Meta-Activity Recognition: A Wearable Approach for Logic Cognition-based Activity Sensing

Lei Xie, Xu Dong, Wei Wang, and Dawei Huang
State Key Laboratory for Novel Software Technology, Nanjing University, China
Email: lxie@nju.edu.cn, dongxu@dislab.nju.edu.cn, ww@nju.edu.cn, huangdw@dislab.nju.edu.cn

Abstract—Activity sensing has become a key technology for many ubiquitous applications, such as exercise monitoring and elder care. Most traditional approaches track the human motions and perform activity recognition based on the waveform matching schemes in the raw data representation level. In regard to the complex activities with relatively large moving range, they usually fail to accurately recognize these activities, due to the inherent variations in human activities. In this paper, we propose a wearable approach for logic cognition-based activity sensing scheme in the logical representation level, by leveraging the meta-activity recognition. Our solution extracts the angle profiles from the raw inertial measurements, to depict the angle variation of limb movement in regard to the consistent body coordinate system. It further extracts the meta-activity profiles to depict the sequence of small-range activity units in the complex activity. By leveraging the least edit distance-based matching scheme, our solution is able to accurately perform the activity sensing. Based on the logic cognition-based activity sensing, our solution achieves lightweight-training recognition, which requires a small quantity of training samples to build the templates, and user-independent recognition, which requires no training from the specific user. The experiment results in real settings shows that our meta-activity recognition achieves an average accuracy of 92% for user-independent activity sensing.

I. INTRODUCTION

Nowadays activity sensing has become a key technology for many ubiquitous applications such as exercise monitoring and elder care. For example, in the daily exercise monitoring, it is essential to figure out what kinds of exercises the human subjects did everyday. The rising of the wearable devices has provided new opportunities for activity sensing during human motion. The wearable devices such as the smart watches are usually embedded with inertial sensors like the accelerometers, gyroscopes and magnetometers. They are able to continuously track the human subject’s movements and classify them into the corresponding activities by matching the waveforms of inertial measurements against the templates. However, a number of common activities, e.g., dumbbell curl and rope skipping, belong to the complex activities. The complex activity refers to an activity which has large range of movement and incurs rotations on multiple joints of the limbs, e.g., the movement has angle change of more than 45° and involves more than 2 joints of the limbs. Moreover, it usually has two complex aspects: the widespread variations in activity details and the large movement range. Due to the user-specific characters like the heights, limb lengths and moving behaviors, there exist obvious deviations in the raw inertial measurements from different human subjects during the process of the complex activity. Therefore, traditional activity sensing schemes [1]–[3] are either based on the user-dependent recognition, which requires to record the training data from the current user to improve the recognition accuracy, or relying on heavy-training, which requires to collect a large quantity of training samples to build the templates. It is essential to propose a brand-new activity sensing scheme, such that the derived recognition models can be scalable to any arbitrary human subjects in a user-independent and light-training approach.

In this paper, we propose a wearable approach for logic cognition-based activity sensing, by leveraging the meta-activity recognition in the logical representation level. We mainly focus on the complex activities from human subjects, as shown in Fig. 1. Our approach is based on the observation that when the human subject is performing an arbitrary activity, he/she is experiencing a very similar sequence of small-range-activity units in the logical aspect, despite of the detailed differences in the waveforms of the raw inertial measurements. We leverage the notion meta-activity to denote the small-range-activity units which compose a common activity of human subject. Given an arbitrary activity, our approach first extracts the angle profiles from the raw measurements to depict the angle variation of limb movement in the consistent body coordinate system. Then, it further extracts the meta-activity profiles to depict the sequence of small-range-activity units in the specific activity. By leveraging the least edit distance-based matching scheme, our solution is able to accurately perform the activity sensing. Since a scalable recognition model is derived from the meta-activity-based templates in the logical representation level, our solution achieves lightweight-training recognition, which requires a small quantity of training samples to build the templates, and user-independent recognition, which requires no training from the specific user.
There are two key technical challenges in realizing the activity sensing scheme. The first challenge is to realize the activity sensing in a user-independent approach, such that the derived recognition model can be extended to recognize the activities of any arbitrary human subjects, regardless of the detailed differences and inherent deviations in the activities from different human subjects. To address this challenge, we propose to leverage the angle profiles, i.e., the angles between the specified limb and the coordinate axes, to depict the limb movements. The angle profiles are able to capture the angle variation of the limb movements relative to the human body, which tackle the deviation details caused by the user-specific characters like the height. Moreover, we propose the method of “meta-activity recognition” to perform activity sensing in the logical representation level, based on the sequence of meta-activity profiles, so as to tackle the variations in the long sequence of small-range activities. Specifically, according to the inertial measurements collected from human motion, instead of performing the waveform-based matching like dynamic time warping, we decompose the complex activity into a sequence of meta-activities, and use this sequence to recognize the complex activity via the least edit distance-based matching.

The second challenge is to build a consistent scheme to depict the human motion according to the inertial measurements from the wearable devices. Since the human subjects may perform the activities towards any arbitrary direction during the human motion, this causes the templates for activity recognition to depend heavily on the actual direction the human body is facing, and further enhances the complexities in performing activity sensing due to the inconsistency. To address this challenge, we depict all the inertial measurements of human motion in terms of a body coordinate system in a consistent approach. Specifically, according to the gravity direction and the magnetic direction extracted from the inertial measurements, we transform the measurements from the watch coordinate system (WCS) to the global coordinate system (GCS). Then, by specifying two signal gestures, i.e., extending the arm to the front and dropping the arm downward, we can figure out the orientation of the human body in the global coordinate system according to the measurements in the signal gestures, thus we further transform the measurements to the body coordinate system (BCS).

To the best of our knowledge, this paper presents the first study of using the method “meta-activity recognition” for logical cognition-based activity sensing. Specifically, we make three key contributions in this paper. 1) Instead of performing waveform matching on the inertial measurements in the raw data level, we extract the angle profiles to depict the angle variation of limb movements, and leverage the meta-activity profiles to depict the complex activities in the logical representation level, such that the derived recognition model is scalable enough for the activity recognition on any arbitrary human subjects. 2) We build a coordinate system transformation scheme to transform the inertial measurement from the watch coordinate system to the body coordinate system, such that the limb movement can be depicted in a consistent approach, regardless of the exact orientation of the human bodies. 3) We have implemented a prototype system to evaluate the real performance, the experiment results in real settings shows that our meta-activity recognition achieves an average accuracy of 92% for user-independent activity sensing.

II. PROBLEM FORMULATION

In this paper, we investigate the wearable approach for activity sensing, i.e., a wearable device is worn by the human subject to continuously collect the inertial measurements of human motion, then an activity sensing scheme is required to accurately recognize the complex activities of limb movements from human subjects. The complex activity refers to the kind of activity with a large range of movement, such as sit-ups and dumbbell lateral raise. Without loss of generality, we leverage the smart watch to sense the human motions, which is embedded with inertial sensors including the accelerometer, gyroscope and magnetometer.

In this paper, we aim to design an activity sensing scheme, by considering the following metrics in system performance: 1) Accuracy: The expected accuracy for the activity sensing scheme to successfully match a specific activity to a correct activity should be greater than a specified threshold, e.g., 85%. 2) Time-efficiency: The time delay of the activity recognition process should be less than a specified threshold, e.g., 500ms. 3) User-independence: When performing activity sensing, no training data should be required from the specified user. 4) Lightweight-training: The essential quantity of the training samples to build the templates should be small enough.

III. MODELING THE HUMAN MOTION

A. Coordinate System Transformation

In regard to activity sensing, as the raw inertial measurements are collected from the embedded inertial sensors in the smart watch, they are measured by reference to the body frame of the smart watch. However, the watch coordinate system is continuously changing with the arm/wrist movement during the process of human motion, thus the measurements from the watch coordinate system cannot be used as a stable reference for the specified activities. In fact, since the human subject may be performing the activity towards any arbitrary direction, the movements should be depicted as the movement of arms or legs relative to the human body, regardless of the absolute moving direction of the limbs. Therefore, in order to perform activity sensing in a scalable approach, it is essential to transform the measurement of limb movements from the watch coordinate system to the body coordinate system.

1) From Watch Coordinate System to Global Coordinate System: Fig. 2(a) shows the three axes of the watch coordinate system. The $X_w$-axis refers to the direction of the lower arm when the watch is worn on the wrist, the $Y_w$-axis refers to the direction of the strap of the watch, and the $Z_w$-axis refers to the direction which is perpendicular to the watch surface.

According to the acceleration measurements from the accelerometer, we can extract a constant gravitational acceleration as a vector $g$ from the low pass filter (such as
the Butterworth filter [4]) in the watch coordinate system. Moreover, according to the magnetic measurements from the magnetometer, we can extract the magnetic force as a vector m in the watch coordinate system. Then, we can build a global coordinate system (GCS) based on the gravity direction and magnetic direction in the watch coordinate system. The procedure is as follows: After we obtain the gravity vector g, we derive its opposite value and normalize this vector as \(\mathbf{z}_g = \frac{-\mathbf{g}}{\|\mathbf{g}\|}\), then we set this vector \(\mathbf{z}_g\) to represent the global Zg-axis as it is in the opposite direction of the gravitational acceleration and it is perpendicular to the horizontal plane. After computing the cross product \(\mathbf{y} = \mathbf{g} \times \mathbf{m}\), we obtain a vector \(\mathbf{y}\) that is perpendicular to the plane determined by the two distinct but intersecting lines corresponding to \(\mathbf{g}\) and \(\mathbf{m}\). We normalize this vector as \(\mathbf{y}_g = \frac{\mathbf{y}}{\|\mathbf{y}\|}\). Since the vector \(\mathbf{y}_g\) is on the horizontal plane, we set this vector \(\mathbf{y}_g\) to represent the global Yg-axis. After that, by computing the cross product \(\mathbf{x} = \mathbf{g} \times \mathbf{y}\), we obtain a vector \(\mathbf{x}\) that is orthogonal to the plane determined by the two distinct but intersecting lines corresponding to \(\mathbf{g}\) and \(\mathbf{y}\). We normalize this vector as \(\mathbf{x}_g = \frac{\mathbf{x}}{\|\mathbf{x}\|}\) to represent the global Xg-axis. Fig. 2(a) further shows the relationship between the three axes (\(\mathbf{x}_w\), \(\mathbf{y}_w\), and \(\mathbf{z}_w\)) of WCS and the three axes (\(\mathbf{x}_g\), \(\mathbf{y}_g\), and \(\mathbf{z}_g\)) of GCS.

To quantify the orientation difference between the watch coordinates and global coordinates, we use the direction cosine representation [5]. In the direction cosine representation, the orientation of the global coordinate relative to the watch coordinate system is specified by a \(3 \times 3\) rotation matrix \(\mathbf{C}'\), in which each column is a unit vector along one of the watch coordinate axes specified in terms of the global coordinate axes. A vector quantity \(v_w\) defined in the watch coordinate system is equivalent to the vector \(\mathbf{v}_g = \mathbf{C}' \cdot \mathbf{v}_w\) defined in the global coordinate system. In this way, we are able to transform any inertial measurement \(v_w\) from WCS to the corresponding inertial measurement \(v_g\) in GCS. During the human motion, the directions of \(\mathbf{g}\) and \(\mathbf{m}\) are continuously updated in WCS to track the three axes of GCS, so as to further update the rotation matrix \(\mathbf{C}'\) in a real-time approach.

2) From Global Coordinate System to Body Coordinate System: During the human motion, the human subject may be facing any arbitrary direction in regard to the global coordinate system. Hence, although we can derive the inertial measurement of limb movements in GCS, these measurements may not be consistent with each other even if they belong to the same activity, due to the differences in the facing directions. Therefore, it is essential to build a body coordinate system (BCS) to depict the limb movements in a consistent approach by reference to the human body.

In regard to the body coordinate system, we set the vector corresponding to the heading direction of the human subject to represent the \(\mathbf{Z}_b\) axis. For the horizontal plane which is orthogonal to the \(\mathbf{Z}_b\) axis, we set the vector which is parallel to the physical plane of the body to represent the \(\mathbf{X}_b\) axis, and set the vector which is perpendicular to the physical plane of the body to represent the \(\mathbf{Y}_b\) axis. Fig. 2(b) shows the three axes \((\mathbf{X}_b, \mathbf{Y}_b, \mathbf{Z}_b)\) of BCS and the three axes \((\mathbf{X}_g, \mathbf{Y}_g, \mathbf{Z}_g)\) of GCS in regard to the physical plane of the human body, respectively. Considering that the human subject can perform the activity with different orientations of the physical plane of the body, e.g., standing on the floor or lying on the floor, in all situations, we can transform any inertial measurement from the GCS to BCS by also using the direction cosine representation. The orientation of the body coordinate system relative to the global coordinate system is specified by a \(3 \times 3\) rotation matrix \(\mathbf{C}'\), in which each column is a unit vector along one of the global coordinate axes specified in terms of the body coordinate axes. A vector quantity \(v_g\) defined in GCS is equivalent to the vector \(v_b = \mathbf{C}' \cdot v_g\) defined in BCS. In this way, we can transform any inertial measurement from the GCS to the BCS. In Section IV, we will introduce the approach to compute the rotation matrix \(\mathbf{C}'\), by leveraging two signal gestures.

In regard to the activities where the physical plane of the human body is continuously changing, e.g., sit-ups, we can set the initial physical plane of the human body as the reference body coordinate system. In this way, each of the following inertial measurements are measured in terms of the reference body coordinate system.

B. Modeling the Human Motion with Meta-Activity

Each complex activity, e.g., dumbbell side raise and bent-over dumbbell laterals, is performed with a large range of movement. So it can be decomposed into a series of small-range movements which are sequentially performed over time. Therefore, we leverage the term meta-activities to denote these small-range movements. Each meta-activity is defined as a unit movement with logically the minimal granularity in regard to the moving range. We can define the whole set of complex activities as a set \(\mathcal{C}\), and the whole set of meta-activities as a set \(\mathcal{M}\). Then, according to the above definition, each complex activity \(c_i \in \mathcal{C}\) can be depicted as a series of meta-activities, i.e., \(c_i = (m_{j_1}, \ldots, m_{j_k})\), where \(m_j \in \mathcal{M}\).

1) Angle Profiles: In regard to the activity sensing, due to the differences in human-specific characters such as the height, arm length, and moving behavior, different human subjects may perform the same activity with different speeds and amplitudes. This causes nonnegligible deviations among the inertial measurements of the same activities in both time domain and space domain. Therefore, the meta-activity should be depicted in a scalable approach, such that the activity sensing scheme can be tolerant to the variances in the limb movements. However, traditional inertial measurements such
as the linear accelerations are very sensitive to the speeds and amplitudes of the limb movements, which fail to depict the meta-activity in a scalable approach. Fortunately, it is found that, during the process of limb movements, the angle variations between the limb and the body are much more stable than the traditional inertial measurements, which are regardless of the human-specific characters such as the height and arm length. Therefore, in this paper, we propose to leverage the angle profiles, i.e., the angles between the lower arm and the three axes in the body coordinate system, to depict the meta-activities of the limb movements. Specifically, since the direction of $X_w$ axis in the watch coordinate system is consistent with the lower arm direction, we can use the vector $x_w$ to depict the lower arm direction in the body coordinate system\(^1\). Fig. 3(a) shows the vector $x_w$ to depict the lower arm direction in the BCS. As shown in Fig. 3(b), we respectively denote the angle profiles, i.e., the angles between the lower arm and the $X$, $Y$ and $Z$ axes in the BCS, as $\alpha$, $\beta$ and $\gamma$.

In order to compute the angle profiles, we take the angle $\alpha$ as an example, suppose the lower arm vector and the vector of the $X$-axis are $v$ ($v = x_w$ ) and $u$, respectively, in the BCS. Then $\alpha$ can be computed according to the cosine value as follows:

$$\cos \alpha = \frac{v \cdot u}{|v||u|} = \frac{v_x u_x + v_y u_y + v_z u_z}{\sqrt{v_x^2 + v_y^2 + v_z^2} \sqrt{u_x^2 + u_y^2 + u_z^2}}. \quad (1)$$

For any specified value of $\cos \alpha$, there exist two solutions of $\alpha$ in the range between $0^\circ$ and $360^\circ$. Hence, we first compute the corresponding solution $\hat{\alpha}$ within the range $[0^\circ, 180^\circ]$, we then further determine the value of $\alpha$ as follows:

$$\alpha = \begin{cases} \hat{\alpha} & \text{if } v_y \geq 0 \\ 360^\circ - \hat{\alpha} & \text{if } v_y < 0. \end{cases} \quad (2)$$

Similarly, we can compute $\cos \beta$ and $\cos \gamma$ accordingly, then the values of $\beta$ and $\gamma$ can be determined as follows:

$$\beta = \begin{cases} \hat{\beta} & \text{if } v_z \geq 0 \\ 360^\circ - \hat{\beta} & \text{if } v_z < 0. \end{cases} \quad (3)$$

$$\gamma = \begin{cases} \hat{\gamma} & \text{if } v_x \geq 0 \\ 360^\circ - \hat{\gamma} & \text{if } v_x < 0. \end{cases} \quad (4)$$

In this way, the angle profiles $\langle \alpha, \beta, \gamma \rangle$ in the BCS can be determined within the range $[0^\circ, 360^\circ]$.

We further conducted empirical studies to validate the above judgement. We invite four human subjects ($a$, $b$, $c$ and $d$) with different heights and genders to perform the specified complex activities, and respectively record the corresponding acceleration measurements and the angle profiles in regard to each of the axes in the body coordinate system. We normalize all the measurements to the range $[0, 1]$ for fair comparison. Fig.4(a) and Fig.4(b) respectively shows the acceleration measurements and angle profiles of the activity Dumbbell Curl. It is found that, among different human subjects, there exist obvious variances in the acceleration measurements, whereas the variances in the angle profiles are relatively small. We further compute the DTW distances between each pair of measurements from different human subjects, and obtain the average distance as the metric to quantify the corresponding variances. Fig.4(c) shows the DTW distances, respectively, for the activity Dumbbell Curl and Sit-Up. It is found that for both cases the angle profiles achieve much smaller distances than the acceleration measurement, which implies that the angle profile is a more stable metric to depict the human motion.

\(^1\)As mentioned in Section III, the vector $x_w$ in BCS can be computed according to the direction cosine representation, it can be continuously updated in a real time approach.
sector to depict the corresponding meta-activity in the specified dimension. Moreover, considering the arm rotation can be anti-clock-wise or clock-wise, it is essential to further label each sector according to the rotation trend. Therefore, suppose the number of sectors is \( m \), we can label each sector with a different ID from 0 to \( m - 1 \) in an anti-clock-wise approach: for the \( i \)th sector, if the rotation direction is anti-clock-wise, then we label it with \( s_j \), otherwise, we label it with \( S_j \). Fig. 5 shows an example of these meta-activity sectors, where each sector has an angle of 30°. These sectors are labeled from \( s_0 \) to \( s_{11} \) if the rotation direction is anti-clock-wise, and they are labeled from \( S_0 \) to \( S_{11} \) otherwise. In this way, we can use these discrete states rather than the continuous waveforms to represent the meta-activities. In comparison to the continuous waveform-based representation in the raw data level, this discrete state-based representation is based on the logic cognition of the human motion, which is more scalable to the inherent variances caused by user specific characters.

Fig. 5. The sectors to depict meta-activity in each dimension of angle profiles

IV. SYSTEM DESIGN

The overall system is composed of three major modules, as shown in Fig.6: Data Acquisition and Preprocessing takes the raw inertial measurements as input. It first performs the coordinate transformation to transform the measurement from WCS to BCS. Then, it extracts the angle profiles and further split the series into separate complex activities. Meta-Activity Segmentation and Classification segments a single complex activity into a series of meta-activities, and classifies the segmented meta-activities into corresponding categories. Complex Activity Recognition performs activity recognition based on the sequences of meta-activities from the test complex activity, by leveraging the least edit distance-based matching scheme.

A. Data Acquisition and Preprocessing

1) Coordinate Transformation: As mentioned in Section III, we can transform the measurement from the WCS to BCS, by using the Direction Cosine method. To figure out the orientation difference, i.e., the rotation matrix \( C' \) between BCS and GCS, before the human subject performs the complex activities, he/she is required to perform the following two signal gestures in advance: 1) Extend the arm to the front: let the human subject extend his/her arm to the front of the body, the arm direction is consistent with the \( Y_b \) axis in the BCS; 2) Drop the arm downward: let the human subject drop the arm downward along his/her legs, the arm direction is opposite to the \( Z_b \) axis in the BCS. Fig.7(a) and Fig.7(b) shows an example of the two signal gestures, respectively. In this way, by computing the corresponding vector of the arm direction in the GCS, we are able to figure out the rotation matrix \( C' \), whatever the human subject is standing or lying on the floor.

Fig. 6. The system framework

(a) Signal gesture 1: extend the arm (b) Signal gesture 2: drop the arm downward

Fig. 7. The signal gestures

2) Angle Profiles Extraction: As aforementioned in Section III, take the smart watch as an example, we use the vector \( x_w, \) i.e., the direction of \( X_w \) axis in the WIS, to depict the arm direction in the BCS. Then, according to the arm vector \( x_w(t) \) at time \( t \), we can extract the angle profiles \( \{\alpha(t), \beta(t), \gamma(t)\} \) over time in the BCS according to Eq. (4)-(7).

3) Segmentation: In practice, the human subject may continuously perform a series of complex activities. Therefore, the recognition system should first split these series of complex activities into separate activities, then we can further identify which activity pattern the current movement belongs to. As the human subject usually takes a short pause between two
We propose to perform the meta-activity segmentation over each dimension of the angle profiles in a separate approach. For each dimension, we leverage a sliding window (which is set to 500ms in our implementation) to continuously store the recent angle profiles \( \langle \alpha(t), \beta(t), \gamma(t) \rangle \) for a single complex activity, where \( t_s \) and \( t_e \) are respectively the start time and the end time of the complex activity. Then, it is essential to further segment the complex activity into a series of meta-activities, according to the properties of the meta-activity mentioned in Section III.B, we can derive the segmentation conditions for any of the angle profiles \( i.e., \alpha, \beta \) or \( \gamma \).

A straightforward solution for meta-activity segmentation is to segment the series of angle profiles by checking all the three dimensions of the angle profiles simultaneously. If the segmentation condition is satisfied in any of the dimensions, the entire series of angle profiles are segmented as a single meta-activity. However, this solution causes the series of angle profiles for a single meta-activity to be too fragmented after segmentation, since the series of angle profiles in the other dimensions are not yet satisfied for the segmentation condition to be a complete meta-activity, which is not suitable for the following meta-activity classification. Therefore, we propose to perform the meta-activity segmentation over each dimension of the angle profiles in a separate approach. For each dimension, we leverage a sliding window (which is set to 500ms in our implementation) to continuously store the recent angle profiles. Our solution scans the series of angle profiles to verify if any of the segmentation condition is satisfied, then we segment the angle profile series as a single meta-activity in the corresponding dimension. After the meta-activity segmentation, we can obtain a separate segmentation for each dimension of the angle profiles.

Fig. 8 shows an example of the meta-activity segmentation for the activity “Dumbbell Curl” in the three dimensions of the angle profiles.

2) Meta-Activity Classification: After the meta-activity segmentation, according to each dimension of the angle profiles, the meta-activity should be further classified into one specific sector based on its moving range and rotation direction. As the meta-activity depicts a process of movement ranging from a start angle \( \theta_s \) to an end angle \( \theta_e \), it usually has some fluctuations in the waveforms of the angle profiles. Hence, when performing meta-activity classification, it is essential to consider the variation trend of angle profiles, instead of only the start and end angles. For example, a test meta-activity may increase the angle profile slowly in part of the sector \( s_5 \), and then increase the angle profile rapidly in part of the sector \( s_6 \), so it should be classified to the sector \( s_5 \) since its major angle profile varies in this area. Therefore, we leverage the method of Dynamic Time Warping (DTW) to match a test meta-activity to a corresponding sector by referring to the variation trend of angle profiles. The procedure is as follows: In regard to any specific sector, considering the rotation range and direction, its start angle profile is \( \eta_s \) and its end angle profile is \( \eta_e \). Then, suppose the corresponding meta-activity is performed with a uniform speed in the time interval of \( \Delta t \) (\( \Delta t = 500ms \) in our implementation), we can use a linear function \( f(t) \) to depict the template of this meta-activity, \( i.e., f(t) = \eta_s + \frac{\eta_e - \eta_s}{\Delta t} \cdot t \). Then, given a test meta-activity with the angle profile \( \eta(t) \), and the templates of meta-activities corresponding to the sectors, we can use a vector \( P_t \) with length \( l \) and a vector \( R' \) with length \( l' \) to denote the test meta-activity’s angle profile and the specified template meta-activity’s angle profile, respectively. For each pair of test meta-activity \( P_t \) and the template meta-activity \( R' \), we construct a distance matrix \( D_{l \times l'} \) as an input to the DTW algorithm, where each element \( D_{l \times l'} \) is defined as the Euclidean distance between each pair of angle profiles \( V_{P, i} \) and \( V_{R, j} \). The output of DTW is a warping path \( \pi = \{ \pi_1, ..., \pi_k \} \) such that the distance \( d \) between the sequences is minimized: \( \arg \min_{\pi} D_{\pi} = \sum_{k=1}^{k} D_{\pi(k), \pi(k+1)} \). Hence, given the test meta-activity, we can enumerate all template meta-activities and leverage DTW to compute the corresponding distance. We then select the template meta-activity with the smallest distance as the classification result.

C. Complex Activity Recognition

After the meta-activity classification, each complex activity can be decomposed into a sequence of meta-activities, respectively, in regard to the three dimensions of angle profiles. We call them meta-activity profiles of the complex activity. In Table I, we show four example meta-activity profiles for the specified complex activities.

In this paper, we leverage the Least Edit Distance-based matching scheme (LED) to perform the complex activity recognition. LED compares the test complex activity against the template complex activities in regard to the corresponding meta-activity profiles. It computes the least edit distance between each pair of test complex activity and template complex activity, and select the template complex activity with the smallest distance as the matching result. By referring to the edit distance [7] for measuring the difference between two
sequences of strings, we leverage this term to denote the difference between two sequences of meta-activity profiles.

In order to compute the edit distance between two pairs of complex activities, it is essential to first consider the distance between two meta-activities. As aforementioned in Section III, for each dimension of the angle profiles, the meta-activity is a process which is performed in a specified sector with a specified rotation direction. When considering the distance between two meta-activities, we should take these two issues into consideration, i.e., the distance between sectors and the distance between rotation directions. Considering the distance between sectors, assume the number of sectors is \( m \), for any two meta-activities \( m_i \) and \( m_j \), suppose their corresponding sector numbers are respectively \( s_i \) and \( s_j \) (\( 0 \leq s_i \leq s_j < m \)), the distance between them is defined as follows:

\[
d_s(m_i, m_j) = \min \{(s_j - s_i) \mod m, (s_i - s_j + m) \mod m\}.
\]

The distance is the minimum distance between the two sectors \( s_i \) and \( s_j \) either clockwise or anti-clockwise. Considering the distance between rotation directions (clock-wise or anti-clock-wise), for any two meta-activities \( m_i \) and \( m_j \), if they have the same rotation direction, then we set the distance \( d_r(m_i, m_j) \) to 0. Otherwise, we set the distance \( d_r(m_i, m_j) \) to \( \Omega \) (\( \Omega = m/4 \) in our implementation). Hence, considering the above two issues, the distance between two meta-activities \( m_i \) and \( m_j \) is as follows:

\[
d(m_i, m_j) = d_s(m_i, m_j) + d_r(m_i, m_j).
\]

Therefore, for each dimension of the angles profiles, a complex activity \( c_i \in C \) can be depicted as a sequence of the meta-activities, i.e., \( c_i = \langle m_{j_1}, \cdots, m_{j_k} \rangle \), where \( m_j \in M \). Then, for a specified dimension, considering any two complex activities, e.g., \( a \) and \( b \), we can compute their distance \( L_{a,b}(\langle a \rangle, \langle b \rangle) \) by referring to the Levenshtein distance [7]:

\[
L_{a,b}(i, j) = \begin{cases} 
\max(i, j) \times \mu & \text{if } \min(i, j) = 0, \\
\min \left\{ L_{a,b}(i-1, j) + \mu, L_{a,b}(i, j-1) + \mu \right\} & \text{otherwise},
\end{cases}
\]

where \( L_{a,b}(i, j) \) is the distance between the first \( i \) meta-activities of \( a \) and the first \( j \) meta-activities of \( b \), \( \mu \) is the average distance between any two meta-activities (\( \mu = 0.75 \times m \) in our implementation), and \( d(a_i, b_j) \) is the distance between the \( i \)th meta-activity of \( a \) and the \( j \)th meta-activity of \( b \).

After that, for any two complex activities \( c_i \) and \( c_j \), we add the distances from all three dimensions together, and obtain the overall distance between the two complex activities as follows:

\[
L_{c_i,c_j} = \sqrt{L_{c_i,a}(\alpha,c_j(\alpha))^2 + L_{c_i,b}(\beta,c_j(\beta))^2 + L_{c_i,c}(\gamma,c_j(\gamma))^2}.
\]

Therefore, given a test complex activity \( c_i \), we enumerate all template meta-activity profiles of all complex activities \( c_j \in C \) and compute their distance \( L_{c_i,c_j} \), then we select the category of the template complex activity with the least distance as the recognition result.

V. PERFORMANCE EVALUATION

A. Experimental Setup

We have implemented a prototype system using the android phone (SAMSUNG Galaxy S5)\(^2\), which is attached to the wrist of the human subject, as shown in Fig.9. The android phone is embedded with inertial sensors including accelerometers and magnetometers. The lower-arm direction is consistent with the Y-axis of the smart phone’s local coordinate system. In the experiment, we let 10 volunteers perform 10 categories of complex activities, they have different heights, genders, and ages. For each category of complex activity, 20 samples of inertial measurements are collected for each subject. In order to evaluate the performance for user-independent activity sensing, we leverage the n-fold cross-validation as follows: for each round of evaluation, we select one human subject as the test case, and obtain the template profiles from \( n - 1 \) of the remaining human subjects. We then evaluate the recognition accuracy and time delay for the three solutions: 1) Acceleration-based Matching (AM): It uses the DTW to perform waveform-based matching in terms of the acceleration measurements. 2) Angle Profiles-based Matching (APM): It uses the DTW to perform waveform-based matching in terms of the angle profiles. 3) Meta-Activity Recognition (MAR): It uses the least edit distance-based matching in terms of the meta-activity profiles.

B. Parameter Selection

For the meta-activity recognition, the angle of a meta-activity sector, i.e., \( \delta \), is very crucial to the performance in terms of recognition accuracy and time efficiency. It directly determines the number of meta-activities within a specified complex activity. Therefore, we conduct experiments to evaluate the performance with different values of \( \delta \). We set the number of human subjects in template construction to 5.

\(^2\)As COTS smart watches are still not embedded with magnetometers to help build the body coordinate system, so in this paper we choose to use the android phone as the testing wearable devices.

<table>
<thead>
<tr>
<th>Complex Activity</th>
<th>Meta-Activity Profiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dumbbell</td>
<td>α : (S0, S8, S7, S5, s7, s8)</td>
</tr>
<tr>
<td>Triceps Extension</td>
<td>β : (s8, S6)</td>
</tr>
<tr>
<td>Upright Barbell</td>
<td>γ : (S2, S1, S0, S11, s1, s2)</td>
</tr>
<tr>
<td>Lateral Raise</td>
<td>β : (s2)</td>
</tr>
<tr>
<td>Butterfly Fly</td>
<td>α : (s2)</td>
</tr>
</tbody>
</table>

Table I: Example meta-activity profiles for the complex activities.
We first evaluate the performance in terms of recognition accuracy when the angle $\delta$ is varied from $5^\circ$ to $45^\circ$, as shown in Fig. 10(a). It is found that all the recognition accuracies are greater than $83\%$ when the angle is varied from $5^\circ$ to $45^\circ$. The highest accuracy is achieved when the angle is set to $10^\circ$, whereas the lowest accuracy is achieved when the angle is set to $45^\circ$. The reason is that, when the angle is fairly large, e.g., $45^\circ$, the granularity of the meta-activity is too coarse to depict the movement of human subjects, thus it leads to many mismatches in activity recognition. However, when the angle is fairly small, e.g., $5^\circ$, the granularity of the meta-activity is too fine to tolerate the detailed deviations due to human-specific characters, thus the performance is also degraded in comparison to the optimum case. Therefore, the parameter of the sector angle should be carefully selected for improving the performance in recognition accuracy.

We then evaluate the performance in terms of time efficiency when the angle $\delta$ is varied from $5^\circ$ to $45^\circ$, as shown in Fig. 10(b). It is found that, as the angle $\delta$ increases from $5^\circ$ to $30^\circ$, the average time delay rapidly decreases from $145\text{ ms}$ to $21\text{ ms}$. The reduced time delay is caused by the increasing granularity of the meta-activity, which reduces the processing time cost. However, when the angle $\delta$ further increases to $45^\circ$, the average time delay slightly increases to $49\text{ ms}$, as the time delay for the meta-activity classification increases due to the increased size of input to the DTW algorithm. Therefore, to achieve an appropriate trade-off between the accuracy and time efficiency, in the following experiment, we set the angle $\delta$ to $10^\circ$ in MAR to achieve the optimized performance.

C. Evaluate the Recognition Accuracy

1) Sensitivity to the number of training samples: Since we aim to achieve the lightweight-training recognition, we require the number of training samples to be as small as possible. Therefore, as we collect 20 samples from each human subject to build the templates for each complex activity, we vary the number of human subjects involved in the template construction from 1 to 8, and evaluate the average recognition accuracy, as shown in Fig. 10(c). It is found that, in almost all situations, MAR achieves the best performance whereas AM achieves the worst performance. Specifically, when the number of human subjects in template construction is 1, the number of training samples is small. AM achieves poor accuracy of $61\%$ as it lacks enough templates for accurate matching. APM leverages the angle profiles to mitigate the impact of variances in the raw inertial measurement, thus it enhances the accuracy from $61\%$ to $81\%$. Due to the character of logical cognition, MAR further improves the accuracy to $87\%$ even if the training samples are so limited. This implies that our meta-activity recognition achieve rather good performance for user-independent recognition while requiring lightweight-training. As the number of templates increases, the accuracies of the three solutions are all increasing to a close value of $92\%$. Nevertheless, MAR always achieves the least variances in recognition accuracy since it leads to very stable performance.

2) The matching ratios among multiple activities: We further investigate the confusion matrices for the three solutions, as shown from Fig.10(d) to Fig.10(f). We set the number of human subjects in template construction to 2, and the activities are listed from $A_1$ to $A_{10}$ according to the order in Fig.1. According to the matching results in the three confusion matrices, it is found that APM is able to reduce most of the mismatches caused by AM, so APM achieves the recognition ratio of $100\%$ for most activities. However, APM still fails to accurately recognize some activities such as $A_6$, $A_9$ and $A_{10}$, since these activities usually have larger movement ranges and more movement variations in details. Fortunately, MAR is able to further reduce these mismatches and improve the recognition accuracy to a fairly high level. Moreover, MAR achieves the least variances in recognizing multiple activities in comparison to the other two solutions.

D. Evaluate the Time Efficiency

We further evaluate the time delay of processing the activity sensing, respectively, for AM, APM and MAR. We vary the number of human subjects in template construction from 1 to
and evaluate the corresponding time delay. It is found that, in all situations, MAR achieves much smaller time delay (all less than 45ms) than AM and APM. Moreover, as the number of human subjects in template construction increases from 1 to 8, the time delay of AM and APM increases rapidly, whereas the time delay of MAR keeps fairly stable. The reason is as follows: As both AM and APM use the DTW for matching, it requires a large amount of time to process the measurements with a small granularity in the raw data level, however, MAR processes the measurements with a much larger granularity in the meta-activity level. It takes most of the processing time on the meta-activity segmentation rather than matching. Thus, MAR achieves the best time efficiency in tens of milliseconds.

VI. RELATED WORK

Wearable Device. Recent researches consider leveraging the inertial sensors embedded in wearable devices to detect and monitor the user’s activities [1], [8]–[13]. Wrist mounted inertial sensors are widely used for arm-based activity sensing [1], [2]. RistQ [1] leverages the accelerations from a wrist strap to detect and recognize smoking gestures. Karatas et al. [2] uses wrist mounted inertial sensors to track steering wheel usage and angle. Foot-mounted inertial sensors are leveraged for indoor localization by sensing the patterns of footsteps [12], [13]. LookUp [3] uses shoe-mounted inertial sensors for location classification based on surface gradient profile and step patterns. Robertson et al. [12] proposes an approach for simultaneous mapping and localization for pedestrians based on odometry with foot mounted inertial sensors.

Wireless Signals. Another branch of activity recognition solutions exploit the change of wireless signals (including WiFi signals, RF-signals, etc.) incurred by the human activities [14]–[19]. FEMO [14] provides a free-weight exercise monitoring scheme by attaching RFID tags on the dumbbells and leveraging the Doppler shift profile of the reflected backscatter signals for activity recognition. Wang et al. [15] propose a CSI based human activity recognition and monitoring system, by quantitatively building the correlation between CSI value dynamics and a specific human activity. E-eyes [16] presents device-free location-oriented activity identification at home through the use of fine-grained WiFi signatures. RF-IDraw [17] can infer a humans writing by tracking a passive RFID tag attached to his/her fingers.

However, most of the above activity recognition schemes leverage the traditional waveform-based matching to process the inertial measurement/wireless signals in the raw data level. In this paper, we propose the meta-activity recognition, which belongs to logic cognition-based activity sensing. Our approach achieves lightweight-training recognition, which requires a small quantity of training samples to build the templates, and user-independent recognition, which requires no training from the specific user.

VII. CONCLUSION

In this paper, we propose a wearable approach for logic cognition-based activity sensing scheme in the logical representation level, by leveraging the meta-activity recognition. Our solution extracts the angle profiles to depict the angle variation of limb movement in the consistent body coordinate system. It further extracts the meta-activity profiles to depict the sequence of small range activities in the complex activity. By leveraging the least edit distance-based matching scheme, the experiment results shows that our solution achieves an average accuracy of 92% for user-independent activity sensing.

ACKNOWLEDGMENTS

This work is supported in part by National Natural Science Foundation of China under Grant Nos. 61472185, 61373129, 61321491, 61502224; JiangSu Natural Science Foundation, No. BK20151390. This work is partially supported by Collaborative Innovation Center of Novel Software Technology and Industrialization.

REFERENCES


