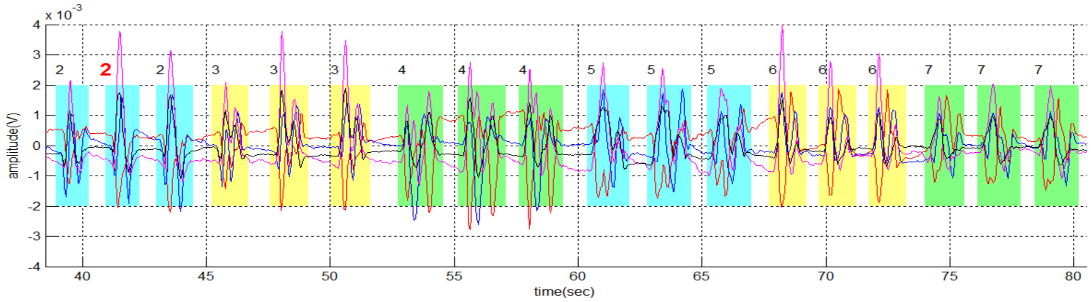

Fig. 4. Sensor signal overview, write number 2 to 7, 3 times each.

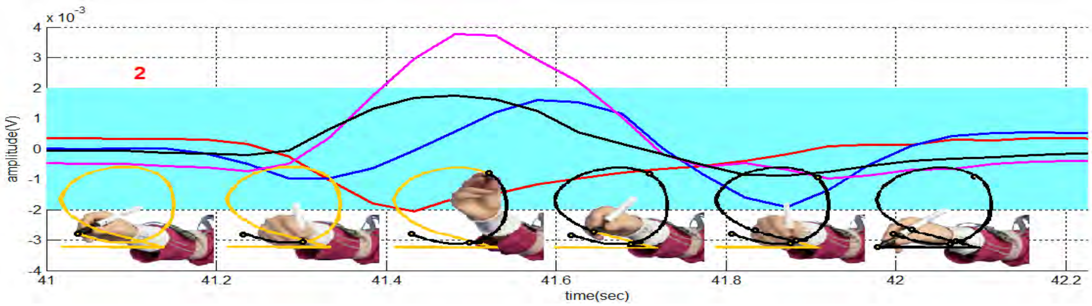

Fig. 5. Sensor signal example for writing the number "2": voltage from the four sensor pads and pictures of the corresponding hand gestures.

IV. Conclusion

In this note, we have shown a novel sensor design capable of capturing a large amount of information about hand and wrist gestures. It has got the major advantage of really being unobtrusive, especially compared to some of the designs shown in related work. The circuit design is based on E-field simulation, which provides a general guideline in designing the sensor pad, not only for hand gestures at the wrist, but also for other activities at other body location. Furthermore, given the large number of classes in our example scenario (random chance is at 2.7%), their inherent complexity and the similarity between some of them, we still achieve a very good 58% recognition accuracy applying only standard machine learning techniques. For domains where a selection of the best 2 or 3 candidates is sufficient (e.g. dictionary based word recognition), we reach an accuracy of 77%.

References

[1] O. Amft, R. Amstutz, A. Smailagic, D. Siewiorek, and G. Tröster. Gesture controlled user input to complete questionnaires on wrist worn watches. In *HCII 2009: Proceedings of the 13th International Conference on Human-Computer Interaction*. Springer, 2009.

[2] O. Amft, H. Junker, P. Lukowicz, G. Troster, and C. Schuster. Sensing muscle activities with body-worn sensors. In *International Workshop on Wearable and Implantable Body Sensor Networks, 2006. BSN 2006.*, pages 4 pp. –141, april 2006.

[3] C. Amma, D. Gehrig, and T. Schultz. Airwriting recognition using wearable motion sensors. In *Proceedings of the 1st Augmented Human International Conference*, AH '10, pages 10:1–10:8, New York, NY, USA, 2010. ACM.

[4] J. Cheng, O. Amft, and P. Lukowicz. Active capacitive sensing: Exploring a new wearable sensing modality for activity recognition. 2010.

[5] J. Cheng, D. Bannach, and P. Lukowicz. On body capacitive sensing for a simple touchless user interface. 2008.

[6] M. Fukumoto and Y. Suenaga. Fingering: a full-time wearable interface. In *Conference companion on Human factors in computing systems*, CHI '94, pages 81–82, New York, NY, USA, 1994. ACM.

[7] S. Gabriel, R. W. Lau, and C. Gabriel. The dielectric properties of biological tissues: II. Measurements in the frequency range 10 Hz to 20 GHz. *Physics in Medicine and Biology*, 41(11):2251–2269, Jan. 1999.

[8] F. Gemperle, C. Kasabach, J. Stivoric, M. Bauer, and R. Martin. Design for wearability. *Second International Symposium on Wearable Computers, 1998. Digest of Papers.*, pages 116–122, 1998.

[9] B. Howard and S. Howard. Lightglove: wrist-worn virtual typing and pointing. In *Wearable Computers, 2001. Proceedings. Fifth International Symposium on*, pages 172–173, 2001.

[10] S. C. Lee, B. Li, and T. Starner. AirTouch: Synchronizing In-air Hand Gesture and On-body Tactile Feedback to Augment Mobile Gesture Interaction. *IEEE International Symposium on Wearable Computers*, 0:3–10, 2011.

[11] T. Maekawa, Y. Yanagisawa, and Y. Kishino. Object-based activity recognition with heterogeneous sensors on wrist. *Proceedings of the 8th International Conference on Pervasive Computing. Helsinki, Finland*, 2010.

[12] J. H. Oum, S. E. Lee, D. W. Kim, and S. Hong. Non-contact heartbeat and respiration detector using capacitive sensor with Colpitts oscillator. *Electronics Letters*, 44(2):87, 2008.

[13] J. Rekimoto. GestureWrist and GesturePad: unobtrusive wearable interaction devices. In *Proceedings of the Fifth International Symposium on Wearable Computers*, pages 21–27, 2001.

[14] T. Starner, J. Auxier, D. Ashbrook, and M. Gandy. The gesture pendant: a self-illuminating, wearable, infrared computer vision system for home automation control and medical monitoring. *Wearable Computers, The Fourth International Symposium on*, pages 87–94, 2000.