Singapore Management University

Institutional Knowledge at Singapore Management University

Research Collection School Of Computing and Information Systems | School of Computing and Information Systems

7-2023

Recognizing hand gestures using solar cells

Dong MA
Singapore Management University, dongma@smu.edu.sg

Guohao LAN

Mahbub HASSAN

Wen HU

B. Mushfika UPAMA

See next page for additional authors

Follow this and additional works at: https://ink.library.smu.edu.sg/sis_research

 Part of the Artificial Intelligence and Robotics Commons

Citation

MA, Dong; LAN, Guohao; HASSAN, Mahbub; HU, Wen; UPAMA, B. Mushfika; UDDIN, Ashraf; YOUSEEF; and Moustafa. Recognizing hand gestures using solar cells. (2023). *IEEE Transactions on Mobile Computing*. 22, (7), 4223-4235.
Available at: https://ink.library.smu.edu.sg/sis_research/7014

This Journal Article is brought to you for free and open access by the School of Computing and Information Systems at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Computing and Information Systems by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email cherylds@smu.edu.sg.

Author

Dong MA, Guohao LAN, Mahbub HASSAN, Wen HU, B. Mushfika UPAMA, Ashraf UDDIN, YOUSEEF, and Moustafa

This journal article is available at Institutional Knowledge at Singapore Management University: https://ink.library.smu.edu.sg/sis_research/7014

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3148143, IEEE Transactions on Mobile Computing

Recognizing Hand Gestures using Solar Cells

Dong Ma, *Member, IEEE,* Guohao Lan, *Member, IEEE,* Changshuo Hu, Mahbub Hassan, *Senior Member, IEEE,* Wen Hu, *Senior Member, IEEE,* Mushfika B. Upama, Ashraf Uddin, *Senior Member, IEEE,* Moustafa Youssef, *Fellow, IEEE*

Abstract—We design a system, SolarGest, which can recognize hand gestures near a solar-powered device by analyzing the patterns of the photocurrent. SolarGest is based on the observation that each gesture interferes with incident light rays on the solar panel in a unique way, leaving its discernible signature in harvested photocurrent. Using solar energy harvesting laws, we develop a model to optimize design and usage of SolarGest. To further improve the robustness of SolarGest under non-deterministic operating conditions, we combine dynamic time warping with Z-score transformation in a signal processing pipeline to pre-process each gesture waveform before it is analyzed for classification. We evaluate SolarGest with both conventional opaque solar cells as well as emerging see-through transparent cells. Our experiments demonstrate that SolarGest achieves 99% for six gestures with a single cell and 95% for fifteen gesture with a 2×2 solar cell array. The power measuement study suggests that SolarGest consume 44% less power compared to light sensor based systems.

Index Terms—Solar Energy Harvesting, Visible Light Sensing, Gesture Recognition.

1 Introduction

As all types of devices around us become smart and capable of taking input from us, there is a necessity to explore practical and efficient ways to interact with them. Gestures, being regarded as one of the most natural ways for human to communicate with anyone or anything, have sparked the integration of gesture recognition to consumer electronics [2, 3]. To recognize human gestures, various sensors and modalities, such as WiFi (electromagnetic) [4, 5, 6], camera (image) [7, 8], microphone (acoustic) [9, 10], accelerometer (motion) [11, 12], and light sensor (ambient light) [13, 14, 15, 16, 17], have been investigated and validated. These modalities, either deployed in the environment (e.g., WiFi, camera), or implemented on the device itself (e.g.,microphone, accelerometer, ad light sensor), can achieve great gesture recognition performance in certain conditions.

With existing approaches, however, gesture recognition would pose two challenges on electronic devices. First, as most consumer devices are usually powered by small batteries with limited energy budget, gesture recognition sensors would draw power from the battery and therefore shorten the continuous operation time of the device. Second, embedding such sensors on devices, especially Internet of Things (IoTs) with limited form factor, requires addition space (e.g., an array of light sensors), thereby making device miniaturization more difficult.

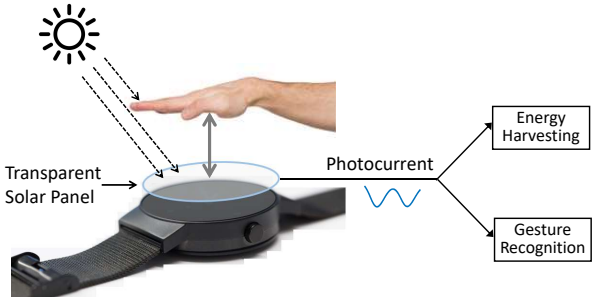


Fig. 1: Illustration of a transparent solar powered smartwatch with solar-based gesture recognition.

To tackle them, we propose the use of solar cells for gesture recognition on consumer devices. On one hand, solar cells are known as energy harvesting materials, which could convert energy from incident light into electricity and therefore supply power to the electronic devices. Further, when employing solar cells as gesture recognition sensors, they are passive elements without consuming any power from the device battery. On the other hand, embedding solar cells on IoTs can be space-free as the conventional silicon solar cells (opaque) can be used as the shell or case of the device. Moreover, the emerging *transparent* solar cells [18, 19] open more opportunities for gesture recognition and power supply on mobile devices. Manufactured with innovative organic materials, transparent cells absorb and harvest energy from infrared and ultraviolet lights, but let the visible lights pass through so we can see through the solar panel like a clear glass. Such property allows the integration of solar panels to the entire device body, especially on top of the screen, thereby eliminating the requirement of

(Corresponding author: Guohao Lan)
Dong Ma is with the School of Computing and Information Systems, Singapore Management University, Singapore (E-mail: dongma@smu.edu.sg).
Guohao Lan is with the Department of Software Technology, Delft University of Technology, 2628CD Delft, the Netherlands (E-mail: g.lan@tudelft.nl).
Changshuo Hu, Mahbub Hassan, and Wen Hu are with the School of Computer Science and Engineering, University of New South Wales (UNSW), Sydney, Australia (E-mail: changshuo.hu@unswalumni.com, mahbub.hassan@unsw.edu.au, wen.hu@unsw.edu.au).
Mushfika B. Upama and Ashraf Uddin are with the School of Photovoltaic and Renewable Energy Engineering, University of New South Wales (UNSW), Sydney, Australia (E-mail: upama.mushfika@gmail.com, a.uddin@unsw.edu.au).
Moustafa Youssef is with the American University in Cairo (AUC) and Alexandria University, Egypt (E-mail: moustafa@alexu.edu.eg).
An early version of this work appeared at ACM MobiCom 2019 [1].

1536-1233 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3148143, IEEE Transactions on Mobile Computing

additional space. Figure 1 depicts how a transparent solar cell fitted on the screen of a smart watch can be used for the dual purpose of energy harvesting and gesture recognition [1]

In this work, we present SolarGest, a hand gesture recognition system with solar cells. It is based on the observation that any hand gesture interferes with incident light rays on the solar panel in a unique way, leaving its discernible signature in the generated photocurrent signal. Combining solar energy harvesting law with geometry, we develop a model that could quantitatively study the impact of factors affecting the gestured-induced photocurrent, such as the incident angles and intensities of ambient lights, the form factor as well as the energy harvesting density of the solar panel, and etc.. In addition, we devise a gesture recognition framework including signal processing and one-dimensional Convolutional Neural Network (CNN) based classification. Particularly, we design a signal alignment stage that combines dynamic time warping (DTW) and Z-score transformation to deal with the temporal and amplitude variations on gesture waveforms incurred by various factors.

We manufacture both transparent and opaque solar cells for experimentation. With 6,960 gesture samples collected from three subjects using a single solar cell, we demonstrate SolarGest can achieve 98% recognition accuracy on six gestures. Further, we extend our previous work [1] by designing a 2×2 solar cell array to collect 3,000 gesture samples and demonstrate it obtains 95% recognition accuracy on fifteen gestures. In addition, our power consumption measurement reveals that SolarGest consumes 44% less power compared to light sensor based systems. As solar cells absorb energy from any form of light, SolarGest has a wide application scenarios both indoor and outdoor. For instance, users can purchase from solar-powered vending machines, configure solar-powered garden lights, or operate solar-powered calculators by simply using gestures.

Key contributions of this paper can be summarized as follows:

- Using solar energy harvesting laws and geometric analysis, we developed a model to simulate photocurrent waveforms produced by hand gestures. With this model, we can analyze the impact of various parameters on the photocurrent waveforms.
- We presented SolarGest, an end-to-end hand gesture recognition framework with solar cells. The framework contains a signal processing pipeline and a CNN-based classifier.
- With the developed transparent and opaque solar cells, we collected 6,960 gesture samples for six gestures under different conditions and demonstrated that even with a single transparent cell, SolarGest can recognize gestures with an accuracy up to 99%.
- We developed a 2×2 opaque solar cell array and expand the gesture set to fifteen gestures. The experimental results show that SolarGest achieves 95% accuracy in recognizing the fifteen gestures.
- Finally, we experimentally demonstrated that SolarGest consumes 44% less power compared to light sensor based gesture recognition systems.

1. Note that the screen light of smartwatch has no impact on the proposed gesture recognition system as: (1) transparent solar cells mainly absorb energy from invisible light (infrared and ultraviolet spectrum), while the screen light is fully visible; (2) transparent solar cells are usually unidirectional and can only absorb energy from the front side.

2 Rationale and Gesture Modelling

SolarGest recognize hand gesture based on the generated photocurrent. In this Section, using fundamentals of solar energy harvesting, we present the operation rationale of SolarGest. Then, with simple geometric arguments, we derive a model to simulate photocurrent waveforms produced by hand gestures in both vertical and horizontal planes relative to the solar panel.

2.1 Photovoltaic Theory

Based on the photovoltaic effect [20], solar cells convert the energy from incident light into electrical current (photocurrent). The amount of generated photocurrent can be denoted as a function of the *form factor* of the solar cell and its *current density*. The current density is defined as the amount of photocurrent generated per unit area (e.g., mA/cm^2), which is a measure of solar energy harvesting efficiency. As current density depends on the light intensity of the operating environment, for fair comparison among different solar cells, it is typically reported under a standard lighting condition, named Global Standard Spectrum (AM1.5g) [21, 22]. Then, the standard current density J^*_{SC} is obtained as [23]:

$$J^*_{SC} = \frac{q}{hc_0} \int_0^\infty a(\lambda) I(\lambda) \lambda d\lambda, \tag{1}$$

where q represents the elementary charge, c_0 represents the speed of light in free space, and h represents the Planck's constant. Symbol λ is the wavelength of incident light. $a(\lambda)$ and $I(\lambda)$ refer to the solar cell absorption efficiency and light intensity at wavelength λ, respectively. Due to the linear relationship between current density and light intensity [24], the current density J_{SC} (mA/cm^2) at any light intensity I can be computed by

$$J_{SC}(I) = \frac{I}{I^*} J^*_{SC}, \tag{2}$$

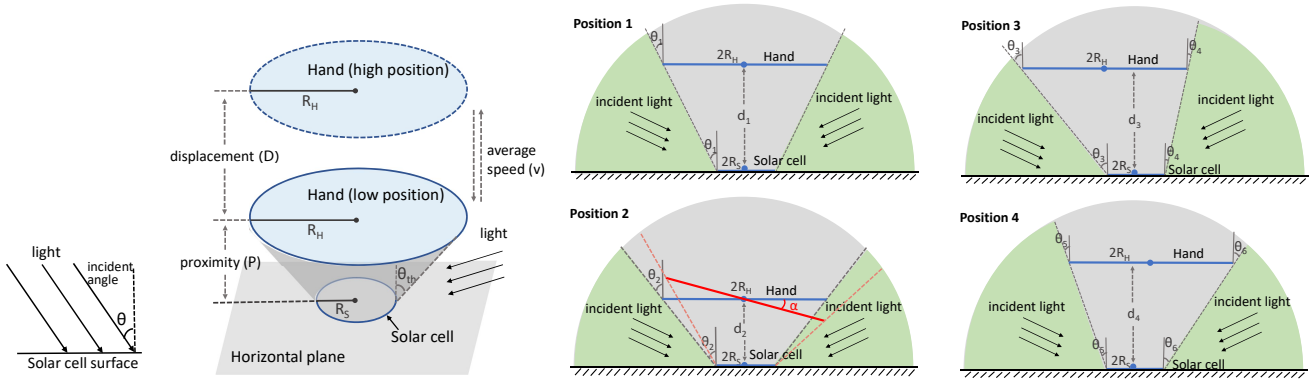
where $I^* (= 100 mW/cm^2)$ represents the light radiance power under Global Standard Spectrum (AM1.5g). Then, using Lambert's cos(θ)-law [25], the generated photocurrent J, is calculated as

$$J = \int_0^{\pi/2} S \cdot J_{SC}(I) \cdot cos(\theta) d\theta, \tag{3}$$

where S refers to the form factor of the solar cell and θ is the *incident angle*, i.e., the angle between light beam and the surface normal (see Figure 2(a)). As light from different sources (such as sun, fluorescent lamp and LED) have a different spectral irradiance profile,the amount of generated photocurrent could be different even under the same light intensity. Although we derive the model based on AM1.5g (dedicated for sunlight), it can be applied to other irradiance spectrum as gestures are differentiated due to their unique patterns (relative changes), rather than the

1536-1233 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3148143, IEEE Transactions on Mobile Computing



(a) Incident light. (b) 3D geometric model. (c) 2D analysis - vertical movement. (d) 2D analysis - horizontal movement.

Fig. 2: (a) Illustration of incident angle. (b) 3D geometric model of SolarGest. (c) 2D geometric analysis of vertical movement. (d) 2D geometric analysis of horizontal movement.

absolute values. We further validate this in Figure 12, which confirms that gesture patterns collected under fluorescent light are consistent with those derived using the model.

2.2 Modelling Solar Gestures

As illustrated in Figure 2(b), we develop a three-dimensional (3D) geometric model to simulate solar photocurrent under hand gestures. The human hand and solar cell are modelled as round surfaces with radius R_H and R_S, respectively. As many IoT devices have small form factors [26], in this paper, we consider the case where solar cell is smaller than hand size (e.g., Lunar Watch [27]), i.e., $R_S < R_H$[2]. The solar cell is assumed to be placed on a horizontal surface and a hand performs different gestures in a parallel plane above it. During a gesture, we define the minimum distance between the solar cell and hand as *proximity*, denoted by P, and define the vertical movement distance as *displacement*, denoted by D. Since $R_S < R_H$, only the light rays with incident angles larger than a certain threshold (θ_{th}) can reach the solar cell.

Figure 2(c) and (d) show the longitudinal section of the 3D model, in which solar cell and hand are represented by two line segments with lengths $2R_S$ and $2R_H$, respectively. The *green area* corresponds to the angular space where light can be absorbed by the solar cell, whereas the light in *gray area* is blocked by human hand. In fact, a gesture is comprised of a sequence of hand positions. Given the initial hand position, moving direction and speed of hand movement, we can compute hand positions at any successive points in time. Taking *Up* gesture as an example, if the initial distance (at time zero) between hand and solar cell is d and the hand moves in a constant speed v, at time t, the distance between hand and solar cell becomes $d + vt$. Thus, the corresponding threshold angles $\theta_{\text{th1}}(t)$ and $\theta_{\text{th2}}(t)$ for the two absorption angular spaces are

$$\theta_{\text{th1}}(t) = \theta_{\text{th2}}(t) = \arctan(\frac{R_H - R_S}{d + vt}). \tag{4}$$

2. Note that the model can be easily extended to $R_S > R_H$. In this case, the inner part of the solar cell, a circle with radius R_H, will be affected by hand movement, but the residual area will generate steady photocurrent during a gesture. Thus, total photocurrent would be obtained as the sum of current from the two parts.

Since only light beams from the two green areas can be absorbed, the photocurrent $J(t)$ can be calculated as

$$\begin{aligned} J(t) = &\int_{\theta_{\text{th1}}(t)}^{\pi/2} S \cdot J_{SC}(I) \cdot cos(\theta) d\theta \\ &+ \int_{\theta_{\text{th2}}(t)}^{\pi/2} S \cdot J_{SC}(I) \cdot cos(\theta) d\theta. \end{aligned} \tag{5}$$

From Eq.5, the complete gesture waveform can be obtained by generating photocurrent values at successive points in time, i.e., $(J(t_1), J(t_2), ..., J(t_n))$, where $J(t_1)$ and $J(t_n)$ represent the start and end of the gesture, respectively. With this model, we study the impact of various parameters (like light intensity, efficiency and form factor of the solar cell and user hand size) on the photocurrent waveforms, as presented in Figure 6.

3 SolarGest Design

3.1 Overview

We present the system architecture and workflow of SolarGest in Figure 3. When performing a gesture at certain lighting condition, a solar-powered device measures time-series of photocurrent and delivers the signal to a gesture recognition system, which could be located on an edge device, such as smartphone, laptop, or home hub (note that such edge-based processing will ensure that latency is minimal), using low-power communications like backscatter or BLE. The gesture recognition system identify the gesture and either sends that information back to the originating device if local control in the device is involved, or communicates with other IoT devices based on the desired action from the gesture. Next, we describe the proposed gesture recognition framework (including signal processing and classification) in detail.

3.2 Recognition Framework

3.2.1 Signal Processing

After acquiring photocurrent signals from the solar cell, we introduce a signal processing pipeline that deals with three

1536-1233 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3148143, IEEE Transactions on Mobile Computing

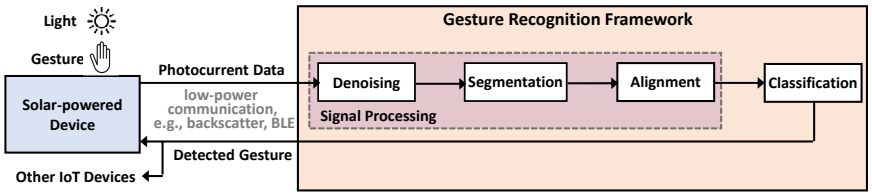
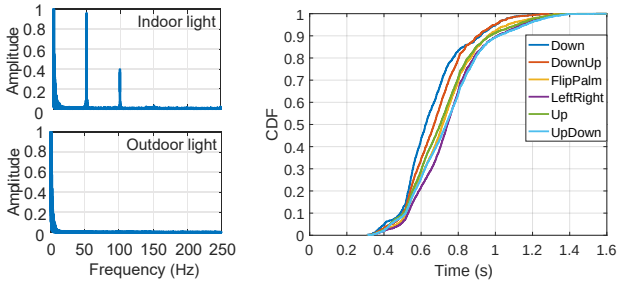


Fig. 3: SolarGest System Architecture.



(a) (b)

Fig. 4: (a) FFT analysis of photocurrent signals collected under indoor fluorescent light (up) and outdoor natural light (bottom), (b) CDFs of gesture duration.

specific issues. First, it removes noise contained in photocurrent signal using discrete wavelet transform (DWT). Then, the boundaries of the gesture are detected using a segmentation algorithm. Finally, a signal alignment module applies a combination of dynamic time warping (DTW) and Z-score transformation on the segmented signal to address specific alignment issues that are caused by variations in operating conditions, such as hand motion speed, lighting conditions etc.

Denoising: Raw photocurrent signals are noisy as shown in the bottom row of Figure 12. The fast Fourier transform (FFT) graphs in Figure 4(a) reveal that there is a 50Hz noise when the signal is collected indoor under a ceiling light powered by 50Hz AC current, but such noise is absent when measured outdoor under the sun. In addition, due to the minor imperfections of the hardware (e.g., micro-controller), Gaussian noise also exists in the photocurrent signal. Discrete wavelet transform (DWT) has been demonstrated to be an affective method to filter noise from both time and frequency domain [13, 28]. By hierarchically decomposing a signal in various resolutions at different frequency range, DWT computes detail coefficients and approximation coefficients at each decomposition level. The main procedure to denoise a signal with DWT is to modify the detail coefficients based on thresholding strategies. In specific, we divide the denoising process into three steps.

First, SolarGest decomposes the photocurrent signal to level 5. The choice of level 5 is based on the sampling rate. Since we sample data at 500Hz, the highest frequency contained in the signal is 250 Hz due to the Nyquist Theorem. As observed from Figure 4(a), the gesture frequency is actually less than around 5Hz. During DWT decomposition, the frequency span at each level is half of the level before it. Thus, at level 5, the frequency range is $[0, 250/2^5]$Hz, i.e., [0, 7.8] Hz, which covers the gesture frequency. Second, a soft thresholding scheme is applied to the detail coefficients at level 5, which shrinks both positive and negative coefficients towards zero. The threshold is adaptively computed using the principle of Stein's Unbiased Risk Estimate (SURE) [29]. Finally, we apply inverse DWT to the altered detail coefficients and unmodified approximation coefficients for reconstruction of the denoised signal.

Gesture Segmentation: After denoising, the next step is to segment exact gestures from the time-series of signal. To detect the start and end of a gesture, previous works either use a preamble scheme [4] or a threshold-based method (i.e., a start is detected once the value is higher than a predefined threshold and an end is detected when the values fall below the threshold) [13, 14]. However, the former requires users to perform an additional gesture every time, which is not user-friendly. The threshold-based method does not work if the amplitude before and after a gesture is different (e.g., *Up* and *Down* shown in Figure 12). To enable detection of gestures, like many other gestures recognition systems [4, 13, 14, 30], SolarGest requires users to take a short pause before and after a gesture. Thus, we proposed a new segmentation algorithm, which can accurately detect the *plateau periods* (i.e., pauses) before and after a gesture.

Specifically, we apply a sliding temporal window on the denoised signal. A gesture start is detected if the following two conditions are satisfied: (1) the standard deviation of the samples in current window is lower than a pre-defined threshold *stdThr*; (2) the difference between the last sample in current window and the mean of all the samples in the window is higher than a threshold *meanThr*. The first condition ensures that the current window is in a plateau, while the second condition determines that a gesture starts right after a pause. Thus, the last sample in current window is regarded as a gesture start. The same principle is utilized to detect the end of a gesture and consecutive samples between start and end are extracted as a gesture. To minimize the probability of falsely extracting an un-occurred gesture, we further apply a gesture length constraint based on our experimental data. Figure 4(b) presents the Cumulative Distribution Functions (CDFs) of gesture durations when three subjects perform the six different gestures. We can observe that around 90% gestures are completed within 1 s. Therefore, we apply a length constraint which ensures gestures less than 0.2 s or greater than 1.4 s are discarded. *meanThr* and *stdThr* are optimized through trial-and-error procedure and the values used in our work is 0.5 and 0.25, respectively. The detected gestures are interpolate to the same length so that they can be fed to the classifier.

Figure 5 shows the gesture segmentation result, where the green dots represent the start points and red squares represent the end points. Note that during a data collection

1536-1233 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3148143, IEEE Transactions on Mobile Computing

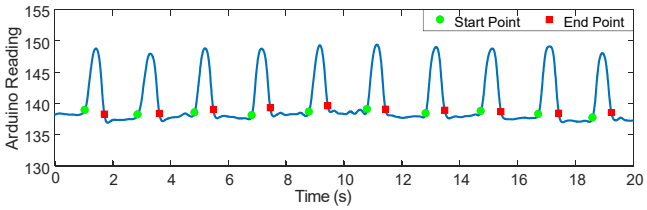


Fig. 5: Segmentation performance. The green dots represent the detected start points and the red squares represent the detected end points.

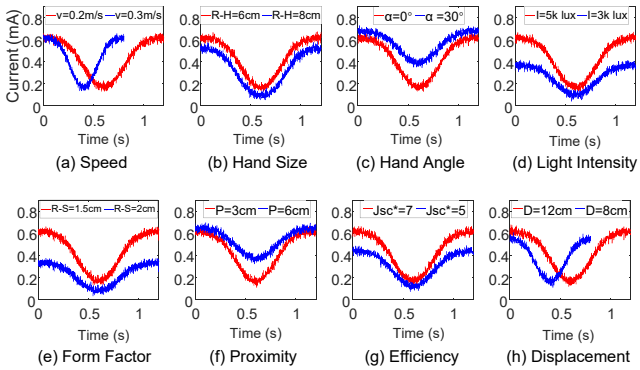


Fig. 6: The impact of different parameters on gesture profile. In each graph, only a specific parameter varies and the rest are in default value.

session, the user always keeps his/her hand within the operating region thus avoiding any *transition effects*, i.e., a slightly descending/ascending signal caused by entering/leaving the operating region. With the proposed segmentation algorithm, SolarGest successfully identifies 96% of gestures in our dataset while incurring no false positives.

Signal Alignment: Using the developed model presented in Section 2, we study the impact of eight practical parameters, i.e., device parameters such as solar cell form factor and efficiency, environment parameters such as light intensity and size, as well as gesture parameters such as speed, hand angle, proximity, and displacement, on the gesture profiles, as presented in Figure 6. In each graph, only a specific parameter varies and the rest are set to default values[3]. It can be observed that each parameter indeed affects the gesture waveform and the impact can be categorized into *temporal variation* (variation in waveform duration) and *amplitude variation*. Specifically, different gesture speeds and displacements lead to varying gesture durations, making the same gestures mismatched in time dimension. Other parameters, such as light intensity and hand angle (refers to the angle between hand and horizontal plane, as shown in Figure 2(c)), result in amplitude shifts.

SolarGest adopts a signal alignment phase to tackle the two issues. Taking two *LeftRight* gesture as an example, Figure 7 illustrate the alignment process. We first apply Z-score transformation, which is known to be an effective function to make multiple signal with different amplitudes

3. Default parameter values: R_H=6cm, R_S=2cm, D=12cm, P=3cm, J^*_{SC}=7mA/cm^2, v= 20cm/s, I=5000lux.

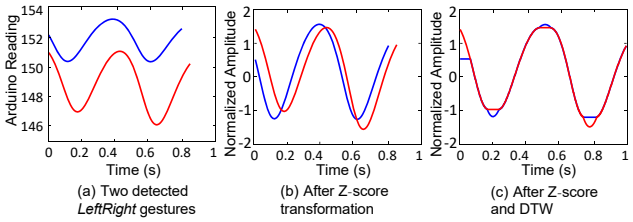


Fig. 7: Illustration of signal alignment using Z-score transformation and DTW.

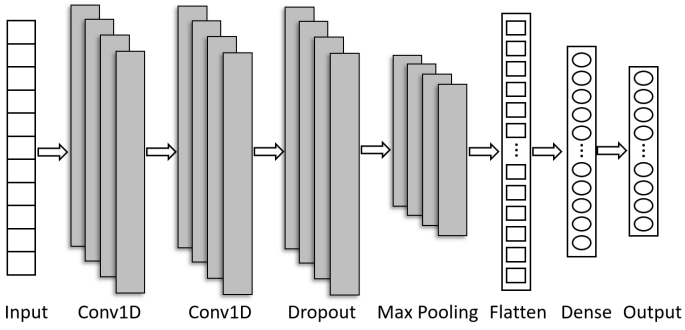


Fig. 8: The designed one-dimensional CNN for gesture classification.

comparable [31], on the gesture signal. From Figure 7(b), we can see that after Z-score, the amplitudes are converted to the same scale between [-2,2] but we can still observe the temporal misalignment issue. Due to the successful application of DTW to cope with temporal mismatch (like in speech recognition [32] and activity recognition [33]), we then apply DTW to the processed gesture signals. The performance is shown in Figure 7(c), from which we can see that the two signals almost overlap, demonstrating the effectiveness of signal alignment.

3.3 Gesture Classification

We design a one-dimensional (1D) convolutional neural network (CNN) to classify the detected gestures, given its effectiveness on mining information from sequential data as well as low computational overhead [34]. As shown in Figure 8, the network contains two 1D convolution layer (with 64 filters and kernel size of 3) as it is common to stack two layers for greater capability of learning features from the input data. A dropout layer (with dropout rate of 0.5) is followed for regularization. Then we add a max pooling layer (with pool size of 2) to consolidate the learned features to the most essential elements. After pooling, the learned features are flattened to one long vector and passed through a fully connected layer (Dense) before making a prediction with the output layer (Dense). We use relu as the activation function and adam as the optimizer. The loss function is set to categorical cross entropy as this is a multi-class classification problem.

We also benchmark the performance of the proposed CNN model with other deep neural networks and machine learning classifiers. Specifically, we consider a long short-term memory (LSTM) neural network based classifier due to its superior performance on mining temporal correlations

1536-1233 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3148143, IEEE Transactions on Mobile Computing

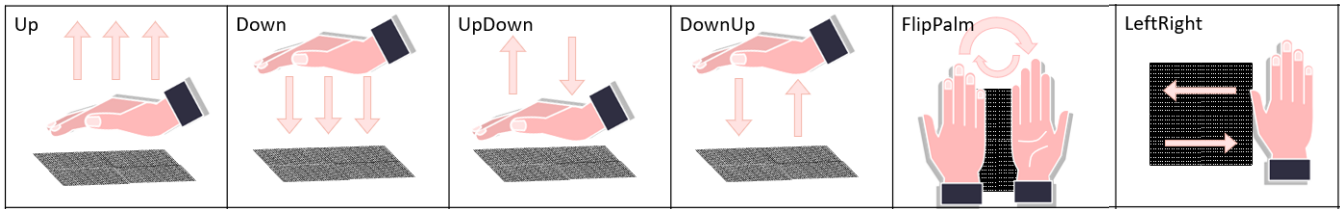
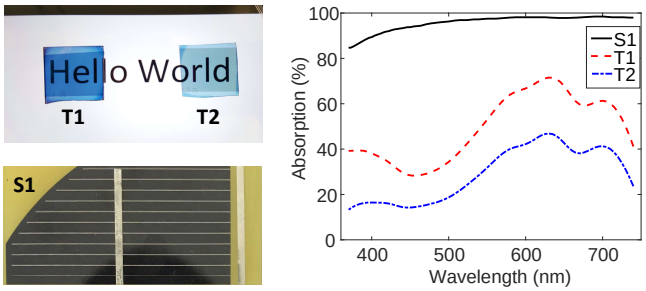


Fig. 9: Illustrations of the six hand gestures conducted over the solar cells.



(a) (b)

Fig. 10: (a) Effect of placing the two transparent solar cells T1 and T2 on an iPhone 7 screen that displays the text 'Hello Word' (up) and the silicon based solar cell S1 (low), (b) Absorption spectra of the three solar cells S1, T1, and T2 within visible light band.

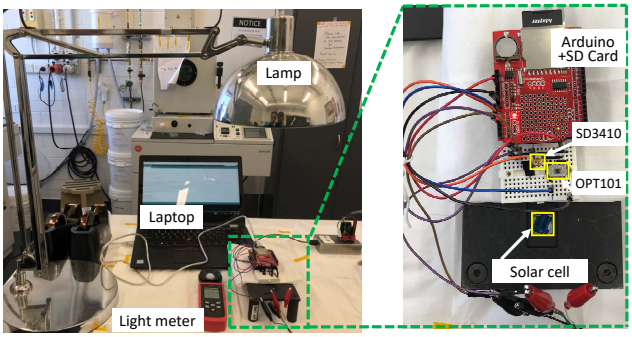


Fig. 11: Data collection setup.

from sequential signals [35]. The LSTM classifier consists of one LSTM layer where each cell contains 32 hidden units, one flatten layer, and one fully connected layer as used in the CNN model. For machine learning based models, we consider five typical classifiers, i.e., Support Vector Machine (SVM), K Nearest Neighbours (KNN), Decision Tree (DT), and Random Forest (RF). The hyper-parameters for each classifier are fine-tuned to provide optimal performance. Given that DWT coefficients are a good representation of the signal [4, 30, 36], we use them as features for the five machine learning models. As described during denoising, we perform five level DWT decomposition on each gesture with Daubechies2 wavelet and extract the detail coefficients in the fifth level as the features. The models are evaluated with the 5-fold cross validation mechanism, i.e., the dataset is randomly split into five sections and we iteratively select one section as the testing data and use the rest as the training data. All the codes are implemented in Tensorflow with Python.

4 Implementation and Data Collection

Given that transparent solar cells are currently not available off-the-shelf, we prototype both transparent and opaque solar cells to collect gesture data under various conditions.

4.1 Solar Cell Prototype

As shown in Figure 10(a), we prototyped three different solar cells, a 10x5cm silicon-based opaque solar cell (S1) and two 1x1cm transparent solar cells (refer to as T1 and T2), in our photovoltaic laboratory. The two transparent solar cells were made from the same organic material (PBDB-T: ITIC) but with different transparency levels (T1 20.2% and T2 35.3%) and thickness (T1 143nm and T2 53nm). To demonstrate the 'see-through' property of the transparent solar cells, we placed them on the screen of an iPhone7. As shown in Figure 10(a), we can clearly see the displayed 'Hello World' context through both T1 and T2. T2 provides a better 'see-through' performance as it has a higher transparency to allow more visible light to pass through. In terms of the energy harvesting efficiency, T1 and T2 provide current densities of $13.82 mA/cm^2$ and $6.85 mA/cm^2$, respectively. More details of our transparent solar cell prototypes are available in [37]. In Figure 10(b), we compare the absorption efficiency of the three solar cells in the visible light spectrum. We can notice that the opaque solar cell S1 achieves nearly 100% absorption efficiency over the entire wavelength range, whereas, the absorption rate of T1 and T2 is only 50% and 30%, respectively, on average.

4.2 Gesture Data Collection

With the prototype solar cells, we have collected a comprehensive gesture dataset for performance evaluation[4]. As shown in Figure 11, the solar cells are connected to an Arduino Uno board. The output of the solar cell is sampled by the Arduino via its onboard ADC at 500Hz and saved in the microSD card. A lamp is used to control the light intensities. For comparison purpose, we also collected the photocurrent signal from two different light sensors, TI OPT101 and Honeywell SD3410, which are widely used in ambient light based gesture recognition systems [15, 16, 17, 38]. Figure 11 illustrates our data collection setup at an indoor environment, which is conducted in our photovoltaic research lab due to the special (and bulky) tools required to connect the transparent cell output to the Arduino.

4. Ethical approval for carrying out this experiment has been granted by the corresponding organization.

1536-1233 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3148143, IEEE Transactions on Mobile Computing

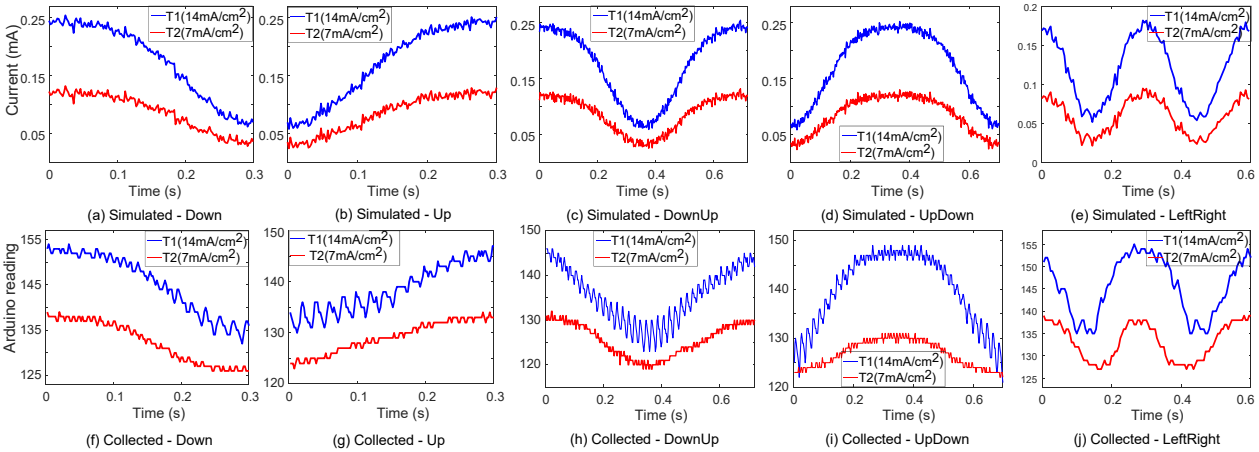


Fig. 12: Comparison of simulated gesture signal with the signal generated by solar cells.

TABLE 1: Experiment setting.

Parameter	Option	Value
Solar cell	3	transparent solar cell: T1, T2 opaque solar cell: S1
Light intensity	4	10, 50, 800, 2600,70klux
Interference	2	with, without
Gesture	6	Down, DownUp, FlipPalm, LeftRight, Up, UpDown
Subject	3	1 male, 2 female
Photodiode	2	TI OPT101, Honeywell SD3410

During data collection, we have considered many different settings, including: (1) three solar cells with different energy harvesting efficiencies/transparencies; (2) five light intensity levels for indoor and outdoor combined (i.e., 800 lux and 2600 lux for transparent solar cells only under indoor lab environment; 10 lux, 50 lux, 800 lux, 2600 lux and 70k lux for the opaque solar cell under different scenarios including indoor and sunny outdoor); (3) six hand gestures as introduced in Figure 9; (4) threes subjects to perform the gestures; and (5) scenarios with/without human interference to investigate the robustness of SolarGest against interference (data collected using the two transparent solar cells only). Specifically, the human interference is introduced by asking one subject to walk around in a half circle with radius of 30cm when another subject is performing gestures. As suggested by [39], light incident angles have little impact on the gesture recognition accuracy. Thus, we consider the case where light source is located at the top of the solar cell.

Table 1 summarizes the considered experiment settings. 3 subjects (including 1 female and 2 male) were invited to collect data. Each subject perform gestures in 13 sessions (i.e., 2 transparent solar cell $\times$ 2 light intensities $\times$ 2 interference conditions + 1 opaque solar cell $\times$ 5 light intensities), where he/she needs to perform each of the 6 gestures 40 times. Only one subject is participated in outdoor experiment and human interference is not considered here. A two-minutes break was set between sessions to avoid human fatigue. The entire data was collected over five days. In total, we created a dataset consisting of $8 \times 3 \times 6 \times 40 + 5 \times 1 \times 6 \times 40$=6,960 gestures.

TABLE 2: Recognition accuracy given different transparencies and classifiers.

Solar Cell	Classifier					
	KNN	DT	SVM	RF	LSTM	CNN
T1	96.1%	94.1%	95.1%	95.9%	98.8%	**98.9%**
T2	95.6%	93.0%	94.5%	95.2%	98.4%	**98.5%**
S1	98.0%	94.9%	96.3%	96.8%	99.4%	**99.5%**

5 Performance Evaluation

5.1 Simulated vs. Real Waveforms

Figure 12 compares simulated gesture waveforms (top row) against actual waveforms (bottom row) collected from prototype transparent solar cells for 5 different gestures. It is clear that even though we model hand and solar cell as circles, the gesture signals simulated by our model are very similar to those generated by real solar cells in terms of signal features and patterns. This demonstrates that our model can be an effective tool to study gesture recognition using solar cells under a variety of scenarios.

5.2 Gesture Recognition Performance of SolarGest

Classifier: Table 2 presents the overall classification accuracy for the three solar cells using different classifiers. We can observe that regardless of solar cells, all the classifiers achieve recognition accuracy greater than 93%. Deep neural network based classifiers outperform the machine learning ones and the proposed CNN model obtains slight higher than LSTM. Moreover, CNN has lower computation overhead which results in short training time.

Transparency Level: Table 2 also compares the gesture recognition capability of solar cells with different transparency levels. The results indicate that despite being transparent with limited energy harvesting capacities, both T1 and T2 prototypes achieved very high accuracies. As expected, the opaque solar cell (S1) results in highest accuracy due to its stronger response to light variations (therefore higher SNR). Overall, transparency level has very limited impact of the recognition performance.

Light Intensity: We select five light intensity levels that correspond to common conditions, 10 lux-dark room, 50 lux-

1536-1233 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3148143, IEEE Transactions on Mobile Computing

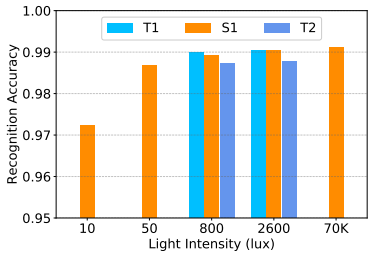
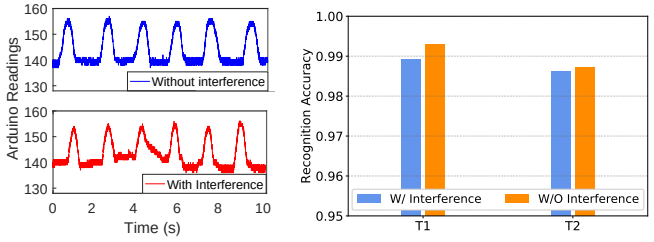


Fig. 13: Recognition accuracy under different light intensity.



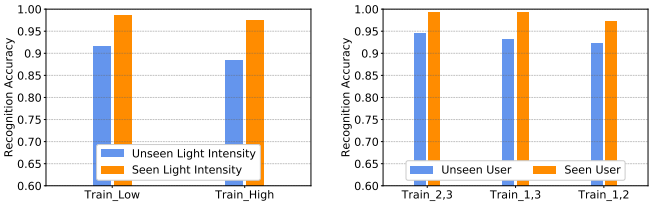
(a) (b)

Fig. 14: (a) Comparison of raw signals with and without interference. (b) Impact of interference on recognition accuracy.

living room, 800 lux-office, 2600 lux-cloudy, and 70k lux-sunny, to access the performance of SolarGest in practical scenarios. As our transparent solar cells are still in laboratory stage without encapsulation, they are very sensitive to environment factors (e.g., humidity). Consequently, we test T1 and T2 under 800 lux and 2600 lux only, while the opaque solar cell is tested under all the five conditions.

The results in Figure 13 show that, for the same solar cell, higher light intensity ensures a higher recognition accuracy, although the improvement is minor. To assess the limit of SolarGest, we create an extremely dark environment (i.e., 10 lux) by turning off all lights in a dark room except a laptop screen. We find that, with our current prototype (Arduino UNO), the collected signal always remains zero, making it impossible to detect any gesture. The reason is that the resolution of Arduino ADC (10 bit) is not sufficient to capture the minor changes in photocurrent. However, we found this issue can be resolved by either using a high resolution ADC (e.g., 16 bit) or amplifying the current. We implemented an amplification circuit and tested two amplification factors: 32× and 64×. The results show that, with both amplification levels, gesture accuracy reaches to around 97%. As a result, SolarGest works well within a broad range of light intensities.

Human Interference: As solar cells can absorb light from all directions, someone walking in vicinity of the solar cell might interfere the amount of incident light, thereby distorting the gesture signals. We investigate the robustness of SolarGest against ambient human interference by asking one subject to walk around when another subject is performing gestures. From Figure 14(a), we can see that human walking near the solar panel indeed introduces some fluctuations in the signals. Figure 14(b) quantifies the impact of such interference by plotting the accuracies with and without human walking. The results suggest that interference reduced accuracy by only 0.5% and SolarGest still achieved over 98% recognition accuracy.



(a) (b)

Fig. 15: Recognition accuracy on (a) unseen lighting environment, (b) unseen user.

TABLE 3: Effect of signal alignment.

Train	Test	Alignment	Accuracy
Low	High	w/o	86.55%
		w/	92.13%
High	Low	w/o	84.95%
		w/	89.07%

We also investigate the impact of sudden and dramatic light intensity change. For example, the global light intensity could increase by tens of thousands lux when walking from indoor to outdoor or decrease when sunlight is blocked by cloud during a gesture. We conduct the experiment using the developed geometric model to simulate gestures and run the recognition pipeline. Specifically, we train the model using gestures simulated under stable light intensity, while test using the distorted gestures (simulated by switching the light intensity in different levels and frequencies) only. Our results indicate that when light intensity changes very fast (e.g., >50Hz), the accuracy is not affected, while almost half of the distorted gestures are wrongly recognized when intensity switching rate is low (e.g., 2Hz). However, as suggested in [39], such low global light change can be effectively filtered out by subtracting it.

Unseen Scenarios: Then, we evaluate the model adaptation capability of SolarGest to unseen scenarios, i.e., the training and testing data are collected under different conditions. Two unseen cases are considered. First, we train the classifier using the data collected under one light intensity and test it by the data collected under another intensity. The performance of training and testing with the same light intensity, i.e., seen scenario, is also obtained. From Figure 15(a), we can see that SolarGest still achieves 92% accuracy even in unseen lightning environment case. Second, we train the classifier using the data collected from two subjects and test it on the remaining one. The results in Figure 15(b) indicate that SolarGest is more robust to subject difference compared to light intensity. Overall, the CNN model achieves 94% accuracy for unseen users.

Effect of Signal Alignment: Since various factors result in the amplitude or temporal shift on the solar signals, we proposed an alignment stage to address it. Here, we assess how the signal alignment affects the recognition accuracy. Specifically, we use signals collected from low and high light intensities to create the amplitude shift. Then, we train the model with data collected at one light intensity and test on the other, with and without applying signal alignment. From Table 3, we can observe that the proposed signal

1536-1233 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3148143, IEEE Transactions on Mobile Computing

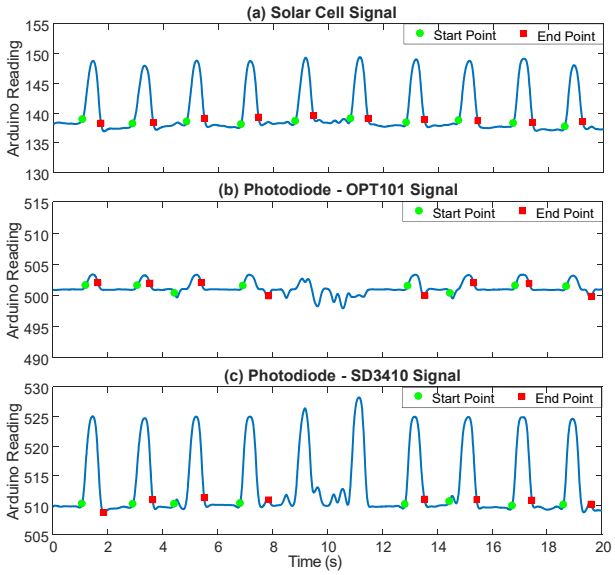


Fig. 16: Segmentation performance comparison using signals from (a) solar cell T1, (b) photodiode OPT101, and (c) SD3410, under gesture FlipPalm. The green dots represent the detected start points and the red squares represent the detected end points.

TABLE 4: Performance comparison between light sensors and solar cells.

Metric	OPT101	SD3410	T1	T2	S1
Segmentation Accuracy	70.2%	84.3%	96.2%	96.1%	96.8%
Recognition Accuracy	55.2%	89.9%	98.9%	98.5%	99.5%

alignment can enhance the recognition accuracy by around 5%.

5.3 Comparison with Light Sensor based Systems

Researchers have demonstrated the feasibility of using light sensors for gesture recognition [13, 14]. To compare SolarGest with such approach, we also collect the data of two light sensors (OPT101 and SD3410) simultaneously during the experiment with transparent solar cells. Figure 16 compares the signal traces from solar cell T1 and the two light sensors. We can observe that signal traces from light sensors are more noisy. With the signal from solar cell, the system can perfectly detect all the ten gestures, whereas, both of the two light sensors can only detect eight of the ten gestures. Table 4 compares the overall performance of light sensors and transparent solar cells in terms of both segmentation and recognition accuracies. We can notice that solar cells achieve 12% to 26% higher segmentation accuracies and at least 9% better gesture recognition accuracy compared to light sensors. These results suggest that compared with a single light sensor, solar cell exhibits stronger robustness to gesture variation and environmental noise due to its larger sensing area. To further improve the segmentation accuracy, the user might need to have a longer and stable pause before and after a gesture. Or a simple wake-up gesture, like completely covering the solar cell so that its outputs are always zeros, can be designed to trigger the gesture recognition system and accurately detect the start of a gesture.

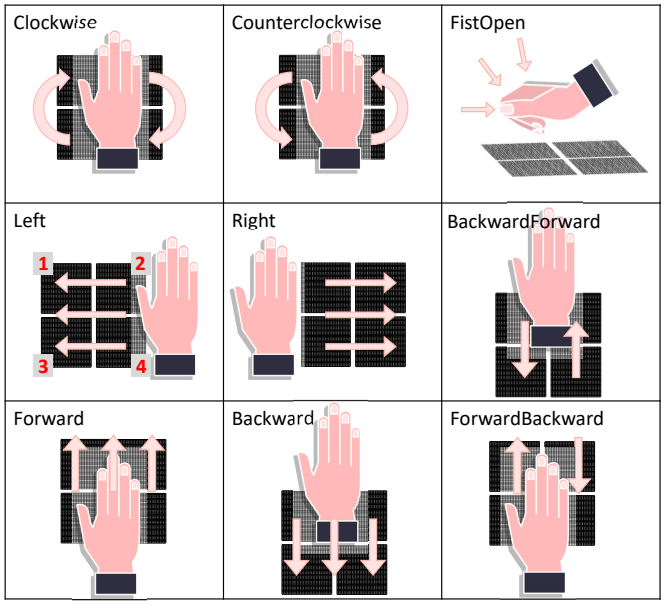


Fig. 17: Illustrations of the nine additional gestures.

TABLE 5: Recognition accuracy given different transparencies and classifiers.

SC Quantity	SC Index	Accuracy	Average Accuracy
1	{1}	31.56%	30.47%
	{2}	28.73%	
	{3}	34.32%	
	{4}	27.27%	
2	{1, 2}	87.38%	84.19%
	{1, 3}	82.84%	
	{1, 4}	87.04%	
	{2, 3}	86.86%	
	{2, 4}	76.35%	
	{3, 4}	84.68%	
3	{1, 2, 3}	94.22%	93.28%
	{1, 2, 4}	91.23%	
	{1, 3, 4}	93.14%	
	{2, 3, 4}	94.51%	
4	{1, 2, 3, 4}	95.07%	95.07%

5.4 Expanding Gesture Sets with Solar Cell Array

To investigate the capability of SolarGest for recognizing more gestures, we design additional nine user-friendly gestures as shown in Figure 17. However, it is obvious that with a single solar cell, some gestures (e.g., Left & Right) can not be distinguished as they interfere with the incident light in the same way. To tackle it, we propose the use of a 2×2 solar cell array, where each solar cell is opaque and has a size of 2.5×2.5 cm. In such case, Left and Right gesture

1536-1233 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3148143, IEEE Transactions on Mobile Computing

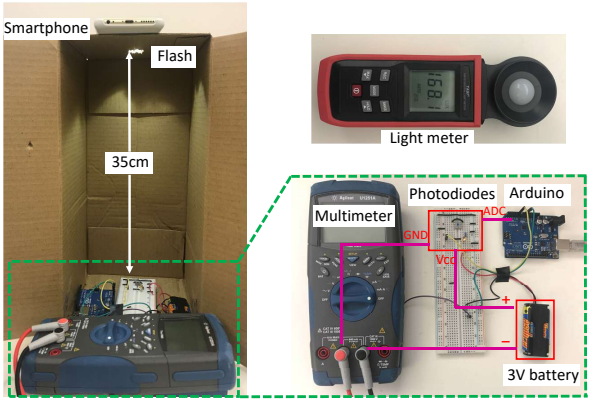


Fig. 18: Light sensor power measurement setup.

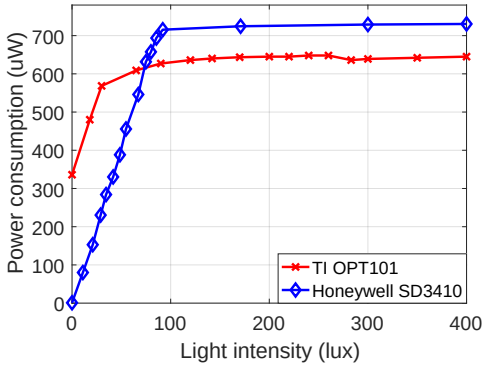


Fig. 19: Power consumption measurement of light sensors under different light intensities.

will have different sequential order when interfering with left and right solar cells of the array. Following the data collection procedures described in Section 4.2, we ask one subject to collect data for the 15 gestures. Each gesture was performed for 200 times and a total of $15 \times 200 = 3,000$ gestures has been collected.

We set the index of the four solar cells from 1 to 4 as shown in Figure 17 (Left). Then, when applying the the gesture recognition pipeline, we consider different number of solar cells with all possible combinations. When multiple solar cells are utilized, we stack the signals from each solar cell in rows, i.e., regard each solar cell as an independent channel. Then, the stacked signal is fed to the CNN classifier. The detailed results are presented in Table 5. We can observe that (1) the recognition accuracy increases when more solar cells are used, (2) a single solar cell obtains very poor performance as it can not distinguish gestures with opposite direction, (3) SolarGest achieve 95% accuracy on 15 gestures when all the four solar cells are used.

6 Power Measurements

After demonstrating the superior gesture recognition performance of SolarGest, we now investigate the power saving advantage of SolarGest against light sensor based systems. From Figure 3, we can see that SolarGest consumes power from two events: MCU sampling and data transmission. In contrast, light sensor based systems consumes additional energy in powering the photodiodes. Next, we perform a conservative comparison which assumes that only one photodiode is required for light sensor based systems (current works usually require an array of sensors [13, 14]).

MCU power measurement: since both solar cell and light sensor are sampled by analog-to-digital converter (ADC), we conducted an experiment to measure the power consumption during ADC sampling. We select the Texas Instrument SensorTag as the target device, which is equipped with an ultra-low power ARM CortexM3 MCU. The SensorTag is running with the Contiki operating system. We duty-cycled the MCU at 50Hz for sampling and applied an oscilloscope to measure the average power consumption of SensorTag during the sampling. According to our measurement, the system consumes 20.28μW in sampling the signal at 50Hz.

TABLE 6: Power consumption comparison.

Sensor	Power Consumption (uW)				Savings
	MCU	Sensor	Backscatter	BLE	
Solar Cell	20.28	0	0.023	31.11	
OPT101	20.28	39.78	0.023	31.11	66.2%/43.6%
SD3410	20.28	42.18	0.023	31.11	67.5%/45.1%

Light sensor power measurement: in addition, we also measure the power consumed by the light sensor itself. We consider two light sensors, namely TI OPT 101 and Honeywell SD 3410, that are widely used in the literature [15, 16, 17, 38]. In particular, we measured the power consumption of the sensors under different light intensities (assuming normal operation scenarios), as the datasheet only gives the power consumption when the sensor is operated in dark environment. Figure 18 illustrates the measurement setup. To minimize the effect of ambient light, we conduct the experiment in a box with one side open. A smartphone is placed on top of the box and its Flash is used as the light source. We create an aperture with a radius of 1cm on the top of the box and place the light sensor right below the aperture to ensure 0° of light incident angle. The light sensor is powered by a 3V battery and a multimeter is used to measure the current. Figure 19 presents the power consumption of the two light sensors under different light intensities. We can observe that the power consumption is not constant. When the light intensity is lower than 100 lux, the power consumption increases linearly with light intensity. Once the light intensity is higher than 100 lux, the energy consumption becomes stable. Since the light intensity of normal environment is usually higher than 100 lux, e.g., 200-800 lux for office environment, it means that, without duty-cycling (sensor always turn on), OPT101 and SD3410 consumes around 650uW and 730uW, respectively. With 50 Hz duty-cycle, the power consumption reduces to 39.78 μW and 42.18μW, respectively. In addition, our results is consistent with the datasheet when light intensity is 0 [40]. In contrast, solar cell is passive and does not require any external power.

Overall system power saving: now, we analyze the overall system power consumption. Considering 50Hz sampling rate and a duty-cycled system, Table 6 presents the power consumption of SolarGest and light sensor (i.e., photodiode) based system. Note that the photodiodes are assumed to

1536-1233 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3148143, IEEE Transactions on Mobile Computing

operate in photoconductive mode, which requires external power supply, in order to provide faster response rate [41]. The recent advancement in Wi-Fi backscattering has demonstrated that 1 Mbps data rate can be achieved with only 14.5 μW power consumption [42]. Given a sampling frequency of 50Hz, SolarGest has 100 Bytes data (2 Bytes for each 12-Bits ADC reading) to be transmitted per second. Thus, it means that 0.023 μW [5] is required for backscattering-based data transmission. Overall, the power consumption of SolarGest will be around 20.3 μW, while the consumption of light sensor based system is about 60.1 μW. Thus, SolarGest is able to save over 66% of the energy. In a more general case where BLE is used for communication, 31.11μW power is required for the data transmission (100 Bytes per second) based on our measurement (using TI SensorTag as the target platform.). In this case, the overall system power consumption for SolarGest and the two light sensors based system increase to 51.39 μW, 91.17 μW, and 93.57 μW, respectively. But SolarGest still saves at least 44% of the energy compares to light sensor based systems. Furthermore, current light sensor based systems implement an array of light sensors (e.g., 9 in [14] and 36 in [13]), which definitely incur much higher power consumption. In summary, our power consumption study demonstrated that the proposed gesture recognition system is lightweight and can be easily integrated to existing mobile devices.

7 Related work

7.1 Gesture Recognition

Gesture recognition is a well-explored area in the research community. A variety of modalities have been utilized for gesture recognition. By capturing images with cameras or depth sensors, vision-based approach utilizes image processing techniques to recognize gestures [7, 8]. Radio frequency (RF) signal can be used for gesture recognition, where the underlying principle is that hand movements interfere with the electromagnetic wave in the space and result in certain variations on the received signal. By analyzing the received RF signals, various information, like received signal strength (RSS) [4, 43], channel state information (CSI) [5], and Doppler shift [6], can be extracted to differentiate gestures. Replacing the RF transceiver with an acoustic (ultrasound) transceiver, the Doppler shift can be extracted from the reflected sound waves to identify gestures [9, 10]. Motion sensor-based gesture recognition leverages an inertial measurement unit (IMU), usually accelerometers and gyroscopes, to track human body/hand movement [11, 12]. Recently, employing light sensors for gesture recognition receives massive attention [13, 14, 15, 16, 17, 39]. When a gesture is performed over an array of light sensors (photodiodes), its shadow will be measured and used recognize the gesture.

Existing approaches can find application in specific scenarios where they can achieve excellent performance. However, each modality suffers from certain limitations. For instance, vision-based method has heavy computation overhead due to the involvement of image processing and it incurs privacy concerns as camera can capture other personal and sensitive data [5, 44]. In addition, all these modalities requires additional energy to power the operating sensor, shortening the operation time of the devices. In contrast, SolarGest recognize gestures with photocurrent signals from solar cells, which not only eliminates sensor-consumed energy but also supplies inexhaustible power to the host device. Interestingly, Li et al. [39] proposed to operate light sensors in the photovoltaic mode for gesture recognition. As a result, the sensor becomes passive without energy consumption and can harvest energy from the incident light as well. However, the transparent solar cells introduced in our work can be seamlessly integrated on top of the screens of electronic devices, resulting in a wider range of applications while without impacting device's appearance.

5. $P_{backscatter} = (100 * 8)/1000000 * 14.5\mu W = 0.023\mu W$.

7.2 Solar Cell based Sensing

In fact, the output photocurrent from a solar cell reflects the light intensity of the surrounding environment. Based on this phenomenon, researchers has explored the sensing potential of solar cells by regarding them as light indicators [45]. Randall et al. [46] equipped a solar panel on human shoulder to detect the received light strength (RLS). The RLS is then used to estimate the distance between celling light and user's shoulder. After generating a RLS map of the room, the instant RLS can be used for indoor positioning. Similarly, as every location on Earth has a unique solar signature, like a unique sunrise and sunset time, Chen et al. [47] proposed to create a sunlight map of the Earth, which can infer a location's longitude and latitude separately using the outputs of solar panels installed on the roof.

There is a prior work that leverages an opaque solar panel for gesture recognition [48]. Our work differs from it in three aspects. First, [48] recognizes three gestures only, Swipe, Two Taps, and Four Taps, which is actually one basic gesture but with different repetitions. However, SolarGest differentiates gestures based on the unique gesture patterns. Second, we prototype the novel transparent solar cell and demonstrated its feasibility for gesture recognition even with much lower energy harvesting capability. Third, with fundamentals in solar technology, we developed a theoretical model which allows us to study the impact of different practical factors in a solar based gesture recognition system.

8 Conclusion

We have proposed SolarGest, a solar-based gesture recognition system for ubiquitous solar-powered IoTs. Using solar energy harvesting fundamentals and geometric analysis, we derived a model that accurately simulates hand gestures and study the impact of various parameters. Employing real solar cells, both opaque and transparent, we have demonstrated that our system can detect six gestures with 96% accuracy using a single cell and detect fifteen gesture with 95% accuracy using a 2×2 cell array. Our power consumption measurement revealed that SolarGest consumes 44% less power compared to light sensor based approach.

Acknowledgements

This work was partially funded by the Australian Research Council Discovery Project DP210100904.

1536-1233 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3148143, IEEE Transactions on Mobile Computing

REFERENCES

[1] D. Ma, G. Lan, M. Hassan, W. Hu, M. B. Upama, A. Uddin, and M. Youssef, "SolarGest: Ubiquitous and battery-free gesture recognition using solar cells," in *Proceedings of the 25th Annual International Conference on Mobile Computing and Networking*, 2019, pp. 1–15.

[2] A. Chaudhary, J. L. Raheja, K. Das, and S. Raheja, "Intelligent approaches to interact with machines using hand gesture recognition in natural way: a survey," *arXiv preprint:1303.2292*, 2013.

[3] Z. Ren, J. Meng, J. Yuan, and Z. Zhang, "Robust hand gesture recognition with kinect sensor," in *Proceedings of the 19th ACM international conference on Multimedia*, 2011, pp. 759–760.

[4] H. Abdelnasser, M. Youssef, and K. A. Harras, "Wigest: A ubiquitous wifi-based gesture recognition system," in *Computer Communications (INFOCOM), 2015 IEEE Conference on*. IEEE, 2015, pp. 1472–1480.

[5] D. Sbîrlea, M. G. Burke, S. Guarnieri, M. Pistoia, and V. Sarkar, "Automatic detection of inter-application permission leaks in android applications," *IBM Journal of Research and Development*, vol. 57, no. 6, pp. 10–1, 2013.

[6] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *Proceedings of the 19th annual international conference on Mobile computing & networking*. ACM, 2013, pp. 27–38.

[7] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison *et al.*, "Kinectfusion: real-time 3d reconstruction and interaction using a moving depth camera," in *Proceedings of the 24th annual ACM symposium on User interface software and technology*. ACM, 2011, pp. 559–568.

[8] N. R. Howe, M. E. Leventon, and W. T. Freeman, "Bayesian reconstruction of 3d human motion from single-camera video," in *Advances in neural information processing systems*, 2000, pp. 820–826.

[9] S. Gupta, D. Morris, S. Patel, and D. Tan, "Soundwave: using the doppler effect to sense gestures," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 1911–1914.

[10] C. Pittman, P. Wisniewski, C. Brooks, and J. J. LaViola Jr, "Multiwave: Doppler effect based gesture recognition in multiple dimensions," in *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2016, pp. 1729–1736.

[11] J. Ruiz, Y. Li, and E. Lank, "User-defined motion gestures for mobile interaction," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2011, pp. 197–206.

[12] Z. Xu, K. Bai, and S. Zhu, "Taplogger: Inferring user inputs on smartphone touchscreens using on-board motion sensors," in *Proceedings of the fifth ACM conference on Security and Privacy in Wireless and Mobile Networks*. ACM, 2012, pp. 113–124.

[13] R. H. Venkatnarayan and M. Shahzad, "Gesture recognition using ambient light," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 1, p. 40, 2018.

[14] M. Kaholokula, "Reusing ambient light to recognize hand gestures," *Dartmouth college*, 2016.

[15] T. Li, X. Xiong, Y. Xie, G. Hito, X.-D. Yang, and X. Zhou, "Reconstructing hand poses using visible light," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 3, p. 71, 2017.

[16] T. Li, C. An, Z. Tian, A. T. Campbell, and X. Zhou, "Human sensing using visible light communication," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. ACM, 2015, pp. 331–344.

[17] T. Li, Q. Liu, and X. Zhou, "Practical human sensing in the light," in *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 2016, pp. 71–84.

[18] X. Wang, L. Zhi, and K. Müllen, "Transparent, conductive graphene electrodes for dye-sensitized solar cells," *Nano letters*, vol. 8, no. 1, pp. 323–327, 2008.

[19] C. J. Traverse, R. Pandey, M. C. Barr, and R. R. Lunt, "Emergence of highly transparent photovoltaics for distributed applications," *Nature Energy*, vol. 2, no. 11, p. 849, 2017.

[20] B. Parida, S. Iniyan, and R. Goic, "A review of solar photovoltaic technologies," *Renewable and sustainable energy reviews*, vol. 15, no. 3, pp. 1625–1636, 2011.

[21] G. P. Smestad, F. C. Krebs, C. M. Lampert, C. G. Granqvist, K. Chopra, X. Mathew, and H. Takakura, "Reporting solar cell efficiencies in solar energy materials and solar cells," 2008.

[22] C. Riordan and R. Hulstron, "What is an air mass 1.5 spectrum?(solar cell performance calculations)," in *Photovoltaic Specialists Conference, 1990., Conference Record of the Twenty First IEEE*. IEEE, 1990, pp. 1085–1088.

[23] M. Wright and A. Uddin, "Organic inorganic hybrid solar cells: A comparative review," *Solar energy materials and solar cells*, vol. 107, pp. 87–111, 2012.

[24] X. Cai, S. Zeng, X. Li, J. Zhang, S. Lin, A. Lin, and B. Zhang, "Effect of light intensity and temperature on the performance of gan-based pin solar cells," in *Electrical and Control Engineering (ICECE), 2011 International Conference on*. IEEE, 2011, pp. 1535–1537.

[25] M. H. Weik, "Lambert's cosine law," in *Computer Science and Communications Dictionary*. Springer, 2000, pp. 868–868.

[26] D. Blaauw, D. Sylvester, P. Dutta, Y. Lee, I. Lee, S. Bang, Y. Kim, G. Kim, P. Pannuto, Y. Kuo *et al.*, "Iot design space challenges: Circuits and systems," in *Symp. VLSI Circuits Dig. Tech. Papers*, 2014, pp. 1–2.

[27] "Lunar watch," https://lunar-smartwatch.com/.

[28] M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells, "Noise reduction using an undecimated discrete wavelet transform," *IEEE Signal Processing Letters*, vol. 3, no. 1, pp. 10–12, 1996.

[29] D. Gradolewski and G. Redlarski, "The use of wavelet analysis to denoising of electrocardiography signal," in *XV International PhD Workshop OWD*, vol. 19, 2013.

[30] A. Virmani and M. Shahzad, "Position and orientation agnostic gesture recognition using wifi," in *Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 2017, pp. 252–264.

1536-1233 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3148143, IEEE Transactions on Mobile Computing

[31] C. Cheadle, M. P. Vawter, W. J. Freed, and K. G. Becker, "Analysis of microarray data using z score transformation," *The Journal of molecular diagnostics*, vol. 5, no. 2, pp. 73–81, 2003.

[32] C. Myers, L. Rabiner, and A. Rosenberg, "Performance tradeoffs in dynamic time warping algorithms for isolated word recognition," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 28, no. 6, pp. 623–635, 1980.

[33] S. Sempena, N. U. Maulidevi, and P. R. Aryan, "Human action recognition using dynamic time warping," in *Electrical Engineering and Informatics (ICEEI), 2011 International Conference on*. IEEE, 2011, pp. 1–5.

[34] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj, and D. J. Inman, "Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks," *Journal of Sound and Vibration*, vol. 388, pp. 154–170, 2017.

[35] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[36] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of wifi signal based human activity recognition," in *Proceedings of the 21st annual international conference on mobile computing and networking*. ACM, 2015, pp. 65–76.

[37] M. B. Upama, M. Wright, N. K. Elumalai, M. A. Mahmud, D. Wang, C. Xu, and A. Uddin, "High-efficiency semitransparent organic solar cells with non-fullerene acceptor for window application," *ACS Photonics*, vol. 4, no. 9, pp. 2327–2334, 2017.

[38] C. An, T. Li, Z. Tian, A. T. Campbell, and X. Zhou, "Visible light knows who you are," in *Proceedings of the 2nd International Workshop on Visible Light Communications Systems*. ACM, 2015, pp. 39–44.

[39] Y. Li, T. Li, R. A. Patel, X.-D. Yang, and X. Zhou, "Self-powered gesture recognition with ambient light," in *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, 2018, pp. 595–608.

[40] "Opt101 datesheet," http://www.ti.com/lit/ds/symlink/opt101.pdf.

[41] Z. Chen, Z. Cheng, J. Wang, X. Wan, C. Shu, H. K. Tsang, H. P. Ho, and J.-B. Xu, "High responsivity, broadband, and fast graphene/silicon photodetector in photoconductor mode," *Advanced Optical Materials*, vol. 3, no. 9, pp. 1207–1214, 2015.

[42] B. Kellogg, V. Talla, S. Gollakota, and J. R. Smith, "Passive wi-fi: Bringing low power to wi-fi transmissions." in *NSDI*, vol. 16, 2016, pp. 151–164.

[43] H. Abdelnasser, K. A. Harras, and M. Youssef, "A ubiquitous wifi-based fine-grained gesture recognition system," *IEEE Transactions on Mobile Computing*, 2018.

[44] Z. Weinberg, E. Y. Chen, P. R. Jayaraman, and C. Jackson, "I still know what you visited last summer: Leaking browsing history via user interaction and side channel attacks," in *Security and Privacy (SP), 2011 IEEE Symposium on*. IEEE, 2011, pp. 147–161.

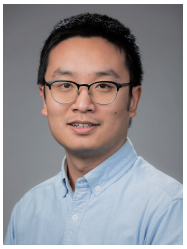
[45] D. Ma, G. Lan, M. Hassan, W. Hu, and S. K. Das, "Sensing, computing, and communications for energy harvesting iots: A survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 1222–1250, 2019.

[46] J. Randall, O. Amft, J. Bohn, and M. Burri, "Luxtrace: indoor positioning using building illumination," *Personal and ubiquitous computing*, vol. 11, no. 6, pp. 417–428, 2007.

[47] D. Chen, S. Iyengar, D. Irwin, and P. Shenoy, "Sunspot: Exposing the location of anonymous solar-powered homes," in *Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments*. ACM, 2016, pp. 85–94.

[48] A. Varshney, A. Soleiman, L. Mottola, and T. Voigt, "Battery-free visible light sensing," in *Proceedings of the 4th ACM Workshop on Visible Light Communication Systems*. ACM, 2017, pp. 3–8.

Dong Ma is currently an Assistant Professor at the School of Computing and Information Systems, Singapore Management Univeristy. Prior to that, he was a Postdoctoral Research Associate with the Department of Computer Science and Technology, the University of Cambridge, Cambridge, UK. He received his Ph.D. degree from the School of Computer Science and Engineering, the University of New South Wales, Sydney, Australia, in 2020. He obtained the B.S. degree in communication engineering from Central South University, China, in 2014. His research interests include Internet of Things, Earable Computing, Mobile Health, and Energy Harvesting. He is a member of the IEEE.



Guohao Lan is an Assistant Professor in the Department of Software Technology, Delft University of Technology, Delft, the Netherlands. In prior, he was a Postdoctoral Research Associate with the Department of Electrical and Computer Engineering, Duke University, USA. He obtained his Ph.D. degree in Computer Science and Engineering from the University of New South Wales, Sydney, Australia. His research interests include wireless sensor networks, wearable computing, and mobile sensing systems. He is a recipient of a Facebook Research Award and the 2020 ACM/IEEE IPSN Best Research Artifacts Award.

Changshuo Hu received the B.S. degrees from Capital Normal University, Beijing, China, in 2016 and the M.S. degrees in School of Computer Science and Engineering, University of New South Wales, Sydney, Australia, in 2020. He is currently preparing for his Ph.D in University of New South Wales. He has authored a workshop paper and a conference poster. His research focuses on ubiquitous computing, pervasive sensing and energy harvesting.

Mahbub Hassan is a Full Professor in the School of Computer Science and Engineering, the University of New South Wales, Sydney, Australia. He is a Senior Member of IEEE and served as a Distinguished Lecturer of IEEE (COMSOC) for 2013-2016. He is currently an Editor of IEEE Communications Surveys and Tutorial and has previously served as Guest Editor for IEEE Network, IEEE Communications Magazine, IEEE Transactions on Multimedia, and Area Editor for Computer Communications. Professor Hassan has earned a PhD from Monash University, Australia, and an MSc from University of Victoria, Canada, both in Computer Science. More information about Professor Hassan is available from http://www.cse.unsw.edu.au/~mahbub.

1536-1233 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TMC.2022.3148143, IEEE Transactions on Mobile Computing

Wen Hu is a professor at School of Computer Science and Engineering, the University of New South Wales (UNSW). Much of his research career has focused on novel applications, low-power communications, security and compressive sensing in sensor network systems and Internet of Things (IoT). Hu published regularly in the top rated sensor network and mobile computing venues such as ACM/IEEE IPSN, ACM SenSys, ACM MobiCOM, ACM UbiCOMP, IEEE PerCOM, ACM transactions on Sensor Networks (TOSN), IEEE Transactions on Mobile Computing (TMC), and Proceedings of the IEEE. Hu was a principal research scientist and research project leader at CSIRO Digital Productivity Flagship, and received his Ph.D from the UNSW. He is a recipient of prestigious CSIRO Office of Chief Executive (OCE) Julius Career Award (2012 - 2015) and multiple research grants from Australian Research Council, CSIRO and industries. Hu is a senior member of ACM and IEEE, and is an associate editor of ACM TOSN and the general chair of CPS-IoT Week 2020. Hu actively commerlizes his research results in smart buildings and IoT, and his endeavors include working as the Chief Technical Officer (part time) and the Chief Scientist (part time) in Parking Spotz (from 2021) and WBS Tech (2016-2020) respectively.

Mushfika B. Upama received her PhD degree in advanced photovoltaics from UNSW, Sydney in 2019. During PhD, she worked on the development of highly efficient and semitransparent organic and perovskite solar cells. She has co-authored 35 journal articles and 10 conference papers in the area of photovoltaics. Currently, she is working as an electrical engineer at Arup, Sydney.

Ashraf Uddin is a Professor working on organic and perovskite solar cells in the School of Photovoltaic and Renewable Energy Engineering, University of New South Wales (UNSW). He has published over 200 journal papers and has 20 patents on semiconductor devices. He obtained his PhD degree in March 1991 in Semiconductor Physics from Osaka University, Osaka, Japan. After completing his PhD, he joined the RD centre of Toshiba Corporation, Japan and worked on opto-electronic devices and on poly-Si thin film transistors to develop a process technology on glass substrates for the fabrication of flat panel displays (LCD type). After that he worked at the School of Materials Science and Engineering, Nanyang Technological University, Singapore as an academic for seven years.

Moustafa Youssef is a professor at Alexandria University and the American University in Cairo, Egypt. He is the founder and director of the Wireless Research Center of Excellence, Egypt. His research interests include mobile wireless networks, mobile computing, location determination technologies, and pervasive computing. He is an associate editor for IEEE TMC and ACM TSAS, served as the lead guest editor of the IEEE Computer Special Issue on Transformative Technologies and an area editor of ACM MC2R as well as on the organizing and technical committees of numerous prestigious conferences. He is the recipient of the 2003 University of Maryland Invention of the Year award, the 2010 TWAS-AAS-Microsoft Award for Young Scientists, the 2013 and 2014 COMESA Innovation Award, the 2013 ACM SIGSpatial GIS Conference Best Paper Award, the 2017 Egyptian State Award, multiple Google Research Awards, among many others. He is an IEEE and ACM Fellow.

1536-1233 (c) 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.