Compressive Representation for Device-Free Activity Recognition with Passive RFID Signal Strength

Lina Yao, *Member, IEEE*, Quan Z. Sheng, *Member, IEEE*, Xue Li, *Member, IEEE*, Tao Gu, *Senior Member, IEEE*, Mingkui Tan, *Member, IEEE*, Xianzhi Wang, *Member, IEEE*, Sen Wang, and Wenjie Ruan

Abstract—Understanding and recognizing human activities is a fundamental research topic for a wide range of important applications such as fall detection and remote health monitoring and intervention. Despite active research in human activity recognition over the past years, existing approaches based on computer vision or wearable sensor technologies present several significant issues such as privacy (e.g., using video camera to monitor the elderly at home) and practicality (e.g., not possible for an older person with dementia to remember wearing devices). In this paper, we present a low-cost, unobtrusive, and robust system that supports independent living of older people. The system interprets what a person is doing by deciphering signal fluctuations using radio-frequency identification (RFID) technology and machine learning algorithms. To deal with noisy, streaming, and unstable RFID signals, we develop a compressive sensing, dictionary-based approach that can learn a set of compact and informative dictionaries of activities using an unsupervised subspace decomposition. In particular, we devise a number of approaches to explore the properties of sparse coefficients of the learned dictionaries for fully utilizing the embodied discriminative information on the activity recognition task. Our approach achieves efficient and robust activity recognition via a more compact and robust representation of activities. Extensive experiments conducted in a real-life residential environment demonstrate that our proposed system offers a good overall performance and shows the promising practical potential to underpin the applications for the independent living of the elderly.

Index Terms—Activity recognition, RFID, compressive sensing, subspace decomposition, feature selection

1 Introduction

The population is aging worldwide due to increasing life expectancy and low birth rate. With the recent developments in cheap sensor and networking technologies, we have seen a wide range of activity recognition applications for remote health monitoring and intervention and behavior analysis. These applications enhance the quality of people's lives, afford a greater sense of security, and facilitate their independent living [1], [2], [3], [4], [5]. For example, by monitoring a person with dementia, it is possible to track how completely and consistently the daily routines are performed, and determine when assistance is needed.

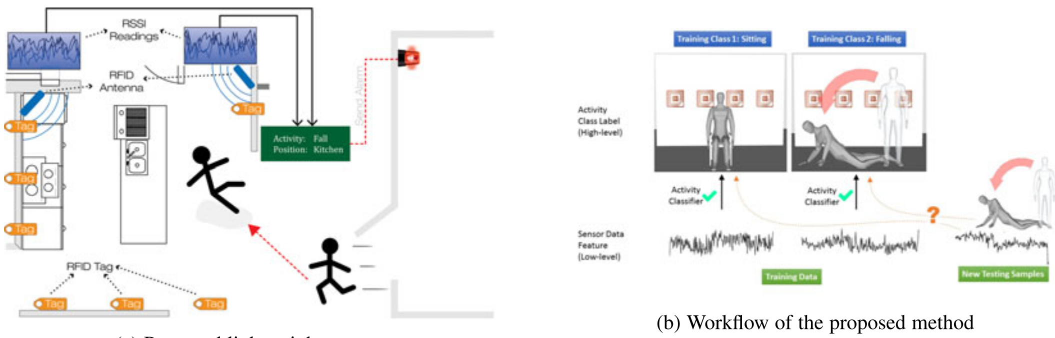
Activity recognition is a core aspect of ubiquitous computing as many application scenarios require an intelligent environment to infer what a person is doing or attempting to do. Essential to realizing these applications is *activity recognition*, which is emerging as an important research area in recent years. In general, activity recognition techniques have mainly focused on the direct observation of people and their behaviors with cameras or wearable sensors (e.g., accelerometer, gyro). To date, many efforts have been made to learn human activities by mining from a broad range of signal sources, such as videos and images [6], radio frequency of wearable or wireless sensors [7], [8], Wi-Fi [9], and even object vibration fluctuations [10].

Recognizing activity from wearable sensors has become a popular research topic in the past few years. This approach typically requires human subjects to wear a number of sensors [11], [12] or RFID tags [13]. Hence, it has two main shortcomings. It may be impractical to require people wearing sensor devices all the time, and the other obstacle is that those sensor devices typically need regular maintenance (e.g., battery replacement). As a result, sensor based activity recognition is not always practical, particularly in monitoring elderly people with cognitive disabilities.

Recently, device-free activity recognition has drawn much attention since it does not require subjects to wear any devices. Instead, sensor devices are placed in the environments, and radio signal fluctuations induced by subject's movements can then be collected and analyzed to recognize activities [14],

- *L. Yao is with the School of Computer Science and Engineering, UNSW, Kensington, NSW 2052, Australia. E-mail: lina.yao@unsw.edu.au.*
- *Q. Z. Sheng is with Macquarie University, Sydney, NSW 2109, Australia. E-mail: michael.sheng@mq.edu.au.*
- *X. Li is with the University of Queensland, St Lucia, QLD 4072, Australia. E-mail: xueli@itee.uq.edu.au.*
- *T. Gu is with RMIT University, Melbourne, VIC 3000, Australia. E-mail: tao.gu@rmit.edu.au.*
- *M. Tan is with the South China University of Technology, Guangzhou Shi 510630, China. E-mail: tanmingkui@gmail.com.*
- *X. Wang is with Singapore Management University, Singapore 188065. E-mail: sandyawang@gmail.com.*
- *S. Wang is with Griffith University, Nathan, QLD 4111, Australia. E-mail: sen.wang@griffith.edu.au.*
- *W. Ruan is with the University of Adelaide, Adelaide, SA 5005, Australia. E-mail: wenjie.ruan@adelaide.edu.au.*

Manuscript received 1 Mar. 2016; revised 9 Apr. 2017; accepted 4 May 2017. Date of publication 5 June 2017; date of current version 5 Jan. 2018.
(Corresponding author: Tao Gu.)
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TMC.2017.2706282



(a) Proposed lightweight setup

(b) Workflow of the proposed method

Fig. 1. (a) The proposed lightweight setup: A person performs different activities between the wall deployed with an RFID array and an RFID antenna. (b) The activities can be recognized by analyzing the corresponding signal strength fluctuation, Received Signal Strength Indicator (RSSI).

[15], [16], [17]. Radio Signal Strength indicator (RSSI) and Channel State Information (CSI) are explored to correlate signal fluctuations to each of the activities. These systems typically require a dense deployment of sensor devices which may incur high cost in both deployment and maintenance. With the advancement of RFID technology, we have seen more and more RFID tags and devices deployment in indoor settings. Recent work [18], [19] suggest exploiting RFID signal for device-free activity recognition. These work typically have stringent requirements on tag placement such as density and distance between tags. For example, a recent study by Wagner et al. [20] shows that an optimal tag placement is needed to alleviate inaccuracy caused by the variability of RSSI. Such optimal tag placement requirement may incur considerable deployment cost, hindering the ease of RFID deployment for practical applications. In addition, existing works use passive tags mixing with active tags, in which active tags need more expenditure in real deployment. They use more advanced signal measurement of RFID like CSI [21], [22], however, not many RFID manufacturers can provide such hardware-level support. Obtaining RSSI is still the easiest way to exploit RFID signal since it can be easily obtained from off-the-shelf RFID devices, which result in more cost-effective solutions. However, RSSI signal suffers from high uncertainties due to the nature of signal fluctuation in real-world conditions such as distraction, diffusion and degradation, and being noise-sensitive. Hence, it is particularly challenging when dealing with fine-grained activities [18].

To overcome signal uncertainty, we exploit sparse representation over RSSI, and study how to learn signal strength fluctuation to improve system robustness and effectiveness. Sparse coding is a common technique to model data vectors as sparse linear combinations (i.e., *sparse representation*) of basis elements, and has been widely used in image processing and computer vision applications [23], [24], [25]. Prior work on classification using sparse representation has mainly dealt with images. There are few work on sparse representation for activity recognition by exploring signal strength fluctuation due to its high uncertainty in terms of physical and deployment in real world. We propose a dictionary learning approach to uncover the structural information between RSSI signals of different activities by learning the compact and discriminative dictionaries per activity. In particular, we model each predefined human activity by learning discriminative dictionaries and its corresponding sparse coefficients using features extracted and selected from raw RSSI streams. The obtained sparse coefficients are systematically examined as enhanced features to better discern different activities. To enhance the robustness, we further design a Canonical Correlation Analysis [26] (CCA)-based greedy feature selection approach to decipher the most informative features from noisy RSSI raw signals.

In this paper, we develop an RFID-based, device-free activity recognition system by leveraging off-the-shelf, *pure passive* RFID tags and exploiting easy-to-obtain RSSI signal. Fig. 1 illustrates the system setup and gives a high-level overview. Passive RFID tags are deployed in an environment (e.g., on the wall in a room) forming a tag array. We design our system in a way that it is insensitive to tag placement such as distance between tags (Section 4.6), lowering the bar for system deployment and making it a more practical solution. We conduct extensive real-world experiments by comparing our system to the state-of-the-arts, and discovering the system bottleneck. The results demonstrate that our system achieves robust performance ($\sim$70 percent accuracy for 12 daily activities in person-independent validation strategy and $\sim$95 percent in person-dependent validation strategy). Previous studies [20] show that tag density has a great impact on system performance. We conduct empirical studies on tag arrangement such as distance between tags. The results show that our system allows arbitrary tag arrangement within a specified distance without significant negative effect on system performance (Section 4.6), alleviating nontrivial tag configuration in real deployment. In general, our system offers several advantages such as easy to deploy, maintenance free, low cost, and lightweight in computational cost. The main contributions of our work are summarized below:

- We develop a compressive sensing dictionary-based learning approach to uncover structural information among RFID signals of different activities. Compared to existing approaches, our approach achieves more compact representation of activities while preserving richer information and uncovering invariant patterns, thereby underpinning an efficient and robust activity recognition system. We show that, even using noisy and uncertain RSSI signals, our algorithm still achieves good performance in terms of both person-independent and person-dependent activities.
- We propose a lightweight but effective feature selection method to assist the extraction of more discriminative signal patterns from noisy RFID streams. We particularly exploit an unsupervised and filter-based feature selection approach based on CCA, which not only retains the natural assignment of feature components, but also uncovers the interdependency between feature components.
- We validate and evaluate our system through prototype applications and conduct extensive experiments

in both office and home settings. Our experimental results demonstrate the effectiveness and efficiency of the proposed techniques.

The remainder of the paper is organized as follows. We present the motivating applications and formulate our research problems in Section 2. The proposed approach and technical details are described in Section 3. In Section 4, we report the experimental results. We overview the related work in Section 5 and wrap up the paper in Section 6.

2 Background

In this section, we first present several representative applications that can be benefited from our device-free activity recognition system, and then discuss two important observations, which hold key groundings for the proposed recognition algorithms.

2.1 Motivating Applications

Fall Detection. With the great progress of medical technologies, many countries are facing the issue of *aging society* where there will be a lower proportion of people providing necessary levels of care to a large portion of elderly people. Meanwhile, the problem of huge nursing cost has a big impact to aged care. The demand for home surveillance systems is rising, and such systems help elderly people stay at their own homes longer and safer, which reduces the necessity for caregivers to oversee individuals (especially seniors).

In particular, falls are the leading cause of fatal injuries for people aged 65 and above [27]. By monitoring the activities of an elderly, we could detect the likely falls (e.g., getting out of bed, going to bathroom), and issue an alert timely. Obviously, it is impractical to require the older people to carry devices all the time.

Ambulatory Monitoring. Posture recognition and monitoring are critical in the medical care, e.g., ambulatory monitoring, because physiological responses, such as changes in heart rate or blood pressure, may result from changes in body position and physical activities [28]. Continuous monitoring and automatic detection of subtle behavioral changes are valuable for physicians and caregivers to estimate the physical well-being of a person.

Sleep Monitoring. Sleep posture recognition is crucial for elderly people as sleep disorders can be associated with some particular diseases, e.g., restless leg syndrome and diabetes [29]. Device-free activity monitoring is an improvement and good supplementary over camera-based monitoring, which suffers from privacy issues and poor performance at low-light conditions.

2.2 Observations and Problem Formulation

To gain better understandings of the groundings about this proposed work, we present two observations from the RFID RSSI data we collected: i) RSSI signal variations are hard to fit in a straightforward way; and ii) there exist invariant underlying patterns of RSSI signal variations which can be explored to design a learning algorithm for identifying different activities.

Observation 1. It is well known that RSSI is quite complicated in real environments due to signal reflection, diffraction, and scattering, especially for the passive RFID tags. It is often severely affected by the propagation environment, the tagged object properties, or human movements in the signal coverage area. Moreover, the signal strength of a passive RFID tag is uncertain and non-linear [18], [30]. As shown in Fig. 2a, the RSSI variations cannot be easily fitted using generic linear and polynomial regressions since the fitting residuals are quite large. It is therefore impossible to directly use raw RSSI signal in activity recognition.

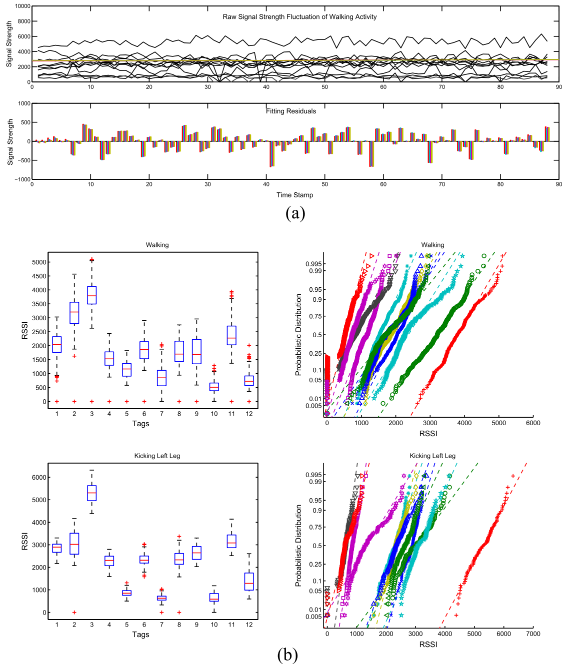


Fig. 2. (a) Signal strength fluctuation of the activity *walking* and its corresponding linear/quadratic/cubic/polynomial fittings and residuals. (b) The signal distribution pattern of activities *walking* (top) and *kicking left leg* (bottom).

Observation 2. Although RSSI reflects more on the uncertainty and non-linear distributed patterns, we can still observe some interesting characteristics of RSSI. More specifically, we discover that the variations of signal strength reflect different patterns, which can be exploited to distinguish different activities. Fig. 2b shows the distinctive fluctuation patterns of signal strength collected from activities *walking* and *kicking left leg*, respectively. From the figure, it is clear that the distribution and accumulative probability of RSSI of these two activities are different and distinguishable.

From above observations, RSSIs of passive RFID tags embody certain patterns for different activities, which can be exploited for effective activity recognition. We therefore formulate our problem as follows.

Let $\mathcal{S} \subset \mathbb{R}^t$ (t is the number of tags) be the domain of observable signal strength fluctuation (RSSI indicator in this work) $\mathbf{s}$, and $\mathcal{L} \in \{1, \ldots, K\} \subset \mathbb{R}$ be the domain of output activity label l (K is the number of activities). Suppose we have n RSSI and activity label pairs $\{(\mathbf{s}_i, l_i) | \mathbf{s}_i \in \mathcal{S}, l_i \in \mathcal{L}, i = 1, \ldots, n\}$. The training dataset can be expressed as

$$\begin{aligned} \mathbf{S} &= [\mathbf{s}_1, \ldots, \mathbf{s}_n] \in \mathbb{R}^{t \times n} \\ \mathbf{l} &= [l_1, \ldots, l_n]^T \in \mathbb{R}^n. \end{aligned} \tag{1}$$

Our goal is to learn a predictor $\mathcal{F} : \mathcal{S} \rightarrow \mathcal{L}$ using the training dataset, to assign the most appropriate activity label for a given query sample.

3 The Proposed System

The overall architecture of our proposed system is shown in Fig. 3. The whole process consists of three main stages:

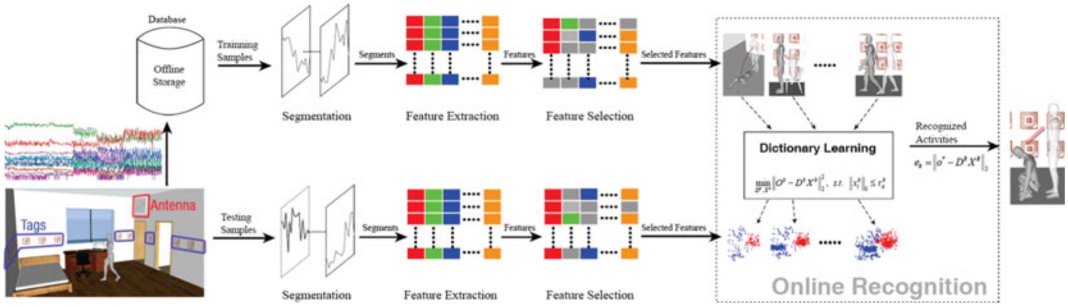


Fig. 3. The architecture of our proposed activity recognition system.

- Processing the noisy raw signal streaming data from various RFID tag inputs into individual segments, and then extracting low-level statistical features from each segment and the salient subset of features are selected (Section 3.1),
- Learning a compact and discriminative dictionary for each activity using the selected features (Section 3.2), and
- Given a new streaming signal, the activity recognition problem is equivalent to finding the dictionary from the learned activity dictionaries that best approximates the testing sample (Section 3.3).

We will describe the technical details of these stages and the algorithms in the rest of this section.

3.1 Feature Representation

The first major task is to divide the continuous sequence of RSSI data stream into a set of individual segments, where each segment corresponds to a specific concept or an activity (e.g., one segment corresponds to *Sitting*, and another segment corresponds to *Standing* etc). Segmentation helps the classifier better understand the underlying activity, by illustrating the temporal dependency, and to compress the streaming data as well.

We incorporate the temporal information during the segmentation process of feature transformation. We divide the raw streaming signal data into segments where each segment is generated by a sliding window based method. So all relevant information can be extracted as features from each single segment.

The continuous $\mathbf{S}$ will be divided into a set of individual segments with equal size $\mathbf{S} = \{\mathbf{S}_1, \ldots, \mathbf{S}_n\}$. We set segment size to 6 in this work. It is well understood that high quality and discriminative features are essential to improve the classification accuracy of any pattern recognition system. After dividing the streaming segments, the information is then transformed by designing 7 types of lightweight statistical feature vectors from each segment, and they are listed in Table 1.

The extraction process in our approach yields a total of m feature vectors $\mathbf{O} = \{\mathbf{o}_1, \ldots, \mathbf{o}_i\}$, where $\mathbf{o} \in \mathbb{R}^m$, with $m = 7 \times t$ where t is the number of tags. However, some features might confuse, rather than help, the classifier to discriminate activities. Also, due to the "curse of dimensionality", the performance may degrade sharply as more features are used when there is not enough training data to reliably learn all the parameters of the activity models. In general, to achieve the best classification performance, the dimensionality of the feature vector should be as small as possible, namely keeping only the most salient and complementary features.

Fig. 4 shows the correlations between features in a 2D space. Fig. 4a shows that although the Mean feature can roughly characterize the *walking* activity from the other four activities, the two features (Max and Mean) cannot well separate the five activities. In Fig. 4b, the Variance feature can help to identify activities *high arm waving (horizontal)* and *bending over*, but these two activities cannot be characterized well along the Min feature due to intersections and the overlapping. In such cases, the Min feature is irrelevant or redundant, and does not provide useful information to improve the classification accuracy. In addition, keeping the dimensionality small could reduce the computational cost such that the recognition algorithms can be implemented and performed on lightweight devices such as mobile phones. Besides, smaller and discriminative feature sets can decrease the latency of recognition system, which is a main concern in activity recognition applications.

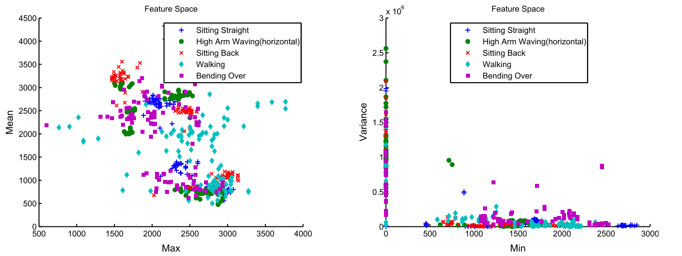
To systematically assess the usefulness and identify the most important features for discriminating different activities, feature selection techniques are needed. In particular, we propose a filter-based unsupervised feature selection method. Compared to the existing feature selection approaches, which treat each component of features independently, we study the correlations between features using Canonical Correlation Analysis [26]. We compute the canonical correlation for each pair of features and generate feature subsets using a greedy algorithm based on computed pairwise canonical correlations.

CCA Ranking. The initial rankings for each pair of features, where two feature vectors are given, and a projection is computed such that they are maximally correlated in the dimensionality-reduced space. We first apply CCA to all pairs of the extracted features. The result is a similarity matrix of canonical correlations. For each pair of feature vectors $\{\mathbf{o}_i, \mathbf{o}_j\}$ that can be linearly mapped into: $\mathbf{o}_i \rightarrow \mathbf{w}_{\mathbf{o}_i}^T \mathbf{o}_i$ and $\mathbf{o}_j \rightarrow \mathbf{w}_{\mathbf{o}_j}^T \mathbf{o}_j$, where $\mathbf{w}_{\mathbf{o}_i} \in \mathbb{R}^m$ and $\mathbf{w}_{\mathbf{o}_j} \in \mathbb{R}^m$, their correlation coefficient ρ_{ij} can be obtained by maximizing the following equation:

$$\rho_{ij} = \frac{\mathbf{w}_{\mathbf{o}_i}^T \mathbf{o}_i \mathbf{o}_j^T \mathbf{w}_{\mathbf{o}_j}}{\sqrt{\mathbf{w}_{\mathbf{o}_i}^T \mathbf{o}_i \mathbf{o}_i^T \mathbf{w}_{\mathbf{o}_i}} \sqrt{\mathbf{w}_{\mathbf{o}_j}^T \mathbf{o}_j \mathbf{o}_j^T \mathbf{w}_{\mathbf{o}_j}}}. \tag{2}$$

TABLE 1
Statistical Features and Brief Descriptions

No.	Feature	Description
1	Min	Minimal value of $\mathbf{S}_i$
2	Max	Maximal value of $\mathbf{S}_i$
3	Mean	Average value of $\mathbf{S}_i$
4	Variance	The square of standard deviation of $\mathbf{S}_i$
5	Root Mean Square	The quadratic mean value of $\mathbf{S}_i$
6	Standard Deviation	Measure of the spreadness of $\mathbf{S}_i$
7	Median	The median $\mathbf{S}_i$



(a) Max vs. Mean (b) Min vs. Variance

Fig. 4. Illustrative examples of feature correlations in 2D space.

After applying to all pairs of features, we can generate an initial ranking for all feature pairs. A higher rank is assigned to those weakly correlated and thus complementary feature pairs. Strongly correlated and thus redundant feature pairs get lower ranks. The initial ranking facilitates the selection of descriptive and complementary features.

Algorithm 1. Activity-Specific Dictionary Learning

Input: Training sample matrix $\mathbf{O} = \{\mathbf{o}_1, \ldots, \mathbf{o}_N\}$, dictionary size d

Output: Dictionary $\mathbf{D}$ and sparse coefficients $\mathbf{X}$

Initialize: Dictionary matrix $\mathbf{D}^{(0)} \in \mathbf{R}^{m \times K}$ with ℓ_2 column normalization and $J = 1$

while (!= stopping criteria) **do**

Use orthogonal matching pursuit to compute the sparse coefficients $\mathbf{x}_i$ for each training sample $\mathbf{o}_i$ by solving the optimization problem.

$$\min_{\mathbf{D},\mathbf{x}_i} ||\mathbf{o}_i - \mathbf{D}\mathbf{x}_i||_2^2, \quad \text{s.t.} \quad ||\mathbf{x}_i||_0 \leq \tau_o \tag{3}$$

Update $\mathbf{d}_j$, the jth column of $\mathbf{D}^{J-1}$

for $j = 1 : N$ **do**

Find a group of vectors:

$$\xi_j \leftarrow \{i : 1 \leq i \leq N, \mathbf{x}_i(j) \neq= 0\} \tag{4}$$

Compute the overall representation error matrix $\mathbf{E}_j$ by:

$$\mathbf{E}_j \leftarrow [\mathbf{o}_|, \ldots, |\mathbf{o}_N] - \sum_{i \neq j} \mathbf{d}_i \mathbf{x}_\tau^i \tag{5}$$

Extract the ith column in $\mathbf{E}_j$ where $i \in \xi_j$ to form $\mathbf{E}_j^R$

Apply SVD to obtain $\mathbf{E}_j^R = \mathbf{U}\Delta\mathbf{V}$, and $\mathbf{d}_i$ is updated with the first column of $\mathbf{U}$. The non-zero elements in $\mathbf{x}_\tau^i$ are updated with the first column of $\mathbf{V} \times \Delta(1, 1)$

end for

$J = J + 1$

end while

Forward Searching. We apply a simple greedy method to find a feature subset based on their pairwise rankings, which traverses the full search space provided by the initial ranking of canonical correlation coefficients of the feature pairs. Forward selection refers to a search that begins at the empty set of features and the features are progressively incorporated into larger and larger subsets. Then, we use the classification performance to evaluate the new feature combinations, and the searching process will be terminated when either the predefined dimensionality of features is reached or all features are already considered.

3.2 Activity Dictionary Learning

A well learned dictionary by fitting overcomplete basis with a collection of training samples can generate more compact and informative representation of given data, thus it helps to achieve better recognition performance. We propose a sparse representation based approach to recognize human activity by investigating RSSI fluctuations. We learn one single dictionary for each activity, which is formed by a set of basis vectors learned by solving a sparse optimization problem. Each basis vector can effectively capture part of the key structural information of the training data from each activity.

There are several advantages in learning activity dictionaries. First, the dictionary for each activity is learned from a collection of training samples via solving a ℓ_1-norm optimization problem [31]. Second, the dictionary learning and training process of each activity is independent from other activities, which makes an activity recognition system flexible and scalable, as no change is needed on the existing activity dictionaries when a new activity is added. Finally, each dictionary can be trained and learned by using only very small number of training samples, which can effectively relax the heavy workload on labeling and annotating training data in activity recognition, as required by the most existing approaches.

Assuming there are K types of activities, we construct K dictionaries (one dictionary for each activity). After that, a new signal is evaluated using the K dictionaries to find the most appropriate activity label. We present the details of the proposed algorithm in the following.

Let $\mathbf{O}^k = \{\mathbf{o}_1^k, \mathbf{o}_2^k, \ldots, \mathbf{o}_i^k\}$ be the training sample from activity class $\mathcal{C}^k$. To learn and encode the information of the testing samples belonging to a particular activity class, we first construct an overcomplete dictionary $\mathcal{D}^k$ for each class $\mathcal{C}^k$. Recall the set of training samples from kth activity as $\mathbf{O}^k = \{\mathbf{o}_1^k, \mathbf{o}_2^k, \ldots, \mathbf{o}_N^k\}$, where $\mathbf{o}_i^k \in \mathbb{R}^m$, m is the feature dimensions. We intend to learn a dictionary matrix $\mathbf{D}^k \in \mathbb{R}^{m \times K}$ (which equals to $K(K > m)$ vectors $\{\mathbf{d}_1^k, \ldots, \mathbf{d}_K^k\}$), over which $\mathbf{O}^k$ has a sparse representation $\mathbf{X}^k = \{\mathbf{x}_1^k, \ldots, \mathbf{x}_N^k\}$. In this case, the original training matrix $\mathbf{O}^k$ can be represented as a linear combination of no more than $\tau_0^k(\tau_0^k << K)$ dictionary vectors. The optimization problem can be formalized as

$$\min_{\mathbf{D},\mathbf{X}} ||\mathbf{O} - \mathbf{D}\mathbf{X}||_2^2, \quad \text{s.t.} \quad ||\mathbf{x}_i||_0 \leq \tau_o. \tag{6}$$

We adopt the K-SVD algorithm [32] to solve this problem, which performs two steps iteratively until converged. The first stage is the sparse coding stage, where $\mathbf{D}$ is kept fixed and the coefficient matrix $\mathbf{X}$ is computed by orthogonal matching pursuit algorithm. In the second stage, the dictionary $\mathbf{D}$ is updated sequentially by allowing the relevant coefficients to be unique to K-SVD, which results in a faster convergence. The dictionary learning algorithm is detailed in Algorithm 1. The complexity is proportional to $N(d^2K + 2mK)$.

3.3 Exploiting Dictionary Coefficients

One advantage of having class-specific dictionaries is that each class is modeled independently from the others, and hence the painful repetition of the training process can be avoided when a new type of activity is added into the system.

After learning K individual activity-specific dictionaries, any new incoming test RFID signal can be represented in terms of its dictionary basis from the learned dictionaries. To calculate the sparse coefficients of an input RFID sample $w.r.t.$ a given dictionary, we use orthogonal matching pursuit [33] to project the testing RFID sample on the subspace spanned by the dictionary basis, in which strong correlates with the signal or its residual are selected and used to calculate the coefficients.

Given multiple activity-specific dictionaries and coefficients obtained in Algorithm 1, a series of different methods can be leveraged to classify a new test signal over these dictionary basis. We propose several strategies of exploiting the learned sparse coefficients, which are detailed as follows:

- Reconstruction error (RE). The reconstruction error for the kth activity ($k \in [1, K]$) can be calculated as

$$e_k = ||\mathbf{o}^* - \mathbf{D}^k\mathbf{X}^k||_2. \tag{7}$$

  Then the activity label of $\mathbf{o}^*$ can be assigned using

$$l_{\mathbf{o}^*} = l(\arg\min_k e_k). \tag{8}$$

- Maximal coefficient (MC). The activity label is associated with the training samples having the largest absolute value of coefficients of $\mathbf{X}^k$

$$l_{\mathbf{o}^*} = l(\arg\max_k d_k^i). \tag{9}$$

- Maximal mean of coefficients (MMC). The activity label is the top label with the maximal sum of coefficients of $\mathbf{X}^k$ divided by dimension of $\mathbf{o}^*$ m

$$l_{\mathbf{o}^*} = l\Bigg(\arg\max\Bigg(\sum_i \mathbf{d}_k^i/m\Bigg). \tag{10}$$

- Maximal sum of coefficients (MSC). The activity label is the top label with the maximal sum of absolute value of coefficients of $\mathbf{X}^k$

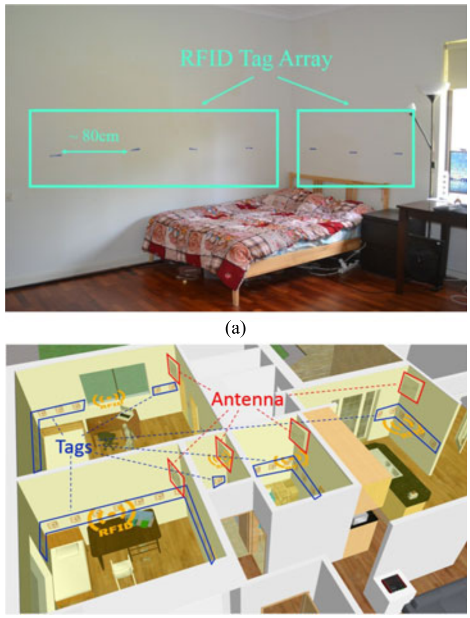
$$l_{\mathbf{o}^*} = l\Bigg(\arg\max\Bigg(\sum_i |d|_k^i\Bigg)\Bigg). \tag{11}$$

- Concatenate coefficients (ConSVM). We stack the learned coefficients with original features to form a new feature vector, and then feed the enhanced features into SVM for classification.

Our proposed activity classification is summarized in Algorithm 2.

Algorithm 2. Overall Algorithm for Activity Classification

Input: Sensor samples $\mathcal{S} = \mathbf{S}_1, \ldots, \mathbf{S}_K$, where K is the number of activity classes; Querying signal samples $\mathbf{S}^* = \{\mathbf{s}_1^*, \ldots, \mathbf{s}_I^*\}$
Output: Activity label $\mathbf{l}^* = \{l_1^*, l_i^*, \ldots, l_I^*\}$ of $\mathbf{S}^*$
Extracting N^k feature vectors of signal samples from each activity class C^k using the proposed feature representation Section 3.1
Constructing K activity-specific dictionaries $\mathbf{D} = \{\mathbf{D}^1, \ldots, \mathbf{D}^K\}$
while $i! = I$ **do**
  Transform $\mathbf{S}^*$ to features $\mathbf{O}^*$ (Section 3.1)
  Computing sparse representation x_i^* of $\mathbf{s}_i^*$ using K dictionaries $\mathbf{D}$ (Section 3.2)
  Outputting activity label by exploiting coefficients $\mathbf{D}^k$ (Section 3.3).
end while



(a)

(b)

Fig. 5. Experimental setup: (a) bedroom setup and (b) whole house setup.

4 Experiments

In this section, we first briefly introduce the experimental settings including hardware setup, tag placement, and data acquisition. We then report our extensive experiment studies on the proposed approach. Our experiments are intended to address the following questions: i) how does our proposed approach compare with other state-of-art methods? ii) what are the optimal settings of our proposed method? iii) how does proposed feature selection affect the activity recognition performance? Sections 4.2, 4.3, and 4.4 devnote to these three questions correspondingly. We also brief analyze the recognition delay of our proposed approach in Section 4.5 and investigate the sensitivity of our approach to indoor environments in Section 4.6.

4.1 Experimental Settings

Hardware Setup. We used one Alien 9,900+ RFID reader, one circular antenna and Squig inlay passive RFID tags in our experiments. Passive tags were placed on the wall with certain distance. The antenna is ~ 1.3 m high from the ground, arranged in an angle of ~ 70 degree to ensure it can catch all the tags' signals. The subjects stood between the wall and the antenna ($\sim$ 1.5 to 1.8 m) and performed different predefined activities. A sequence of RSSI signals were collected at a sampling rate of 0.5 second. The overall set up is shown in Fig. 5.

It may be a concern that RFID-based activity recognition systems may pose a potential risk to people's health. Commercial RFID readers and tags operate at electromagnetic frequencies in the low-energy range, effectively eliminating the risk of interaction with human cells. Furthermore, a passive tag itself has no baseline electromagnetic activity and only produces a signal in response to the interrogation from an RFID reader. The tags themselves even have been approved for implantation in humans and have shown no negative health effects [34].

Sampling Rate. Passive RFID tags tend to be noisy. For example, one of the challenges in existing RFID systems is false negative readings, caused by missed detections (i.e., a tag in the antenna's reading range may not be detected).

Meanwhile, RSSI data is sensitive to environments, e.g., some disturbance from an environment can cause RSSI fluctuations. Appropriate sampling rates can reduce the aforementioned problems. However, too small sampling rates make our method more sensitive to the noise of RFID readings, while too big sampling rates blur the inter-class activity boundaries. In our implementation, we collected the continuous RSSI data streams every ≈ 0.5 second.*Data Collection.* The data acquisition process involves six subjects (five males and one female), and the set of 23 fine-grained, orientation-sensitive activities (including 6 postures and 17 actions, as shown in Fig. 11). These activities are the most common ones in people's daily lives. Identifying orientation of postures can be valuable when combined with the layout of the place in practice. For instance, if we know that a table is on the left side of an elderly person, based on the layout, when the orientation of a fall is detected, it is possible to estimate how severe the fall would be (e.g., she may hit the table if she is falling to her left). In our experimental study, each subject performed each activity for 120 seconds and all 23 different activities performed sequentially by one subject were regarded as one set of activity data. We make the dataset publicly available[1] for reproducing our results and support other researchers in the area.

Validation Strategy. We validate our approach using two strategies: *person-dependent* and *person-independent*. The former uses the partial samples of each subject for testing and the remaining samples of the same participant are used for training. The final result is the averaged value of all subjects. This is reasonable since elderly people often live alone. The latter applies the *one subject out* strategy, where we use the data from five subjects as training examples to train our algorithm and build activity recognition models. The data from the left-out subject is used for testing. This process iterates for every subject. The final result is the averaged value across all the subjects.

Performance Metrics. Instead of using the overall classification accuracy, we evaluate our proposed approaches using F1 score, which is a harmonic mean of precision and recall scores

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}. \tag{12}$$

4.2 Overall Comparison

In this section, we report our experiments that focus on two aspects: i) the performance comparison of a set of sparse representation dictionary-based methods, and ii) the performance comparison of our proposed method with other five widely used generic classifiers in sensor-based activity recognition.

Dictionary-Based Approaches Comparison. We first compared the series of proposed dictionary-based recognition methods with varying number of selected features (Section 3.3) to discern the best strategy out of them before conducting comparable experiments with other methods. Figs. 6a and 6b show the results. From the figure, we can see that the reconstruction error based method produces the worst performance in both person-dependent and person-independent scenarios, with only less than 60 percent in accuracy. The other four methods demonstrate the similar performance, all of which can obtain nearly 96 percent F1-score under person-dependent and over 60 percent F1-score under person-independent strategy. We selected MC as the best strategy as it performs the best and shows stable recognition among all proposed dictionary based approaches compared to MCC and MSE, and its light computational cost compared to ConSVM, which requires a full spectrum of stacked coefficients as extra features. We used MC in the following experiments.

Comparison with State-of-the-Arts. To evaluate the performance of activity recognition, we further compared MC with a few state-of-the-art methods widely used in the activity recognition community such as Multinomial Logistic Regression with ℓ_1 (MLGL1), SVM with linear kernel (LSVM), k nearest neighbor (kNN), random forest (RF), and Naive Bayes (NB). We selected these methods since they have already been successfully applied for sensor-based activity recognition applications in the recent literature.

- Multinomial Logistic Regression with ℓ_1 (*MLGL1*) is a modification of linear regression that is able to predict dependent variables based on the logistic function. Multinomial (or multivariate) computations are solved by the decomposition into a series of binary variables. In this work, we integrated the ℓ_1 regularization into linear classifier in the objective term. Given our multi-class posture recognition problem, we combined the ℓ_1 regularization with multinomial logistic regression, which models the conditional probability $P_{\mathbf{w}}(l_j = \mp 1|\mathbf{o})$. The prime problem with ℓ_1 regularization can be calculated by optimizing the log likelihood

$$\begin{aligned} l_k = \arg\min_{\mathbf{w}} \sum_{k=1}^{K} ||\mathbf{w}_k||_1 - \sum_{i=1}^{n} \sum_{k=1}^{K} l_{ik} \mathbf{w}_k^T \mathbf{o}_i \\ + \sum_{i=1}^{n} \log \Big(\sum_{k=1}^{K} \exp(\mathbf{w}_k^T \mathbf{o}_i) \Big) \end{aligned}. \tag{13}$$

- k-Nearest Neighbor (k**NN**) is a common classifier for a variety of classification problems. It predicts the class of a sample by a majority voting of the class labels of the K nearest training instances. We set $k = 3$.
- Linear Support Vector Machine (*LSVM*) aims at finding the best separation of binary-labeled instances by determining a hyperplane which maximizes the margin between support vectors of different classes. We set $C = 1$.
- Random Forest (*RF*) builds a forest of decision trees that have the same distribution but independent output classes. It is based on a random selection of features for each tree and construction of a combination of the individual tree outputs. We set the number of trees as 1,000.
- Naive Bayes (*NB*) classifier finds the most posterior probability $Pr(l_k|\mathbf{o}^*)$ for a given testing RSSI sample $\mathbf{o}^*$ as its predicted activity label $l(\mathbf{o}^*)$

$$\begin{aligned} l_k &= \arg\max_{l_k} \frac{Pr(l_k)\prod_j^D Pr(\mathbf{o}_j^*|l_k)}{\sum_j^k Pr(l_k)\prod_j^D Pr(\mathbf{o}_j^*|l_k)} \\ &= \arg\max_{l_k} Pr(l_k) \prod_j^D Pr(\mathbf{o}_j^*|l_k), \end{aligned} \tag{14}$$

where prior probability $Pr(l_j)$ is proportional to the size of training samples in each posture class, which is obtained via dividing number of samples belonging to posture l_j by total number of training samples,

1. http://linayao.com/data/rssi-activity.zip

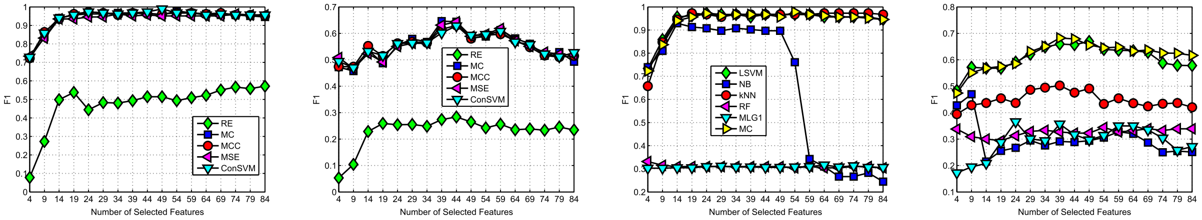


Fig. 6. (a) Dictionary-based methods comparison under person dependent validation. (b) Dictionary-based methods comparison under person independent validation. (c) Comparison with other methods under person dependent validation. (d) Comparison with other methods under person independent validation (the legend of (d) is same as (c)).

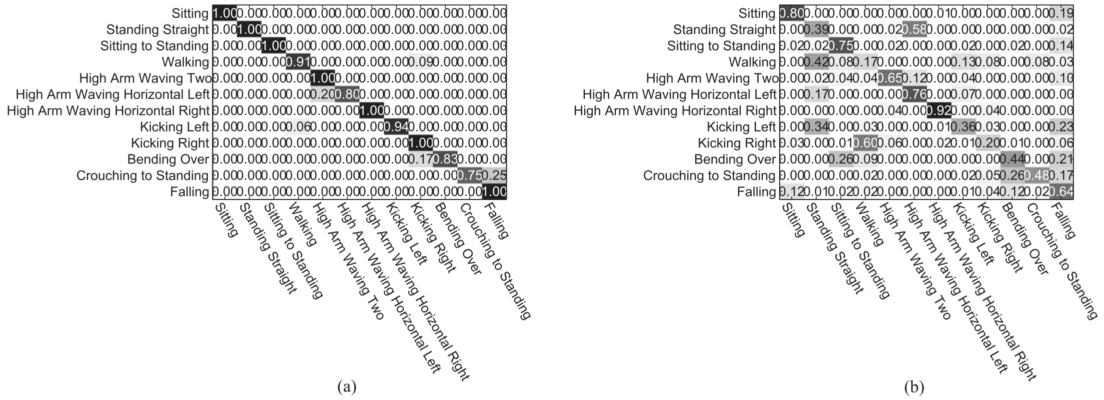


Fig. 7. (a) Confusion matrix under person dependent validation. (b) Confusion matrix under person independent validation.

i.e., $Pr(l_j) = \frac{|\mathbf{o}_{l_j}|}{|\mathcal{O}|}$. K is the number of posture classes. Conditional probability on each dimensional RSSI $Pr(\mathbf{o}_j^*|l_k) \sim N(\mu_j, \sigma_j^2)$ can be obtained from the training dataset.

We can clearly draw the following observations from the results shown in Figs. 6c and 6d:

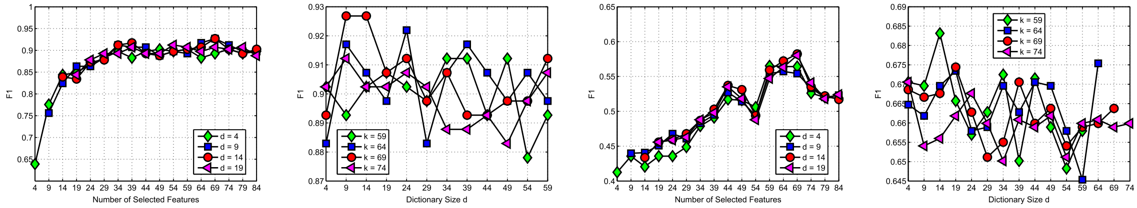
- The performance of all the methods are gradually improved when more features are selected, and the improvement are not obviously when a certain number of features are selected. Specially, the performance may decrease with larger size of the feature size.
- Our dictionary-based method *MC* consistently outperforms all the state-of-the-arts. LSVM shows the comparable performance with MC under person-dependent validation, but MC's recognition accuracy is more competitive compared with LSVM. This result shows the better ability of our proposed method in dealing with the intra-class variability across different persons.

Taking a closer look at the accuracy for each activity in terms of person-dependent and person-independent validations from Fig. 7, we can see that the results under person dependent validation shows robust outcomes, where almost all the activities can be correctly identified. For the person-independent validation, our method can distinguish *sitting*, *sitting to standing* and *arm weaving* with reasonable good performance, also can recognize *falling* with over 60 percent accuracy across different persons. But it fails to identify some lower body activities (e.g., *kicking*) and also confuses *walking* with *sitting to standing* due to failure to capture the activity signatures of different persons. The possible reasons might lie in that i) from the methodology perspective, intra-class variability is still a big challenge for activity recognition community, and more informative and discriminative patterns discerning different persons should be developed from the RSSI fluctuations, and ii) from the hardware setting perspective, according to our preliminary research on tag placement related to tag density, single-line tag placement is capable of capturing signal variations, but it may fail to detect fine-grained body movements, such as sitting leaning right or left. Furthermore, it is also hard to capture the signal variations caused by subjects with different heights.

To achieve better accuracy and higher sensitivity, we tried to increase single-line tag placement to multiple lines, eventually forming an *array*. Different lines correspond to different parts of human body. For instance, the upper line of tags would be expected to reflect the variations from upper human body like waving arms or shaking head; the middle line of tags would be more sensitive to movements of torso; and the bottom line of tags are supposed to have more response to lower body movements such as falling. In this way, we may perform more robust activity recognition with the collected *full spectrum* of RSSI variations. More technical details can refer to our previous work in [30].

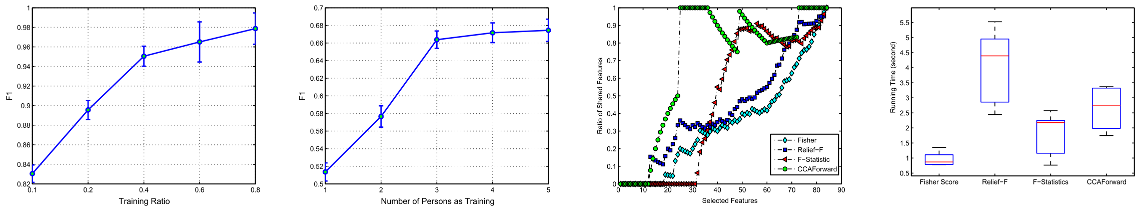
4.3 Parameter Tuning

Impact of Selected Feature Size k. The top k features control how many of the top effective features are used to feed into the classification algorithm. We varied the value of k from 4 to 84 (full feature set) with 5 stepsize under some fixed dictionary size d. The results are shown in Figs. 8a and 8c. For all the feature selection tests, we kept the dictionary size fixed. The result shows that in most cases feature selection improves classification performance in comparison with the full feature set. We also can observe that the increasing number of features selected, the performance increases as well until k reaches around 64 in both person-dependent and person-independent validations, at that point our classification algorithm performs the best. After that, the performance shows some slight degradation, especially for the person-independent scenario.



(a) (b) (c) (d)

Fig. 8. Experiments on parameter tuning: (a) Impact of different feature size k with a set of fixed dictionary size d under person dependent validation, (b) impact of different dictionary size d with a set of fixed feature size k under person independent validation, (c) impact of different feature size k with a set of fixed dictionary size d under person independent validation, and (d) impact of different dictionary size d with a set of fixed feature size k under person independent invalidation.



(a) (b) (c) (d)

Fig. 9. (a) Impact of training ratio under the person dependent validation, (b) impact of number of persons as training data under the person independent validation, (c) robustness evaluation, and (d) running time comparison.

Impact of Dictionary Size d. The activity dictionary is an overcomplete set of vectors and the number of vectors indicates the size of the dictionary. Similar to the experiment of studying the impact of k, we varied the dictionary size d from 4 to 59 with a set of fixed feature size k. The results are reported in Figs. 8b and 8d. From the figure we can see that the classification performance reaches the highest at a certain point (e.g., $d = 9$ for Fig. 8b). After that, the performance stays stable and even slightly decreases when d gets larger.

Impact of Training Size. The third important factor of affecting the activity recognition performance is how much training data should be involved in our proposed method. We conducted the evaluation with fixed $k = 69$ and $d = 9$ by varying the training ratio of the whole dataset from 0.1 to 0.2 with stepsize 0.2 for the person-dependent scenario. The results are shown in Fig. 9a. We can observe that only using 10 percent samples for the training, our proposed method reaches over 80 percent accuracy, and it reaches over 90 percent with only 20 percent data as the training data. The performance keeps improving along with more training samples. In our experiments, we set 0.2 as our default training percentage. Fig. 9b shows the result under the person-independent scenario, where we used p $(p = 1, 2, 3, 4, 5)$ person data for training and 1 person's data for testing. The performance keeps increasing from over 50 percent with 1 person's data as the training data, and stabilizes over 66 percent when we used 3 persons' data as training data. The improvement is not quite significant after that, thus we set it as our default setting under person-independent validation.

4.4 Comparison on Feature Selection

In this experiment, we evaluated our proposed CCA based feature selection method with three widely adopted feature selection methods in terms of efficiency (e.g., running time) and effectiveness (e.g., precision/recall/F1). Specially, we compared the proposed CCA-based forward selection with *fisher score*, *sequential forward with relief-f score* and *forward selection with F-statistics score* based methods.

- Fisher Score. It is for quantifying the score of ith feature $\mathbf{o}_i$

$$S_i = \frac{\sum_{k=1}^{K} n_k (\bar{\mathbf{o}}_{ik} - \bar{\mathbf{o}}_i)^T (\bar{\mathbf{o}}_{ik} - \bar{\mathbf{o}}_i)}{\sum_{k=1}^{K} n_k v_{ik}}, \tag{15}$$

  where n_k is the number of samples in the kth activity class, $\bar{\mathbf{o}}_{ik}$ and v_{ik} are the mean and the variance of the ith feature, and $\bar{\mathbf{o}}_i$ is the mean of the ith feature.
- Sequential Forward with Relief-F Score (SFRF). This technique estimates the relevance of features according to how well their values distinguish between the data points of the same and different activity classes that are close to each other. It computes a weight for each feature to quantify its merit. Its weight is updated for the signal samples presented in each activity class, according to the evaluation function

$$\begin{aligned} S_i = S_i + \sum_{j \in \mathcal{L}, j \neq l(\mathbf{o}_i)} \frac{P(l_j)}{1 - P(l_j)} |\mathbf{o}_i - nearmiss_i^j(\mathbf{o}_i)| \\ - |\mathbf{o}_i - nearhit_i(\mathbf{o}_i)|, \end{aligned} \tag{16}$$

  where $nearmiss_j(\mathbf{o}_i)$ and $nearhit_i(\mathbf{o}_i)$ denote the nearest RSSI samples to $\mathbf{o}_i$ from the same and different activity classes, respectively.
- Forward Selection with F-Statistics Score (SFSS). This method measures the discrimination of multiple sets of real numbers, which can be calculated using

$$S_i = \frac{\sum_{j=1}^{l} (\bar{\mathbf{o}}_i^j - \bar{\mathbf{o}}_i)^2}{\sum_{j=1}^{l} \frac{1}{n_j - 1} \sum_{k=1}^{n_j} (\mathbf{o}_{k,i}^j - \bar{\mathbf{o}}_i^j)^2}, \tag{17}$$

  where n_j is the number of samples in the jth activity class, $\bar{\mathbf{o}}_i$ denotes the mean value of tag i in the training dataset, and $\bar{\mathbf{o}}_i^j$ is the mean value of the ith tag in the jth activity class. The numerator indicates the discrimination between positive and negative sets, and the denominator indicates the one within each of the two sets. The larger the F-score is, the more likely this feature is discriminative in the activity recognition.

A robust feature selection method should generate consistent feature selections for a given task and a given set of features independently from the input data. In order to investigate the robustness of the proposed feature selection method, we employed feature selection on all six subjects, and expected the evaluated feature selection methods to generate similar feature types and subsets. For all rank-based methods, we generated complete feature rankings from 1 to 7 features for CCA based feature selection and from 1 to all 84 feature components for F-statistics, Relief-F, Fisher. Rankings were generated for 10 cross validation runs. The robustness of feature selections was computed by the intersection of the generated rankings as follows:

$$I(i) = \frac{\cap_{j=1}^{m} F_j^p(i)}{i}, \tag{18}$$

where F^p is the set of corresponding features in the generated ranking list, $F^p(i) = \{f(1), \ldots, f(i)\}, i \leq n$, and n is the total number of features.

Fig. 9c shows the portion of shared features for all compared feature selection methods and all the target dimensions. We can clearly observe that the robustness of CCA based method is significantly higher compared with other generic selection methods. Besides, based on our previous study, our method reaches the best performance when the size of feature subset is around 69, at which point the ratio of shared features are also the highest.

We then compared the running time of the feature selection process since recognition delay is a critical concern for activity recognition applications. The running time can be split into the *feature ranking* time and the *feature set evaluation* time according to the nature of the algorithms. Our rank-based feature selections perform an initial ranking for the whole feature set and therefore have constant running time independent of the actual amount of features to be selected. Fig. 9d plots the running time of three classic feature selection based methods over the whole dataset. We observe that our proposed CCA-based feature selection has competitive performance in comparison to the other two feature selection based methods even though the fisher score method uses the least time.

4.5 System Latency Analysis

Fast detection of activities is critical, particularly for applications such as aged care. For example, we should send an alert to notify care givers as quickly as possible to offer medical assistance for the elderly people when a fall occurs.

Our system has about $4 \sim 4.5$ seconds recognition latency, which results from two main factors, namely i) data collection and ii) feature selection. The latency caused by feature selection can be referred to our previous experiment in Section 4.4, particularly Fig. 9d. The latency from data collection comes from two aspects. First, our system evaluates subject's postures every 0.5 second by using the RSSI stream of the latest two seconds. Second, the RSSI collector is programmed with a timer to poll RSSI with a predefined order of transmission, taking around one second to complete a new measurement with no workarounds.

4.6 Sensitivity to Indoor Environments

Activity recognition from the RSSI changes remains a challenging task in complex indoor environments due to the diffraction and reflection effects from furniture, layout and subject's differences on performing activities. In this section, we report some empirical results of several experiments regarding practical issues of our proposed system. These experimental studies aimed at

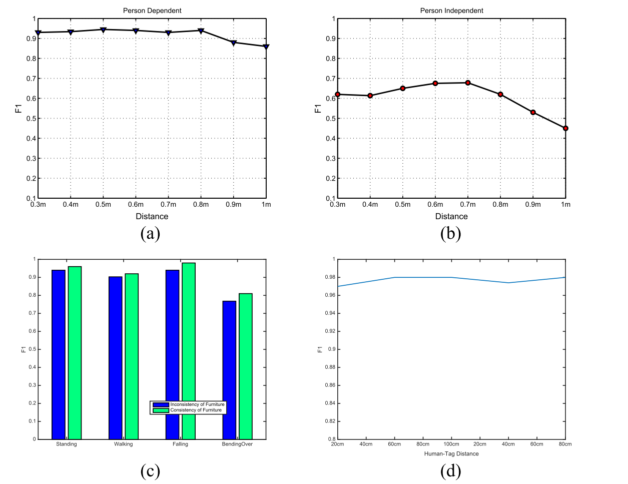


Fig. 10. Performance comparison of different tag density (distance between tags): (a) Person-dependent, (b) person-independent, (c) performance comparison on furniture changes, and (d) performance comparison on human-tag distance.

- Evaluating the effect of the distances between RFID tags on the system performance (Section 4.6.1);
- Examining the performance of the proposed system on object changes (e.g., moving a chair) (Section 4.6.2);
- Studying the effect of the distances between subject and tag array on the system performance (Section 4.6.3); and
- Investigating how well the system deals with orientation-sensitive activities (e.g., sitting leaning back or sitting leaning forward) (Section 4.6.4).

4.6.1 *Sensitivity of Tag Density*

We claim one of advantages of our proposed system is to relax the deployment from time-consuming and complex tag placement problem due to the robust feature selection and compact dictionaries. As long as the tags and the reader can form a signal field, the tags can be arbitrarily arranged without significant negative effect on the system performance. To show this declaration, we systematically studied the sensitivity of RFID tag density. Specifically, we varied the distance between two tags from 0.3 to 1 m and ran our system. In general, smaller distances between tags cause some high correlations and redundancy. Whilst, too sparse tag arrangements cannot capture useful and complete RSSI patterns. Interestingly, the results of our system (shown in Fig. 10) shows the insensitivity to the varying tag density, of which the results are more stable for person-dependent setting compared to the person-independent setting. This advantage makes our system more practical and attractive in real deployment.

4.6.2 *Sensitivity to Room Furniture Changes*

To evaluate the sensitivity of the proposed system to room furniture change, we conducted an experiment where subjects performed activities with and without a chair. In our experiment, we evaluated four activities: *Standing*, *Walking*, *Falling*, and *Bending Over*. Fig. 10c shows the results. From the figure we can observe that the recognition performance slightly drops, without any significant change. The results indicate that the furniture changes affect on the recognition accuracy, but not in a significant way. The degree of the effects might be dependent on the material, size and location of the object. More sophisticated investigations on this will be part of our future work.

4.6.3 Sensitivity to Human-Tag Distance

We varied the distance between the subject and the tags from 20 to 100 cm with 20 cm intervals due the spatial constraints of the testing area, and recorded the system recognition performance. From the result in Fig. 10d, we can clearly see that the performance generally remains stable in our experiment. The reason might lie in the fact that we used the industry level high-performing RFID reader that covers bigger area. The small changes in the centimeter level do not bring significant influence.

4.6.4 Sensitivity to Activity Orientations

We also explored the potential of identifying the *orientation-sensitive* activities (shown in Fig. 11). From the experimental results (see Fig. 12), we can see that the most errors happen when identifying the activities with a similar intra-class gap (e.g., *falling left* and *falling right*). From the results, we can see that our method can accurately recognize most of orientation sensitive activities in a cluttered indoor environment under person-dependent validation.

5 Related Work

The goal of activity recognition is to detect human physical activities from the data collected via various sensors [6], [35]. There are generally two main ways for activity recognition: i) to instrument people, where sensors and RFID tags are attached to human bodies, and ii) to instrument the environment, where sensors are deployed inside an environment.

5.1 Activity Recognition by Instrumenting People

Wearable sensors such as accelerometers and gyros are commonly used for recognizing activities [36]. For example, Kern et al. [37] design a network of three-axis accelerometers distributed over a user's body. The user's activities can then be inferred by learning the data provided by these accelerometers about the orientation and movement of the corresponding body parts. However, such approaches have obvious disadvantages including discomfort of wires attached to the body as well as the irritability that comes from wearing sensors for a long duration. For example, Krishnan et al. [38] propose an activity inference approach based on motion sensors installed in a home environment to avoid such problems.

Recently, researchers are exploring smart phones equipped with accelerometers and gyroscopes to recognize activities and gesture patterns. For example, Brezmes et al. [39] have implemented a real-time classification system for some basic human movements using a conventional mobile phone equipped with an accelerometer. The results show that the capacity of conventional mobile phones in executing in real-time all the necessary pattern recognition algorithms to classify the corresponding human movements. Kwapisz et al. [40] describe a different implementation that uses phone-based accelerometers to perform activity recognition. The authors use labeled accelerometer data from twenty-nine users for daily activities such as walking, jogging, climbing stairs, sitting, and standing, and induce a predictive model for activity recognition. To improve the robustness of activity recognition using mobile sensors, Henpraserttae et al. [41] address two major issues in using a tri-axial accelerometer-embedded mobile phone for continuous activity monitoring, i.e., the difference in orientations and locations of the device. Their algorithms are suitable for accurate activity recognition using a mobile phone regardless of device orientation and location. An extensive survey on sensor-based activity recognition can be found in [42].

Apart from sensors, RFID has been increasingly explored in the area of human activity recognition. Some research efforts propose to realize human activity recognition by combining passive RFID tags with traditional sensors (e.g., accelerometers). Daily activities can be inferred from the traces of object usage via various classification algorithms such as Hidden Markov Model, boosting and Bayesian networks [13], [43]. Other efforts dedicate to exploit "pure" RFID techniques for activity recognition. For example, Wang et al. [44] use RFID radio patterns to extract both spatial and temporal features, which are in turn used to characterize various activities. However, such solutions require people to carry RFID tags or even readers (e.g., wearing a bracelet).

5.2 Activity Recognition by Instrumenting Environment

Recently, there have emerged research efforts focusing on exploring *device-free* activity recognition. Such approaches require one or more radio transmitters, but people are free from carrying any receiver or transmitter. Most device-free approaches concentrate on analyzing and learning distribution of radio signal strength (RSSI) or radio links. The main idea is to exploit the phenomenon that RSSI changes significantly when an object is passing by. For instance, Liu et al. [45] introduce a novel application that uses RF tag arrays for activity monitoring to provide an economically attractive solution to the traditional image analysis-based approaches. Youssef et al. [14] propose to pinpoint people's locations by analyzing the moving average and variance of wireless signal strength. Zhang et al. [18] develop a sensing approach using an RFID tag array. Different from previous schemes, this work is more cost-effective because it uses passive tag arrays together with a few active RFID tags. Another advantage is that it proposes several algorithms to reduce noise in the readings of passive RFID tags and achieves better accuracy. Zhu et al. [46] further develop a novel approach for RFID reader localization using passive RFID tags. However, most of these efforts focus on localization and tracking. There are not much work on study device-free activity recognition. Sigg et al. [17] propose a device-free activity recognition system based on a sensor array. Compared to this work, we develop a robust dictionary-based algorithm for identifying larger set of daily activities, which extensively exploits the handy and low-cost radio signals of passive RFID devices.

5.3 Sparse Representation

The theory of sparse representation aims at finding efficient and compact representations for signals in signal processing [31], which is primarily suitable for problems like denoising, compression, inpainting. Sparse representation in general refers to the process of choosing a good subset of dictionary elements along with the corresponding coefficients to represent a signal.

Sparse representation has been widely used in video tracking, e.g., the monitored object is modeled as a sparse linear combination of a series of templates [47], [48]. The employed dictionary plays an important role in sparse

Fig. 11. Our proposed algorithm can detect and classify 23 postures and actions with an average accuracy of over 96 percent.

	SittingStraight	SittingLeft	SittingRight	SittingBack	SittingForward	StandingStraight	SittingtoStanding	Walking	HighArmWavingHorizontalTwo	HighArmWavingBackForthTwo	HighArmWavingHorizontalLeft	HighWavingHorizontalRight	KickingLeftForward	KickingLeftLeft	KickingLeftBack	KickingRightForward	KickingRightRight	KickingRightBackt	BendingOver	CrouchingToStanding	FallingForward	FallingLeft	FallingRight
SittingStraight	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SittingLeft	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SittingRight	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SittingBack	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SittingForward	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
StandingStraight	0.00	0.00	0.00	0.00	0.09	0.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SittingtoStanding	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Walking	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00
HighArmWavingHorizontalTwo	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
HighArmWavingBackForthTwo	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
HighArmWavingHorizontalLeft	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.91	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
HighWavingHorizontalRight	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.82	0.00	0.09	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.00
KickingLeftForward	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.00	0.00	0.09	0.73	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
KickingLeftLeft	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
KickingLeftBack	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
KickingRightForward	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
KickingRightRight	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.91	0.00	0.00	0.00	0.00	0.00	0.00
KickingRightBack	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.00	0.00	0.00	0.00	0.82	0.00	0.00	0.00	0.00	0.00
BendingOver	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.91	0.00	0.09	0.00	0.00
CrouchingToStanding	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.91	0.00	0.00	0.00
FallingForward	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.91	0.00	0.00
FallingLeft	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.09	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.09	0.09	0.00	0.55	0.00
FallingRight	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.32

Fig. 12. The confusion matrix of our proposed approach for orientation-sensitive activity recognition.

representation or sparse coding based image reconstruction and classification, while learning dictionaries from the training data has led to state-of-the-art results in image classification tasks. It has several successful applications, such as face detection and image classification [23], [24], [25]. For example, [25] directly took the training samples of all classes as the dictionary to represent the query face image, and classified it by evaluating which class leads to the minimal reconstruction error of it.

To the best of our knowledge, our work is the very first of few on investigating the dictionary-based sparse representation in human activity recognition by learning signal strength stream. Compared to our previous work in [30], [49], we further develop the dictionary-based sparse learning algorithm for constructing activity dictionary, and explore multiple strategies of using the learned sparse coefficients of dictionaries under person-independent scenario. Moreover, we have conducted extensive and thorough evaluations in terms of person-independent along with person-dependent scenarios.

6 Conclusion

We have presented in this paper the technical details of a device-free, unobtrusive human activity recognition system that holds the potential to support independent living of older people, which is a critical research and development area given the significant challenges presented by the ageing population in most countries nowadays. We particularly investigate a dictionary-based approach for sparse representation of noisy and unstable radio frequency identification

(RFID) streaming signals. Our approach achieves a more compact representation of activities while preserves richer information, thereby supporting efficient and robust recognition of human activities. The ideas proposed in this paper are generic and applicable in many other applications. In particular, we adopt the orthogonal matching pursuit to solve the sparse optimization problem. We have implemented our system to validate the proposed techniques and some demonstration video clips are available from the first author's homepage.[2] We have conducted extensive experiments using real datasets collected in both office and home settings, and the experimental results demonstrate effectiveness, efficiency, and robustness of our proposed approach. We have also investigated the way of extracting robust features from raw signal strength stream by designing a simple but highly rank-based feature selection method. Our dataset is publicly available to other researchers in the community.[3]

Our future work will concentrate on validating and further developing this system in more complex and dynamic environments, e.g., what if the locations of furniture change and what if there are different ways when performing activities. The work presented in this paper is the first step to recognize high-level, complex human activities. While we only focus on atomic postures in this paper, there are widely recognized three types of human activities: i) actions, which consist of multiple postures for a single person with temporal dimension, e.g., "walking"; ii) interactions, which are activities that involve two or more persons, e.g., shaking hands with others; and iii) group activities, which are activities performed by conceptual groups of people, e.g., having a meeting with a group of people. Identifying and recognizing more complex human activities is one main goal of our future work, e.g., inferring concurrent activities, eating and watching TV.

REFERENCES

[1] S. Mennicken, J. Vermeulen, and E. M. Huang, "From today's augmented houses to tomorrow's smart homes: New directions for home automation research," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2014, pp. 105–115.

[2] E. Kim, S. Helal, and D. Cook, "Human activity recognition and pattern discovery," *IEEE Pervasive Comput.*, vol. 9, no. 1, pp. 48–53, Jan.–Mar. 2010.

[3] C. Xin and M. Song, "Detection of PUE attacks in cognitive radio networks based on signal activity pattern," *IEEE Trans. Mobile Comput.*, vol. 13, no. 5, pp. 1022–1034, May 2014.

[4] J. Tung, H. Snyder, J. Hoey, A. Mihailidis, M. Carrillo, and J. Favela, "Everyday patient-care technologies for alzheimer's disease," *IEEE Pervasive Comput.*, vol. 12, no. 4, pp. 80–83, Oct.–Dec. 2013.

[5] Y. Lee, S. Lee, B. Kim, J. Kim, Y. Rhee, and J. Song, "Scalable activity-travel pattern monitoring framework for large-scale city environment," *IEEE Trans. Mobile Comput.*, vol. 11, no. 4, pp. 644–662, Apr. 2012.

[6] J. K. Aggarwal and M. S. Ryoo, "Human activity analysis: A review," *ACM Comput. Surveys*, vol. 43, no. 3, 2011, Art. no. 16.

[7] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in *Proc. 17th Conf. Innovative Appl. Artif. Intell.*, 2005, pp. 1541–1546.

[8] Y. Zhao, N. Patwari, J. M. Phillips, and S. Venkatasubramanian, "Radio tomographic imaging and tracking of stationary and moving people via kernel distance," in *Proc. 12th Int. Conf. Inf. Process. Sensor Netw.*, 2013, pp. 229–240.

[9] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *Proc. 19th Annu. Intl. Conf. Mobile Comput. Netw.*, 2013, pp. 27–38.

2. http://linayao.com/demo.html
3. http://linayao.com/data/rssi-activity.zip

[10] A. Davis, M. Rubinstein, N. Wadhwa, G. J. Mysore, F. Durand, and W. T. Freeman, "The visual microphone: Passive recovery of sound from video," *ACM Trans. Graph.*, vol. 33, no. 4, 2014, Art. no. 79.

[11] H. Cao, M. N. Nguyen, C. Phua, S. Krishnaswamy, and X. Li, "An integrated framework for human activity classification," in *Proc. ACM Conf. Ubiquitous Comput.*, 2012, pp. 331–340.

[12] J. A. Ward, P. Lukowicz, G. Troster, and T. E. Starner, "Activity recognition of assembly tasks using body-worn microphones and accelerometers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1553–1567, Oct. 2006.

[13] M. Buettner, R. Prasad, M. Philipose, and D. Wetherall, "Recognizing daily activities with RFID-based sensors," in *Proc. 11th Int. Conf. Ubiquitous Comput.*, 2009, pp. 51–60.

[14] M. Youssef, M. Mah, and A. Agrawala, "Challenges: Device-free passive localization for wireless environments," in *Proc. 13th ACM Int. Conf. Mobile Comput. Netw.*, 2007, pp. 222–229.

[15] M. Seifeldin, A. Saeed, A. E. Kosba, A. El-Keyi, and M. Youssef, "Nuzzer: A large-scale device-free passive localization system for wireless environments," *IEEE Trans. Mobile Comput.*, vol. 12, no. 7, pp. 1321–1334, Jul. 2013.

[16] D. Zhang, Y. Liu, and L. M. Ni, "RASS: A real-time, accurate and scalable system for tracking transceiver-free objects," in *Proc. IEEE Conf. Pervasive Comput. Commun.*, 2011, pp. 197–204.

[17] S. Sigg, M. Scholz, S. Shi, Y. Ji, and M. Beigl, "RF-sensing of activities from non-cooperative subjects in device-free recognition systems using ambient and local signals," *IEEE Trans. Mobile Comput.*, vol. 13, no. 4, pp. 907–920, Apr. 2014.

[18] D. Zhang, J. Zhou, M. Guo, J. Cao, and T. Li, "TASA: Tag-free activity sensing using RFID tag arrays," *IEEE Trans. Parallel Distrib. Syst.*, vol. 22, no. 4, pp. 558–570, Apr. 2011.

[19] J. Han, et al., "Twins: Device-free object tracking using passive tags," in *Proc. IEEE Int. Conf. Comput. Commun.*, 2014, pp. 469–476.

[20] S. Wagner, M. Handte, M. Zuniga, and P. J. Marrón, "On optimal tag placement for indoor localization," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun.*, 2012, pp. 162–170.

[21] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI: Indoor localization via channel response," *ACM Comput. Surveys*, vol. 46, no. 2, 2013, Art. no. 25.

[22] B. Wei, W. Hu, M. Yang, and C. T. Chou, "Radio-based device-free activity recognition with radio frequency interference," in *Proc. 14th Int. Conf. Inf. Process. Sensor Netw.*, 2015, pp. 154–165.

[23] S. Gao, L.-T. Chia, and I. W. Tsang, "Multi-layer group sparse coding—for concurrent image classification and annotation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2011, pp. 2809–2816.

[24] M. Elad and M. Aharon, "Image denoising via learned dictionaries and sparse representation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2006, pp. 895–900.

[25] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[26] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, pp. 321–377, 1936.

[27] M. Mubashir, L. Shao, and L. Seed, "A survey on fall detection: Principles and approaches," *Neurocomputing*, vol. 100, pp. 144–152, 2013.

[28] B. Najafi, K. Aminian, A. Paraschiv-Ionescu, F. Loew, C. J. Bula, and P. Robert, "Ambulatory system for human motion analysis using a kinematic sensor: Monitoring of daily physical activity in the elderly," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 6, pp. 711–723, Jun. 2003.

[29] J. J. Liu, et al., "A dense pressure sensitive bedsheet design for unobtrusive sleep posture monitoring," in *Proc. Pervasive Comput. Commun.*, 2013, pp. 207–215.

[30] L. Yao, et al., "RF-care: Device-free posture recognition for elderly people using a passive RFID tag array," in *Proc. 12th Int. Conf. Mobile Ubiquitous Syst.*, 2015, pp. 120–129.

[31] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[32] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006.

[33] J. Tropp, "Greed is good: Algorithmic results for sparse approximation," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.

[34] R. Acierno, M. Maffia, L. Mainetti, L. Patrono, and E. Urso, "RFID-based tracing systems for drugs: Technological aspects and potential exposure risks," in *Proc. IEEE Top. Conf. Biomed. Wireless Technol. Netw. Sens. Syst.*, 2011, pp. 87–90.

[35] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Proc. 2nd Int. Conf. Pervasive Comput.*, 2004, pp. 1–17.
[36] T. Gu, L. Wang, Z. Wu, X. Tao, and J. Lu, "Recognizing multiuser activities using wireless body sensor networks," *IEEE Trans. Mobile Comput.*, vol. 10, no. 11, pp. 1618–1631, Nov. 2011.
[37] N. Kern, B. Schiele, H. Junker, P. Lukowicz, and G. Tröster, "Wearable sensing to annotate meeting recordings," *Pers. Ubiquitous Comput.*, vol. 7, no. 5, pp. 263–274, 2003.
[38] N. C. Krishnan and D. J. Cook, "Activity recognition on streaming sensor data," *Pervasive Mobile Comput.*, vol. 10, pp. 138–154, 2014.
[39] T. Brezmes, J.-L. Gorricho, and J. Cotrina, "Activity recognition from accelerometer data on a mobile phone," in *Distributed Computing, Artificial Intelligence, Bioinformatics, Soft Computing, and Ambient Assisted Living*. Berlin, Germany: Springer, 2009, pp. 796–799.
[40] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SigKDD Explorations Newslett.*, vol. 12, no. 2, pp. 74–82, 2011.
[41] L. Sun, D. Zhang, B. Li, B. Guo, and S. Li, "Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations," in *Proc. 7th Int. Conf. Ubiquitous Intell. Comput.*, 2010, pp. 548–562.
[42] L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, and Z. Yu, "Sensor-based activity recognition," *IEEE Trans. Syst. Man Cybern. Part C: Appl. Rev.*, vol. 42, no. 6, pp. 790–808, Nov. 2012.
[43] M. Stikic, T. Huynh, K. Van Laerhoven, and B. Schiele, "ADL recognition based on the combination of RFID and accelerometer sensing," in *Proc. 2nd IEEE Int. Conf. Pervasive Comput. Technol. Healthcare*, 2008, pp. 258–263.
[44] L. Wang, T. Gu, H. Xie, X. Tao, J. Lu, and Y. Huang, "A wearable RFID system for real-time activity recognition using radio patterns," in *Proc. 10th Int. Conf. Mobile Ubiquitous Syst.*, 2013, pp. 370–383.
[45] Y. Liu, Y. Zhao, L. Chen, J. Pei, and J. Han, "Mining frequent trajectory patterns for activity monitoring using radio frequency tag arrays," *IEEE Trans. Parallel Distrib. Syst.*, vol. 23, no. 11, pp. 2138–2149, Nov. 2012.
[46] W. Zhu, J. Cao, Y. Xu, L. Yang, and J. Kong, "Fault-tolerant RFID reader localization based on passive RFID tags," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 8, pp. 2065–2076, Aug. 2014.
[47] X. Mei and H. Ling, "Robust visual tracking and vehicle classification via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 11, pp. 2259–2272, Nov. 2011.
[48] J. Ho, K.-C. Lee, M.-H. Yang, and D. Kriegman, "Visual tracking using learned linear subspaces," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2004, pp. I-782–I-789.
[49] L. Yao, et al., "Freedom: Online activity recognition via dictionary-based sparse representation of RFID sensing data," in *Proc. 15th IEEE Int. Conf. Data Mining*, 2015, pp. 1087–1092.

Lina Yao received the PhD degree in computer science from the University of Adelaide, Australia. She is currently a lecturer in the School of Computer Science and Engineering, UNSW. Her research interests include the Internet of Things, information filtering and recommending, and human activity recognition. She is the author of more than 60 publications. She is a member of the IEEE and the ACM.

Quan Z. Sheng received the PhD degree in computer science from the University of New South Wales, in 2006. He is a full professor and head of Department of Computing, Macquarie University. His research interests include big data analytics, Internet of Things, and Web science. He is the recipient of the ARC Future Fellowship in 2014, Chris Wallace Award for Outstanding Research Contribution in 2012, and Microsoft Research Fellowship in 2003. He is the author of more than 280 publications. He is a member of the ACM and the IEEE.

Xue Li received the PhD degree in information systems from the Queensland University of Technology, in 1997. He is a full professor in the School of Information Technology and Electrical Engineering, University of Queensland, Australia. He is an adjunct professor with Chongqing University, China. His research areas are big data analytics, pattern recognition, and intelligent information systems. He is a member of the ACM and the IEEE.

Tao Gu is currently an associate professor in computer science with RMIT University, Australia. His current research interests include mobile computing, ubiquitous/pervasive computing, wireless sensor networks, distributed network systems, sensor data analytics, cyber physical system, Internet of Things, and online social networks. He is a senior member of the IEEE and a member of the ACM.

Mingkui Tan received the bachelor's degree in environmental science and engineering in 2006, the master's degree in control science and engineering in 2009, both from Hunan University, China, and the PhD degree in computer science from Nanyang Technological University, Singapore, in 2014. He is a full professor in the School of Software Engineering, South China University of Technology. His research interests include compressive sensing, big data learning, and large-scale optimization. He is a member of the IEEE.

Xianzhi Wang received the BE degree in computer science from Xi'an Jiaotong University, China, in 2007, and the master's and PhD degrees in computer science from the Harbin Institute of Technology, China, in 2009 and 2014, respectively. He is a research fellow in the School of Information Systems, Singapore Management University. His research interests include data management and service-oriented computing. He is a member of the IEEE and the ACM.

Sen Wang received the PhD degree from the University of Queensland (UQ), Australia, in 2014. During Jan. 2014 and Jul. 2016, he was an ARC post-doctoral research fellow in DKE Group, UQ. Currently, he is a lecturer in the School of Information and Communication Technology, Griffith University. His major areas of research interests include: Medical data analysis, signal and image processing, pattern recognition and machine learning algorithms, big data analytics.

Wenjie Ruan received the BS degree in automation and the MSc degree in control science and engineering from Central South University, China, in 2010 and 2013, respectively. He is working toward the PhD degree in the School of Computer Science, University of Adelaide, Australia. His research interests include indoor localization and tracking, pervasive computing, and mobile computing.