Recognition of Human Behavior for Assisted Living Using Dictionary Learning Approach

Pubali De, Amitava Chatterjee, *Senior Member, IEEE*, and Anjan Rakshit

***Abstract*— The ambient assisted living (AAL) technology aims to provide more safety and self-sufficiency, while permitting older persons to live self-dependently in their homes. Monitoring of activities of daily living (ADL) is one of the key ideas of AAL. This becomes an interesting research idea in modern world, where condition monitoring of various ADLs and their automatic classification is a big challenge. This paper proposes a new approach for activity recognition of motion primitives relying on the sparse representation of signals, where signals are represented using a sparse combination of atoms from an over-complete dictionary. This paper intends to investigate the suitability of applying dictionary learning algorithms like K-singular value decomposition (K-SVD), which is usually used to construct an over-complete dictionary, for the effective progress of the ADL monitoring system. This paper proposes to formulate the classification approach by using SRC classifiers, based on the dictionaries learned using K-SVD algorithm. We have validated our proposed approach on a publicly available ADL data set of wrist-worn accelerometer sensor for activity recognition. Performance evaluations demonstrate that the proposed method outperforms several other competing methods.**

***Index Terms*— ADL problem, data clustering, dictionary learning, K-singular value decomposition, sparse representation classifier.**

I. Introduction

THE main goal of Ambient Assisted Living (AAL) is to check the self-care of aging people still living in their community [1] and to help them if the need arises. Observing the activities of daily living (ADLs) of aging people and recognition of deviations from preceding normal patterns enable us to identify whether any critical situation is imminent [2]. Smart homes that can implement these facilities, encourage the elderly population to live self-sufficiently as continuous monitoring of assisted living (ADL) data can provide early recognition of imminent acute situations of a human [2].

As manual identification of ADL problem is almost impossible in real life, this necessitates the development of sensor-based systems for automatic acquisition and classification of ADL data. Thus, it is pertinent to develop ambient intelligence based AL tools, especially for smart homes that can help elderly people to live independently in their homes and more and more research efforts, all over the world, are presently directed to solve these problems.

The current developments in sensor technology and ever reducing sensor costs have encouraged researchers to place various sensors in various combinations to solve this problem, including static setups as well as wearable sensors. ADL classifications based on sensory signal processing and machine learning algorithms are becoming more and more popular in a modern world scenario. Among the popular machine learning algorithms explored so far in this context, hidden markov models (HMMs) [3] and conditional random fields (CRFs) [4], [5] have enjoyed some initial success for activity and behavior recognition purposes. Threshold-based decision-making methods [6], [7] and classical machine learning algorithms, like SVM, are also common [8], [9] for activity recognition purposes. On the other hand, some probability based classification methods have also experienced initial success for activity recognition problems [10]–[13]. Some interesting human behaviour or activity recognition systems have also been developed using motion modelling [14], [15] and mahalanobis distance method [16], [17].

In recent years, the sparse representation (SR) of signals and task-driven dictionary learning algorithms are enjoying ever increasing interest among the researchers due to their effective applications in the domains of face recognition, image classification [18], [19], image denoising [20], etc. The basic spirit behind sparse representation lies in the fact that a signal can be effectively represented using a sparse combination of atoms from an over-complete dictionary. While using a fixed dictionary, comprising several training atoms from several classes, have shown to provide effective solutions for sparse representation based classification (SRC), it has also been shown that, on many occasions, learning or suitably adapting an initial dictionary, from a larger set of data signals, can actually enhance the performances that can be achieved in sparse representation based solutions. The choice of an effective over-complete dictionary **D** that may culminate into an effective sparse representation can either be given comprising a set of pre-specified functions or by developing a suitable adaptation algorithm that can adapt this dictionary to model a set of training signal exemplars [21]. In recent years, two such effective algorithms proposed are K-singular value decomposition (K-SVD) [21] and MOD [22] algorithms. Because of the above mentioned effectiveness, dictionary learning based sparse representation approaches have drawn increased attention in face recognition, image classification, video signal classification [22], [23] etc.

Manuscript received October 15, 2017; revised December 3, 2017; accepted December 11, 2017. Date of publication December 27, 2017; date of current version February 21, 2018. The associate editor coordinating the review of this paper and approving it for publication was Prof. Nitaigour P. Mahalik. *(Corresponding author: Pubali De.)*

P. De is with the Department of Electrical Engineering, Techno India-Batanagar, Kolkata 700141, India, and also with the Department of Electrical Engineering, Jadavpur University, Kolkata 700032, India (e-mail: de.pubali19@gmail.com).

A. Chatterjee is with the Department of Electrical Engineering, Jadavpur University, Kolkata 700032, India (e-mail: cha_ami@yahoo.co.in).

A. Rakshit, retired, was with the Department of Electrical Engineering, Jadavpur University, Kolkata 700032, India (e-mail: anjan_rakshit@yahoo.com).

Digital Object Identifier 10.1109/JSEN.2017.2787616

Inspired by the success of dictionary learning based algorithms, mainly used in image processing domain, this paper proposes to utilize dictionary learning based SRC that can be suitably applied for recognition of several ADL. The results achieved with our proposed approach have been demonstrated to outperform other state-of-the-art competing algorithms proposed for this purpose e.g. dynamic time warping and mahalanobis distance.

The rest of this paper is presented as follows. Section, II and III present the macroscopic view of the activity recognition problem and the signal acquisition procedure from the sensor based instrumentation system. Section IV presents the dictionary learning algorithm. Section V presents the classification approach that has been proposed in this work using sparse representation classifier, based on the learned dictionary obtained from dictionary learning algorithm. Section VI to VIII present the implementation of the proposed approach, performance evaluation and conclusion and discussions regarding future scope of research work.

II. Activity Recognition Problem

One of the most vital signal processing problems in Ambient Assisted Living (AAL) is the recognition of human activities from the signals acquired by various types of sensor combination. Consequently, various studies on AAL technologies emphasize on automatic recognition of human activities that draw parallel to ADLs like bathing, cooking, sleeping etc., which provide the capabilities to detect any deviation in their patterns. There are mainly two broad categories of sensors, i.e. wearable and non-wearable sensors, used for the purpose of solving ADL problems. Wearable sensors, attached directly to a human body, are most commonly used for activity recognition of different human posture. Among these, accelerometer sensors and wrist-worn smartphones are most often used for detection of ADLs like running, walking, falling etc. On the other hand, non-wearable sensors are deployed in a static place in the home or in a community to detect a human's movement and activities. Non-wearable sensors like infrared sensor (IR), ultrasonic sensor etc. are most commonly used for ADL classification. The complete block diagram of our proposed approach utilizing dictionary learning for activity recognition is shown in Fig. 1.The block diagram signifies the four major steps of our approach to recognize human activities of ADL problem. Each step will be discussed in detail in following sections.

III. Development of ADL System and Signal Acquisition

A successful practical system for monitoring ADL activities is one that needs less training and configuration effort as well as it should be attachable with minimum effort in the household. The two major types of challenges in the ADL classification systems is choice of sensors and their installation, depending upon the type of data to be collected. In an ADL system, the installation cost of the sensors should be as minimum as possible. The sensors should also respect privacy and simultaneously strive to attain enhanced classification accuracy. An ADL system should ideally give a self-maintainable environment whether utilized for the purpose of helping or monitoring aging community or for intruder detection or for the purpose of fall detection etc., depending on the sensor technology deployed. A "Smart Home" augmented with ADL technology may comprise sensor based instrumentation systems fitted in the roof mounting places, wall, and floor-mounting places for non-wearable sensors or wearable sensor based instrumentation systems fitted in the human cloth or a human body. In many situations, the original analog sensor outputs are converted to digital signals by analog to digital converters. Very often, feature vectors are extracted from such digital signals for different ADLs for further processing to draw meaningful conclusions.

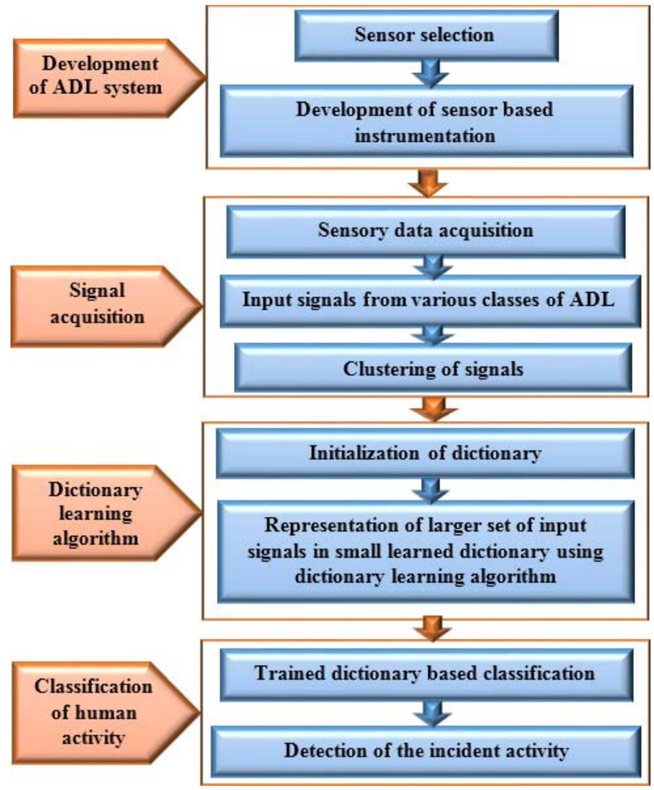

Fig. 1. Overall block diagram for activity recognition problem.

IV. Dictionary Learning Algorithm

In this section dictionary learning approach is explained in detail which is used with sparse representation, that can be suitably utilized for classification of human behavior. In sparse representation, each signal is assumed to be composed of a sparse, linear combination of a few atoms from a large, over-complete dictionary. Each atom refers to an input training signal in the training database. In a dictionary learning algorithm the objective is to adapt or learn the dictionary in such a fashion that can achieve best suitable representation of each member within that dictionary, strictly maintaining the sparsity constraints [21]. In dictionary learning approach, the dictionary learns from a larger training set, instead of using a predefined basis, such as fourier or wavelet basis [16] or uses the larger training data set itself [17]. In the latter case, using the entire set of the input training signals as a complete dictionary can potentially give improved performance, but it may involve large computational burden. A dictionary, learned from a very large training database and smaller in size than that, suitably maintaining sparsity constraints, can overcome this problem of huge computational burden, yet maintaining desired performance accuracy.

Our present approach utilized in this work is based on K-SVD algorithm, a prominent alternative among popular dictionary learning algorithms known [21]. K-SVD is known to effectively learn over-complete dictionaries from larger sets of input training signals and has been applied suitably for other engineering problems before. The main spirit of sparse representation algorithm can be viewed as a generalization of vector quantization (VQ) objective, where K-SVD can be viewed as a generalization of K-means algorithm that is utilized to satisfy the VQ objective. The main objective of K-SVD dictionary learning algorithm is to minimize the following objective function [21]:

$$\arg\min_{\mathbf{D},\mathbf{X}} \|\mathbf{Y}-\mathbf{D}\mathbf{X}\|_F^2 \text{ subject to } \forall\, i, \|\mathbf{x}_i\|_0 \le T_0 \tag{1}$$

where T_0 is called the sparsity prior. The number of non-zero elements in each $\mathbf{x}_i$ should be less than T_0. Basic K-SVD algorithm produces a reconstructive dictionary where the dictionary is composed of K prototype atoms, each of dimension n. We assume that the input signal matrix $\mathbf{Y}$ comprises N input signals, each of dimension n.

$$\mathbf{Y} = \left[\mathbf{y_1}, \mathbf{y}_2, \ldots, \mathbf{y}_N\right] \in \Re^{n\times N} \tag{2}$$

$$\text{and dictionary } \mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_K] \in \Re^{n\times K} \tag{3}$$

where $K \gg n$ which makes the dictionary over-complete. Also, it is so chosen that $N \gg K$, so that the dictionary, playing the role of a codebook, comprises K code words, that can efficiently represent the original training database.

The sparse coding of input signal $\mathbf{Y}$ produces

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N] \in \Re^{K\times N} \tag{4}$$

There are two basic stages of KSVD algorithm [21]:

i. sparse coding stage
ii. dictionary update stage

In sparse coding stage, the best possible sparse coefficient matrix $\mathbf{X}$ is determined by keeping the dictionary $\mathbf{D}$ static. Here, the objective function in (1) can be presented in the modified form, given as [21]:

$$\|\mathbf{Y}-\mathbf{D}\mathbf{X}\|_F^2 = \sum_{i=1}^{N} \left\|\mathbf{y}_i - \mathbf{D}\mathbf{x}_i\right\|_2^2 \tag{5}$$

Orthogonal matching pursuit (OMP) [18] algorithm is a popular choice for the sparse coding stage, although other pursuit algorithms can also be effectively employed there. In the dictionary update stage, keeping $\mathbf{X}$ fixed, dictionary $\mathbf{D}$ is restructured iteratively by updating one atom at a time. These two basic stages are used in an iterative fashion, for minimizing the objective function in K-SVD algorithm, given in (1). In the dictionary update stage, at any point of time, we focus on a specific atom $\mathbf{d}_k$ of the dictionary and its corresponding coefficients in $\mathbf{X}$ in its kth row, denoted as $\mathbf{x}_T^k$ (it is not the transpose of vector $\mathbf{x}^k$ i.e. the kth column of $\mathbf{X}$). Then the objective function in (1) can be given as [21]:

$$\|\mathbf{Y} - \mathbf{D}\mathbf{X}\|_F^2 = \left\|\mathbf{E}_k - \mathbf{d}_k\mathbf{x}_T^k\right\|_F^2 \tag{6}$$

where E_k presents the error involved for all N samples, after removal of the kth atom $\mathbf{d}_k$. However, a straightforward use of singular value decomposition (SVD) algorithm to determine $\mathbf{d}_k$ and corresponding $\mathbf{x}_T^k$ by approximating $\mathbf{E}_k$ will not serve the purpose, as the solution produced will not satisfy the sparsity constraints in the first place. So, to obtain a solution respecting the sparsity constraints, a set of indices w_k is formed which denotes the positions where the array elements of $\mathbf{x}_T^k$ are nonzero, given as:

$$w_k = \left\{(i \left| 1 \le i \le N, \mathbf{x}_T^k\,(i) \ne 0\right.\right\} \tag{7}$$

where $\mathbf{x}_T^k$ is the kth row in **X.**

This array w_k is utilized to form a matrix $\boldsymbol{\Omega}_k$ which can be utilized to build a restricted array $\mathbf{x}_R^k$ from $\mathbf{x}_T^k$ utilizing the matrix operation $\mathbf{x}_R^k = \mathbf{x}_R^k\boldsymbol{\Omega}_k$. This row vector $\mathbf{x}_R^k$ is essentially formed from $\mathbf{x}_T^k$ by removing zero entries from $\mathbf{x}_T^k$. This ensures that, for subsequent operations, any adaptation carried out for $\mathbf{x}_R^k$ array means we are essentially adapting the nonzero entries in $\mathbf{x}_T^k$ and all zero entries in $\mathbf{x}_T^k$ remain unchanged i.e. the sparsity constraints are fully respected. This means that the objective function given in (6) can now be transformed to the equivalent form, respecting sparsity constraints, given as:

$$\left\|\mathbf{E}_k\boldsymbol{\Omega}_k - \mathbf{d}_k\mathbf{x}_T^k\boldsymbol{\Omega}_k\right\|_F^2 = \left\|\mathbf{E}_k^R - \mathbf{d}_k\mathbf{x}_R^k\right\|_F^2 \tag{8}$$

Now, the minimization of (8) can be directly obtained using SVD, which decomposes the restricted matrix $\mathbf{E}_k^R$ as $\mathbf{E}_k^R = \mathbf{U}\Delta\mathbf{V}^T$. Consequently, the kth column of the dictionary is updated by choosing the first column of $\mathbf{U}$ i.e. $\mathbf{U}(:,1)$ as the solution $\tilde{d}_k$. The corresponding coefficient vector $\mathbf{x}_R^k$ is updated as $\mathbf{x}_R^k = \Delta\,(1,1) * V\,(:,1)$ where $\mathbf{V}(:,1)$ is the first column of $\mathbf{V}$. This procedure is followed to update each atom of the dictionary (along with its corresponding $\mathbf{x}_T^k$), one at a time, until all atoms are updated. Algorithm 1 presents the complete procedure of the K-SVD algorithm employed in this work.

V. Classification Based on Sparse Representation and Learned Dictionary

This section presents the classification of human activities carried out using the philosophy of sparse representation, utilizing the trained dictionary obtained from K-SVD algorithm. For a given test sample $\mathbf{y}$, the coefficient matrix $\hat{\mathbf{x}}_{\mathrm{AL}}$ is computed by using the objective function given in (9):

$$\hat{\mathbf{x}}_{\mathrm{AL}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{subject to } \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_2 \le \varepsilon \tag{9}$$

In this paper, the SRC problem is solved using matrix $\mathbf{A}$ as the trained dictionary $\mathbf{D}$ which is achieved from K-SVD algorithm. This $\mathrm{A} = [\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_C] \in \Re^{n\times m}$ is composed of a total of m atoms associated with C classes, each of dimension n and the unknown test sample $\mathbf{y} \in \Re^n$ is identified based upon the trained dictionary. The non-zero entries in the estimated $\hat{\mathbf{x}}_{\mathrm{AL}}$ associated with the columns of the trained dictionary $\mathbf{A}$ from

Algorithm 1 K-SVD Algorithm

Step 1:
- **Initialization of Dictionary**: $\mathbf{D}^0 \in \Re^{n\times K}$ with input signals and the columns of $\mathbf{D}^0$ are normalized using l_2 norm.
- **Input:** $\mathbf{Y} \in \Re^{n\times N}$, T_0, $\mathbf{D}^0 \in \Re^{n\times K}$
- **Output:** $\mathbf{D} \in \Re^{n\times K}, \mathbf{X} \in \Re^{K\times N}$
- **Set** $J = 1$
- **Repeat** until convergence (stopping Criterion)

Step 2:
- **Sparse coding stage:** $i = 1, \ldots, N$

$$\min_{\mathbf{x}_i} \left\{ \|\mathbf{y}_i - \mathbf{D}\mathbf{x}_i\|_2^2 \right\} \quad \text{subject to} \quad \|\mathbf{x}_i\|_0 \le T_0$$

Step 3:
- **Dictionary update stage:**

For each Column $k = 1, 2, ..., K$ in $\mathbf{D}^{J-1}$ Update it by
- Determine indices w_k using (7)
- Determine $\mathbf{E}_k = \mathbf{Y} - \sum_{j\neq k} \mathbf{d}_j \mathbf{x}_T^j$
- Restrict $\mathbf{E}_k$ to obtain $\mathbf{E}_k^R$ by selecting only the columns corresponding to w_k.
- SVD decomposition is applied as $\mathbf{E}_k^R = \mathbf{U}\Delta\mathbf{V}^T$
- Update $\tilde{\mathbf{d}}_k = \mathbf{U}(:,1)$ and $\mathbf{x}_R^k$ is updated as $\mathbf{x}_R^k = \Delta\,(1,1) * \mathbf{V}\,(:,1)$

- **Set** J=J+1

Algorithm 2 Sparse Representation Based Classification

Step 1: Training samples matrix represented as $\mathbf{A} = [\mathbf{A}_1, \mathbf{A}_2, \ldots\ldots, \mathbf{A}_C] \in \Re^{n\times m}$, for C classes.
Step 2: Normalize the columns of training matrix to have unit l_2-norm.
Step 3: Solve the l_1-minimization problem to determine coefficient vector $\hat{\mathbf{x}}_{\mathrm{AL}}$:

$$\hat{\mathbf{x}}_{\mathrm{AL}} = \arg\min_{\mathbf{x}} \|\mathbf{x}\|_1 \text{ subject to } \|\mathbf{Ax} - \mathbf{y}\|_2 \le \varepsilon$$

Step 4: For $i = 1, \ldots, C$
Residuals are computed
$r_i(\mathbf{y}) = \left\|\mathbf{y} - \mathbf{A}\delta_i(\hat{\mathbf{x}}_{\mathrm{AL}})\right\|_2$
End
Step 5: Output: identity($\mathbf{y}$) $= \arg\min_i r_i(\mathbf{y})$

an individual class i help us to assign the test sample $\mathbf{y}$ to that specified class of entity. In this process, dictionary learning based SRC becomes more effective as we can accommodate a larger set of training database using a smaller size dictionary, and yet strictly maintaining the sparsity constraints, which consequently reduces the computational burden. Let us adopt that $\delta_i : \Re^m \to \Re^m$ (for each individual class i) be the specific function that chooses the coefficients linked with the ith class.

For $\hat{\mathbf{x}} \in \Re^m, \delta_i \in \Re^m$ is defined as a new vector for every class $(i = 1, \ldots, C)$ with all zero entries except for the ones associated with class i which are equal to the corresponding ones in the estimated vector $\hat{\mathbf{x}}_{\mathrm{AL}}$ [18]. After obtaining the associated coefficients of the specified object class, one can approximate the test sample $\mathbf{y}$ for a single class entity using $\hat{\mathbf{y}}_{\mathrm{AL}} = \mathbf{A}\delta_i(\hat{\mathbf{x}}_{\mathrm{AL}})$. Then, the classification of test sample $\mathbf{y}$ can be obtained by determining the minimum among all class specific residuals, calculated using (10), between the test sample $\mathbf{y}$ and the approximated test sample $\hat{\mathbf{y}}_{\mathrm{AL}}$ [18].

$$\min_i r_i(\mathbf{y}) = \|\mathbf{y} - \mathbf{A}\delta_i(\hat{\mathbf{x}}_{\mathrm{AL}})\|_2 \quad \text{for } i = 1, \cdots, C \tag{10}$$

Algorithm 2 presents the sparse representation based classification scheme. For our approach, the learned dictionary $\mathbf{D}$, plays the role of training samples matrix $\mathbf{A}$.

VI. Application of Proposed Approach for Classification of Human Behavior

As mentioned before, the main purpose of this work is to propose a dictionary learning based SRC approach that can be suitably applied for human behaviour activity recognition problems. In this section, Fig. 2 shows how our proposed approach can be implemented for activity recognition problem.

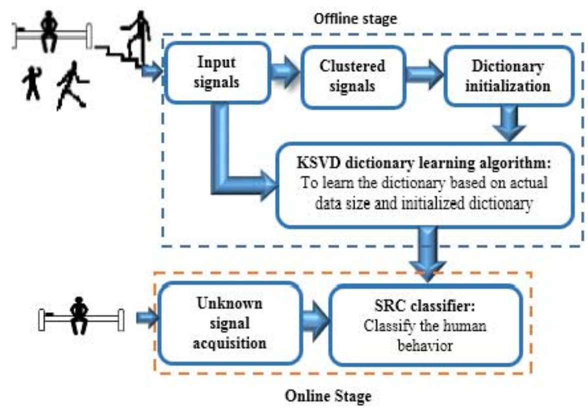

Fig. 2. Block diagram for implementation of the proposed approach.

In the first step, input training signals are acquired from volunteers or subjects to build the database and the database comprising input training signals of ADL problem are clustered to initialize the dictionary. The dictionary initialization can be carried out with different sizes of the dictionary. Specifically speaking, we have utilized 5%, 10%, 20%, and 50% clustering (i.e., for example, for 5% clustering, we use 5% of actual data belonging to each class to build the initial dictionary and so on) data, which undergoes further adaptation using a dictionary learning algorithm. In case of initialization of the dictionary, we choose $\mathbf{D}^0 \in \Re^{n\times K}$ from original signal database, which is the starting point of the K-SVD dictionary learning algorithm. In the next step, the learning of the dictionary is carried out using K-SVD algorithm, as described in section IV. Once the dictionary is learned, this trained dictionary is used as the dictionary $\mathbf{A}$ (i.e. much smaller in size compared to the input matrix of original training samples of all classes) in algorithm 2 to implement the SRC to identify different human behavior. In this work, we have solved both bi-class classification and multi-class classification problems, utilizing our proposed approach, for a benchmark, publicly available ADL dataset.

VII. Performance Evaluation

In this section, we present the results obtained using our proposed approach for the benchmark, wrist-worn accelerometer sensor database [12], [17], [24], utilized for recognition of various types of ADL behavior. These results are also compared with other state-of-the-art, competing algorithms to show the effectiveness of our proposed approach. As per dataset description [12], [24], the sensing bracelet had a sensing range of 3G, used 6 bits/axis for coding information, and the sampling frequency was 32 Hz and the sensor was mounted in such a way that x-axis points towards the hand, y-axis points towards the left and the z-axis points perpendicular to the plane of the hand. In [12], 700 trials of eight motion primitives from 16 volunteers (including 11 men and five women) from the database [24] was considered, out of which seven motion primitives are utilized here for the recognition purpose. For each motion primitive, we have selected 80% data signal for training purpose and 20% data signal for testing purpose. In our analysis, the proposed dictionary learning-based classifier uses a trained dictionary of K atoms. Here, the classification accuracy is analyzed for initialization of the dictionary using different sizes of clusters, like $K = 5\%$, 10%, 20% and 50% of the size of the input data set is used to determine which is the most suitable cluster size that can achieve superior classification accuracy. Here the seven classes of ADL activities considered for analysis of classification accuracy in activity recognition problem are in conformation with those seven classes utilized in [12] that belong to selected motion primitives like getting up from the bed (GUB), sitting down on a chair (SDC) and standing up from a chair (SUC), reiterated actions (climbing stairs (CS) and walking (WK)) and complex actions (drinking water (DW) and pouring water from glasses (PWG)). At first, the problem is solved as a bi-class classification problem where we try to determine whether an unknown behaviour belongs to a specific behaviour or not. For this purpose the specific behaviour in question constitutes the main event class and a mixture of signal from all other six classes constitutes the other class. The dictionary learning algorithm can be accordingly utilized for this purpose, reducing the computational burden. Table I shows corresponding results obtained. Fig. 3 shows these results in graphical form.

These results show that with increase in cluster size chosen, recognition error starts to decrease first, and after a certain cluster size is attained, the error starts to increase.

Here, TP denotes true positive rate (i.e. percentage of correct identification of the event activity under investigation) and TN denotes true negative rate (i.e. percentage of correct identification of those event activities which are not under investigation). It can be seen that the choice of a cluster size beyond an optimum size of the cluster can degrade the accuracy, whereas, the choice of a very small cluster size is also not desirable. Hence, one has to carefully choose the size of the clusters to obtain the desired accuracy.

Table II shows performance comparison of our results with the contemporary, competing work [12] for the recognition accuracies obtained for bi-class classification of human behaviour recognition. The results show that our method could better identify the true activity under investigation (given by the TP values) in case of five out of seven classes under consideration. Only for the class of 'pouring water', the results

TABLE I
Recognition Accuracy of Bi-Class Classification for Different Cluster Sizes

Behavior class	TP (%)	TN (%)	Clustering (%)
WK	72.81	82.35	5
	81.46	83.35	10
	85.73	83.31	**20**
	82.60	63.04	50
SUC	63.90	58.11	5
	69.00	56.78	**10**
	64.80	63.78	20
	63.60	57.78	50
DW	89.30	84.53	5
	100.00	80.20	**10**
	86.60	69.72	20
	83.33	70.50	50
SDC	72.00	64.68	5
	75.67	66.68	**10**
	73.10	64.56	20
	74.20	57.00	50
PWG	93.20	78.00	5
	94.10	75.00	**10**
	85.60	79.00	20
	83.40	70.67	50
GUB	62.70	61.33	5
	69.99	62.00	**10**
	65.80	62.44	20
	67.00	65.11	50
CS	74.60	71.22	5
	83.70	73.44	10
	85.00	67.11	**20**
	81.50	63.22	50

TABLE II
Comparison of Recognition Accuracy With Competing Work

Behavior class	Adopted approach		Competing work [12]	
	TP (%)	TN (%)	TP (%)	TN (%)
DW	**100.00**	80.20	**100.00**	83.34
CS	**85.00**	67.10	20.00	93.34
SUC	**69.00**	56.78	60.00	83.34
SDC	**75.67**	66.68	0.00	93.34
WK	**85.73**	83.31	40.00	70.00
PWG	94.01	75.00	**100.00**	80.00
GUB	**69.99**	62.00	60.00	66.67

TABLE III

COMPARISON OF RECOGNITION ACCURACY WITH OTHER STATE-OF-THE-ART APPROACHES

Method	Behavior class						
	DW	**CS**	**GUB**	**PWG**	**SDC**	**SUB**	**WK**
KNN	82.15%	75%	64.29%	73.17%	66.67%	62.16%	78.57%
Linear SVM	98.21%	76.83%	**78.50%**	91.14%	76%	64.84%	85.71%
CRC-RLS	97%	71.60%	30%	10%	**76.67%**	**80%**	78.33%
SRC	**100%**	39%	15%	8%	67%	70%	70%
Competing work [12]	**100%**	20%	60%	**100%**	0%	60%	40%
Our approach	**100%**	**85%**	69.99%	94.01%	75.67%	69%	**85.73%**

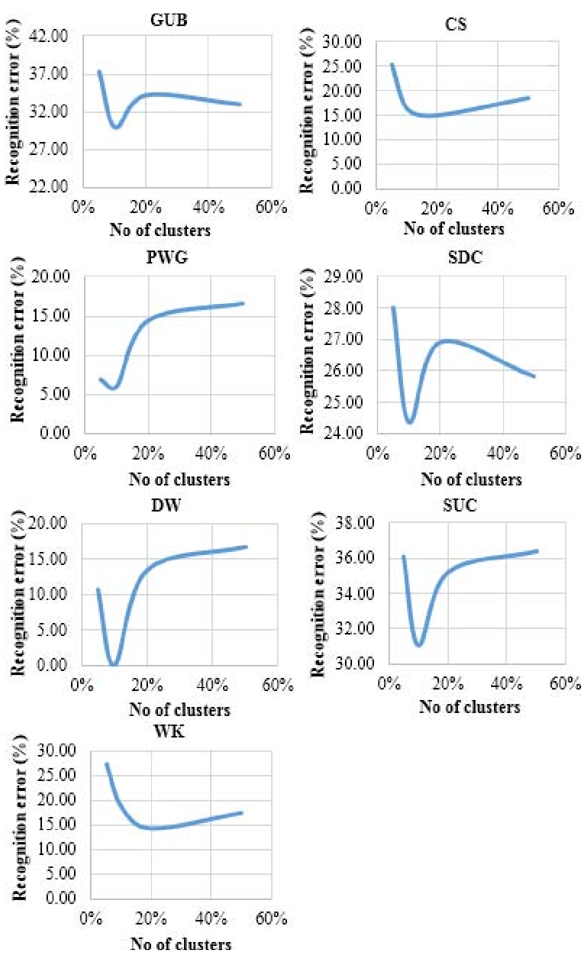

Fig. 3. Recognition error for bi-class classification of the ADL dataset.

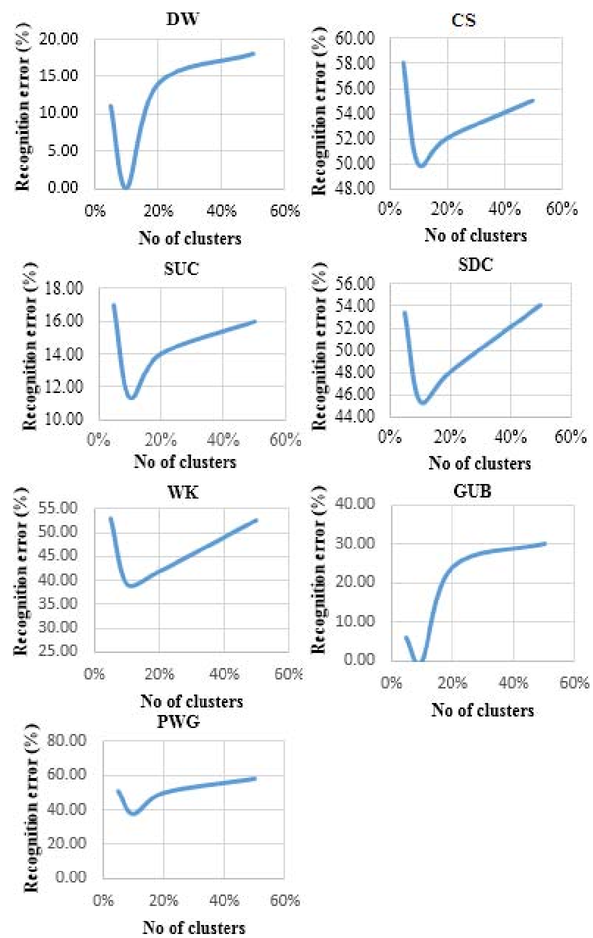

Fig. 4. Recognition error for multi-class classification of the ADL dataset.

reported in [12] are superior. Overall, these results aptly demonstrate the superiority of our proposed approach. Our proposed approach can also identify the class of sitting down on a chair and the class of climbing stairs with satisfactory accuracy although each of these two classes has more challenging composite and reiterated action respectively.

Table III shows the overall performance comparison of our results, for bi-class classification, with other state-of-the-art methods based on the recognition accuracies obtained for each class of human behaviour recognition. The results show varied performances for this challenging problem. Among those, our approach was able to achieve the most superior result for three classes, while no other method was able to achieve superior results for more than two classes. Also, our method was the most consistent one, as the worst accuracy result, for any class, in our case was 69%, while for other algorithms it was always less than 65% and even as low as 0% or 8% or 10%. These suitably validate the appropriateness of our proposed approach in the field of human behaviour classification.

Next, multi-class classification is performed to solve the human behavior classification problem as a true segregation problem, for the same benchmark ADL problem dataset.

TABLE IV

RECOGNITION ACCURACY OBTAINED FOR MULTI-CLASS CLASSIFICATION

Behavior class	**Recognition accuracy (%)**
DW	100.00
CS	50.00
SUC	88.67
SDC	54.67
WK	60.67
PWG	62.00
GUB	100.00

Table IV presents the recognition accuracy of multi-class classification based on our proposed approach. Solving the multiclass classification problem is considered as a much more difficult problem and yet the results that could be achieved are quite encouraging. Fig. 4 shows similar trends of error variations with variation in cluster sizes for multi-class classification as we obtained for bi-class classification.

VIII. DISCUSSION AND CONCLUSION

In this research, the suitability of the dictionary learning algorithm in conjunction with SRC has been explored in the field of human behavior recognition. Here it has also been established that the sparse representation based classification approach can be advantageously utilized for the purpose of human behavior recognition. The results achieved in this work demonstrate that our proposed approach yields high recognition accuracy with respect to other competing approaches. The proposed approach can be applied for both bi-class and multi-class behaviour classification/segregation purposes efficiently. Additionally, this work has shown that a suitable choice of cluster size has a significant impact on the identification accuracy of the algorithm. Hence, it is justified that our proposed approach opens another door of application for dictionary learning based classification approaches in the field of ADL recognition problem. Recognition of more challenging reiterated and composite actions in human behavior more efficiently may be a probable extension of our proposed approach in future.

REFERENCES

[1] C. Debes, A. Merentitis, S. Sukhanov, M. Niessen, N. Frangiadakis, and A. Bauer, "Monitoring activities of daily living in smart homes: Understanding human behavior," *IEEE Signal Process. Mag.*, vol. 33, no. 2, pp. 81–94, Mar. 2016.

[2] F. Erden, S. Velipasalar, A. E. Cetin, and A. Z. Alkar, "Sensors in assisted living: A survey of signal and image processing methods," *IEEE Signal Process. Mag.*, vol. 33, no. 2, pp. 36–44, Mar. 2016.

[3] R. Amstutz, O. Amft, B. French, A. Smailagic, D. Siewiorek, and G. Troster, "Performance analysis of an HMM-based gesture recognition using a wristwatch device," in *Proc. Int. Conf. Comput. Sci. Eng. (CSE)*, Vancouver, BC, Canada, Aug. 2009, pp. 303–309.

[4] K. Bousmalis, S. Zafeiriou, L. Morency, and M. Pantic, "Infinite hidden conditional random fields for human behavior analysis," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 1, pp. 170–177, Jan. 2013.

[5] E. Nazerfard, B. Das, L. B. Holder, and D. J. Cook, "Conditional random fields for activity recognition in smart environments," in *Proc. 1st ACM Int. Health Inform. Symp.*, Nov. 2010, pp. 282–286.

[6] O. Amft and G. Tröster, "Recognition of dietary activity events using on-body sensors," *Artif. Intell. Med.*, vol. 42, no. 2, pp. 121–136, Feb. 2008.

[7] B. Najafi, K. Aminian, F. Loew, Y. Blanc, and P. A. Robert, "Measurement of stand-sit and sit-stand transitions using a miniature gyroscope and its application in fall risk evaluation in the elderly," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 8, pp. 843–851, Aug. 2002.

[8] S. Chernbumroong, S. Cang, A. Atkins, and H. Yu, "Elderly activities recognition and classification for applications in assisted living," *Expert Syst. Appl.*, vol. 40, no. 5, pp. 1662–1674, Apr. 2013.

[9] V. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer, 1995.

[10] F. Martinez-Contreras, C. Orrite-Urunuela, H. Ragheb, S. A. Velastin, and E. Herrero-Jaraba, "Recognizing human actions using silhouette-based HMM," in *Proc. 6th IEEE Int. Conf. Adv. Video Signal Based Surveill. (AVSS)*, Genova, Italy, Sep. 2009, pp. 43–48.

[11] S. Hongeng, R. Nevatia, and F. Bremond, "Video-based event recognition: Activity representation and probabilistic recognition methods," *Comput. Vis. Image Understand.*, vol. 96, no. 2, pp. 129–162, Nov. 2004.

[12] B. Bruno, F. Mastrogiovanni, A. Sgorbissa, T. Vernazza, and R. Zaccaria, "Analysis of human behavior recognition algorithms based on acceleration data," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Karlsruhe, Germany, May 2013, pp. 1602–1607.

[13] H. Wang, M. M. Ullah, A. Klaser, I. Laptev, and C. Schmid, "Evaluation of local spatio-temporal features for action recognition," in *Proc. Brit. Mach. Vis. Conf.*, Sep. 2009, pp. 124.1–124.11.

[14] R. Messing, C. Pal, and H. Kautz, "Activity recognition using the velocity histories of tracked keypoints," in *Proc. 12th IEEE Int. Conf. Comput. Vis.*, Kyoto, Japan, Sep./Oct. 2009, pp. 104–111.

[15] F. I. Bashir, A. A. Khokhar, and D. Schonfeld, "Object trajectory-based activity classification and recognition using hidden Markov models," *IEEE Trans. Image Process.*, vol. 16, no. 7, pp. 1912–1919, Jul. 2007.

[16] Y.-S. Lee and S.-B. Cho, "Activity recognition using hierarchical hidden Markov models on a smartphone with 3D accelerometer," in *Proc. Int. Conf. Hybrid Artif. Intell. Syst. (HAIS)*, 2011, pp. 460–467.

[17] B. Bruno, F. Mastrogiovanni, A. Sgorbissa, T. Vernazza, and R. Zaccaria, "Human motion modelling and recognition: A computational approach," in *Proc. IEEE Int. Conf. Autom. Sci. Eng. (CASE)*, Aug. 2012, pp. 156–161.

[18] 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.

[19] J. Yang, J. Wang, and T. Huang, "Learning the sparse representation for classification," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2011, pp. 1–6.

[20] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.

[21] 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.

[22] K. Engan, S. O. Aase, and J. Husøy, "Multi-frame compression: Theory and design," *Signal Process.*, vol. 80, no. 10, pp. 2121–2140, Oct. 2000.

[23] R. Rosas-Romero, "Remote detection of forest fires from video signals with classifiers based on K-SVD learned dictionaries," *Eng. Appl. Artif. Intell.*, vol. 33, pp. 1–11, Aug. 2014.

[24] *UCI Machine Learning Repository. Dataset for ADL Recognition With Wrist-Worn Accelerometer Data Set*. Accessed: Oct. 9, 2016. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Dataset+for+ADL+Recognition+with+Wrist-worn+Accelerometer

Pubali De received the B.Tech. degree in electrical engineering from the Gurunanak Institute of Technology, West Bengal University of Technology, Kolkata, India, in 2012, and the M.Eng. degree in control system from Jadavpur University, Kolkata, in 2014. She is currently an Assistant Professor with Techno India Batanagar, Kolkata. Her current research interests include signal processing, instrumentation, and control.

Amitava Chatterjee (SM'13) received the B.E., M.E., and Ph.D. degrees in electrical engineering from Jadavpur University, Kolkata, India, in 1991, 1994, and 2002, respectively. He joined the Department of Electrical Engineering, Jadavpur University, in 1997, where he is currently a Professor. In 2003, he visited Saga University, Saga, Japan. From 2004 to 2005, he was with the University of Electro-Communications, Tokyo, Japan, where he was involved in the Japan Society for the Promotion of Science Post-Doctoral Fellowship for foreign researchers. In 2004 and 2009, he visited the University of Paris XII, Val de Marne, France, as an Invited Teacher. He has authored or co-authored over 125 technical papers, including 78 international journal papers. His current research interests include fuzzy-based nonlinear control, intelligent instrumentation, robotics, computer vision, image processing and pattern recognition, signal processing, and stochastic optimization techniques.

Anjan Rakshit received the M.E. and Ph.D. degrees in electrical engineering from Jadavpur University, Kolkata, India, in 1978 and 1987, respectively. He is a Retired Professor with the Department of Electrical Engineering, Jadavpur University. He has authored almost 70 technical papers. His current research interests include digital signal processing, Internet-based instrumentation, smart instrumentation, intelligent control, and design of real-time systems.