Energy-Optimal Gesture Recognition using Self-Powered Wearable Devices

Jaehyun Park[1], Ganapati Bhat[2], Cemil S. Geyik[2], Hyung Gyu Lee[3], Umit Y. Ogras[2]
[1]School of EE, University of Ulsan, Ulsan, Korea, jaehyun@ulsan.ac.kr
[2]School of ECEE, Arizona State University, Tempe, AZ, {gmbhat, cemil.geyik, umit}@asu.edu
[3]School of CCE, Daegu University, Gyeongsan-si, Korea, hglee@daegu.ac.kr

***Abstract*— Small form factor and low-cost wearable devices enable a variety of applications including gesture recognition, health monitoring, and activity tracking. Energy harvesting and optimal energy management are critical for the adoption of these devices, since they are severely constrained by battery capacity. This paper considers optimal gesture recognition using self-powered devices. We propose an approach to maximize the number of gestures that can be recognized under energy budget and accuracy constraints. We construct a computationally efficient optimization algorithm with the help of analytical models derived using the energy consumption breakdown of a wearable device. Our empirical evaluations demonstrate up to 2.4× increase in the number of recognized gestures compared to a manually optimized solution.**

I. Introduction

Wearable internet of things (IoT) devices are becoming popular due to their small form factor and low cost [3]. Small form factor enables interesting applications including gesture-based control, health monitoring, and activity tracking [2, 12]. However, it also limits the battery capacity, which is one of the major obstacles for widespread adoption of wearable IoT devices [13].

Wearable devices cannot rely on high capacity batteries used in smartphones due to their relatively large size and weight (2100 mAh @ 42 g) [7]. Lighter flexible batteries cannot be used alone either, since they have modest capacities (200 mAh @ 1.2 g) [6]. Therefore, harvesting energy from ambient sources is crucial to relieve from the dependence on batteries [5]. Recent research shows that photovoltaic cells (PV-cells) can provide 10–100 mW/cm^2 density [10]. Wearable devices can greatly benefit from this harvesting potential, since they can be personalized for each user. For example, the device can learn the usage patterns, and adapt the operating points to its user.

In this work, we consider wearable devices powered primarily through ambient energy sources, as illustrated in Figure 1. Since the amount of the harvested energy sets the available energy budget, *the device has to maximize the work performed under this energy budget*. To this end, we employ gesture recognition as the target domain, because it has a wide range of biomedical applications, such as gesture-based control and interaction with assistive devices. More precisely, we dynamically maximize the number of gestures that can be recognized under energy budget and accuracy constraints.

Energy-optimal gesture recognition problem is convoluted by three major challenges. First, accurate energy consumption and gesture recognition accuracy models are needed to guide this optimization. Second, this problem should be solved at runtime with minimum implementation overhead. Finally, the optimization methodology has to be validated using an energy harvesting device and user studies. To address these challenges, we employ a custom wearable device whose components are illustrated in Figure 1. Using this prototype, we characterize the power consumption of the accelerometer, microprocessor and Bluetooth Low Energy (BLE) while performing gesture recognition. Then, we develop a compact energy model that can be used at runtime by the proposed optimization approach. Similarly, we analyze the recognition accuracy as a function of the gesture recognition duration by performing user studies. Finally, we present a computationally efficient algorithm to perform gesture recognition under the energy budget and accuracy constraints. We show that the proposed approach increases the number of recognized gestures by up to 2.4× compared to manual optimization.

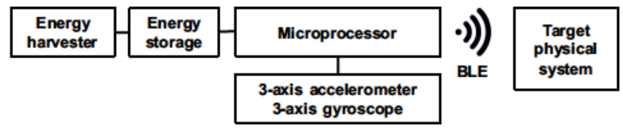


Fig. 1. Wearable gesture recognition system.

In summary, the major contributions of this paper are:

- An algorithm to maximize the number of gestures that can be recognized under energy budget and accuracy constraints,
- Empirical evaluations on a wearable device prototype that demonstrate up to 2.4× increase in the number of recognized gestures compared to manual optimization.

The rest of the paper is organized as follows. We review the related work in Section II. We present the proposed energy-optimal gesture recognition framework in Section III. Finally, we discuss the experimental results in Section IV, and summarize the conclusions in Section V.

II. Related Work

Wearable IoT devices have been studied extensively due to their form factor and cost advantages. Researchers have proposed gesture-based control, health monitoring, and activity monitoring as applications of IoT devices [8]. Significant amount of research has also focused on wearable devices with energy harvesting [5]. For instance, a jacket with solar and thermal energy harvesting is proposed in [5]. Energy harvesting in IoT devices has necessitated the development of energy management and allocation algorithms for wearable IoT devices [4]. For example, the algorithm proposed in [4] uses a dynamic programming approach to perform a near-optimal energy allocation for self-powered wearable devices. In this work, we assume that the energy budget for each time horizon is provided by a similar algorithm.

Power-aware computing is critical for wearable devices due to limited energy budget. Therefore, recent research has focused on an accuracy-power trade-off in wearable devices [9, 14]. For instance, the technique presented in [14] uses a dynamic sensor selection to minimize the power consumption of a gesture recognition body area network. This leads to a maximization of the network lifetime. The work in [9] proposes an algorithm

978-1-5386-3603-9/18/$31.00 ©2018 IEEE

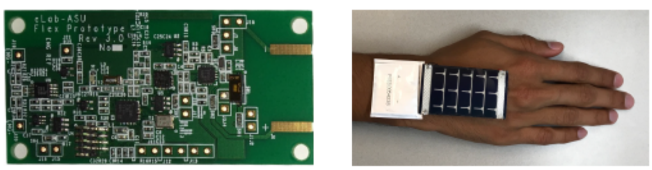
for offloading gesture recognition to a smartphone for certain gestures, such that the energy consumption of the system is minimized. In contrast to these approaches, we propose a novel runtime algorithm that optimizes gesture recognition on the wearable device under energy budget and accuracy constraints. To this end, we first formulate a nonlinear optimization problem to maximize the number of gestures. Then, we use experimental measurements on a wearable device to derive a low complexity solution to the problem.

III. Energy-Optimal Gesture Recognition

A. Self-Powered Wearable Device Prototype

The target wearable device shown in Figure 2 harvests energy using PV-cells and an energy harvesting circuit. Since the harvested energy is intermittent and exhibits significant variations over a day [4], we also employ a 45 mAh Lithium-polymer battery with 1g weight. The power circuitry employs a PV-cell SP3-37 and a MPPT charger TI BQ25504. The target system includes a microprocessor TI CC2650 and a motion processing unit Invensense MPU-9250.

When attached to user's hand, the wearable device captures the hand motion using the 3-axis accelerometer. Then, the microprocessor processes the captured data to recognize the intended gesture. Finally, the decoded gesture is transmitted to the target physical system through the BLE interface. The energy budget available for gesture recognition is determined by the harvested energy. To be practical, the system has to maximize the number of intended operations (i.e., recognized gestures) *under this budget*, while maintaining a minimum level of recognition accuracy. Therefore, we propose a methodology to achieve this goal.



(a) Mounted prototype **(b) Prototype attached on the hand**

Fig. 2. Gesture recognition prototype used in this work.

B. Problem Formulation

Given the characteristics of the energy harvesting system, we can determine the energy that can be harvested over a finite horizon t_h [4]. We use this amount as the energy budget E_b available for the wearable device. We define the gesture recognition duration t_g as the time spent by the device to infer a single gesture. The wearable device actively senses the hand motion and processes the data during this period, which takes a portion of t_h. We denote the number of gestures recognized within the finite horizon by $N_g(t_g)$, since it is a function of the gesture recognition duration. The energy consumption per gesture $E_g(t_g)$ is a function of t_g, because t_g determines the active time of the processor and sensor. Similarly, the energy consumption of the device during the idle time is denoted by $E_i(t_g)$. Finally, the energy consumed for transmitting the recognized gesture is denoted by E_{comm}. With this notation, we formulate the optimization problem as:

$$\text{maximize} \quad N_g(t_g) \quad \textit{such that} \tag{1}$$

$$E_{total}(t_g) = E_g(t_g) \cdot N_g(t_g) + E_i(t_g) + E_{comm} \leq E_b \tag{2}$$

$$G_{acc}(t_g) \geq G_{acc,min} \tag{3}$$

The first constraint in this formulation ensures that the total system energy consumption is less than the energy budget. The second constraint guarantees that the accuracy $G_{acc}(t_g)$ is greater than a minimum accuracy $G_{acc,min}$. Note that $G_{acc}(t_g)$ is a function of t_g, since t_g determines the number of data points used for gesture recognition given the sampling frequency.

Solving the optimization problem given by Equations 1–3 at runtime is not easy, since both the objective and constraints are nonlinear. Moreover, system dependencies make it hard to model the behavior of $E_g(t_g)$ and $E_i(t_g)$.

C. Gesture Recognition Classifier Design

We define five gestures made by hand – backward, forward, left, right, and wave – as shown in Figure 3. In addition, we include a stationary gesture to detect when the device is inactive. To classify these gestures, one can use a variety of supervised learning algorithms, such as support vector machine (SVM), decision tree, logistic regression and neural network (NN). Selecting the appropriate algorithm depends on the input data size, accuracy and latency requirements, as well as available computational power and memory. In our application, the input is provided by a 3-axis accelerometer with 50 Hz sampling rate. Since common gestures take approximately 0.8 s [1], the number of data points for each gesture is 120.

We target 90% or higher accuracy on a small IoT device. Using user subject experiments, we verified that SVM, decision tree and NN classifiers can meet this accuracy requirement. In addition to accuracy, we also aim at a flexible solution that can be easily extended to more number of gestures, input features, and other applications. Furthermore, personal assistive devices can learn its user's behavior and adapt the algorithm at runtime to maximize the recognition accuracy. In this work, we adopt a NN classifier, since it facilitates the two additional constraints with the help of a programmable solution and reinforcement learning.

The input layer in our NN classifier is connected to the input features, i.e., the sensor data. The output layer has 6 neurons, one corresponding to each gesture shown in Figure 3. Each output neuron evaluates the probability of the corresponding gesture. Our programmable solution takes the number of hidden layers, the number of neurons in each layer, and the weights as inputs. Then, it instantiates the classifier that is ready to run on the MCU. In order to choose the number of neurons in our NN, we perform design space exploration by varying the number of hidden layers and neurons in each layer. We observe that the increase in accuracy diminishes once the number of neurons exceeds three. In our experiments, we employ four neurons in the hidden layer, since further increase does not improve the accuracy, and it leads due to a lower variation in accuracy compared to using 3 neurons. We employ two versions of the NN for gesture recognition:

- **Baseline NN** uses all 120 accelerometer samples collected by the 3-axis accelerometer during t_g as input features.
- **Reduced NN** employs transformed features derived from the raw accelerometer data. We utilize the minimum, maximum and mean values of each axis (x, y, z) over t_g. Hence, this amounts to a total of 9 input features. Since the number of transformed features does not depend on t_g, we can change it at runtime.

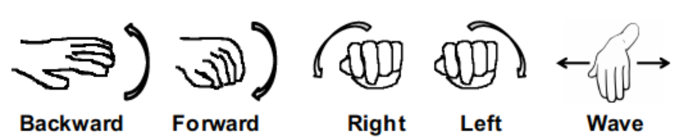


Fig. 3. Illustration of the target gestures.

D. Energy Consumption Characterization

As mentioned before, solving the optimization problem given in Equations 1-3 requires modeling energy per gesture $E_g(t_g)$ and idle energy $E_i(t_g)$ as a function of gesture recognition duration t_g. Therefore, we analyze the power consumption of the microprocessor and the sensor (i.e., accelerometer) while processing one gesture. The dashed blue and solid red lines in Figure 4 represent the measured power consumption of the microprocessor and the sensor, respectively. Initially, the system waits for user motion in the idle state. When the user makes a gesture, the accelerometer sensor wakes the system up, and performs a preprocessing routine to prepare the accelerometer and microprocessor. Then, the accelerometer starts sampling the motion data for a duration of t_g. We observe two different levels of power consumption in the sensor. The sensor power consumption is close to zero during idle state, while it consumes around 2 mW power in the active state. The power consumption also exhibits peaks during the state transitions because of pre- and post- processing of data acquisition. Once the data acquisition is completed, the microprocessor processes the sensor data, and transmits the recognized gesture using BLE. Unlike the sensor, the power consumption of the microprocessor shows periodic peaks, which are caused by the BLE module to maintain an active connection. In addition, pre-processing and post-processing tasks cause larger peaks in the microprocessor power.

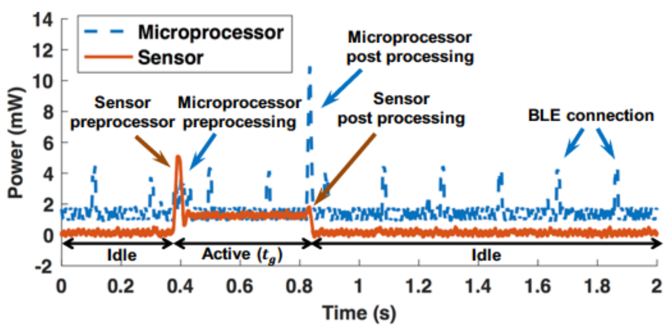


Fig. 4. Power consumption during a gesture recognition when $t_g = 400$ ms.

E. Solution to the Optimization Problem

This section presents a technique to maximize the number of gestures that can be recognized under the energy budget and accuracy constraints given in Equations 1–3.

Longer gesture recognition duration t_g implies longer active time for both the sensor (*more data samples*) and microprocessor (*more processing*), as shown in Figure 4. Consequently, the total energy consumption is an increasing function of the gesture recognition duration t_g. We characterize the energy consumption thoroughly and construct a detailed energy model to capture this relation [11]. With the help of this characterization, we illustrate the energy consumption as a function of t_g in Figure 5 (the left axis). The energy budget, specified by the horizontal dotted line, limits the energy consumed by the system. Hence, the gesture recognition duration t_g is *bounded from above by the given energy budget* E_b. Similarly, the gesture recognition accuracy is expected to improve, if larger number of data samples are used. Its precise behavior can be found after user studies, but we can conceptualize it as a non-decreasing function of the gesture recognition duration, as illustrated by the right axis in Figure 5. Consequently, the minimum accuracy requirement bounds the gesture recognition duration t_g from below regardless of the shape of the curve. As

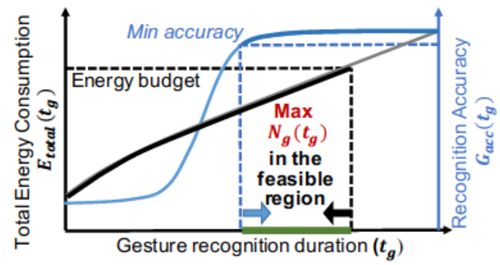


Fig. 5. Energy budget and minimum accuracy requirements constrain the gesture recognition duration t_g from above and below, respectively. Hence, we maximize the number of recognized gestures within the feasible region.

a result, the feasible region for the optimization problem is the intersection of the regions for energy and accuracy, as highlighted in Figure 5.

To maximize the number of recognized gestures $N_g(t_g)$, we express the total energy consumption as a function of t_g, i.e., we model $E_g(t_g)$ and $E_i(t_g)$. Then, we derive an expression for $N_g(t_g)$ such that it can be maximized within the feasible region. Finally, we prove that $N_g(t_g)$ is maximized when t_g is chosen as the minimum value within the feasible region. Detailed models and the proof of this result is provided in the technical report [11] due to page limitations. In what follows, we adopt the minimum t_g that satisfies the accuracy constraints based on our user studies. The effectiveness of this optimization is demonstrated empirically in the following section.

IV. Experimental Evaluation

A. Experimental Setup

Wearable system: We use the in-house wearable prototype described in Section III-A. It features test ports to measure the power consumption of the microprocessor and the MPU separately. Power measurements are performed using NI PXIe-4081 and PXIe-4080 digital multimeter systems with 5 kHz sampling frequency.

Gesture recognition: The wearable device uses the NN to detect the gesture and transmits it to a host device. The host device stores the detected and the reference gesture. We test the accuracy of the gesture recognition system using 30 data sets from seven users. Each set has a series of 50 gestures performed in random order. Ten data sets are reserved for training the NN. The training data is further divided into 80% training, 10% cross-validation and 10% test data, which is used to analyze the accuracy of the NN. We obtain 96.5%, 97.4%, and 98.4% accuracy for the training, cross-validation and test data, respectively. The remaining 20 data sets are used for testing the accuracy of the NN after the training is completed. They are never seen by the NN during the training to reliably test the robustness of our gesture recognition system.

B. Gesture Recognition Accuracy Analysis

The recognition accuracy with the baseline NN is 98.0%, thus it is higher than 90% we target. Similarly, we observe greater than 90% recognition accuracy for the reduced NN when the gesture recognition duration $t_g > 380$ ms, as shown in Figure 6. There is a significant degradation in accuracy when t_g is reduced below 380 ms. We observe this behavior because a lower t_g does not allow sufficient time for the NN to sufficiently differentiate between the gestures. Moreover, there may not be sufficient time to complete a gesture, when t_g is not long enough. For example, the accuracy for the wave gesture degrades faster than the rest of the gestures, since a larger number of samples is required to

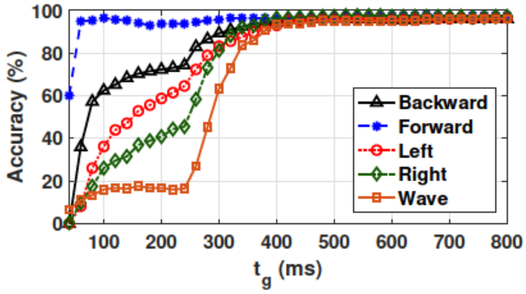


Fig. 6. Accuracy of gesture recognition for all the users.

extract its signature. Hence, we use $t_g = 380$ ms as the lower bound of the gesture recognition time. We also note that the accuracy of the baseline NN degrades more rapidly with reducing t_g, since it uses all the data points.

C. Optimization Results

Inputs to the proposed optimization methodology are the time horizon and corresponding energy budget. Since the harvested energy can fluctuate rapidly due to environmental conditions, we assume t_h as 1 minute and analyze the optimization results for energy budget E_b={120 mJ, 180 mJ, 240 mJ}. We also note that a larger time horizon does not change the percentage savings significantly, as it does not change the proposed algorithm. For comparisons, we use the baseline NN and a manually optimized version of the baseline NN by increasing the BLE connection interval t_{conn} to reduce the BLE overhead. Our solution (labeled as *Reduced*) uses the proposed optimization algorithm. Throughout the experiments, we enforce a minimum gesture recognition accuracy of 90%.

When the energy budget is 120 mJ, the baseline NN is able to recognize only 4 gestures, since the static energy and BLE communication consume 72.5% and 21.6% of the energy budget, respectively. The baseline NN can recognize 15 gestures by reducing BLE communication energy with longer t_{conn}. The proposed optimization provides an additional 2× boost and increases N_g to 31, as shown in Figure 7. Increasing the energy budget to 180 mJ and 240 mJ benefits the baseline NN significantly. Nevertheless, our optimization approach still provides 2× and 2.4× improvement over the optimized baseline, respectively. In particular, when the energy budget is 240 mJ, the maximum N_g for our approach is limited by the 1-minute time horizon, *not the energy budget*.

We illustrate the optimization results in more detail in Figure 8. The implicit upper bound induced by t_h is shown with the dotted curve, while the vertical dashed line illustrates the accuracy constraint. The result obtained with the baseline NN is the point labeled with the △ marker. In contrast, our optimization approach enables us to vary the number of gestures $N_g(t_g)$ along the solid curve. This curve is a decreasing function of t_g, as shown by the proof in [11]. Hence, the optimal point is determined as the minimum t_g, as stated in Section III-E.

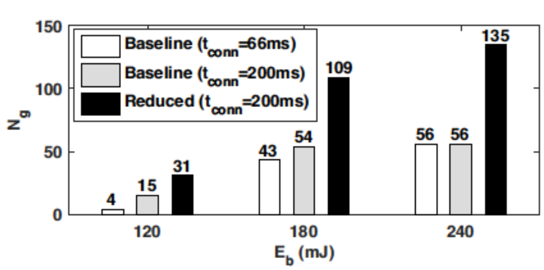


Fig. 7. Comparison of the number of recognized gestures for various energy budgets.

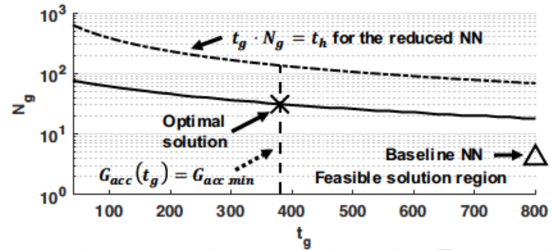


Fig. 8. Illustration of the optimal solution when $E_b = 120mJ$.

V. Conclusions

Wearable IoT devices are becoming popular in interesting applications such as gesture-based control due to their small form factor and low cost. Battery life limitation is one of the major issues of wearable devices. Hence, energy harvesting and optimal use of the harvested energy are critical. We presented an optimization approach to maximize the number of gestures can be recognized under the energy budget and accuracy constraints. We show that the proposed algorithm shows up to 2.4× improvement in the number of recognized gestures over the optimized baseline.

Acknowledgment: This work was supported partially by NSF CAREER Award CNS-1651624, and NRF grant funded by the Korea government (MSIT) (NRF-2017R1D1A1B03032382 and 2018R1C1B5047150).

References

[1] A. Akl, C. Feng, and S. Valaee, "A Novel Accelerometer-Based Gesture Recognition System," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6197–6205, 2011.

[2] H. Banaee, M. U. Ahmed, and A. Loutfi, "Data Mining for Wearable Sensors in Health Monitoring Systems: A Review of Recent Trends and Challenges," *Sensors*, vol. 13, no. 12, pp. 17 472–17 500, 2013.

[3] D. Bandyopadhyay and J. Sen, "Internet of Things: Applications and Challenges in Technology and Standardization," *Wireless Personal Commun.*, vol. 58, no. 1, pp. 49–69, 2011.

[4] G. Bhat, J. Park, and U. Y. Ogras, "Near Optimal Energy Allocation for Self-Powered Wearable Systems," in *Proc. Int. Conf. on Comput.-Aided Design*, 2017, pp. 368–375.

[5] Q. Brogan, T. O'Connor, and D. S. Ha, "Solar and Thermal Energy Harvesting with a Wearable Jacket," in *Proc. IEEE Intl. Symp. Circuits and Syst.*, 2014.

[6] DMI International Distribution Ltd. (2017) Curved lithium thin cells. http://www.dmi-international.com/data%20sheets/Curved%20Li%20Polymer.pdf.

[7] Guangzhou Markyn Battery Co., Ltd. (2011) GM555375. https://www.powerstream.com/poly/GM555375.pdf.

[8] S. Hiremath, G. Yang, and K. Mankodiya, "Wearable Internet of Things: Concept, Architectural Components and Promises for Person-Centered Healthcare," in *Proc. Int. Conf. Wireless Mobile Commun. and Healthcare*, 2014, pp. 304–307.

[9] J. Korpela *et al.*, "An Energy-Aware Method for the Joint Recognition of Activities and Gestures Using Wearable Sensors," in *Proc. Int. Symp. on Wearable Comput.*, 2015, pp. 101–108.

[10] J. A. Paradiso and T. Starner, "Energy Scavenging for Mobile and Wireless Electronics," *IEEE Pervasive Computing*, vol. 4, no. 1, pp. 18–27, 2005.

[11] J. Park *et al.*, "Optimizing Operations per Joule for Energy Harvesting IoT Devices," Technical Report, Arizona State University, [Online], https://elab.engineering.asu.edu/wp-content/uploads/2018/06/energy_gesture.pdf.

[12] M. Patel and J. Wang, "Applications, Challenges, and Prospective in Emerging Body Area Networking Technologies," *IEEE Wireless Commun.*, vol. 17, no. 1, 2010.

[13] R. Rawassizadeh, B. A. Price, and M. Petre, "Wearables: Has the Age of Smartwatches Finally Arrived?" *Commun. ACM*, vol. 58, no. 1, pp. 45–47, 2014.

[14] P. Zappi *et al.*, "Activity Recognition from On-Body Sensors: Accuracy-Power Trade-Off by Dynamic Sensor Selection," *Lecture Notes in Comput. Sci.*, vol. 4913, p. 17, 2008.