

Article

# Abnormal Activity Detection Using Pyroelectric Infrared Sensors

Xiaomu Luo <sup>1,\*</sup>, Huoyuan Tan <sup>1</sup>, Qiuju Guan <sup>2</sup>, Tong Liu <sup>3</sup>, Hankz Hankui Zhuo <sup>4</sup> and Baihua Shen <sup>5</sup>

<sup>1</sup> School of Medical Information Engineering, Guangzhou University of Chinese Medicine, Guangzhou 510006, China; tanhuoyuan@gmail.com

<sup>2</sup> College of Mechanical and Electrical Engineering, Zhongkai University of Agriculture Engineering, Guangzhou 5102256, China; qiuju95@126.com

<sup>3</sup> Department of Electronic Science, Huizhou University, Huizhou 516007, China; liutongboss@163.com

<sup>4</sup> School of Data and Computer Science, Sun Yat-sen University, Guangzhou 510006, China; zhuohank@mail.sysu.edu.cn

<sup>5</sup> School of Information Engineering, Guangdong University of Technology, Guangzhou 510006, China; bhshen@sohu.com

\* Correspondence: woodwood2000@163.com; Tel.: +86-135-7038-9403

Academic Editors: Octavian Adrian Postolache, Alex Casson and Subhas Mukhopadhyay

Received: 23 March 2016; Accepted: 31 May 2016; Published: 3 June 2016

**Abstract:** Healthy aging is one of the most important social issues. In this paper, we propose a method for abnormal activity detection without any manual labeling of the training samples. By leveraging the Field of View (FOV) modulation, the spatio-temporal characteristic of human activity is encoded into low-dimension data stream generated by the ceiling-mounted Pyroelectric Infrared (PIR) sensors. The similarity between normal training samples are measured based on Kullback-Leibler (KL) divergence of each pair of them. The natural clustering of normal activities is discovered through a self-tuning spectral clustering algorithm with unsupervised model selection on the eigenvectors of a modified similarity matrix. Hidden Markov Models (HMMs) are employed to model each cluster of normal activities and form feature vectors. One-Class Support Vector Machines (OSVMs) are used to profile the normal activities and detect abnormal activities. To validate the efficacy of our method, we conducted experiments in real indoor environments. The encouraging results show that our method is able to detect abnormal activities given only the normal training samples, which aims to avoid the laborious and inconsistent data labeling process.

**Keywords:** pyroelectric infrared (PIR) sensor; abnormal activity detection; wireless sensor network

## 1. Introduction

The world population is aging rapidly. As time goes on, the proportion of the elderly relative to the total population has increased, and continues to increase, especially in developed countries [1]. Thus, helping seniors live a better life is crucial and has great societal benefits. Although the elders have the option of going to nursing homes or hospice care, most of them would prefer to stay in their own houses where they feel more familiar and comfortable. Limited funding for public healthcare services and the shortage of registered nurses are also driving factors for the adoption of a home-based assisted living paradigm. Therefore, healthy aging at home has become one of the most active research areas [1], especially the problem of abnormal activity detection [2–5]. The elderly living alone in isolated areas have been in need of emergency attention, and in the worst cases, some were found dead in their homes [6].

Traditionally, abnormal activity detection approaches use cameras to obtain the data of full human body movements [7]. However, there are challenging issues in vision-based methods,

such as computational complexity in image processing, data consistency under different illumination conditions, and privacy infringement of the human target [8]. These problems make the practical deployment of vision-based systems difficult. An alternative method is to collect sensing data from wearable motion sensors and detect abnormal activities based on the collected sensing data [2]. Although motion sensors worn on the human body or integrated into human clothing can collect motion data with much less volume of data compared to those from vision-based systems, such wearable devices may make the human subject feel obtrusive. In addition, the elders are prone to forget wearing the devices after they change clothes. In addition, having to recharge the wearable devices regularly, even after deliberate design of power management units, is inconvenient for the user [2]. In order to serve as a reliable and robust abnormal activity detection system for the elderly living alone, the following factors should be considered:

1. robust to the change of environment, especially the light illumination;
2. protective to the residents' privacy;
3. convenient to use, especially for the elderly.

Bearing those factors in mind, a Pyroelectric Infrared (PIR) sensing paradigm offers a promising alternative to the optical and wearable counterparts [3]. PIR sensors are non-intrusive sensors and only sensitive to the infrared radiation changes induced by human motion, which makes them robust to interference caused by clustered background and illumination variance. In addition, as PIR sensors are relatively cheap and can be embedded within home environments, such as ceiling-mounted deployment, they are suitable to be used for home-based assisted living.

However, there are some challenges facing the PIR based sensing system for abnormal behavior detection. The first one is the design of sensor nodes, which are required to capture the spatio-temporal characteristic of the human motion. Furthermore, as the data generated continuously by ambient PIR sensors, there is an increasing demand to analyze those ever-growing sensing data automatically, with little human intervention. Most importantly, the problem of abnormal activity detection is computationally challenging [4]. Here, we define abnormal activities as events that they have not been expected in advance. Unlike normal activities, the abnormal samples are extremely scarce, or even non-existent. It is impossible to acquire or simulate all kinds of abnormal samples to train the system beforehand.

In this paper, we propose a PIR-based sensing system for anomaly detection. We design a PIR sensor node that can capture the spatio-temporal feature of human motion effectively. The key to achieve this target is by leveraging the visibility modulation of each sensor to enhance spatial resolution. We employ the reference structure tomography (RST) paradigm [9] to segment the Field of View (FOV) of each PIR sensor into sampling cells. Thus, different human activities will generate discriminative spatio-temporal signals under the monitor region. Next, we use the Hidden Markov Models (HMMs) [10], one kind of generative models, to profile normal activities. Each training sample is modeled by an HMM, and their dissimilarity is calculated based on the Kullback–Leibler (KL) divergence [11]. A self-tuning spectral clustering algorithm is used to cluster similar training samples, without the need to specify the number of cluster and the distance kernel width manually [12]. Finally, One-Class Support Vector Machines (OSVM) [13] are setup to profile the normal activities, and any unexpected activities will be classified as anomaly. It is worth pointing out that our system is trained in an unsupervised manner, which aims to avoid the laborious and inconsistent manual data labeling process.

The rest of the paper is structured as follows: Section 2 introduces the related work. Section 3 provides the design and implement of the PIR sensors. Section 4 offers the overview of our proposed method. Section 5 presents the framework of human activity representation, dissimilarity calculation and clustering. Section 6 depicts the usage of the OSVM algorithm for abnormal activity detection. The experimental results are provided in Section 7. Conclusions and future work are given in Section 8.

## 2. Related Works

Traditionally, cameras have been used for human activity classification and abnormal activity detection [14]. The processing of video includes background subtraction, human motion extraction and activity modeling [15]. The video data streams usually contain tens of thousands of pixels in each frame, and their intensity is easily effected by the change of illumination [5]. The excessive computational burden and the feeling of privacy intrusion make it difficult to be employed massively in real home environment.

Wearable sensors is another paradigm for anomaly detection. Yin *et al.* [4] propose a two-stage approach for detecting abnormal activities. The first stage is to train an OSVM to model the commonly normal activities, and the second stage is to derive abnormal activity models from the suspicious activities filtered by the first stage. Zhu *et al.* [2] propose using the wearable sensors together with location information provided by the camera to detect abnormal activities. A probabilistic framework is used to model different anomalies, including spatial anomalies, timing anomalies, duration anomalies and sequence anomalies. However, the biggest inconvenience of wearable sensing is that the sensors have to be recharged regularly even after deliberate design of power control units [2].

PIR sensing is a promising choice besides camera-based and wearable-based sensing. In [16], three PIR sensor models which are deployed in a hallway are used to detect the movement of eight human targets, including the two moving directions, three distance intervals and three speed levels. PIR sensor models can also be used to construct wireless sensor networks, which are intended to track and recognize multiple human targets [17]. Using only the binary information obtained by infrared sensors attached to the ceiling of a room, the human positions can be estimated, and even the number of humans in the room changes dynamically [18].

To capture the spatio-temporal feature of human movement, the monitored region is segmented into discrete sampling cells [19]. By leveraging the idea of compressive infrared sampling, the FOV of each PIR sensor is modulated by reference structure [9]. In [20], PIR sensors were used to extract the spatio-temporal feature of the human motion from the infrared radiation domain. Ten aerobic exercises, 360 examples in total, which were performed in front of the PIR sensor node were recorded, and the nearest neighbor classifier was then used to classify different exercises. Furthermore, they presented a PIR-based compressive classification method for recognizing six typical physical activities [21]. Similar to the experimental setup in [16], three sensor models were located on the ceiling, with opposite tripods facing each other. SVM and HMMs were used to evaluate the performance of their system [21].

PIR sensors are also used to detect abnormal activities, especially fall detection. In [22], PIR sensors were deployed in a distributed sensing paradigm, which aimed at capturing the synergistic motion patterns of head, upper-limb and lower-limb. The experiment results of fall detection were encouraging. However, their system was side-view, which means it was easily occluded by other objects, and the falls had to occur perpendicular to the FOV of the PIR sensors. In other words, it was view-dependent. To overcome these limitations, Luo *et al.* [3] proposed using the ceiling-mounted PIR sensor array to implement a fall detection system, *SensFall*. To achieve fall detection, the normal and abnormal training samples had to be collected beforehand. In other words, it was the supervised machine learning paradigm [23]. However, the abnormal detection is clearly a cost sensitive problem [4] because the samples of abnormal behavior were rare, or even non-existent. All types of the anomalies can not be elaborated on in advance. If we can only acquire the sensing data generated from normal activities, how can we train the system for abnormal activity detection automatically? This is the motivation of our study.

In this paper, we propose a PIR sensor based sensing paradigm for abnormal activity detection in an unsupervised fashion. To avoid the laborious and inconsistent manual data labeling process, we propose using the self-tuning spectral clustering algorithm to discover the number of normal activities automatically. The KL divergence is employed to measure the similarity between each pair of normal training samples and construct the similarity matrix. HMMs are then utilized to profile each cluster of

training samples, and OSVM is trained to detect abnormal activities. The details of our method will be elaborated on in the following sections.

### 3. Sensing System

#### 3.1. Sensing Model

In this subsection, we review the design of our sensing model. The task of the sensing model is to capture the discriminative spatio-temporal feature of the human activities.

The schematic diagram of our sensing model is shown in Figure 1a. Our model springs from the reference structure tomography (RST), which uses multidimensional modulations to encode the mapping between radiating objects and sensor measurements [9]. The object space refers to the space where humans perform different activities. It is the 3D physical space where human motion will generate varying radiation patterns. The measurement space refers to the space where the PIR sensors are located. Before visibility modulation, the output of all the PIR sensors are the same; they can not be used for activity classification. To capture the spatio-temporal feature of human motion, we segment the object space into discrete cells. The projection of these sampling cells on the ground is shown in Figure 1b. Assume that the object space  $\Omega$  is divided into  $L$  discrete non-overlapping sampling cells, denoted as  $\Omega_i$ , then  $\Omega = \cup_i \Omega_i, \Omega_i \cap \Omega_j = \emptyset$ , where  $1 \leq i, j \leq L$ .

Assume that there are  $M$  sensors located in the measurement space. The visibility function  $v_{ji}$  is binary valued, depending on whether the sampling cell  $\Omega_i$  is visible to the  $j$ th PIR sensor:

$$v_{ji} = \begin{cases} 1 & \Omega_i \text{ is visible to the } j\text{th PIR} \\ 0 & \text{otherwise} \end{cases}$$

The output of the  $j$ th PIR sensor is given by

$$\begin{aligned} m_j(t) &= h(t) * \sum_{i=1}^L v_{ji} \int_{\Omega_i} s(\mathbf{r}, t) d\mathbf{r} \\ &= \sum_{i=1}^L v_{ji} [h(t) * \int_{\Omega_i} s(\mathbf{r}, t) d\mathbf{r}] \\ &= \sum_{i=1}^L v_{ji} s_i(t) \end{aligned} \quad (1)$$

where “\*” denotes convolution,  $h(t)$  is the impulse response of the PIR sensor, and  $\Omega_i \in \mathbb{R}^3$  is the  $i$ th sampling cell.  $s(\mathbf{r}, t)$  is the thermal density function in the object space, and then  $s_i(t) = h(t) * \int_{\Omega_i} s(\mathbf{r}, t) d\mathbf{r}$  is the sensor measurement of sampling cell  $\Omega_i$ .

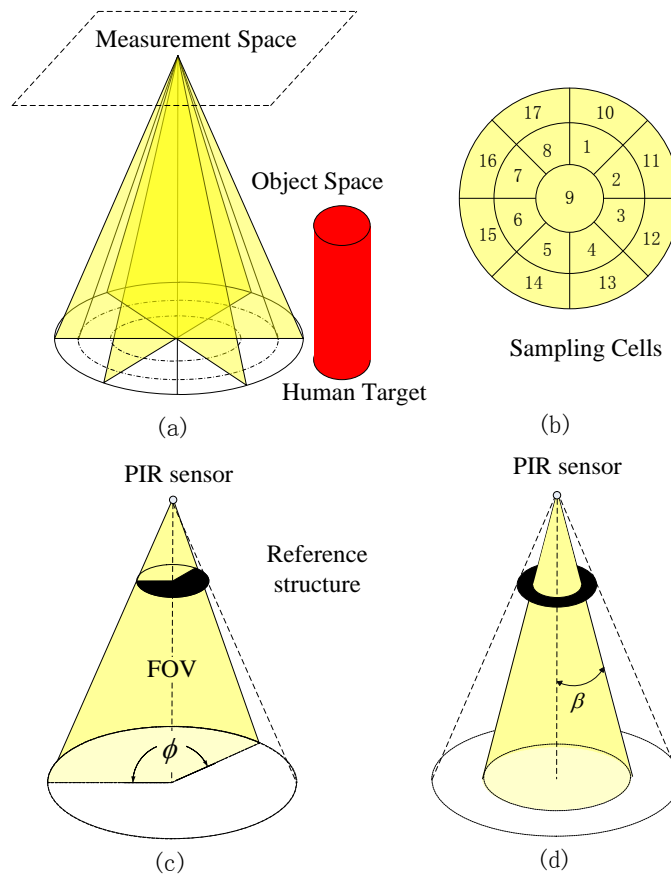
Equation (1) can be equivalently represented in a matrix form as

$$\mathbf{M} = \mathbf{V}\mathbf{S} \quad (2)$$

where  $\mathbf{M} = [m_j(t)] \in \mathbb{R}^{M \times 1}$  is the measurement vector of PIR sensors,  $\mathbf{V} = [v_{ji}] \in \mathbb{R}^{M \times L}$  is the measure matrix determined by the visibility modulation scheme, and  $\mathbf{S} = [s_i(t)] \in \mathbb{R}^{L \times 1}$  is the sensor measurement of the sampling cells before visibility modulation.  $\mathbf{M}$  can be regarded as a linear measurement of the radiation variation within all cells.

The human body can be regarded as an infrared radiative source to the surrounding environment. In comparison with the whole object space, the human body is sparsely distributed. Thus, the change of infrared radiation induced by human motion takes place only in a few sampling cells. This can be regarded as a compressive sensing problem [24], and the activity classification in the object space is projected into an analogous problem in the measurement space. When the signal is sparse or

compressible, learning and classification directly in the compressive measurement domain are possible in the compressed sensing framework [21,25].



**Figure 1.** Sensing model design: (a) Measurement space, object space and the human thermal target; (b) The projection of sampling cells on the ground; (c) Type I mask; (d) Type II mask.

### 3.2. Reference Structure Implementation

To implement the sensing model described in the previous subsection, which segments the object space into discrete sampling cells, we employ two kinds of masks. These masks play the role of reference structure, which modulates the FOV of PIR sensors. The first type of mask, Type I, is a fan shape, as shown in Figure 1c. After applying such mask, the FOV of the PIR sensor is no longer a full cone, but a partial cone shape, called a fan cone. The second type of mask, Type II, is a ring shape, as shown in Figure 1d. The FOV of the PIR sensor after masked is still a full cone, but its cone angle  $\beta$  is less than that of the original cone. These two types of masks provide two degrees of freedom (DOF) spatial partitions.

In our system implementation, the performance of our system will improve as the number of PIR sensors increases [20]. Because of the hardware constraint of our sensor node, seven PIR sensors with masks are multiplexing to segment the object space into sampling cells, as shown in Figure 1b. Four PIR sensors are masked by Type I mask, and the remaining three PIR sensors are masked by the Type II mask. In such a configuration, the object space is segmented into 17 sampling cells. Referring to Equation (2),  $M = 7$ ,  $L = 17$ , and the measurement matrix  $V$  is shown in Figure 2.

PIR ID \ Cell ID	Cell ID																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	0	1	1	1	0	0	0	0	0	0	1	1	1	0	0	0	0
2	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	0	0
3	0	0	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1
4	1	1	0	0	0	0	0	1	0	1	1	0	0	0	0	0	1
5	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
6	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Figure 2. Measurement matrix: seven pyroelectric infrared (PIR) sensors for 17 sampling cells.

#### 4. Proposed Algorithm

Based on the implementation of our sensing model, human activity under the object space will generate PIR data streams correspondingly. The measurement of PIR sensors are segmented by the Short-Time Energy method automatically [3,26]. Given a collection of normal samples  $\{Y_1, Y_2, \dots, Y_N\}$ , our abnormal activity detection method works in two phases. Figure 3 shows a diagrammatic illustration of our method. In the first phase, by applying the self-tuning spectral clustering, the number of activities classes  $C$  is determined automatically, and the normal traces are grouped accordingly. In the second phase, each class of activity is modeled by an HMM. The equal-length feature vectors are constructed based on the likelihood output of training samples generated by these  $C$  HMMs. The OSVM is then trained for abnormal activity detection. It shows clearly that the spectral clustering algorithm is the core of our approach. The key components of our approach are explained in detail in the following sections.

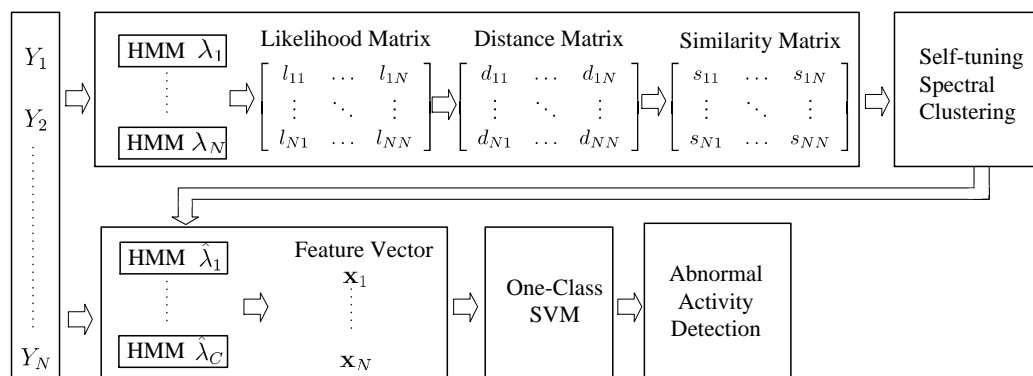


Figure 3. A block diagram illustrating our approach.

#### 5. Spectral Clustering

To profile the normal activities in an unsupervised fashion, we employ the spectral clustering method to cluster similar sequences. However, because the lengths of these sequential data are different and vary greatly in value, it is a challenging issue to model these data for better similarity measures.

##### 5.1. Likelihood Matrix Construction

Since the training sequences are generated by a hidden mechanism associated with human's underlying activities, it is reasonable to model these sequences using a generative model [27]. In this work, we adopt a set of HMMs to model the training sequences. HMMs are a type of non-deterministic stochastic finite state automata, which are widely employed in signal processing



and pattern recognition [28]. The parameters of a continuous HMM with Gaussian mixture emissions can be represented in the following compact form:

$$\lambda = \{\pi, A, \mu, \Sigma\} \quad (3)$$

where  $\pi$  is the initial state probability distribution,  $A$  is the state transition probability distribution,  $\mu$  is the mean vector, and  $\Sigma$  is the covariance matrix.

The  $i$ th training sequence  $Y_i$  can be presented as the output of  $M$  PIR sensors,

$$Y_i = \begin{bmatrix} m_1(1) & m_1(2) & \dots & m_1(T_i) \\ \vdots & \vdots & \ddots & \vdots \\ m_M(1) & m_M(2) & \dots & m_M(T_i) \end{bmatrix} \quad (4)$$

where  $m_j(t)$  is the output of the  $j$ th PIR sensor at time  $t$ ,  $t = 1 \dots T_i$ . By using the Baum–Welch algorithm [10], we fit  $N$  HMMs, one for each individual sequence  $Y_i$ ,  $1 \leq i \leq N$ .

To calculate the distance between each pair of these sequences, a probabilistic model-based framework for sequence clustering is proposed in [29]. The likelihood matrix  $\mathbf{L} = \{l_{ij}\}$ , whose  $ij$ th element is defined as

$$l_{ij} = \log p_{ij} = \frac{1}{\text{length}(Y_j)} \log P(Y_j; \lambda_i), 1 \leq i, j \leq N \quad (5)$$

where  $Y_j$  is the  $j$ th sequence,  $\lambda_i$  is the model trained for the  $i$ th sequence, and  $P(Y_j; \lambda_i)$  is the likelihood of  $Y_j$  generated by model  $\lambda_i$ .

## 5.2. Sequence Distance Measures

After the likelihood matrix  $\mathbf{L}$  is constructed, the original variable-length sequence clustering problem is transformed to a typical similarity-based one. The  $j$ th column of  $\mathbf{L}$  represents the likelihood of sequence  $Y_j$  under each of the trained models. The next step is to define a meaningful distance measure for these sequences.

A popular paradigm is to obtain likelihood-based distances between each pair of sequences [29]. Based on this work, several other distance measures have been proposed under a similar philosophy [27,30,31]. However, the main limitation of these methods is that they only consider the distance between two sequences each time, not including the global information of the whole set of data. Hence, we propose using the definition of distance measurement from a probabilistic perspective [32].

We regard the likelihood of each of the sequences under the trained models as samples from the conditional likelihoods of the models given the data, which embeds information from the whole data set [32]. This gives rise to highly structured distance matrices to give a better performance in comparison with aforementioned distance-based methods [27,29–31].

According to the definition of likelihood matrix  $\mathbf{L}$ , the  $j$ th column of  $\mathbf{L}$  can be regarded as the likelihood of the sequence  $Y_j$  under each of the trained models  $\lambda_i$ ,  $1 \leq i \leq N$ . These  $N$  models can be regarded as a set of “sampled points” from the model space  $\Lambda$  surrounding the HMMs that actually span the data space. Thus, these  $N$  trained models become a good discrete approximation  $\bar{\Lambda} = \{\lambda_1, \dots, \lambda_N\}$  to the model space of interest.

If we normalize the likelihood matrix  $\mathbf{L}$ , which means each column adds up to one, we get a new matrix  $\mathbf{L}_N$  whose columns can be regarded as the probability density functions (pdfs) over the approximated model space conditioned on each of the individual sequences:

$$\mathbf{L}_N = [f_{\bar{\Lambda}}(Y_1), \dots, f_{\bar{\Lambda}}(Y_N)] \quad (6)$$

This interpretation leads to the Kullback–Leibler (KL) divergence, which is a natural choice for the measurement of the dissimilarity between two pdfs. The discrete case of the KL divergence formulation is as follows:

$$D_{KL}(f_P||f_Q) = \sum_i f_P(i) \log \frac{f_P(i)}{f_Q(i)} \quad (7)$$

where  $f_P$  and  $f_Q$  are two discrete pdfs. Obviously, the KL divergence is not a proper distance because of its asymmetry; a symmetrized version is used as:

$$D_{KL}^{sym}(f_P||f_Q) = \frac{1}{2}[D_{KL}(f_P||f_Q) + D_{KL}(f_Q||f_P)] \quad (8)$$

Thus, the distance between the sequences  $Y_i$  and  $Y_j$  can be defined as:

$$d_{ij} = D_{KL}^{sym}(f_{\bar{\Lambda}}(Y_i)||f_{\bar{\Lambda}}(Y_j)) \quad (9)$$

Distances defined this way are obtained according to the patterns created by each sequence in the probability space spanned by different models, and the distance measured between two sequences  $Y_i$  and  $Y_j$  involves information related to the rest of the data sequences.

### 5.3. Similarity Matrix Construction

Before applying a spectral clustering algorithm, the distance matrix  $\mathbf{D} = \{d_{ij}\}$  should be transformed into a similarity matrix  $\mathbf{S} = \{s_{ij}\}$ . A commonly used procedure is to apply a Gaussian kernel,

$$s_{ij} = \begin{cases} \exp(-\frac{d_{ij}^2}{2\sigma^2}) & \text{for } i \neq j \\ 0 & \text{for } i = j \end{cases} \quad (10)$$

where  $\sigma$  is the scaling parameter controlling the kernel width.

The value of  $\sigma$  is commonly specified manually, or numerous iteration has to be run for a range of  $\sigma$  [33]. However, when the input data includes clusters with different local statistics, a single value of  $\sigma$  may not work well for all the data. Thus, instead of selecting a single scaling parameter, we propose calculating a local scaling  $\sigma_i$  for each data point  $d_i$  [12]. The similarity between  $Y_i$  to  $Y_j$  can be revised as  $d_{ij}/\sigma_i$  while the converse is  $d_{ji}/\sigma_j$ . Therefore,  $d_{ij}$  is symmetry, and the Equation (10) can be generalized as:

$$\hat{s}_{ij} = \begin{cases} \exp(-\frac{d_{ij}^2}{\sigma_i\sigma_j}) & \text{for } i \neq j \\ 0 & \text{for } i = j \end{cases} \quad (11)$$

where  $\sigma_i = d(Y_i, Y_K)$ ,  $Y_K$  is the  $K$ 'th neighbor of  $Y_i$ . The selection of  $K$  is independent of scale and is a function of the data dimension of the embedding space.

Thus, the scaling parameters for each pair of  $Y_i$  and  $Y_j$  are not fixed; they are determined automatically according to the local statistics of the neighborhoods.

### 5.4. Self-Tuning Spectral Clustering

After the similarity matrix  $\hat{\mathbf{S}} = \{\hat{s}_{ij}\}$  is constructed, we apply spectral clustering methods to partition the training sequences into clusters. For an undirected graph  $G$  with vertices  $v_i$  and edges  $s_{ij}$ , the matrix  $\mathbf{S}$  could be considered as an adjacent matrix for  $G$ , where each element  $s_{ij}$  can be viewed as the similarity between the vector  $v_i$  and  $v_j$ . The target of spectral clustering is to partition the  $G$  into a distinct sub-graph.



It is a tricky problem to specify the number of clusters  $C$ . One method to discover the number of clusters is to analyze the eigenvalues of the normalized Laplacian matrix  $\mathbf{L}_a$ , which is based on the similarity matrix  $\hat{\mathbf{S}}$ . The analysis given in [33] shows the number of repeated eigenvalues of magnitude 0 with multiplicity equal to the number of clusters  $C$ . However, eigenvalues depend on the structure of the individual clusters, and no assumptions can be placed on their values [12]. Once noise is introduced, the eigenvalues deviate from the ideal case, and it is difficult to decide the number of clusters.

An alternative approach to discover the number of clusters  $C$  automatically is to analyze the eigenvectors of Laplacian matrix  $\mathbf{L}_a$  [12]. Assume the matrix  $\mathbf{X} = [x_1, \dots, x_C] \in \mathbb{R}^{N \times C}$  is constructed by stacking the largest eigenvectors of  $\mathbf{L}_a$  in columns. In the ideal case where the data points could be separated distinctly,  $\mathbf{X}$  will be strictly block diagonal after sorting the eigenvectors of  $\mathbf{L}_a$ . Nevertheless, in the general case, the  $\mathbf{X}$ 's off-diagonal blocks are non-zero, and the eigensolver could just pick any other set of the orthogonal vectors;  $\mathbf{X}$  could have been replaced by  $\hat{\mathbf{X}} = \mathbf{X}\mathbf{R}$  for any orthogonal matrix  $\mathbf{R} \in \mathbb{R}^{C \times C}$ . Now, we have to recover the rotation which best aligns  $\mathbf{X}$ 's columns with the canonical system with the minimum cost.

Let  $\mathbf{Z} \in \mathbb{R}^{N \times C}$  be the matrix obtained after rotating the eigenvector matrix  $\mathbf{X}$ ; that is,  $\mathbf{Z} = \mathbf{X}\mathbf{R}$ . We wish to recover the rotation  $\mathbf{R}$  for which, in every row in  $\mathbf{Z}$ , there will be at most one non-zero entry. We thus define the cost function:

$$J = \sum_{i=1}^N \sum_{j=1}^C \frac{Z_{ij}^2}{M_i^2} \quad (12)$$

where  $M_i = \max_j Z_{ij}$ . Minimizing this cost function over all possible rotations will provide the best alignment with the canonical coordinate system. The number of clusters,  $C$ , is taken as the one providing the minimal cost.

The spectral clustering algorithm that we apply is similar to the one proposed in [12]. The algorithm works as follows:

1. Define a diagonal degree matrix  $\mathbf{D} = \{d_{ij}\}$  with  $d_{ii} = \sum_{j=1}^N \hat{s}_{ij}$ , and then construct the normalized Laplacian matrix  $\mathbf{L}_a = \mathbf{D}^{-1/2} \hat{\mathbf{S}} \mathbf{D}^{-1/2}$ .
2. Find  $C'$  principal eigenvectors  $x_1, x_2, \dots, x_{C'}$  and form the matrix  $\mathbf{X} = [x_1, \dots, x_{C'}] \in \mathbb{R}^{N \times C'}$  by stacking the eigenvectors in columns, where  $C'$  is the largest possible cluster number.
3. Recover the rotation  $\mathbf{R}$  which best aligns  $\mathbf{X}$ 's columns with the canonical coordinate system using the incremental gradient descent algorithm [12].
4. According to Equation (12), grade the cost of the alignment for each group number up to  $C'$ . Set the final group number  $C_{\text{best}}$  to be the largest group number with minimal alignment cost.
5. Take the alignment result  $\mathbf{Z}$  of the  $C_{\text{best}}$  eigenvectors, and assign the original point  $s_i$  to cluster  $c$  if and only if  $\max_j (Z_{ij}^2) = Z_{ic}^2$ .

In our experiments, because the volunteers will emulate five kinds of activities,  $C'$  is set to 10, and the self-tuning spectral clustering will determine  $C_{\text{best}}$  automatically.

## 6. One-Class SVM Classifier

### 6.1. Feature Extraction

After applying the spectral clustering algorithm, we can group the  $N$  training traces into  $C$  clusters, which correspond to  $C$  different types of activities.

To train an OSVM, we need to transform the training samples that are of variable lengths into a set of fixed-length feature vectors. Again, we apply HMMs to model these normal activities, one for each cluster. For each learned model with the corresponding parameters  $\hat{\lambda}_i$ ,  $1 \leq i \leq C$ , we calculate

the log-likelihood of each of the  $N$  normal traces given the model parameters  $\hat{\lambda}_j$ . The log-likelihood value for each pair of trace and HMMs is computed as follows:

$$L(Y_i; \hat{\lambda}_j) = \log P(Y_i; \hat{\lambda}_j), 1 \leq i \leq N, 1 \leq j \leq C \quad (13)$$

This is calculated by applying the standard forward-backward algorithm [10]. In this way, for each training trace  $Y_i, 1 \leq i \leq N$ , we can obtain a  $C$ -dimensional feature vector  $\mathbf{x}_i = \langle L(Y_i; \hat{\lambda}_1), \dots, L(Y_i; \hat{\lambda}_C) \rangle$ .

## 6.2. One-Class SVM Training

After transforming the  $N$  training traces into a set of feature vector  $\mathbf{x}_1, \dots, \mathbf{x}_N$ , we can train the one-class SVM for normal activities. The basic idea is to find a sphere that contains most of the normal data such that the corresponding radius  $R$  can be minimized:

$$\begin{aligned} \min_{R, \xi, \mathbf{a}} \quad & R^2 + \nu \sum_{i=1}^n \xi_i \\ \text{s.t.} \quad & \|\mathbf{x}_i - \mathbf{a}\|^2 \leq R^2 + \xi_i \\ & \xi_i \geq 0 \end{aligned} \quad (14)$$

The slack variables  $\xi_i$  are introduced to allow some data points to lie outside the sphere, and the parameter  $\nu \geq 0$  controls the tradeoff between the volume of the sphere and the number of errors. Using the dual representation of the Lagrangian [34], the objection function is equivalent to

$$\begin{aligned} \min_{\alpha} \quad & \sum_{i,j=1}^n \alpha_i \alpha_j \langle \mathbf{x}_i \cdot \mathbf{x}_j \rangle - \sum_{i=1}^n \alpha_i \langle \mathbf{x}_i \cdot \mathbf{x}_i \rangle \\ \text{s.t.} \quad & 0 \leq \alpha_i \leq \nu, \sum_{i=1}^n \alpha_i = 1 \end{aligned} \quad (15)$$

This quadratic programming (QP) problem can be solved using standard optimization techniques [35]. To determine if a testing sample is within the sphere, the distance to the center of the sphere has to be calculated. If the distance is larger than the radius  $R$ , the testing sample is considered abnormal.

Typically, the training samples are not spherically distributed in the input space. Thus, the original data points are first mapped into a feature space so that a better data description can be obtained. Instead of requiring an explicit mapping function from the input space to the feature space, the solution can be obtained by replacing all the inner products  $\langle \cdot, \cdot \rangle$  in Equation (15) by a kernel function  $k(\cdot, \cdot)$ :

$$\min_{\alpha} \quad k(\mathbf{x}_i \cdot \mathbf{x}_j) - \sum_{i=1}^n \alpha_i k(\mathbf{x}_i, \mathbf{x}_i) \quad (16)$$

In our context, due to the noisy and nonlinear characteristic of the PIR sensors, the decision boundary of the OSVM is quite complex. Thus, we apply the Gaussian Radial Basis Function (RBF) kernel for the OSVM, which is defined as follows:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2) \quad (17)$$

where  $\gamma > 0$  is a scaling factor that controls the width of the kernel function.

## 7. Experimental Evaluation

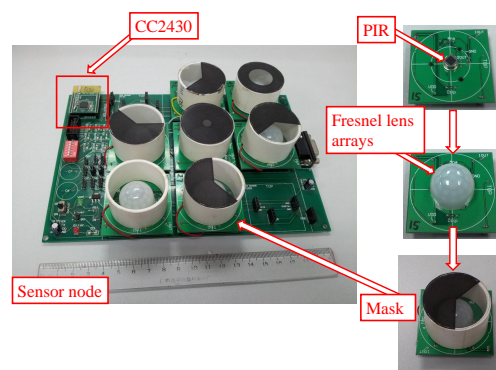
In order to evaluate the performance of our proposed method, experiments were carried out on a real data set collected from a wireless sensor network. Our proposed method is referred to as

SC+OSVM in the experiments. Two other approaches were also used for comparison. We list these three methods as the following:

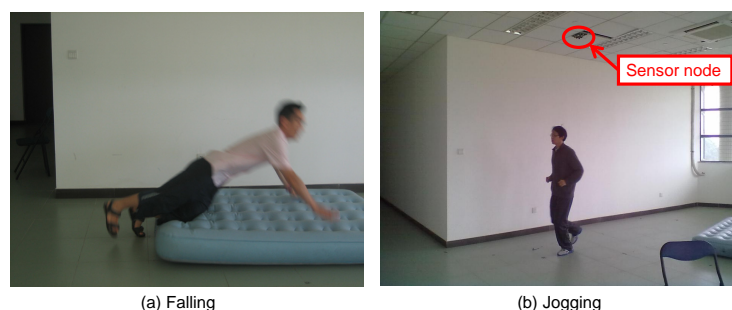
- SC+OSVM—The method proposed in our study, mainly including self-turning spectral clustering and One-Class SVM.
- SC+iForest—The difference between this method and the SC+OSVM is that we use isolation forest to replace One-Class SVM for abnormal detection. Isolation forest is an alternative algorithm for abnormal detection [36].
- OneHMM—All the normal training samples are modeled by only one HMM, which corresponds to not applying spectral clustering to the unlabeled samples. A threshold is set to distinguish normal and abnormal activities.

### 7.1. Experimental Setup

The experiments were carried out in a real indoor environment. The monitored region covered by the sensor node was a cone with a 3 m radius. The sensor node is with seven PIR sensors, each of them equipped with fresnel lens arrays and a mask, as shown in Figure 4. The CC2430 module is used for data transmission between the sensor node and the sink based on ZigBee protocol. More details of the sensor node could be found in our previous work [3].



**Figure 4.** The wireless sensor node: there are seven PIR sensors on one sensor node. The sensor node is mounted at a height of 3m from the floor, looking down to classify human activities. Each PIR sensor is equipped with its own fresnel lens arrays and mask. The sampling frequency of PIR sensor is 25 Hz, and the resolution of the A/D converter is 8-bit. The CC2430 is the Radio Frequency (RF) transmission module based on ZigBee protocol, the transmission rate of which is 250 Kbps at 2.4 GHz.



**Figure 5.** Different volunteers emulated different activities with their own speed and strategy: (a) Falling; (b) Jogging.

There were eight volunteers that participated in our experiments, including three females and five males. The height of them ranges from 1.64 m to 1.80 m, and the weight of them ranges from 50 kg

to 70 kg. Each volunteer emulated five kinds of activities, including falling, sitting down, standing up from a chair, walking and jogging. Every activity was emulated ten times by each volunteer at a self-select speed and strategy, as shown in Figure 5. Totally, we obtained 400 samples, including 80 fall-simulated samples and 320 normal activity samples. In our experiments, fall-simulated samples are regarded as abnormal activities, and other samples are regarded as normal activities. Each sample is segmented automatically by thresholding Short-Time Energy, and all the normal training samples are unlabeled.

### 7.2. Evaluation Metrics

The performance of the abnormal human-activity detection methods can be evaluated in terms of two rates, detection rate and false alarm rate. The detection rate is computed as the ratio of the number of correctly detected abnormal activities to the total number of abnormal activities. The false alarm rate is computed as the ration of the number of normal activities that are incorrectly detected as abnormal activities to the total number of normal activities.

Based on the confusion matrix shown in Table 1, the two metrics can be defined as follows:

$$\text{Detection Rate} = \frac{TP}{TP + FN} \quad (18)$$

$$\text{False Alarm Rate} = \frac{FP}{FP + TN} \quad (19)$$

The performance of an ideal abnormal human-activity detection algorithm should have a high detection rate and a low false alarm rate. Therefore, we evaluate the performance of the algorithms using an Receiver Operating Characteristic (ROC) curve, which plots the detection rate against the false alarm rate. In addition, we compute the area under the ROC curve (AUC) to compare these algorithms. AUC is a better measurement than accuracy in the evaluation of learning algorithms [37,38], especially in the cost-sensitive problems. A desirable algorithm with a high detection rate and a low false alarm rate should have an AUC value closer to one.

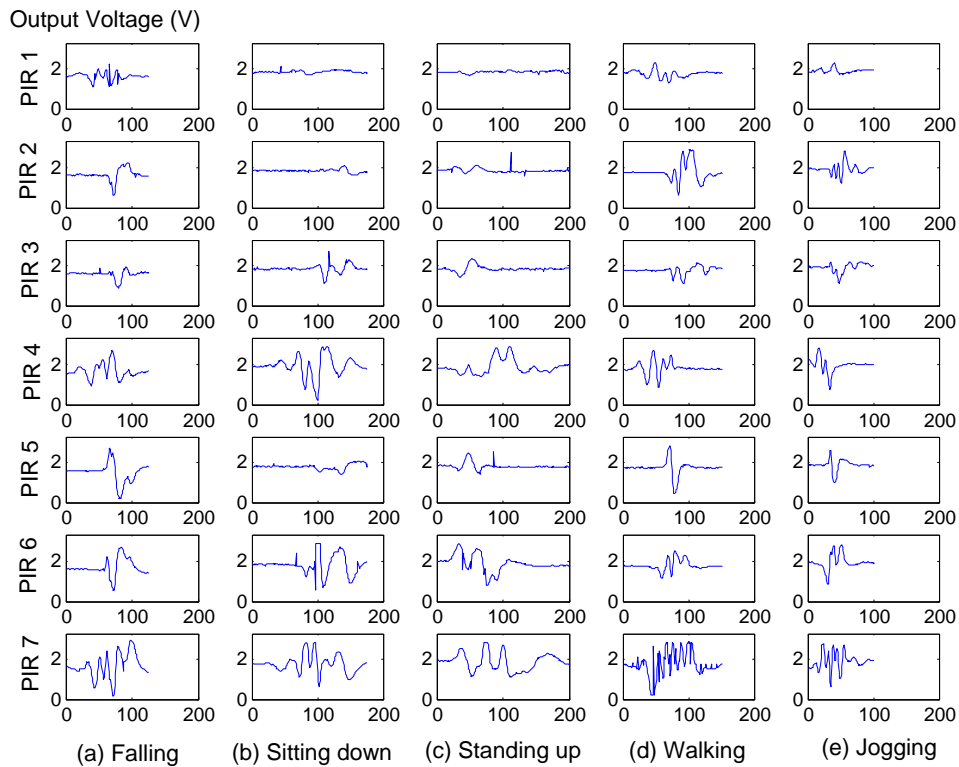
**Table 1.** Confusion matrix.

		Actual Activity	
		Abnormal	Normal
Predicted Label	Abnormal	True Positive (TP)	False Postivitive (FP)
	Normal	False Negative (FN)	True Negative (TN)

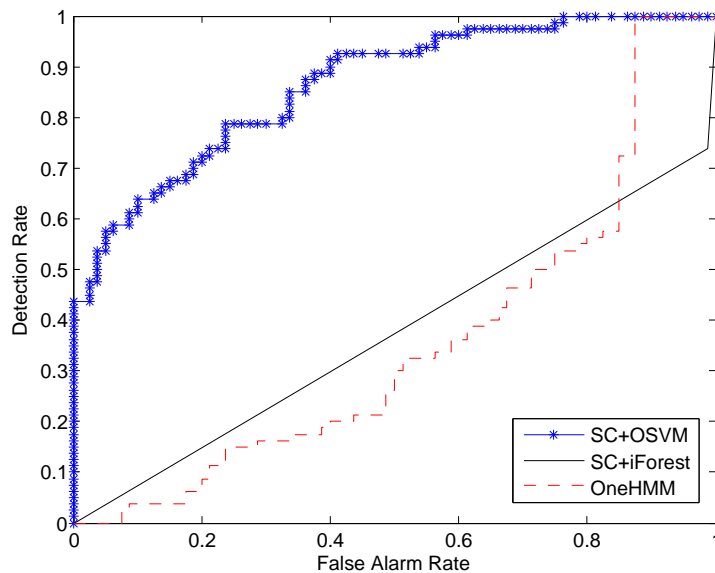
### 7.3. Experimental Results

In our experiments, because the output of the sensor node was a seven-dimensional data stream with continuous values, as shown in Figure 6, we employed HMMs with Gaussian observation density. The number of hidden states and the number of Gaussian models are determined by the log-likelihood of the training samples [3]. Specifically, we use HMMs with two Gaussians and eight states to profile the general human normal activity and each cluster of normal activities. For OSVM, the parameter  $\nu$  was set to 0.01.

Experiments were conducted to compare the performance of all three of the algorithms. We randomly selected 240 normal samples for training. The other 80 normal samples and all the 80 abnormal samples were randomly mixed together for testing. Figure 7 shows the ROC curve with respect to the detection rate and the false alarm rate. We can see from the figure that SC+OSVM gives the best detection result. This is mainly because the self-turning spectral clustering can estimate the number of different types of activities automatically and then cluster the similar activities accordingly. The discriminative feature vectors are helpful to improve the accuracy of OSVM.



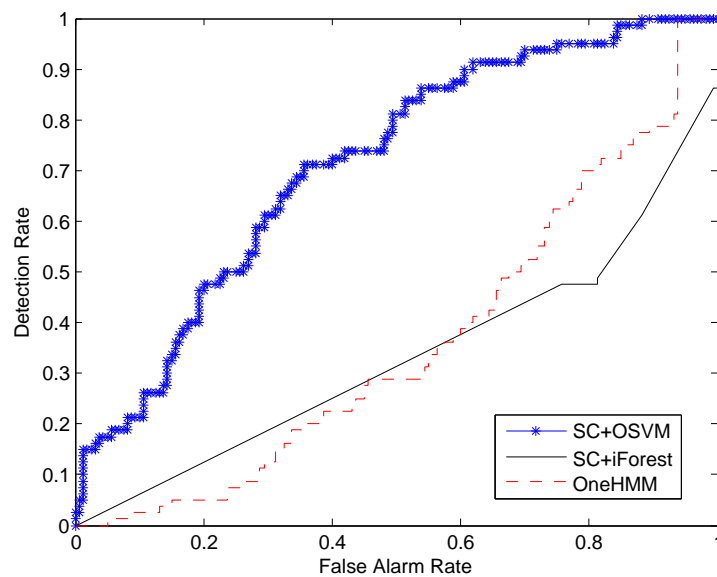
**Figure 6.** The original output of the sensor node with seven PIR sensors for different activities: (a) Falling; (b) Sitting down; (c) Standing up; (d) Walking; (e) Jogging. The horizontal axis represents the sampling points (25 Hz) of the activity.



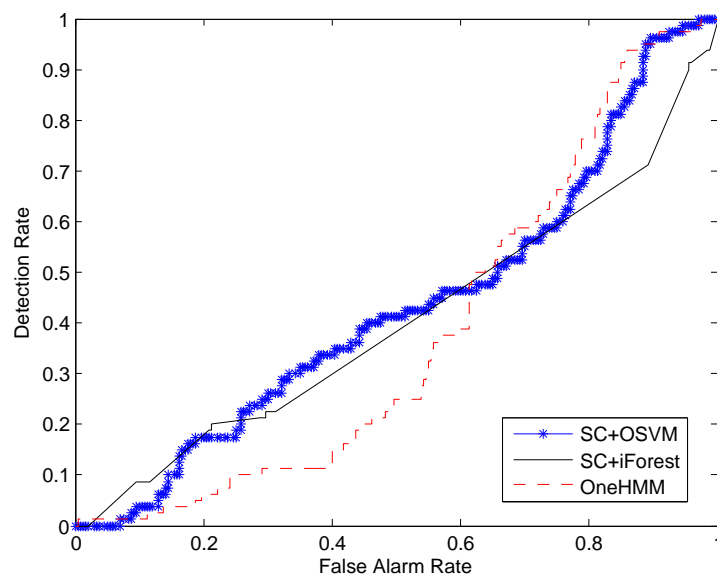
**Figure 7.** Comparison of the detection rate and the false alarm rate *versus* different numbers of training samples: training on 240 normal samples.

We also conducted experiments to investigate the effect of varying the number of training samples on the performance of the three algorithms. In these experiments, we kept the amount of testing data unchanged and reduced the number of training samples. Figures 8 and 9 show the experimental results using 160 and 80 normal traces for training, respectively. We can see from the figure that, when the number of normal traces for training decreases, the performance of the three algorithms decreases

correspondingly. The reason lies in two aspects. First, for HMMs, which are generative models [3,9], sparse training samples affect the accuracy of parameter estimation of the models. Second, for OSVM, when the training samples are sparse, the calculated decision boundary may not exactly capture the characteristics of normal activities. Therefore, the abilities of three algorithms to distinguish normal and abnormal activities degrade. As shown in Figure 9, when there are only 80 normal samples for training, the performance of OneHMM and SC+iForest are comparable to that of SC+OSVM because all of them could not model the normal activities well with less training samples. When the training samples are less than 80, the performance of three algorithms will be worse and unsatisfactory. Again, as shown in Figure 8, when we have 160 normal samples for training, SC+OSVM still performs the best among the three algorithms.



**Figure 8.** Comparison of the detection rate and the false alarm rate *versus* different numbers of training samples: training on 160 normal samples.



**Figure 9.** Comparison of the detection rate and the false alarm rate *versus* different numbers of training samples: training on 80 normal samples.

To explicitly compare the performance of the three algorithms, we computed the AUC values by calculating the area under the ROC curves depicted in Figures 7–9. The results are summarized in Table 2. We can see from the first column of the table that, when 240 normal samples are used for training, the AUC value for SC+OSVM, SC+iForest and OneHMM are 0.863, 0.375 and 0.354, respectively. SC+OSVM performs better than the other two algorithms. In addition, SC+OSVM is the best among the tree algorithms when we consider 80 and 160 normal samples for training as well.

**Table 2.** Area under the ROC curve (AUC) values with different algorithms and different numbers of training samples.

	240 Samples	160 Samples	80 Samples
SC+OSVM	0.863	0.433	0.710
SC+iForest	0.375	0.404	0.332
OneHMM	0.354	0.379	0.360

Another observation is that the performance of OneHMM will not be improved as the number of training samples increases. This is because the dissimilarity between different activities will degrade the classification ability of the HMMs; a single HMM to model all normal activities is not discriminative enough. By contrast, after applying the self-turning spectral clustering, each type of normal activities will be profiled by an HMM; it will obviously improve the performance of the OSVM.

## 8. Conclusions

In this paper, we propose a novel approach for detecting a human's abnormal activities. By employing the FOV modulation for PIR sensors, the human activity is encoded into low-dimensional data streams, which can be used to extract the tempo-spatial feature of the human motion. A self-turning spectral clustering algorithm is used to cluster the training samples with modified KL distance. The HMMs are trained to profile each cluster of activity. One-class SVM is setup to classify whether the testing samples are abnormal or not. A major advantage of our approach is that it does not need abnormal samples for training in advance. This is critical in real deployment because it is unrealistic to train the system by providing all kinds of abnormal activities. Another advantage is that our training procedure is an unsupervised learning fashion, which does not need to specify the number of kinds of normal activities. In other words, it is not necessary to label the training samples. This advantage will facilitate the mass deployment in different locations, and it can greatly reduce the human labor spent on training sample preprocessing. We demonstrate the effectiveness of our approach using real data collected from PIR sensors attached to the ceiling of the monitoring region. It shows that, as the number of training samples increases, the performance of our system will improve accordingly.

In the future, we wish to continue in the direction of detecting abnormal activities from continuous data streams. We will investigate how to incorporate the location information with the motion information to improve the performance of our system. A robust and reliable abnormal activity detection system is the most important prerequisite of the home-based assisted living paradigm.

**Acknowledgments:** This work is supported by the National Natural Science Foundation of China (NSFC) under Grant No. 61301294 and 61401174, the Natural Science Foundation of Guangdong under Grant No. 2014A030310462, and the Youth Elite Project of Guangzhou University of Chinese Medicine.

**Author Contributions:** Study concept and design: Xiaomu Luo. Acquisition of data: Xiaomu Luo, Qiuju Guan, Tong Liu and Baihua Shen. Analysis and interpretation of data: Xiaomu Luo and Huoyuan Tan. Drafting of the manuscript: Xiaomu Luo. Discussion and reviewing the manuscript: Qiuju Guan and Hankz Hankui Zhuo. All authors read and approved the final manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.



## References

1. Bloom, D.E.; Chatterji, S.; Kowal, P.; Lloyd-Sherlock, P.; McKee, M.; Rechel, B.; Rosenberg, L.; Smith, J.P. Macroeconomic implications of population ageing and selected policy responses. *Lancet* **2015**, *385*, 649–657.
2. Zhu, C.; Sheng, W.; Liu, M. Wearable sensor-based behavioral anomaly detection in smart assisted living systems. *IEEE Trans. Acoust. Speech Signal Process.* **2015**, *12*, 1225–1234.
3. Luo, X.; Liu, T.; Liu, J.; Guo, X.; Wang, G. Design and implementation of a distributed fall detection system based on wireless sensor networks. *EURASIP J. Wirel. Commun. Netw.* **2012**, *2012*, 1–13.
4. Yin, J.; Yang, Q.; Pan, J.J. Sensor-based abnormal human-activity detection. *IEEE Trans. Knowl. Data Eng.* **2008**, *20*, 1082–1090.
5. Stone, E.E.; Skubic, M. Fall detection in homes of older adults using the Microsoft Kinect. *IEEE J. Biomed. Health Inform.* **2015**, *19*, 290–301.
6. Eriksen, M.D.; Greenhalgh-Stanley, N.; Engelhardt, G.V. Home safety, accessibility, and elderly health: Evidence from falls. *J. Urban Econ.* **2015**, *87*, 14–24.
7. Turaga, P.; Chellappa, R.; Subrahmanian, V.; Udrea, O. Machine recognition of human activities: A survey. *IEEE Trans. Circuits Syst. Video Technol.* **2008**, *18*, 1473–1488.
8. Dai, J.; Wu, J.; Saghafi, B.; Konrad, J.; Ishwar, P. Towards privacy-preserving activity recognition using extremely low temporal and spatial resolution cameras. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, Boston, MA, USA, 7–12 June 2015; pp. 68–76.
9. Brady, D.; Pitsianis, N.; Sun, X. Reference structure tomography. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **2004**, *21*, 1140–1147.
10. Rabiner, L. A tutorial on hidden Markov models and selected applications in speech recognition. *IEEE Proc.* **1989**, *77*, 257–286.
11. Kullback, S.; Leibler, R.A. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86.
12. Zelnik-Manor, L. Self-tuning spectral clustering. *Adv. Neural Inf. Process. Syst.* **2004**, *17*, 1601–1608.
13. Schölkopf, B.; Williamson, R.C.; Smola, A.J.; Shawe-Taylor, J.; Platt, J.C. Support vector method for novelty detection. *Adv. Neural Inf. Process. Syst.* **1999**, *12*, 582–588.
14. Xiang, T.; Gong, S. Video behavior profiling for anomaly detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *30*, 893–908.
15. Duong, T.V.; Bui, H.H.; Phung, D.Q.; Venkatesh, S. Activity recognition and abnormality detection with the switching hidden semi-markov model. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), San Diego, CA, USA, 20–26 June 2005; pp. 838–845.
16. Yun, J.; Lee, S.S. Human movement detection and identification using pyroelectric infrared sensors. *Sensors* **2014**, *14*, 8057–8081.
17. Xiong, J.; Li, F.; Zhao, N.; Jiang, N. Tracking and recognition of multiple human targets moving in a wireless pyroelectric infrared sensor network. *Sensors* **2014**, *14*, 7209–7228.
18. Miyazaki, T.; Kasama, Y. Multiple human tracking using binary infrared sensors. *Sensors* **2015**, *15*, 13459–13476.
19. Liu, T.; Liu, J. Design and implementation of a compressive infrared sampling for motion acquisition. *EURASIP J. Adv. Signal Process.* **2014**, *2014*, 1–15.
20. Guan, Q.; Li, C.; Guo, X.; Wang, G. Compressive classification of human motion using pyroelectric infrared sensors. *Pattern Recognit. Lett.* **2014**, *49*, 231–237.
21. Guan, Q.; Yin, X.; Guo, X.; Wang, G. A novel infrared motion sensing system for compressive classification of physical activity. *IEEE Sens. J.* **2016**, *16*, 2251–2259.
22. Liu, T.; Guo, X.; Wang, G. Elderly-falling detection using distributed direction-sensitive pyroelectric infrared sensor arrays. *Multidimens. Syst. Signal Process.* **2012**, *23*, 451–467.
23. Barber, D. *Bayesian Reasoning and Machine Learning*; Cambridge University Press: Cambridge, UK, 2012.
24. Duarte, M.; Eldar, Y. Structured compressed sensing: from theory to applications. *IEEE Trans. Signal Process.* **2011**, *59*, 4053–4085.
25. Davenport, M.; Boufounos, P.; Wakin, M.; Baraniuk, R. Signal processing with compressive measurements. *IEEE J. Biomed. Health Inform.* **2010**, *4*, 445–460.
26. Lu, L.; Zhang, H.; Jiang, H. Content analysis for audio classification and segmentation. *IEEE Trans. Speech Audio Process.* **2002**, *10*, 504–516.

27. Yin, J.; Yang, Q. Integrating hidden Markov models and spectral analysis for sensory time series clustering. In Proceedings of the 5th IEEE International Conference on Data Mining (ICDM'05). Hoston, TX, USA, 27–30 November 2005, pp. 506–513.
28. Ordonez, F.J.; Englebienne, G.; de Toledo, P.; van Kasteren, T.; Sanchis, A.; Krose, B. In-home activity recognition: Bayesian inference for hidden Markov models. *IEEE Pervasive Comput.* **2014**, *13*, 67–75.
29. Smyth, P. Clustering sequences with hidden Markov models. *Adv. Neural Inf. Process. Syst.* **1997**, *12*, 648–654.
30. Panuccio, A.; Bicego, M.; Murino, V. A Hidden Markov Model-based approach to sequential data clustering. In *Structural, Syntactic, and Statistical Pattern Recognition*; Springer: Berlin, Germany, 2002; pp. 734–743.
31. Porikli, F. Clustering variable length sequences by eigenvector decomposition using HMM. In *Structural, Syntactic, and Statistical Pattern Recognition*; Springer: Berlin, Germany, 2004; pp. 352–360.
32. Dario, G.G.; Emilio, P.H.; Fernando, D.D.M. A new distance measure for model-based sequence clustering. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 1325–1331.
33. Ng, A.Y.; Jordan, M.I.; Weiss, Y. On spectral clustering: Analysis and an algorithm. *Adv. Neural Inf. Process. Syst.* **2002**, *2*, 849–856.
34. Schölkopf, B.; Platt, J.C.; Shawe-Taylor, J.C.; Smola, A.J.; Williamson, R.C. Estimating the support of a high-dimensional distribution. *Neural Comput.* **2001**, *13*, 1443–1471.
35. Chang, C.C.; Lin, C.J. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* **2011**, *2*, 27:1–27:27.
36. Liu, F.T.; Ting, K.M.; Zhou, Z.H. Isolation forest. In Proceedings of the 8th IEEE International Conference on Data Mining (ICDM'08), Pisa, Italy, 15–19 December 2008; pp. 413–422.
37. Ling, C.X.; Huang, J.; Zhang, H. AUC: A Statistically consistent and more discriminating measure than accuracy. In Proceedings of the 8th International Joint Conference on Artificial Intelligence (IJCAI), Acapulco, Mexico, 9–15 August 2003; pp. 519–524.
38. Davis, J.; Goadrich, M. The relationship between Precision-Recall and ROC curves. In Proceedings of the 23rd International Conference on Machine Learning (ICML), Pittsburgh, PA, USA, 25–29 June 2006; pp. 233–240.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).