We develop a dictionary-based learning approach to un-
We propose a lightweight yet effective feature selection
tification (RFID), and wireless sensor network technologies,
greatly restricts its applicability in real environments. Cameras
strongly on the lighting conditions (e.g., poor performance at
high computational cost for machine interpretation. In addi-
tion, the performance of such vision-based approaches depends
strongly on the lighting conditions (e.g., poor performance at
night), camera facing angles (e.g., uncovered areas) etc, which
fluctuations, our system can successfully recognize different
activities performed by a person in the environment.
Extensive experiments conducted in a real-life residential environment
demonstrate that our proposed system offers a good overall
performance (e.g., achieving over 96% accuracy in recognizing
23 activities) and has the potential to be further developed to
support the independent living of elderly people.

**Keywords**—Activity recognition, RFID, sparse coding, dictio-
nary, feature selection, sensing data

I. INTRODUCTION

Worldwide, the population is aging due to increasing life
expectancy and low birth rate. With recent developments in
cheap sensor and networking technologies, it has become
possible to develop a wide range of valuable applications
such as the remote health monitoring and intervention. These
applications offer the potential to enhance the quality of life
for the elderly, afford them a greater sense of security, and
facilitate independent living [1]. For example, by monitoring
the daily routines of a person with dementia, it is possible to
track how completely and consistently the daily routines are
performed, and determine when the resident needs assistance.
Central to realizing these applications is *activity recognition*,
which is emerging as an important area of research and
development in recent years [2].

Computer vision related human activity recognition is one of
the directions, but unfortunately, such solutions demand
high computational cost for machine interpretation. In addition,
the performance of such vision-based approaches depends
strongly on the lighting conditions (e.g., poor performance at
night), camera facing angles (e.g., uncovered areas) etc, which
greatly restricts its applicability in real environments. Cameras
are generally considered to be intrusive to people’s privacy.

With the growing maturity of sensor, radio-frequency iden-
tification (RFID), and wireless sensor network technologies,
activity recognition from inertial, unobtrusive sensor readings
has become a popular research area in last few years. Inertial
sensors are the most frequently used wearable sensors for
human activity recognition [3], [2]. Although sensor-based
activity recognition can better address issues such as privacy
than conventional computer vision-based approaches, most
work from sensor-based activity recognition require people
to wear the inertial sensors [4], [3] and RFID tags [2]. The
main drawbacks of such solutions is that they need users’
cooperation. As a result, these approaches are not always
practical, particularly for monitoring elderly persons with
cognitive disabilities.

To overcome the aforementioned issues, we have developed
an effective and unobtrusive activity recognition system to
support independent living of older people by leveraging
passive RFID tags—which are maintenance free (no batteries
needed for tags) and inexpensive (about 10 cents each and still
dropping quickly)—and data mining techniques. Our approach
is lightweight in computational cost, and people do not need
to wear any devices. Passive RFID tags are deployed in an
environment (e.g., on the wall in a room) forming a tag array.
By analyzing the Received Signal Strength Indicator (RSSI)
fluctuations, our system can successfully recognize different
activities performed by a person in the environment.

Recognizing human actions from streaming RFID signals
is a challenging problem due to the nature of signal changes
in real-world conditions like distraction, diffusion and degra-
dation [5]. Sparse coding aims at modeling 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 [6]. In this paper, we explore
sparse representation for robust activity recognition. The main
contributions of our work are summarized below:

- We develop a dictionary-based learning approach to un-
cover structural information between RSSI signals of dif-
frent activities. The dictionaries are learned by a sparse
coding based algorithm. Compared to the existing activity
recognition approaches, our dictionary-based approach
achieves more compact representation of activities while
preserving richer information, thereby underpinning effi-
cient and robust recognition of human activities.

- We propose a lightweight yet effective feature selection
method to extract signal patterns. We particularly exploit
an unsupervised and filter-based feature selection ap-

**Abstract**—Understanding and recognizing the activities per-
formed by people is a fundamental research topic for a wide range
of important applications such as fall detection of elderly people.
In this paper, we present the technical details behind Freedom, a
low-cost, unobtrusive system that supports independent living
of the older people. The Freedom system interprets what a
person is doing by leveraging machine learning algorithms
and radio-frequency identification (RFID) technology. To deal
with noisy, streaming, unstable RFID signals, we particularly
develop a dictionary-based approach that can learn dictionaries
for activities using an unsupervised sparse coding algorithm.
Our approach achieves efficient and robust activity recognition
via a more compact representation of the activities. Extensive
experiments conducted in a real-life residential environment
demonstrate that our proposed system offers a good overall
performance (e.g., achieving over 96% accuracy in recognizing
23 activities) and has the potential to be further developed to
support the independent living of elderly people.

**Keywords**—Activity recognition, RFID, sparse coding, dictio-
nary, feature selection, sensing data

1550-4786/15 $31.00 © 2015 IEEE
DOI 10.1109/ICDM.2015.102

Lina Yao†, Quan Z. Sheng†, Xue Li‡, Sen Wang†, Tao Gu†, Wenjie Ruan†, and Wan Zou†
†School of Computer Science, The University of Adelaide, Australia; Email: {lina.yao, michael.sheng}@adelaide.edu.au
‡School of ITEE, The University of Queensland, Australia; Email: {xueli, sen.wang}@itee.uq.edu.au
§School of CS&IT, RMIT University, Australia; Email: tao.gu@rmit.edu.au
1088

From above observations, we believe that 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 $\mathcal{S} \subseteq \mathbb{R}^t$ ($t$ is the number of tags) be the domain of observable signal strength fluctuation (RSSI indicator in this work), $\mathcal{S}$ and $\mathcal{L} \subseteq \{1, ..., K\} \subseteq \mathbb{R}$ be the domain of output activity label $l$ ($K$ is the number of activities). Suppose we have $n$ RSSI and activity label pairs $\{(s_i, l_i) | s_i \in \mathcal{S}, l_i \in \mathcal{L}, i = 1, ..., n\}$. The training dataset can be represented as:

$$\mathbf{S} = [s_1, ..., s_n] \in \mathbb{R}^{t \times n}$$

$$\mathbf{L} = [l_1, ..., l_n]^T \in \mathbb{R}^n$$

(1)

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

III. THE SYSTEM AND THE TECHNIQUES

Our system (see Figure 2) consists of three main stages:

- Processing the raw signal streaming data from various RFID tag inputs into individual segments, and then extracting low-level statistical features from each segment,
- Learning the overcomplete dictionary for each activity using the extracted features, and
- Given a new signal streaming data, finding the dictionary from the learned activity dictionaries that best approximates the testing sample.

A. Segmentation

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). We incorporate the temporal information during the segmentation process of feature transformation. We particularly divide the raw streaming signal data into segments where each segment is generated by a sliding window based method. Relevant information can be extracted as features from each single segment. It should be noted that different activities have different mean temporal lengths. To incorporate the activity duration, we use a one-sliding-window-per-activity strategy and the length of each window corresponds to the activity duration as carried out by the subject. In this way, our method takes the temporal dependency into account for constructing each activity’s dictionary.

Some existing segmentation methods (e.g., online time series segmentation) apply an adaptive model on one dimensional data [8], which are not applicable to build segmentation for multi-dimensional streaming data (e.g., N-dimensional RFID RSSI data in our case). In our approach, we design a slope variation based sliding window segmentation algorithm by analyzing a distance criterion and the geometrical content of the RSSI data. The algorithm aims to follow the shape of the RSSI fluctuation curve. We first read the samples into a buffer and for each buffer, we mark points with slope change on each dimensional data. In other words, they indicate the changes from peak to trough of the curve, which represent the trend of changes from one posture to the other.
We adopt the majority voting rules to integrate the segmentation points which are calculated from each dimensional data within each buffer. In terms of all potential segment points of each dimension in one buffer, we count times of each potential point and select the points with the most repeating times as the final segmentation point for the multi-dimensional RSSI in each buffer. If some points have the same repeating times, these points could be regarded as the center points and then we respectively calculate the sum of horizontal geometric distances between other potential points and these center points in one buffer. The potential points with the lowest sum of horizontal geometric distances between itself and others would be set as the best segmentation point. We also design a backtrack mechanism to eliminate the troughs or peaks resulted from outliers. After the segmentation process, the continuous $S$ will be divided as a set of individual segments.

B. Feature Extraction

The feature extraction process focuses on transforming each segment into feature vectors. In our approach, we design seven types of lightweight statistical feature vectors for this purpose, as listed in Table I.

![Fig. 2. The architecture of our proposed activity recognition system](image)

**TABLE I**

<table>
<thead>
<tr>
<th>No.</th>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Min</td>
<td>Minimal value of $S_i$</td>
</tr>
<tr>
<td>2</td>
<td>Max</td>
<td>Maximal value of $S_i$</td>
</tr>
<tr>
<td>3</td>
<td>Mean</td>
<td>Average value of $S_i$</td>
</tr>
<tr>
<td>4</td>
<td>Variance</td>
<td>The square of standard deviation of $S_i$</td>
</tr>
<tr>
<td>5</td>
<td>Root Mean Square</td>
<td>The quadratic mean value of $S_i$</td>
</tr>
<tr>
<td>6</td>
<td>Standard Deviation</td>
<td>Measure of the spreadness of $S_i$</td>
</tr>
<tr>
<td>7</td>
<td>Median</td>
<td>The median $S_i$</td>
</tr>
</tbody>
</table>

The extraction process in our approach yields a total of $m$ feature vectors $O = \{o_1, \ldots, o_t\}$, where $o \in \mathbb{R}^m$, here $m = 7 \times t$ 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 added when there is no enough training data to reliably learn all the parameters of the activity models. To achieve the best classification performance, the dimensionality of the feature vector should be as small as possible, keeping only the most salient and complementary features.

We develop a canonical correlation analysis (CCA) based method to compute initial rankings for each pair of features, where two feature vectors are given and a projection is computed for feature vector such that they are maximally correlated in the dimensionality-reduced space. We first apply CCA for all pairs of the extracted features. The result is a similarity matrix of canonical correlations. For each pair of feature vectors $\{o_i, o_j\}$ that can be linearly mapped into: $o_i \rightarrow w_{o_i}^T o_i$ and $o_j \rightarrow w_{o_j}^T o_j$, where $w_{o_i} \in \mathbb{R}^m$ and $w_{o_j} \in \mathbb{R}^m$, their correlation coefficient $\rho$ can be obtained by maximizing the following equation:

$$\rho = \frac{w_{o_i}^T o_i w_{o_j}^T o_j}{\sqrt{w_{o_i}^T o_i w_{o_i} \cdot w_{o_j}^T o_j w_{o_j}}}$$

After we deploy CCA for 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.

We then employ a forward searching strategy to search the feature subsets based on their pairwise rankings, which traverses the full search space provided by the initial ranking of canonical correlation coefficients of the feature pairs to find the optimal feature set, and the searching process will be terminated when either the predefined dimensionality of features is reached or all features are already considered.

C. Activity Dictionary Learning

Recent research shows that learning a dictionary by fitting a set of overcomplete basis vectors to a collection of training samples can generate more compact and informative representation from given data and achieve better recognition results [9]. We propose a sparse representation based approach to recognize human activities 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 key structural information of given 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 $\ell_1$ optimization problem [10], which represents structural information of RSSI data in a more compact and informative way. Second, the dictionary learning 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. Third, each dictionary can be trained and learned by using only very small training samples, which can effectively relax the heavy workload on labeling and annotating training data in the activity recognition, as required by most existing approaches.

Assuming we have \( K \) types of activities, and we construct \( K \) dictionaries (one dictionary for each activity). After that, a new signal strength vector is reconstructed using the \( K \) dictionaries. The reconstruction errors using different dictionaries are compared and the smallest reconstruction error indicates that the new testing signal sample fits better to the specific corresponding dictionary than others. In what follows, we present the details of the proposed algorithm.

Let \( \mathbf{O}^k = \{ \mathbf{o}_1^k, \mathbf{o}_2^k, \ldots, \mathbf{o}_m^k \} \) be the training sample from activity class \( C^k \), to learn and encode the information of the testing samples belonging to a particular activity class, we first construct an overcomplete dictionary \( \mathbf{D}^k \) for each class \( C^k \). Recall the set of training samples from the \( k \)-th activity as \( \mathbf{O}^k = \{ \mathbf{o}_1^k, \mathbf{o}_2^k, \ldots, \mathbf{o}_m^k \} \), where \( \mathbf{o}_k^m \in \mathbb{R}^m \) is the feature dimensions. We intend to find a dictionary matrix \( \mathbf{D}^k \in \mathbb{R}^{m \times K} \) having \( K \) \((K > m) \) vectors \( \{ \mathbf{d}_1^k, \ldots, \mathbf{d}_K^k \} \), over which \( \mathbf{O}^k \) has a sparse representation \( \mathbf{X}^k = \{ \mathbf{x}_1^k, \ldots, \mathbf{x}_K^k \} \), where \( \mathbf{x}_k^i \in \mathbb{R}^K \). In this case, the original training matrix \( \mathbf{O}^k \) can be represented as a linear combination of no more than \( \tau_0^k \approx \frac{K}{2} \) dictionary vectors. The optimization problem can be formalized as:

\[
\min_{\mathbf{D}^k, \mathbf{X}^k} \|\mathbf{O}^k - \mathbf{D}^k \mathbf{X}^k\|_2^2, \text{ s.t. } \|\mathbf{x}_k^i\|_0 \leq \tau_0^k
\]

We adopt the K-SVD algorithm [9] to solve this problem, which performs two steps iteratively until converged. The first stage is the sparse coding stage, \( \mathbf{D}^k \) is kept fixed and the coefficient matrix \( \mathbf{X}^k \) is computed by orthogonal matching pursuit algorithm, and then the dictionary \( \mathbf{D}^k \) is updated sequentially allowing the relevant coefficients to be unique to K-SVD and resulting in a faster convergence. The dictionary learning algorithm is detailed in Algorithm 1.

D. Reconstruction-based Classification

As mentioned above, one advantage of having class-specific dictionaries is that each class is modeled independently and hence the painful repetition of the training process can be avoided when a new type of activity is added to the system. After constructing the dictionary for each activity, for a given query feature vector of signal samples \( \mathbf{o}^* \), its reconstruction error for the \( k \)-th activity (\( k \in [1, K] \)) can be calculated as:

\[
e_k = \| \mathbf{o}^* - \mathbf{D}^k \mathbf{X}^k \|_2
\]

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

\[
l_{\mathbf{o}^*} = \arg\min_k e_k
\]

Our proposed activity classification is summarized in Algorithm 2.

IV. EXPERIMENTS

A. Experimental Settings

Hardware Setup. We used one Alien 9900+ RFID reader, four circular antennas (each antenna for one room) and multiple Squig inlay passive RFID tags in our study. The tags were placed along the walls, where each grid is roughly 0.8m \( \times \) 0.8m. The antennas were arranged between \( \approx 1.3m \sim 1.6m \) height with angle \( \approx 70^\circ \). Figure 3 shows part of the setup and hardware used in our experimental environment.

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 is in the antenna’s reading range, but not detected). Meanwhile, RSSI data is much sensitive to environments. 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 at the sampling rate of \( \approx 0.5 \) 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). The 23 postures and activities are the most common ones in people’s daily lives. Each subject performed each posture or action for 120 seconds and all 23 different activities performed sequentially by one subject were regarded as one set of activity data.
Validation Strategy. We validated our approach using a subject-dependent strategy, where partial samples of a subject were used for testing and the remaining samples of the same participant were used for training. This is reasonable since that elderly people often live alone.

B. Comparison With Other Methods

We report our experiments that focus on two aspects. First, we implemented four heuristics based on sparse representation proposed in [11], in which the training samples are directly used to construct the dictionaries. Suppose the matrix $\mathbf{B}$ formed by training feature vectors is transformed from $K$ activities, $\mathbf{B} = \{\mathbf{B}_1, ..., \mathbf{B}_K\}$, where $\mathbf{B}_i$ is the subset of training samples from activity $i$. Given query samples $\mathbf{o}^*$, the overall process includes two main steps. The first step is about finding sparse representation of $\mathbf{o}^*$ on B. We computed its sparse representation, i.e., the coefficient vector $\mathbf{w}_i$ associated with different activities which can be used to recover $\mathbf{o}^*$.

In other words, the querying sample $\mathbf{o}^*$ can be represented as a linear span of $\mathbf{B} \in \mathbb{R}^{n \times d}$,

$$\mathbf{o}^* = \mathbf{b}_{11}\hat{\mathbf{w}}_1 + \mathbf{b}_{12}\hat{\mathbf{w}}_2 + ... + \mathbf{b}_{1n}\hat{\mathbf{w}}_n = \mathbf{B}\hat{\mathbf{w}}$$  \hspace{1cm} (9)

The equation can be reformulated as:

$$\hat{\mathbf{w}} = \arg \min_{\mathbf{w}} ||\mathbf{w}||, \text{ s.t. } \mathbf{o}^* = \mathbf{B}\hat{\mathbf{w}}$$  \hspace{1cm} (10)

where the sparse solution of $\hat{\mathbf{w}}$ can be found via $\ell_1$ minimization using truncated Newton interior point method [12].

The second step is about classification. Given a query sample $\mathbf{o}^*$, the four heuristics utilize the reconstruction coefficients to perform activities classification by harnessing the subspace structure of coefficients $\hat{\mathbf{w}}$, detailed as the following:

- Maximal Coefficients (MC). The testing sample label associates with the largest coefficient of $\mathbf{o}_i$. So the predicted activity label $l_i$ for the query sample is:

$$l_{o^*} = \arg \max_k \delta_k(\mathbf{o}^*)$$  \hspace{1cm} (11)

where $\delta_k(\mathbf{o}^*)$ represents the coefficients in $\mathbf{o}^*$ only associated with label $l_k$.

- Maximal Coefficients Sum (MCS). The predicted label of the query sample is the label whose sum of coefficients of $\mathbf{o}^*$ is maximized:

$$l_{o^*} = \arg \max_k \left( \sum_k \mathbf{o}^* \right)$$  \hspace{1cm} (12)

For a given testing RSSI, the label with the largest sum value is the predicted activity.

- Minimal Residual (MR). For each activity $k$, we define a characteristic function $\delta_k : \mathbb{R}^n \to \mathbb{R}^n$, which selects the coefficients associated with the $k^{th}$ posture class. For $\hat{\mathbf{w}} \in \mathbb{R}^d$, $\delta_k(\hat{\mathbf{w}})$ is a new vector whose only non-zero entries are the entries in $\hat{\mathbf{w}}$ that are associated with $k^{th}$ posture class. We can reconstruct the given query sample $\mathbf{o}^*$ as $\mathbf{o}^* = \mathbf{B}\delta_k(\hat{\mathbf{w}})$. Thus, we can classify $\mathbf{o}^*$ to the posture class based on reconstruction approximations from each activity class that has the minimal residual between the real $\mathbf{o}^*$ and the estimated $\mathbf{o}^*$:

$$r_k(\mathbf{o}^*) = ||\mathbf{o}^* - \mathbf{B}\delta_k(\hat{\mathbf{w}})||_2$$  \hspace{1cm} (13)

Then $\mathbf{o}_i$ can be classified to activity $k$ that has the minimal residual value using the following equation:

$$l_{o^*} = \arg \min_k r_k(\mathbf{o}^*)$$  \hspace{1cm} (14)

- Maximal Number of Nonzero Coefficients (NonZero). The predicted activity label of sample $\mathbf{o}^*$ is denoted as:

$$l_{o^*} = \arg \max_k \delta_k(\mathbf{o}^*)$$  \hspace{1cm} (15)

where $\delta_k(\mathbf{o}^*)$ represents the coefficients in $\mathbf{o}^*$ only associated with label $l_k$. \| \cdot \| denotes the length of $\delta_k(\mathbf{o}^*)$.

Second, we employed multiple classifiers to evaluate the quality of the generated feature subsets. The following classification techniques were selected since they have already been successfully applied for activity recognition applications.

- Multinomial Logistic Regression with $\ell_1$ (MLGL1) is a modification of linear regression that is able to predict dependent variables based on the logistic function. Given our multi-class activity recognition problem, we combine the $\ell_1$ regularization with multinomial logistic regression, which models the conditional probability $P_{\mathbf{w}}(l_j = \pm 1 | \alpha)$

The prime problem with $\ell_1$ regularization is that it can lead to some variables being assigned zero values, which is a desired property in many applications, especially when the number of samples is much smaller than the number of features. However, this can also lead to overfitting if the number of samples is too small. Therefore, it is important to carefully choose the regularization parameter $\lambda$. In our experiments, we used cross-validation to select the optimal $\lambda$.

- k-Nearest Neighbor (kNN) 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.

- 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. Given the
sequence of training RSSI and corresponding posture labels \( \Omega = \{(\mathbf{o}_1, l_1), \ldots, (\mathbf{o}_n, l_n)\} \), where \( \mathbf{o} \in \mathbb{R}^D \) and \( l \in \{1, \ldots, k\} \). The objective function is formulated as:

\[
\min_{w, b, \xi} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^{n} \xi_i
\]

s.t. \( l_i (\mathbf{w}^T \phi(\mathbf{o}_i) + b) \geq 1 - \xi_i, i = 1, 2, \ldots, n \)

\( \xi_i \geq 0, i = 1, 2, \ldots, n \)

where \( \xi_i \) is the slack variables, \( C \) is the penalty of error term, \( K(\mathbf{o}_i, \mathbf{o}_j) = \phi(\mathbf{o}_i)^T \phi(\mathbf{o}_j) \) is the kernel function.

- Random Forest (RF) builds a forest of decision trees that have the same distribution but independent output classes.

Figure 4 shows the overall performance comparison results. We can observe that our method (RFM, RFID-based Activity Monitoring) significantly outperforms all of the other approaches, which shows a good potential and effectiveness in activity classification. From the results, our method can accurately recognize most of orientation sensitive activities in a cluttered real living environment. Figure 5 shows a detailed example of a sequences of activities, our proposed method can identify a series of activities only with minor misclassification during the activity transitions.

V. RELATED WORK

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). In this way, daily activities are inferred from the traces of object usage via various classification algorithms such as Hidden Markov Model, boosting and Bayesian networks [2]. Other efforts dedicate to exploit the potential of using “pure” RFID techniques for activity recognition [5]. For example, Wang et al. [13] 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).

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. For instance, Youssef et al. [14] propose to pinpoint people’s locations by analyzing the moving average and variance of wireless signal strength. Zhang et al. [7] develop a sensing approach using an RFID tag array. However, most of these efforts focus on localization and tracking. There are not much work on study device-free activity recognition.

VI. CONCLUSION AND FUTURE RESEARCH

We have presented in this paper the technical details underneath a device-free, unobtrusive human activity recognition system with location support 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 aging population in most countries. 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 the activities while preserves richer information, thereby supporting efficient and robust recognition of human activities. Our future work will focus on two main directions: (1) we plan to evaluate our system on larger datasets along with more subjects via incorporating with local aged care centers; (2) we will work on identifying complex human activities is one main goal of our future work, e.g., inferring a person is eating or watching TV.

REFERENCES


