WhereWear: Calibration-free Wearable Device Identification through Ambient Sensing

Carlos Ruiz

Carnegie Mellon University
Electrical & Computer Eng.
Moffett Field, California
carlosrd@cmu.edu

Shijia Pan

Carnegie Mellon University
Electrical & Computer Eng.
Pittsburgh, Pennsylvania
shijiapan@cmu.edu

Hae Young Noh

Carnegie Mellon University
Civil & Environmental Eng.
Pittsburgh, Pennsylvania
noh@cmu.edu

Pei Zhang

Carnegie Mellon University
Electrical & Computer Eng.
Moffett Field, California
peizhang@cmu.edu

ABSTRACT

With the growth of wearable devices, numerous health and smart building applications are enabled. As a result, many people wear multiple devices for different applications, such as fitness tracking. Being able to match devices' physical identity (e.g., smartwatch on Bob's left forearm) to their virtual identity (e.g., IP 192.168.0.22) is often a requirement for these applications, especially when different people could use the same device –e.g., home, gym equipment– or be worn in different places. Context-aware sensing, utilizing the insight that co-located sensors detect the same events, has been used to establish such association ubiquitously. However, challenges arise with the growing number of on-body devices, and existing approaches fail to distinguish where the devices are on the body.

In this paper, we present *WhereWear*, a human pose estimation based calibration-free wearable device identification mechanism through ambient sensing of vision (camera in the environment) and motion (IMU in wearable devices). Our system utilizes the key fact that the orientation change of the devices is related to the orientation change of the body part where the device is on, and matches each device's physical and virtual identities at the *bodypart* –any part of the body between two joints, e.g. forearm, thigh, etc.– resolution. We evaluate the proposed mechanism on a publicly available dataset, where multiple human subjects perform activities with 13 devices on their body. Our system demonstrates up to 64% device identification accuracy on average and up to 2.7× improvement over existing baselines with five simultaneous users. With only one user, it achieves up to 92.2% accuracy.

CCS CONCEPTS

• **Computer systems organization** → **Embedded and cyber-physical systems**; • **Networks** → *Cyber-physical networks*.

KEYWORDS

Wearables; ID; Multi-Modal; Heterogeneous Sensing; Camera; IMU; Inertial; Motion Matching; Pairing; On-body Device Localization; Human Pose

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

WearSys'19, June 21, 2019, Seoul, Republic of Korea

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6775-2/19/06...\$15.00

<https://doi.org/10.1145/3325424.3329667>

ACM Reference Format:

Carlos Ruiz, Shijia Pan, Hae Young Noh, and Pei Zhang. 2019. *WhereWear: Calibration-free Wearable Device Identification through Ambient Sensing*. In *The 5th ACM Workshop on Wearable Systems and Applications (WearSys'19)*, June 21, 2019, Seoul, Republic of Korea. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3325424.3329667>

1 INTRODUCTION

Wearable devices have become ubiquitous in people's lives, where they provide continuous information about the user's physical condition and activities. They enable various applications, such as e.g., fitness tracking [3], heart rate monitoring [3, 13], muscular fatigue monitoring [15, 16], sweat sensing[6]. However, with the growing number of wearable devices, it becomes cumbersome to manually calibrate or match the device's physical identity and location (e.g., accelerometer on your right thigh) to its virtual identity (e.g., Sensor 7, with MAC address AA:BB:CC:DD). Traditionally, this association was performed by the user by, for instance, manually pressing a button on each sensor worn or pairing them through a 3rd-party app. When there are multiple people sharing devices (a home or a gym), performing this procedure before each use can limit the adoption rate of such devices and applications. Furthermore, a successful device identification can enable the fusion of multiple sources of information about the user's activity or performance. For example, heart-rate sensor data could be enhanced by sports analytics –like stride length, ground contact time – coming from a gym camera if the camera can figure out how to communicate (virtual identity) with the device or the user it sees. Therefore, associating the physical and virtual identities of these devices through different systems quickly and automatically is crucial for enabling and exploring more personalized services with shared smart infrastructures. Prior works on calibration-free or interaction-free matching schemes often rely on the shared context sensed by wearable devices on the same body [9, 17, 27]. However, due to the noisy sensing data and the dynamic sensing environment, it is difficult if not impossible to achieve sub-body level –*bodypart*¹ level– identification resolution.

We present *WhereWear*, a calibration-free wearable device identification scheme through ambient sensing of bodypart motions. In particular, a camera in the environment captures the pose of every person (within the field of view) and analyzes the motion of each bodypart by estimating its 3D orientation. Simultaneously, the wearable devices' inertial sensors track the changes in their orientation over time. Then, *WhereWear* estimates the alignment between the camera's and each device's coordinate frames and compares

¹We employ the term *bodypart* to refer to any part of the human body between two "consecutive" joints, such as the forearm, torso, thigh, etc.

Figure 1: *WhereWear* system overview.

the orientation signals of each bodypart-device combination, associating the ones with highest similarity. Finally, the contributions of this work can be summarized as follows:

- A system for wearable device identification with bodypart level resolution using on-device inertial sensors and a camera. The identification is possible even under occlusion (e.g., device covered by clothing) or when several users simultaneously wear multiple devices.
- A novel algorithm to estimate the likelihood that the motion sensed by a 3D inertial sensor corresponds to the motion of a bodypart (e.g., person 1's left forearm) observed by a 2D (traditional RGB) camera.
- Evaluation in a public dataset containing five human subjects wearing inertial sensors on 13 different parts of their bodies while they perform different types of activities.

The rest of the paper is organized as follows. We first introduce our system and detail the functionality of each component in Section 2. Then, we evaluate *WhereWear* in a publicly available dataset and compare our results to existing device identification baselines in Section 3. Next, we discuss the prior works in this area and how our work distinguishes from them in Section 4. Finally, we conclude the paper in Section 5.

2 WHEREWEAR SYSTEM OVERVIEW

Our system, *WhereWear*, targets the wearable device identification problem, i.e., associating devices' physical ID –where people wear them– to their virtual ID, such as their MAC address. *WhereWear*

can work with any traditional 2D camera to capture the physical motions of a user, which it then matches to the motion detected by the 3D IMU (Inertial Measurement Unit: 3-axis accelerometer, gyroscope and magnetometer) on wearable devices. We specifically address the following challenges in this work:

- ✓ combining the different sensing modalities of 2D image and 3D orientation sensing,
- ✓ multiple devices worn by the same/different user, and
- ✓ non-line-of-sight to the device(s).

To achieve that, *WhereWear* relies on the insight that wearables must be located on some part of the body. Therefore, even without directly seeing a device, its motion can be estimated by analyzing the motion of every part of the body of every person detected in an image. Figure 1 reflects the architecture of our system. The camera images are first processed through a Human Pose Estimation module (Sec. 2.1). Then, for each pair of connected joints (e.g., wrist and elbow determine the forearm bodypart), the Vision-based Bodypart Orientation Estimation module (Sec. 2.2) computes the plane of possible 3D orientations that correspond to the observed 2D bodypart coordinates. Section 2.3 explains how to obtain 3D orientation measurements from inertial sensor readings. Next, the Bodypart-IMU Alignment module (Sec. 2.4) details our optimization formulation to find the rotation offset between the two joints that determine a bodypart and the IMU reference axes, and then between those two and the camera coordinate frame. Finally, the most likely location of each wearable is output by the Device Identification via Hungarian Assignment module (Sec. 2.5).

2.1 Human Pose Estimation

Leveraging human pose information for device identification serves two purposes. On the one hand, human visual appearance and poses have a well-studied distribution, which leads to rather highly accurate models, as opposed to visually identifying each possible wearable device in an image. On the other hand, it allows our system to eliminate the line-of-sight constraint to the device: as long as the user can be seen, a device covered by clothing (e.g., under a shirt) may still be localized and identified. As previously discussed, *WhereWear* relies on the intuition that the motion of any device worn by a person should be closely related to the motion of the bodypart where it is placed (e.g., chest for a heart-rate monitor). In order to analyze that motion, the Human Pose Estimation module is responsible for determining the position of each joint for all people in the image (see Figure 3b for an example output). With recent advancements in this Computer Vision task, these can be robustly estimated in real-time. Furthermore, *WhereWear*'s algorithm does not rely on any particular implementation of this module, hence any improvement in the state-of-the-art of Human Pose Estimation would contribute to even higher device identification accuracy.

2.2 Vision-based Bodypart Orientation Estimation

Once the joints of every person in the image have been localized, Figure 2 shows the concept of our orientation-based device identification scheme. With a single (monocular RGB) camera, projected 2D joint locations ($\mathbf{P}_A, \mathbf{P}_B$) cannot uniquely determine the actual 3D positions ($\mathbf{J}_A, \mathbf{J}_B$) since depth information is missing –i.e., $\mathbf{J}_A$

Figure 2: 3D human pose (joints J_A and J_B , e.g. elbow and wrist) is mapped to the 2D camera (P_A , P_B). These projected points allow us to compute the visual orientation plane (red), determined by its normal vector $\mathbf{n}_{AB}$. The wearable device (orange circle) is carried somewhere along the limb $\overline{J_A J_B}$. We can find a vector $\mathbf{f}$ (which rotates with the device's orientation) such that in the absence of noise $\mathbf{f}$ always lies on the visual orientation plane (or at least $\mathbf{f} \cdot \mathbf{n}_{AB}$ is minimized).

could be anywhere along the line l_A . However, they do define a visual *orientation* plane (shown in red in Figure 2) that specifies valid orientations for the device. The intuition behind *WhereWear* is that, over time, the wearables' IMU orientations should match those of the bodyparts on which they are worn, such as the forearm for a watch. This plane is defined by the lines l_A and l_B , therefore:

$$\mathbf{n}_{AB} = \mathbf{j}_A \times \mathbf{j}_B \quad (1)$$

Figure 2 demonstrates the image plane in the pinhole camera model. The image plane is perpendicular to the camera's viewing direction $\mathbf{z}^C$ axis. It is also situated at a distance of focal length f_L in front. The image plane axes $\mathbf{u}$ and $\mathbf{v}$ are parallel to $\mathbf{x}^C$ and $\mathbf{y}^C$ respectively, and offset by c_u and c_v (image center). We define the pixel coordinates of joint J_A projected on the camera as $\mathbf{P}_A = (u_A, v_A)$. Then we can denote the vector $\mathbf{j}_A$ in camera coordinates as: $\mathbf{j}_A = (u_A - c_u, v_A - c_v, f_L)$ (same for $\mathbf{j}_B$). With these denotation, we rewrite Eq. 1 as:

$$\mathbf{n}_{AB}^C(t) = \begin{pmatrix} f_L(v_A^t - v_B^t), \\ f_L(u_B^t - u_A^t), \\ (u_A^t - c_u)(v_B^t - c_v) - (u_B^t - c_u)(v_A^t - c_v) \end{pmatrix} \quad (2)$$

2.3 Device Orientation Acquisition

Estimating device orientation by fusing accelerometer, gyroscope and magnetometer readings has been long studied [4, 11]. The orientation is typically output as a rotation with respect to a fixed frame of reference, which most of the times is determined based on the gravitational and magnetic fields, such as the North-East-Down convention. However, this may not be case since the magnetic field tends to be unreliable in indoor environments. Therefore, *WhereWear* still has to align each device's orientation to the bodypart orientations obtained through Eq. 2.

2.4 Bodypart-IMU Alignment

Since the IMU's and camera's axes are not aligned, this module solves an optimization formulation to find the best alignment. Figure 2 shows the IMU's local frame (IL superscript), which rotates with the wearable device (in orange). The IMU's fixed reference, denoted IMU world frame (IW superscript), is used for reporting

orientation as a 3x3 rotation matrix $R^{IL-IW}(t)$. However, not only is the time-invariant rotation from the IMU's world frame to the camera axes (R^{IW-C}) unknown; *WhereWear* also needs to find the vector $\mathbf{f}$ that determines the relationship or offset between the IMU's axes and the bodypart vector $\overline{J_A J_B}$. Once everything is aligned, $\mathbf{f}$ should always lie on the visual orientation plane. Therefore, we formulate the bodypart-IMU alignment as the following optimization:

$$\begin{aligned} \arg \min_{R^{IW-C}, \mathbf{f}^{IL}} \sum_t & \left| \mathbf{f}^C(t)^\top \cdot \mathbf{n}_{AB}^C(t) \right| = \\ & = \sum_t \left| \left(R^{IW-C} R^{IL-IW}(t) \mathbf{f}^{IL} \right)^\top \cdot \mathbf{n}_{AB}^C(t) \right| \\ \text{subj. to } & (R^{IW-C})^T = (R^{IW-C})^{-1} \\ & \det(R^{IW-C}) = 1 \\ & \|\mathbf{f}^{IL}\| = 1 \end{aligned} \quad (3)$$

where $\mathbf{a}^\top \cdot \mathbf{b}$ denotes the dot product between vectors $\mathbf{a}$ and $\mathbf{b}$. We take the absolute value inside the summation because we are only interested in the magnitude of the angle between $\mathbf{f}$ and $\mathbf{n}_{AB}^C$.

2.5 Device Identification via Hungarian Assignment

Finally, once we find the optimal coordinate alignment R^{IW-C} for each bodypart-IMU pair, we define a score to measure how well both orientations match. The intuition is that, if we are observing the same bodypart where the IMU is worn, both orientations should follow each other tightly. On the contrary, IMU orientation should be uncorrelated to all bodyparts other than the one where it's carried, so they should differ over time. We chose the mean dot product and call this metric the *pairwise matching score*, which is therefore defined for each bodypart-IMU combination as:

$$\text{score} = \frac{1}{t} \sum_t \left| \mathbf{f}^C(t)^\top \cdot \mathbf{n}_{AB}^C(t) \right| \quad (4)$$

Last, *WhereWear* computes an $N_{\text{IMU}} \times N_{\text{bodyparts}}$ score matrix S based on Eq. 4 and uses the Hungarian method for the linear assignment problem [8] to produce the IMU to bodypart mapping that minimizes the total sum of assigned scores in polynomial time, that is, it outputs the most likely location for each wearable device.

3 EVALUATION

WhereWear associates the wearable ID to the most-likely bodypart they are on. To understand the factors that affect the system performance, we conducted experiments on a publicly available dataset, where a camera and on-body IMUs capture body motions, and evaluate the wearable device identity accuracy.

In this section, we first describe the public dataset used in this work (Section 3.1). Next, we introduce the experimental implementation and the baseline algorithm (Section 3.2). Finally, we compare the results from both *WhereWear* and the baseline (Sections 3.3).

3.1 Dataset Description

We re-used a public dataset [24] to evaluate our system, where cameras were installed to capture occupants moving within a 6×4 m area in a room. Each human subject had 13 IMUs placed on the body

Figure 3: Public dataset experimental settings. (a) In total 13 IMU devices are placed on the test subject. (b) An example of movement and joints detected by human pose estimation.

as pointed out in Figure 3a. The dataset contains data from five human subjects. Figure 3b shows another example of the frames in the dataset, the type of movements conducted by the subjects, and the pose detected by the Human Pose Estimation module.

Moreover, the dataset only provides single-person data. In order to evaluate the case where there are N people in the same scene, we extract N different segments from the whole dataset, and feed all of the detected joint locations as well as all $\#Devices \text{ per person} \times N$ IMUs into our algorithm, as if all N people were simultaneously present in the room.

3.2 System Implementation

We compared *WhereWear* to our prior work on IoT device pairing via heterogeneous sensing introduced in *UniverSense* [18] (adapted for the 3D scenario). *UniverSense* compares the acceleration of the device sensed by its IMU and the displacement of the device sensed by the camera (double-differentiating it to obtain acceleration) and computes the IoT device-camera matching score from these (Section 3.2.1). Then, Section 3.2.2 discusses the details of our *WhereWear* implementation.

3.2.1 Baseline: 3D Acceleration Matching. We established our baseline based on *UniverSense*, which constrained user motion to a 2D plane parallel to the camera at a fixed depth, and implemented a 3D version. We use the state-of-the-art 2D pose estimation framework *OpenPose* [2, 26] followed by a 2D-to-3D human pose estimation model [12]. We define a hyperparameter $\alpha_j \in [0, 1]$ to specify the relative placement of the IMU along a given bodypart by interpolating between two joints. We differentiate twice using a Savitzky-Golay filter [18, 23], converting displacements to accelerations.

3.2.2 WhereWear. Our *WhereWear* implementation is based on the system design in Section 2. For video processing we use Python with OpenCV and OpenPose [2, 12, 26] for 2D human pose estimation. Since this model runs frame-by-frame, we added a simple joint-based tracking layer on top to correctly follow the same person's bodyparts over time. Finally we implement our bodypart-IMU alignment algorithm described in Section 2.4 in Matlab, using its Optimization Toolbox to solve Equation 3. Unless otherwise noted, we use a window length of $T_w = 10s$ and only consider the IMU and visual data within that segment for device identification.

Figure 4: Average device identification accuracy for 13 body-parts when there are one, three or five people in the scene.

3.3 Results and Analysis

We analyze our system with different parameters to demonstrate its performance under different movements, number of people, and bodyparts. We compare the average device identification accuracy for each case investigated.

3.3.1 Bodyparts. The placement of the sensor affects the system performance significantly. Our system holds the assumption that the target wearable devices are on rigid bodyparts (otherwise the orientation of the device might differ from the joint-joint detected orientation), therefore we evaluate on all 13 IMU locations in the dataset to verify our system assumption. Figure 4 demonstrates the average device identification accuracy when there are one, three, and five people detected by camera simultaneously. The x-axis is the bodypart where the IMU is located, and the y-axis is the accuracy, averaged out over 50 trials of $T_w = 10s$ each. Figure 4a shows the

Figure 5: Wearable device identification accuracy as a function of the number of people simultaneously present in a scene. Note *WhereWear*’s robustness to increasing people density compared to the baseline.

results when there is only one person present in the scene. The average accuracy of 13 devices using the baseline method is 24%. The lower arms performed the best, with accuracies of 40% and 52%. *WhereWear* achieved an average of 47%. The sensors that are placed on a rigid bodypart, i.e., eight sensors on arms and legs, achieved the highest identification accuracy, and their accuracy reaches 61% on average with only 10 seconds of data.

We further evaluate the cases where there are three and five people detected in the same scene in Figures 4b and 4c. The average device identification accuracy for baseline methods are 12.4% and 9% respectively (random guess is 7.7%). And our system achieved 40% and 37% identification accuracy, which is 3 to 4× improvement compared to the baseline. The eight sensors on rigid bodyparts achieved an average accuracy of 57.5% and 54.5% respectively, which is comparable to the case where there is only one person in the scene. For the rest of the evaluation, we are going to focus on these eight sensors whose placement fits our system’s assumption.

3.3.2 Number of People. Number of people is another important aspect that affects our system performance. When there are more people in the scene, more bodyparts that may have similar motion may get confused with each other. From the results in Section 3.3.1, we observe that the accuracies of both the baseline and our methods decreased when more than one person are in the scene. We further evaluate this factor more systematically with the eight sensors placed on the rigid bodyparts. Figure 5 shows the average accuracy when from one to five persons appear in the scene at the same time. The x-axis is the number of people. We observe that the baseline method shows a consistent decreasing trend (42%, 25%, 22%, 16%, and 15%, overall 24%). Our method, on the other hand, achieves 70% accuracy for the one person case, and achieved 61%, 64%, 62%, and 61% for the two to five person cases, which demonstrates the robustness against higher people density compared to the baseline with an overall 64% accuracy, a 2.7× improvement.

3.3.3 Window Length and Activity Types. The longer the tracking window, the more distinguished or unique a device’s movement is—in that window—compared to other devices. Therefore, we further evaluate our system using different window lengths. We selected a shorter (5 seconds) and a longer (20 seconds) window length

Figure 6: Wearable device identification accuracy using different window lengths (5s, 10s, 20s). As more data is observed, the system is able to make better predictions. Also, activities with more unique bodypart motions yield higher accuracy.

compared to the default $T_w = 10s$ used in the rest of the evaluation. Figure 6 shows the average accuracy when different activities (labeled from the public dataset) are conducted. We observe that for the activity of acting, where the motion of the body is richer than in the rest of the activities, the accuracy for all three window lengths is higher than other activities, which are 64%, 80%, and 92% respectively. RoM shows the lowest accuracy (28%, 39%, and 59%) due to the lack of leg motion involved. Even with that, as the window length increases, the accuracy increases as well thanks to the accumulated uniqueness from small movements.

4 RELATED WORK

We focused on two types of related work here: 1) pairing systems that utilize shared sensing information; and 2) vision-based device/object identification.

4.1 IoT Pairing through Sensing

We summarize the sensing-based pairing schemes in two categories: active and passive. Yang et al. proposed to pair devices by sensing unique muscle EMG signals [28]. Our prior works pair devices by detecting IoT device motion through their IMU and a camera [18, 21, 22]. These are active pairing schemes that are fast and straightforward, even without traditional input such as keyboard or QR codes. On the other hand, these proactive pairing methods still require human involvement, which can be cumbersome. Passive pairing schemes often rely on contextual sensing to achieve pairing. Lester et al. are pioneers in identifying two co-located devices on the same body [9]. Miettinen et al. achieve secure pairing through shared context information [14]. Han et al. then further extend this concept to heterogeneous sensing [5].

WhereWear achieves device association in a semi-proactive way—it is also based on the intuition that the shared sensing events between an ambient camera and on-body IMU(s) can be used to identify the physical location of each device. Unlike prior works, *WhereWear* does not require prior knowledge of the relative position between wearable devices and the camera to achieve fast identification, nor does it make any assumptions about their motion.

4.2 Vision-based Device/Object Identification

Vision-based device or object identification methods mainly fall into two categories: object recognition and device motion matching.

Object/marker recognition. If an object/device has a unique appearance, vision-based object recognition methods can distinguish the object/device in the image frame [1, 7, 10, 19]. For example, Lowe et al. introduced scale-invariant features for object recognition [10]. Belongie et al. presented their approach using shape contexts [1]. These object/marker recognition based methods are active, meaning that the user would have to manually scan every single wearable to associate their physical and virtual IDs. Therefore, they do not meet our application requirements.

Object/device motion matching. Prior works on human motion matching through different sensing modalities have been explored for various applications. Trumble et al. combine multiple cameras and inertial sensors to improve the human pose tracking accuracy [24]. However, requiring multiple cameras may not be a realistic assumption for one sensing area in everyday life. Marcard et al., on the other hand, achieved 3D human pose estimation with IMUs and cameras on mobile devices [25]. However, their goal is to obtain accurate human information such as posture and does not address the challenges of identity association. Prior works have also studied motion matching for device identity association. Ruiz et al. utilized the shared motion information of an object detected by different sensors to identify the device, with the assumption of accurate visual object recognition [20]. However, they focused on the actuation of the device to increase the uniqueness of each devices' motion. Nguyen et al. leverage the shared motion detected by the camera and the fluctuation reflected in the WiFi signal to associate the device to the user holding it [17]. However, due to the detection resolution of the WiFi signal, the system cannot achieve body-part resolution in terms of device identity association.

5 CONCLUSION

In this paper we presented *WhereWear*, a wearable device identity association scheme using multi-modal sensing. Our system leverages the shared bodypart orientation detected by wearable devices as well as a camera in the ambient environment to achieve the association. *WhereWear* shows robustness towards different number of users in the scene. Compared to the state-of-the-art baselines, our system achieved up to 64% device identification accuracy on average, which is a 2.7× improvement with five users in the scene, evaluated on a large-scale public dataset. When there is only one user, *WhereWear* achieves up to 92% device identification accuracy.

REFERENCES

- [1] Serge Belongie, Jitendra Malik, and Jan Puzicha. 2002. *Shape matching and object recognition using shape contexts*. Technical Report. California University San Diego La Jolla, Department of Computer Science and Engineering.
- [2] Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2017. Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. In *CVPR*. IEEE, 7291–7299.
- [3] Toon De Pessemier and Luc Martens. 2018. Heart rate monitoring, activity recognition, and recommendation for e-coaching. *Multimedia Tools and Applications* 77, 18 (2018), 23317–23334.
- [4] Bingfei Fan, Qingguo Li, and Tao Liu. 2018. Improving the accuracy of wearable sensor orientation using a two-step complementary filter with state machine-based adaptive strategy. *Measurement Science and Technology* 29, 11 (2018).
- [5] Jun Han, Albert Jin Chung, Manal Kumar Sinha, Madhumitha Harishankar, Shijia Pan, Hae Young Noh, Pei Zhang, and Patrick Tague. 2018. Do you feel what I hear? Enabling autonomous IoT device pairing using different sensor types. In *2018 IEEE Symposium on Security and Privacy (SP)*. IEEE, 0.
- [6] Ji Jia, Chengtian Xu, Shijia Pan, Stephen Xia, Peter Wei, Hae Noh, Pei Zhang, and Xiaofan Jiang. 2018. Conductive Thread-Based Textile Sensor for Continuous Perspiration Level Monitoring. *Sensors* 18, 11 (2018), 3775.
- [7] Andrew E Johnson and Martial Hebert. 1999. Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 5 (1999), 433–449.
- [8] Harold W Kuhn. 1955. The Hungarian method for the assignment problem. *Naval research logistics quarterly* 2, 1-2 (1955), 83–97.
- [9] Jonathan Lester, Blake Hannaford, and Gaetano Borriello. 2004. Are you with me?—using accelerometers to determine if two devices are carried by the same person. In *International Conference on Pervasive Computing*. Springer, 33–50.
- [10] David G Lowe. 1999. Object recognition from local scale-invariant features. In *Computer vision, 1999. The proceedings of the seventh IEEE international conference on*, Vol. 2. IEEE, 1150–1157.
- [11] Sebastian OH Madgwick, Andrew JL Harrison, and Ravi Vaidyanathan. 2011. Estimation of IMU and MARG orientation using a gradient descent algorithm. In *Rehabilitation Robotics (ICORR), 2011 IEEE International Conference on*. IEEE, 1–7.
- [12] Julieta Martinez, Rayat Hossain, Javier Romero, and James J. Little. 2017. A simple yet effective baseline for 3d human pose estimation. In *ICCV*. IEEE, 2640–2649.
- [13] Medtronic. [n.d.]. Zephyr Performance Systems. <https://www.zephyranywhere.com/>. Accessed: 2019-04-07.
- [14] Markus Miettinen, N Asokan, Thien Duc Nguyen, Ahmad-Reza Sadeghi, and Majid Sobhani. 2014. Context-based zero-interaction pairing and key evolution for advanced personal devices. In *Proceedings of the 2014 ACM SIGSAC conference on computer and communications security*. ACM, 880–891.
- [15] Frank Mokaya, Roland Lucas, Hae Young Noh, and Pei Zhang. 2016. Burnout: a wearable system for unobtrusive skeletal muscle fatigue estimation. In *2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE, 1–12.
- [16] Frank O Mokaya, Brian Nguyen, Cynthia Kuo, Quinn Jacobson, Anthony Rowe, and Pei Zhang. 2013. MARS: a muscle activity recognition system enabling self-configuring musculoskeletal sensor networks. In *Proceedings of the 12th international conference on Information processing in sensor networks*. ACM, 191–202.
- [17] Le T Nguyen, Yu Seung Kim, Patrick Tague, and Joy Zhang. 2014. IdentityLink: user-device linking through visual and RF-signal cues. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 529–539.
- [18] Shijia Pan, Carlos Ruiz, Jun Han, Adeola Bannis, Patrick Tague, Hae Young Noh, and Pei Zhang. 2018. UniverSense: IoT Device Pairing through Heterogeneous Sensing Signals. In *Proceedings of the 19th International Workshop on Mobile Computing Systems & Applications*. ACM, 55–60.
- [19] Maximilian Riesenhuber and Tomaso Poggio. 1999. Hierarchical models of object recognition in cortex. *Nature neuroscience* 2, 11 (1999), 1019.
- [20] Carlos Ruiz, Shijia Pan, Adeola Bannis, Xinlei Chen, Carlee Joe-Wong, Hae Young Noh, and Pei Zhang. 2018. IDrone: Robust Drone Identification through Motion Actuation Feedback. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (2018), 80.
- [21] Carlos Ruiz, Shijia Pan, Hae Young Noh, Pei Zhang, and Jun Han. 2019. Secure pairing via video and IMU verification: demo abstract. In *Proceedings of the 18th International Conference on Information Processing in Sensor Networks*. ACM, 333–334.
- [22] Carlos Ruiz, Shijia Pan, Alberto Sadde, Hae Young Noh, and Pei Zhang. 2018. PosePair: pairing IoT devices through visual human pose analysis: demo abstract. In *Proceedings of the 17th ACM/IEEE International Conference on Information Processing in Sensor Networks*. IEEE Press, 144–145.
- [23] Abraham Savitzky and Marcel JE Golay. 1964. Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry* 36, 8 (1964), 1627–1639.
- [24] Matt Trumble, Andrew Gilbert, Charles Malleon, Adrian Hilton, and John Colomosse. 2017. Total Capture: 3D Human Pose Estimation Fusing Video and Inertial Sensors. In *2017 British Machine Vision Conference (BMVC)*. 1–13.
- [25] Timo von Marcard, Roberto Henschel, Michael Black, Bodo Rosenhahn, and Gerard Pons-Moll. 2018. Recovering Accurate 3D Human Pose in The Wild Using IMUs and a Moving Camera. In *European Conference on Computer Vision (ECCV)*. 1–17.
- [26] Shih-En Wei, Varun Ramakrishna, Takeo Kanade, and Yaser Sheikh. 2016. Convolutional pose machines. In *CVPR*. IEEE, 4724–4732.
- [27] Yuezong Wu, Wen Hu, and Mahbub Hassan. 2018. Learning for Device Pairing in Body Area Networks. In *Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems*. ACM, 321–322.
- [28] Lin Yang, Wei Wang, and Qian Zhang. 2016. Secret from muscle: Enabling secure pairing with electromyography. In *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM*. ACM, 28–41.