

# A hierarchical abnormal human activity recognition system based on R-transform and kernel discriminant analysis for elderly health care

Zafar Ali Khan · Won Sohn

Received: 15 January 2012 / Accepted: 29 August 2012 / Published online: 23 September 2012  
© Springer-Verlag 2012

**Abstract** A hierarchical human activity recognition (HAR) system is proposed to recognize abnormal activities from the daily life activities of elderly people living alone. The system is structured to have two-levels of feature extraction and activity recognition. The first level consists of R-transform, kernel discriminant analysis (KDA), *k*-means algorithm and HMM to recognize the video activity. The second level consists of KDA, *k*-means algorithm and HMM, and is selectively applied to the recognized activities from the first level when it belongs to the specified group. The proposed hierarchical approach is useful in increasing the recognition rate for the highly similar activities. System performance is analyzed by selecting the optimized number of features, number of HMM states and the number of frames per second to achieve maximum recognition rate. The system is validated by a novel set of six abnormal activities; falling backward, falling forward, chest pain, headache, vomiting, and fainting and a normal activity walking. Experimental results show an average recognition rate of 97.1 % for all the activities by using the proposed hierarchical HAR system.

**Keywords** Abnormal human activity recognition · Hierarchical model · Feature extraction · Kernel discriminant analysis · R-transform

**Mathematics Subject Classification (2000)** 68T45 · 68T10 · 62H30

---

Z. A. Khan · W. Sohn (✉)  
Kyung Hee University, Yongin, Republic of Korea  
e-mail: wsohn@khu.ac.kr

Z. A. Khan  
e-mail: zafaryaar@yahoo.com

## 1 Introduction

Human activity recognition (HAR) is an area of immense interest for researchers due to its applications in smart homes, elderly healthcare, sports and automated surveillance systems [1]. The growing population of elderly people result in the increasing number of elderly people depending on others for their daily care. The proportion of elderly people is expected to reach 22 % of the world population by 2050 [2]. Elderly people want to live independently but the increasing age adversely affects their health and brings certain disabilities in their lives e.g., weakness, vision and hearing impairments. The result is poor mobility and frequent falls causing severe injuries. Therefore, regular monitoring is required for elderly people while performing the daily life activities. Automated HAR systems are needed to ensure the safety of elderly people and allow them to live an independent life in their homes as long as possible. Automated HAR systems will improve the quality of life for elderly people living alone and decrease the cost of health care.

Humans find it difficult to efficiently perform the task of activity recognition for long durations because abnormal activities occur infrequently, and waiting for an abnormal activity turn out to be a boring task for humans. Automatic HAR system overcomes the limitations of human monitoring, provide continuous health monitoring and generate an alert in case of abnormal HAR. In healthcare, HAR systems can be used to recognize daily health monitoring of dementia patients, monitoring daily life activities of elderly people and recognizing abnormal activities e.g., falling, fainting [3–5]. The researchers commonly use optical sensor systems e.g., video sensors [1,6,7] and non-optical sensor systems e.g., accelerometer or motion sensors attached to different parts of human body for HAR [8–10]. In this paper, vision-based approach is used to recognize activities of elderly people at home due to practical reasons. The use of wearable sensors to recognize activities of elderly people has certain disadvantages; elderly people often forget to wear the wearable sensors, elderly people often feel it frustrating to wear sensors on different parts of the body for long durations, and it may restrain the movement of elderly people.

Abnormal activity recognition is an important research area in computer vision and pattern recognition due to its usefulness and cost-effectiveness in providing efficient care for elderly people [11]. Abnormal activity in this research is defined as a potentially hurtful activity that requires emergency medical help for elderly people. Most of the previous researches for abnormal HAR are related to falling activity due to the fact that falling is the highest cause of accidental deaths and paralysis in elderly people [12]. Two third of the accidents in the elderly people over the age of 65 years are falling accidents [13,14]. This paper defined and generated a novel set of abnormal activities consisting of falling backward, falling forward, fainting, vomiting, headache, and chest pain. To get realistic idea of frequent emergency situations faced by elderly people, abnormal activities are defined and selected after reviewing medical journals and consulting with doctors. The literature review for abnormal activities include; falling [15–19], vomiting [20,21], chest pain [22,23], headache [24,25], and fainting activities [26,27].

## 2 Related work

The majority of abnormal HAR systems presented in the past considered falling or fainting activities only [28–30]. In [28], abnormal and normal activities including falling forward, falling backward, and walking are recognized by the combination of eigenspace technique, and integrated time and motion images. Multiclass support vector machine is used for activity recognition. Only the best view point is utilized for each activity in order to minimize ambiguities, whereas in real world the activities can be performed from different view point. In [29], vision based elderly activities are recognized in a home environment. The features from binary silhouettes are extracted by the best fit approximated ellipse around the human silhouette, projection histogram of segmented silhouette, and change of head position tracking with time. Neural network classifier is used to recognize the activities. The use of ellipse around the silhouette to exploit orientation, length of major and minor axes features for activity recognition is an interesting approach but it is not helpful in distinguishing the highly similar activities in our dataset. In [30], video based HAR system is presented to recognize the falling activity as an abnormal activity. The system consists of fall detection and fall confirmation steps. Aspect ratio, and the horizontal and vertical gradient values of the person are used for fall detection and the fall angle is used for fall confirmation. A two state finite state machine is used to continuously monitor human behavior. For a standing or walking person, a vertical angle between  $45^\circ$  to  $90^\circ$  is assumed around the person using the centroid of a bounding box. When the angle is less than  $45^\circ$ , then the activity is considered as a falling activity. This approach is not feasible for vomiting and fainting activities in our dataset due to their falling like sequences. Therefore, vomiting and fainting activities will be wrongly recognized as falling activities most of the time. In [31], human activities resting, walking, walk upstairs, walk downstairs, running, and cycling are recognized by using tri-axial accelerometer sensor embedded in smart phone. A three level hierarchical model is used to recognize static and dynamic activities. In first part spectral entropy, autoregressive coefficients, signal magnitude area with tilt angle are used to differentiate static and dynamic activities. LDA is used to increase the discrimination between the activities. Three different artificial neural networks (ANN) are used for activity recognition at three levels of the hierarchical model i.e., static, dynamic (upper body) and dynamic (lower body). The hierarchical approach used in our system achieves excellent recognition rate for the highly similar activities within less number of classification steps. In [32], wearable tri-axial accelerometer sensors are used to recognize different types of walking activity e.g., level walk, down stairs walk and upstairs walk. PCA, ICA and wavelet transforms are used for the feature extraction and multilayer perceptron neural networks are used for activity recognition. Our research is focused to recognize the activities of elderly people. They feel uncomfortable and sometimes forget to wear the wearable sensors. Therefore, video based HAR system is preferable for the activity recognition of elderly people at home. In [33], human walking patterns are recognized from three directions by using 3-D wavelet transform and kernel LDA. The recognition rate is not very promising for a simple activity like walking. Also, the use of 3-D wavelet transform increases the complexity of the system. In [34], a HAR system is presented for the behavior recognition of elderly people in a nursing home.

The system is based on the hierarchical HMM architecture called hierarchical context HMM. The behavior recognition process is divided into three steps. In the first step, behaviors with spatial differences are recognized. In the second step, the activities based on selected behavior in the first step are recognized and in the third step, the activities with temporal differences are recognized to infer a particular behavior. In [35], human activities from Weizmann dataset are recognized by using R-transform with Fourier-Mellin transform. PCA is used to reduce dimensions and earth mover's distance is used to recognize the activities. Rotation in space will increase the ambiguities in different activities because in real world human movements are usually performed in horizontal plane, therefore in-plane rotation will be more realistic in HAR systems. We produced the dataset with different horizontal in-plane rotations ( $90^\circ$ ,  $-90^\circ$ ,  $45^\circ$ ,  $-45^\circ$ ). In [36], HAR system is proposed by using shape and motion features. Optical flow features are used to represent motion and eigen-shape features are used to represent shape features. An activity is represented by a HMM for each view direction for view invariance and voting-based approach is used for activity recognition. The use of an HMM for each view angle increases the complexity of the system. Our system use only one HMM for each activity performed from different view angles in order to increase efficiency and reduce the complexity of the classification task.

In [37], multiple hierarchical HMMs are integrated to recognize the behavior of multiple persons. Joint probabilistic data association filters are used for data association and Rao-Blackwellised particle filters are used for approximate inference. Improved recognition is achieved as compared to the kalman filter approach. The system will not be suitable for the highly similar activity sequences presented in our dataset. In [38], layered HMMs are utilized to recognize real time human activities in office environment e.g., phone conversation, face to face conversation, presentation, other activity, nobody around, and distant conversation. Each layer of hierarchical HMM model is independently trained to use different levels of abstraction and time granularities linked to different levels of human behaviours. The highest likelihood for each layer is used as an input to the next level HMMs and all the activities are recognized. In [39], binary silhouettes are used to recognize walking, jumping, crouching and climbing activities by posture matching and fuzzy rule based reasoning. The features are transformed and recognized in canonical space. Fuzzy rule approach combine temporal sequences for activity recognition but also increase complexity of the system. In [40] multiple cameras are used to generate and train the 3-D visual hulls from silhouettes. Recognition is performed by projecting each visual hull to a 2-D image which represents the best match for a given silhouette. The problems with this approach is the high computational cost in generating the 3-D visual hulls and searching for a 2-D image that best matches the silhouette. The use of multiple cameras to generate the visual hulls also increases the cost and complexity of the system. Our approach minimize the cost and complexity of the system by using a single video camera to recognize the activities performed from different view angles.

Our main contribution is related to the hierarchical abnormal HAR system based on R-transform and KDA features having a two level feature extraction and activity recognition process to recognize the complex and highly similar abnormal activities performed from different view angles. R-transform is used to extract symmetric, scale

and translation invariant features. KDA is implemented on the R-transform features to increase the discrimination between different classes of highly similar activities based on non-linear KDA representations. A novel abnormal human activity dataset is defined and produced for the validation of the proposed system.

In this paper, a basic HAR system is primarily used to recognize all the activities and identify the highly similar activities. Then hierarchical HAR system is implemented. First level of hierarchical HAR system group the highly similar activities from different view angles and use R-transform and KDA methods for features extraction from silhouette sequences. The second level of hierarchical HAR system is applied to grouped activities only. KDA is applied again to further increase discrimination between the highly similar activities. The system is evaluated on a novel dataset of six abnormal activities; falling backward, falling forward, chest pain, headache, vomiting, and fainting.

The rest of the paper is organized as follows. Section 3 presents the overall system design, including problem statement, dataset generation and the system model. Section 4 describes feature extraction and activity recognition methodologies. Section 5 presents experimental results and analysis for the basic and hierarchical HAR systems. Section 6 provides conclusion including limitations and the future research directions.

### 3 System design

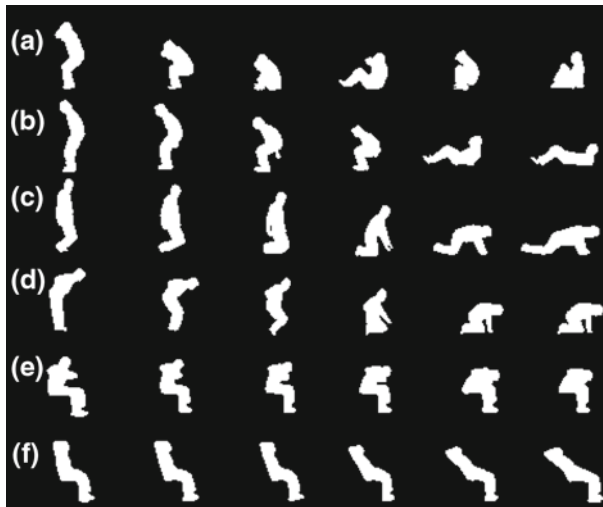
#### 3.1 Problem statement

The problems encountered in video based HAR systems include inherent similarities in the human postures, and the symmetric, scale, translation variations due to different persons performing the activities with changing view angle and distance to the video camera. Figure 1a–f illustrates the similarity in the postures of some activities, particularly between falling forward/vomiting, and between falling backward/fainting activities.

We present a hierarchical HAR system that utilize R-transform and KDA algorithms to solve the above mentioned problems. For the first problem, R-transform is applied to extract geometrical (symmetric, scale and translation) invariant features. For the second problem, KDA is applied to extract non-linear features. KDA increase the variation between different classes of activities and decrease the variation within the same class of activity using non-linear approach. However, there is a limit observed to improve the recognition accuracy by KDA. Therefore, in this paper we present a hierarchical HAR system to further improve the recognition rate by applying KDA in a hierarchical approach and improve the discrimination between different classes of highly similar activities.

#### 3.2 Dataset generation

The activities recognized by vision-based systems can be categorized into static and dynamic activities based on the movements involved. Static activities may involve



**Fig. 1** Selected postures for the six abnormal activities: **a** fainting, **b** falling backward, **c** falling forward, **d** vomiting, **e** chest pain, and **f** headache

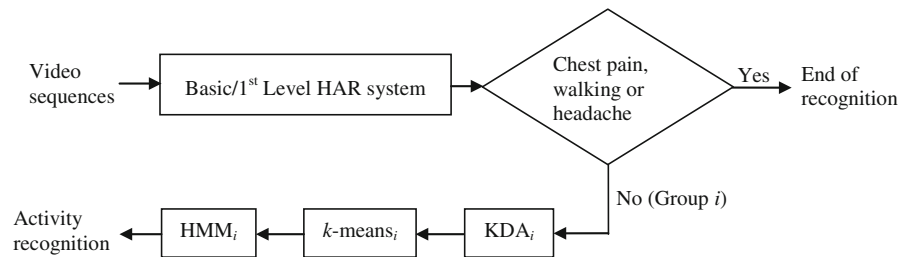
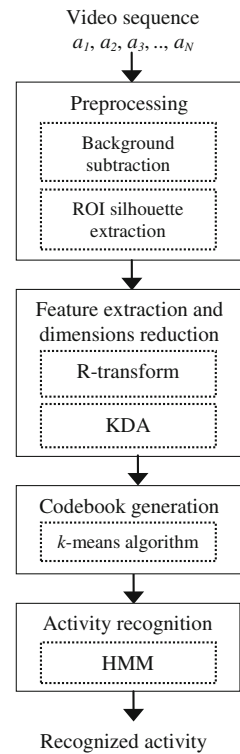
minor movement whereas dynamic activities involve major movement of the human body. Static activities include standing, sitting, lying, and dynamic activities include walking, running and falling. In this study, we concentrate on dynamic activities because abnormal activities usually involve the major movements of human body. We produced a dataset of abnormal activities by considering the daily life activities of elderly people at home as shown in Fig. 1.

### 3.3 System model

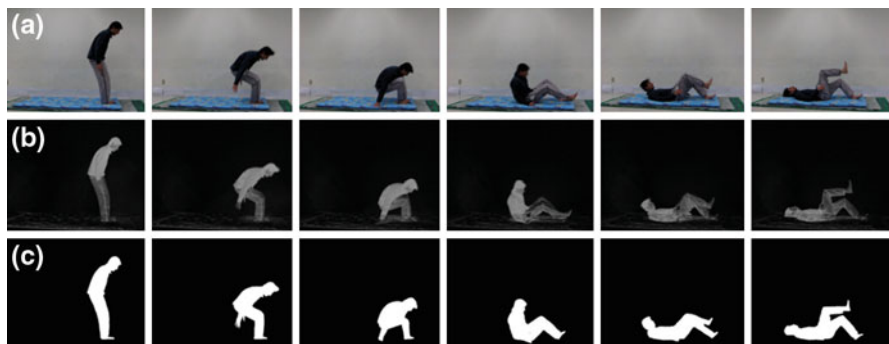
The selected numbers of uniformly sampled silhouettes are used to represent an activity sequence and silhouettes are normalized. In the basic HAR system, R-transform is used for symmetric, scale and translation invariant feature extraction and dimensions reduction. KDA is applied on R-transformed features to increase the discrimination among different classes of highly similar activities. Symbol sequences are generated from extracted features by implementing  $k$ -means clustering algorithm. These sequences of symbols are utilized by discrete HMM classifier for the training and recognition of activities. The system recognizes human activities with good recognition rate but some activities with highly similar posture sequences still have lower recognition rate which should be further improved for a better HAR system. Figure 2 illustrates the block diagram of the basic HAR system model.

A hierarchical HAR system is recommended to further improve the discrimination between the different classes of highly similar activities; falling forward/vomiting, and falling backward/fainting. Figure 3 shows the block diagram for the proposed hierarchical HAR system model.

The basic HAR system and the first level of the hierarchical HAR system have same models. The basic HAR system actually becomes the first level of the hierarchical HAR

**Fig. 2** Block diagram of the basic HAR system**Fig. 3** Block diagram of the hierarchical HAR system

system when the system is applied on the highly similar activities grouped together in different groups. The proposed hierarchical HAR system improved discrimination between different classes of activities. The basic level of hierarchical HAR system is applied to all the activities. Chest pain, headache and walking are recognized without difficulty but fainting/falling backward, and vomiting/falling forward activities result in misclassifications due to high similarity in postures. Therefore, these activities are grouped together. The second level of hierarchical HAR system is applied to the grouped activities only.  $(KDA)_i$ ,  $(k\text{-means})_i$  and  $(HMM)_i$  of second level of hierarchical HAR system are configured for each group  $i$ .



**Fig. 4** Preprocessing steps to extract binary images: **a** sample frames from an original falling backward sequence, **b** background subtracted images, and **c** extracted binary silhouettes

## 4 Feature extraction and activity recognition

This section describes the preprocessing, feature extraction and activity recognition methodologies, including R-transform, KDA, *k*-means and HMM.

### 4.1 Preprocessing

The preprocessing process is performed on raw video data to reduce noise and redundancy. The sequences of video frames are extracted from each activity and background subtraction process is performed to get foreground region of interest. The background subtracted images are then converted to binary based on threshold values determined by experiments using algorithm [41]. The shape vectors are transformed to zero mean prior to applying the feature extraction algorithms. Figure 4 shows the preprocessing steps to extract binary silhouettes from a falling backward activity.

### 4.2 R-transform

The normalized R-transform is symmetric, scale and translation invariant. It has low computational complexity, and it is robust to frame loss and noisy images [42]. R-transform is a shape descriptor defined on the Radon transform. The Radon transform of a silhouette  $f(x, y)$  is defined as

$$T_{Radon}(\rho, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - \rho) dx dy, \quad (1)$$

where  $\rho \in [-\infty, \infty]$  represents distance from a Radon line to origin given by  $\rho = x \cos \theta + y \sin \theta$ ,  $\theta \in [0, \pi)$  is an inclination angle along the line to compute projections, and  $\delta(\cdot)$  is Dirac delta function. R-transform is defined as the integral of the squared values of Radon transform along the Radon line at a certain angle  $\theta$ , and it is given as

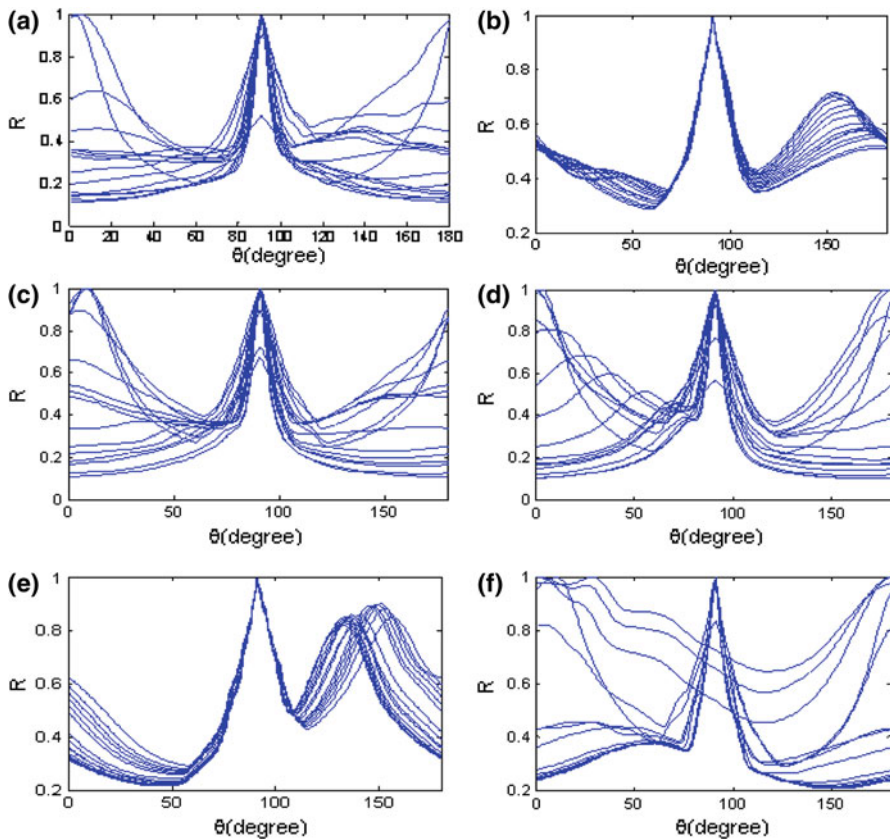


$$T_R(\theta) = \int_{-\infty}^{\infty} T_{Radon}^2(\rho, \theta) d\rho. \quad (2)$$

R-transform captures dominant features from silhouette sequences of different activities and compact the 2-D shape features in 180 dimensions in the range of  $0^\circ$ – $179^\circ$  [43]. The dimensions for each silhouette are reduced to  $1 \times 180$ . Figure 5 shows the plots of normalized R-transform for a sequence of 15 uniformly sampled frames for falling backward, chest pain, fainting, falling forward, headache, and vomiting activities. High similarities in falling backward/fainting and falling forward/vomiting activities can be observed.

#### 4.3 Kernel discriminant analysis

KDA is a non-linear extension of LDA and obtains non-linear discriminating features by the kernel technique [44]. KDA maps the transformed data to the high dimensional



**Fig. 5** The R-transform plots for a sequence of 15 frames from **a** falling backward, **b** chest pain, **c** fainting, **d** falling forward, **e** headache, and **f** vomiting activities

feature space by non-linear mapping [45]. In KDA,  $S_B^\varphi$  represents the variation between different activities and  $S_W^\varphi$  represents the variation within similar activities.  $S_B^\varphi$  and  $S_W^\varphi$  are defined as [45]

$$S_B^\varphi = \frac{1}{n} \sum_{i=1}^c n_i (\mu_i^\varphi - \mu^\varphi) (\mu_i^\varphi - \mu^\varphi)^T \quad (3)$$

$$S_W^\varphi = \frac{1}{n} \sum_{i=1}^c \sum_{x \in X_i} (\varphi(\mathbf{x}) - \mu_i^\varphi) (\varphi(\mathbf{x}) - \mu_i^\varphi)^T \quad (4)$$

where  $\mu_i^\varphi$  is the centroid of  $i$ th class and  $\mu^\varphi$  is the global centroid.  $S_B^\varphi$  should to be maximized and  $S_W^\varphi$  should to be minimized to achieve better activity recognition.

For the basic HAR system, KDA is applied to the features of all activities obtained by R-transform. KDA further improves the discrimination between different classes of abnormal activities. At the first level of the hierarchical system, KDA is applied to two groups (falling forward/vomiting, and falling backward/fainting) and three individual activities (headache, chest pain, walking). At the second level of the hierarchical system, KDA is implemented for each group and results in  $1 \times 1$  dimensional feature vector for each silhouette.

#### 4.4 Activity recognition

The extracted features after KDA are transformed into discrete symbol sequences and represented in the form of codebook by  $k$ -means algorithm.  $k$ -means is a simple center based clustering algorithm [46]. A codebook size of 32 is used to generate the discrete symbols for the basic system and the first level of the hierarchical system. For the second level of hierarchical system, the 16-sized codebook provided the optimal results. The extracted features for each activity sequence are transformed into corresponding symbol sequences by assigning the closest codeword from the codebook between 1 and 32 or 1 and 16. After  $k$ -means implementation, each silhouette is represented by a symbol and each silhouette sequence is represented by a sequence of symbols. The symbols generated by  $k$ -means are used by HMM for activity training and recognition [47]. A generic HMM model can be explained by the set of parameters represented as

$$\lambda = \{\mathbf{A}, \mathbf{B}, \boldsymbol{\pi}\} \quad (5)$$

where  $\lambda$  is the HMM model,  $\mathbf{A}$  is the state transition probability matrix between hidden states,  $\mathbf{B}$  is the emission probability matrix of the observation symbols, and  $\boldsymbol{\pi}$  is the initial state transitions probability [47]. To test an activity sequence, its symbol sequence is matched with the observed symbol sequences  $O = O_1, O_2, \dots, O_T$  from the trained HMM models for each activity, and the one with the highest likelihood is selected to be a recognized activity as

$$decision = \arg \max_{i=1,2,\dots,N} \{P_i\}, \quad (6)$$

$$P_i = (Pr(O|\lambda_i)) \quad (7)$$

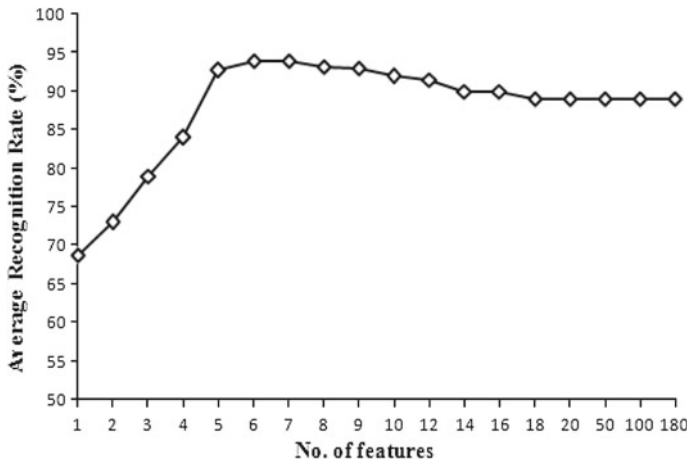
where  $P_i$  is the likelihood of the  $i$ th HMM  $\lambda_i$ , and  $N$  represents the number of activities.

In this research, we trained a separate discrete HMM for each activity and computed the overall likelihood probability for each activity. In the testing phase, the likelihood probability of the testing sequences are matched with the likelihood probabilities of the trained HMM for each activity, and the one with the nearest match is considered to be the most probable representation of the testing activity. Figures 7 and 11, shows the results of the preliminary stage of the experimental setup to acquire the optimal parameters for the proposed system. The number of HMM states are not dynamically changed during the performance of hierarchical HAR system. The four state HMM model provide the optimal results for the basic level HAR system and three state HMM model provide the optimal results in the hierarchical HAR system. The discrete HMM model in this research does not represent an activity by the particular HMM state but compute the overall likelihood probabilities for training of each activity. For the hierarchical model, the highly similar activities are grouped together which simplifies the task of classifier. The overall likelihood probabilities with three states HMM model provide the highest recognition rate as shown in Fig. 11. The four state HMM model can be used as it also provides the maximum recognition results as shown in Fig. 11. We selected the three state HMM model for the hierarchical HAR system to minimize cost and complexity of the system.

## 5 Experimental results

The video activity dataset is generated for the six abnormal activities (falling backward, falling forward, chest pain, headache, vomiting, fainting) and a normal activity (walking) based on the daily life healthcare problems often encountered by elderly. The activities are performed from different view angles ( $90^\circ$ ,  $-90^\circ$ ,  $45^\circ$ , and  $-45^\circ$ ). A video camera is used to capture the activity videos with a frame size of  $320 \times 240$  pixels at 25 fps (frames per second). Ten persons (six males, four females; age  $35.5 \pm 6.8$  years; range 28–55 years) performed the activities and repeated each activity fifty times. The activities are performed in a studio apartment under good lighting conditions and limited background variations. Leave one out cross validation approach is implemented for activity recognition. For each testing activity, the activities from nine persons are used for training and the remaining one person activities are used for testing. The testing person activities are then replaced in the dataset and next person activities are selected for testing manually. This process is repeated until all the persons in the dataset participated in the testing process. The final recognition rate for each activity is represented by averaging the recognition results obtained for each activity. The recognition results are presented in Tables 1–5.

Prior to implementing basic and hierarchical HAR systems, extensive experiments are performed to select the optimal number of features, optimal number of HMM states and optimal number of frames to achieve the maximum recognition rate. The frame rate per second is selected by segmenting all the video activities into sequences



**Fig. 6** Average recognition rate versus the number of features for basic HAR system

of approximately 3 second. Then, we uniformly select different number of frames per second for each sequence and perform the recognition process. The numbers of frames that provide the maximum recognition rate for the basic and hierarchical HAR systems are used in the system as shown in Figs. 8 and 9.

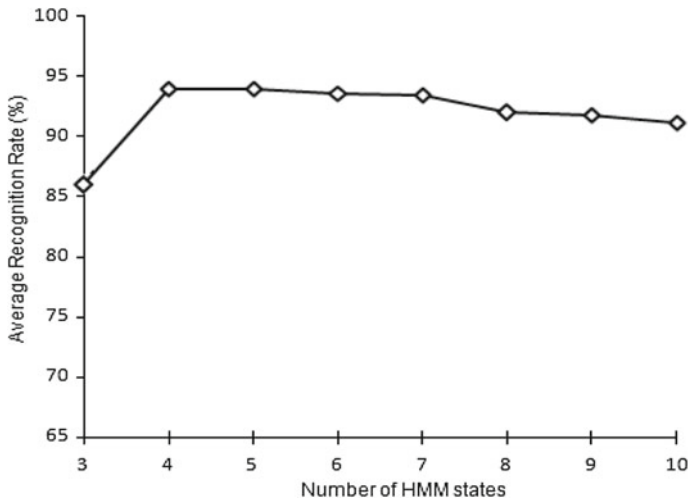
### 5.1 Basic HAR system

In the basic HAR system, R-transform is applied on the binary silhouettes and KDA is applied on the R-transformed features. *k*-means algorithm generates discrete symbol sequences and HMM is used for the training and recognition of activities. The activities (headache, chest pain and walking) are well-recognized by the basic HAR system but the activities (falling forward/vomiting, and falling backward/fainting) are not well-recognized due to the remaining confusions between their sequences. The maximum recognition rate is achieved when the six features are selected for KDA as shown in Fig. 6.

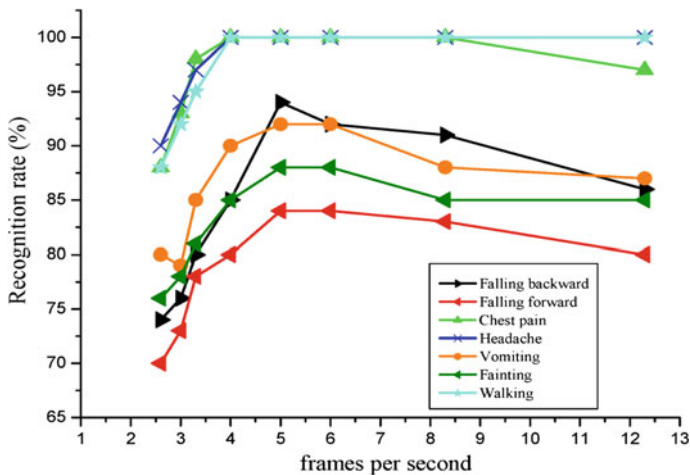
The four-state HMM model is selected for the training and recognition of activities after experimenting with different states HMM models as shown in Fig. 7.

Each individual activity provided the maximum recognition rate for different number of frames per second. The overall maximum recognition rate is achieved when the five frames per second are used as shown in Fig. 8.

The average recognition rate using basic HAR system for all activities is 92.8 % as shown in Table 1. The recognition rate for headache, chest pain and walking is 100 % because these activities are clearly differentiated while falling forward/vomiting and falling backward/fainting have a lower recognition rate due to the high similarities in their postures as shown in Table 1. The between-class variation for these activities should be increased in order to improve the recognition rate.



**Fig. 7** Average recognition rate versus number of states for HMM using the basic HAR system



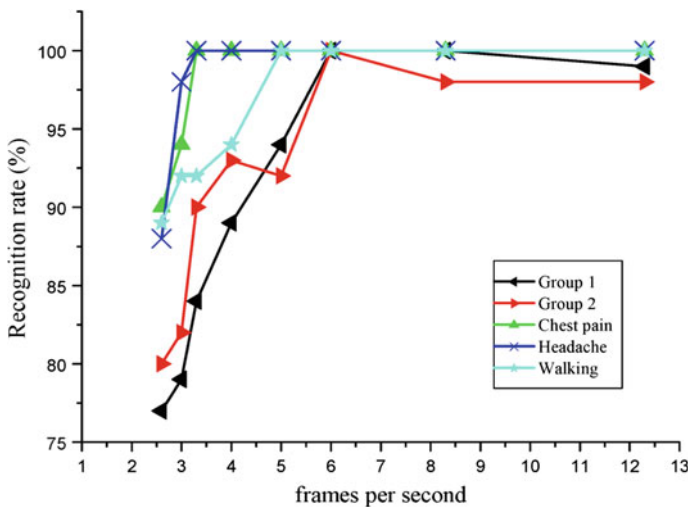
**Fig. 8** Recognition rate versus number of frames per second for basic HAR system

## 5.2 Hierarchical HAR system

The highly similar activities from basic HAR system are grouped in the first level of hierarchical HAR system. Group 1 consists of falling forward/vomiting activities and group 2 consists of falling backward/fainting activities. First level of the hierarchical HAR system consisting of KDA,  $k$ -means and HMM is applied to the group 1, 2, headache chest pain and walking activities. For hierarchical system, six frames per second yields maximum average recognition rate for all the activities as shown in Fig. 9.

**Table 1** Confusion matrix for the basic HAR system

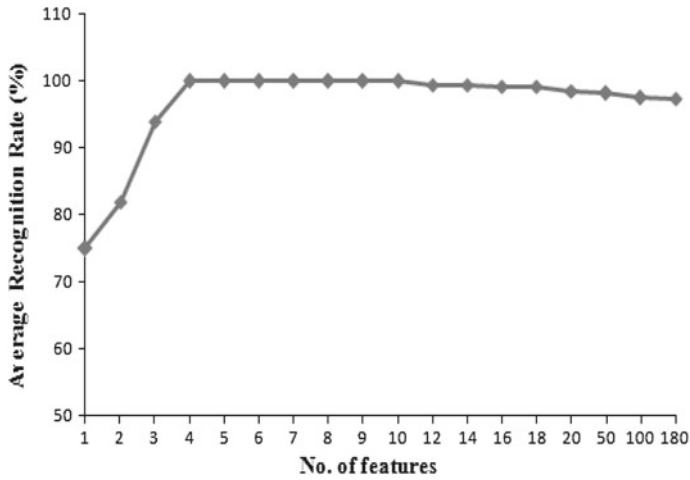
Activity	Recognition rate (%)
Falling backward	91
Falling forward	84.5
Chest pain	100
Headache	100
Vomiting	90
Fainting	84
Walking	100
Average recognition rate	92.8

**Fig. 9** Recognition rate versus the number of frames per second for hierarchical HAR system

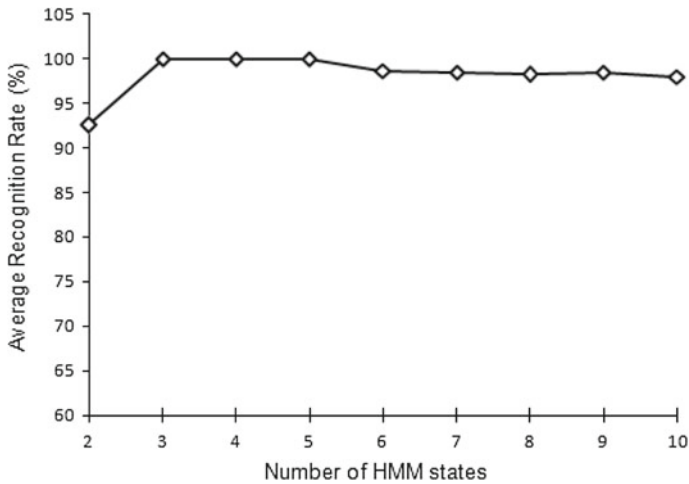
For the hierarchical HAR system,  $1 \times 4$  dimensional feature vector for each silhouette achieved the maximum recognition rate. Figure 10 shows average recognition rate for different number of features. In the hierarchical HAR system, three-state HMM model is selected after experimenting with different number of states for HMM as shown in Fig. 11.

The average recognition rate of 100 % is achieved for the first level of hierarchical HAR system as shown in Table 2.

For group 1 or 2 activities, the second level of hierarchical HAR system is applied, which includes KDA,  $k$ -means and HMM as shown in Fig. 3. The second level of hierarchical HAR system is configured for each group  $i$ . It consists of  $(KDA)_i$ ,  $(k\text{-means})_i$  and  $(HMM)_i$ . The KDA for each group results in  $1 \times 1$  dimensional feature vector for each silhouette. Average recognition rate of 95 % is achieved for group 1 activities as shown in Table 3. Average recognition rate of 94.8 % is achieved for group 2 activities as shown in Table 4.



**Fig. 10** Average recognition rate versus the number of features for hierarchical HAR system



**Fig. 11** Average recognition rate versus the number of states for HMM using hierarchical HAR system

**Table 2** Confusion matrix for the first level of hierarchical HAR system

Activity	Recognition rate (%)
Group 1	100
Group 2	100
Chest pain	100
Headache	100
Walking	100
Average recognition rate	100

**Table 3** Confusion matrix for the second level of hierarchical HAR system (group 1)

Activity	Recognition rate (%)
Vomiting	97.5
Falling forward	92.5
Average recognition rate	95.0

**Table 4** Confusion matrix for the second level of hierarchical HAR system (group 2)

Activity	Recognition rate (%)
Fainting	95.0
Falling backward	94.5
Average recognition rate	94.8

**Table 5** Average recognition rate for all the activities using hierarchical HAR system

Activity	Falling backward	Falling forward	Chest pain	Headache	Vomiting	Fainting	Walking
Average recognition rate (%)	94.5	92.5	100	100	97.5	95	100
Overall recognition rate (%)	97.1						

Table 5 shows overall recognition rate of 97.1 % for all the activities using hierarchical HAR system.

The high similarities in the postures of different activities are effectively reduced by the proposed hierarchical HAR system. The experimental results yield high recognition rate for all the activities using the hierarchical approach.

The basic HAR system is utilized as an inspection system, before implementing the hierarchical HAR system to analyze the different activities with high similarities or misclassifications. The basic HAR system is part of the procedure but basic and hierarchical HAR systems are different systems. In hierarchical HAR system, the basic HAR system becomes the first level of hierarchical HAR system. The highly similar activities detected in the basic level are grouped in the first level of the hierarchical HAR system to analyze if we can achieve higher recognition for the highly similar activities.

A vision-based, hierarchical abnormal HAR system is presented utilizing R-transform and KDA to extract geometrical invariant and discriminating features from different view angles. The two level hierarchical feature extraction and activity recognition approach increased the discrimination between the highly similar activities and improved the recognition accuracy. The advantage of our proposed approach is that simple activities can be recognized in the basic level HAR system only, without the need to go to the first or second level of the hierarchical system. The more complex activities are recognized in the second level of hierar-



chical HAR system. The effectiveness of our proposed hierarchical HAR system is demonstrated by recognizing some complex and highly similar abnormal activities; falling forward/vomiting and falling backward/fainting.

The aim of this research is to present a HAR system using efficient feature extraction methodologies to improve the health care of elderly people living alone. In case of abnormal HAR, an alert will be generated for providing emergency care to save the lives of elderly people. The feasibility of the system is established from methodological and performance view points.

## 6 Conclusion

In this paper, we presented a novel vision-based, hierarchical HAR system to recognize the abnormal activities of elderly people. The system is validated by using six abnormal activities and a normal activity selected from the daily life activities of elderly people. The system successfully distinguished the activities with high recognition rate even for complex and highly similar activities in our abnormal human activity dataset. The aim is to build an automatic HAR system for monitoring the daily activities of elderly people in order to provide them urgent help in potential abnormal situations. This research will particularly provide improved health care for elderly people living alone at home. The excellent recognition results show that our proposed system has the potential to be extended to real-life health care applications for elderly care at home. The research will improve the quality of life of elderly people and provide them confidence to live an independent life.

The limitations of video based HAR system include the privacy concerns raised by some elderly people. In our approach, we will minimize this concern by extracting the silhouettes from video activities directly without human intervention and only the recognition result will be transmitted to the paramedical staff for help. RGB image sequences are not utilized and silhouette are used for feature extraction because we do not have to recognize a particular person as in face recognition systems. In fact, RGB images will further increase the complexity of the system. The proposed HAR system focus on recognizing abnormal activities from the daily life activities of elderly people and achieved high recognition rate even for the complex abnormal activities from different view angles and high similarities.

In future, datasets from real elderly people will be used. The dataset used in this research has good quality silhouettes because the activities are performed in a room with good lighting conditions and limited background variations. To analyze the further effects of noise, datasets with increased noise will be generated and utilized. The activities from more view angles will be added to the dataset. The binary silhouettes usually show confusion in some postures e.g., when the activities are performed in front of the camera ( $0^\circ$  view angle). Then the features of hand, head or some other part of the body cannot be easily distinguished from the body silhouettes. This creates ambiguities in differentiating the particular postures and may results in decreasing the recognition rate of the system. This problem will be addressed by extracting 2-D stick features for the activities [48]. The system will be implemented in real time for long term activity recognition with the increased number of abnormal and normal activities

in the dataset. It will be interesting to analyze the robustness and reliability of the system in real time.

## References

1. Moeslund T-B, Hilton A, Krüger V (2006) A survey of advances in vision based human motion capture and analysis. *Comput Vis Image Underst J* 104(2):90–126
2. UN population estimates and projections report (2010) World population prospects, the 2010 revision. <http://esa.un.org/unpd/wpp/>. Accessed Mar 2012
3. Chen D, Bharucha A-J, and Wactlar H-D (2007) Intelligent video monitoring to improve safety of older persons. In: Proceedings of the IEEE international conference on engineering in medicine and biology society (EMBS), 22–26 Aug 2007, pp 3814–3817
4. Fleck S, Strasser W (2008) Smart camera based monitoring system and its application to assisted living. *IEEE J* 96(10):1698–1714
5. Cambul G, Sahin M-R, and Karshgail M-E (2006) Automatic fall detection in real time video based applications. In: Proceedings of IEEE signal processing and communications applications, 17–19 Apr 2006, pp 1–4
6. Turaga P, Chellappa R, Subrahmanian V-S, Udrea O (2008) Machine recognition of human activities: a survey. *IEEE Trans Circuits Systems Video Tech* 18(11):1473–1488
7. Aggarwal J-K, Cai Q (1999) Human motion analysis: a review. *Comput Vision Image Underst J* 73(3):428–440
8. Aminian K, Robert P, Buchser E, Rutschmann B, Hayoz D, Depairon M (1999) Physical activity monitoring based on accelerometry: validation and comparison with video observation. *Med Biol Eng Comput J* 37(3):304–308
9. Fahrenberg J, Foerster F, Smeja M, Muller W (1997) Assessment of posture and motion by multichannel piezoresistive accelerometer recordings. *Psychophysiol J* 34(5):607–612
10. Foerster F, Fahrenberg J (2000) Motion pattern and posture: correctly assessed by calibrated accelerometers. *Behav Res Methods Instrum Comput J* 32(3):450–457
11. Chappell N-L, Dlott B-H, Hollander M-J, Miller J-A, McWilliam C (2004) Comparative costs of home care and residential care. *The Gerontol J* 44(3):389–400
12. Tinetti M-E, Doucette J, Claus E, Marottoli R (1995) Risk factors for serious injury during falls by older persons in the community. *Am Geriatr Soc J* 43(11):1214–1221
13. Dunn J-E, Rudberg M-A, Furner S-E, Cassel C-K (1992) Mortality, disability, and falls in older persons: the role of underlying disease and disability. *Am Public Health J* 82(3):395–400
14. McCaig L-F, Burt C-W (2001) National hospital ambulatory medical care survey: 1999 emergency department summary. *Adv Data J* 320:1–34
15. Rougier C, Meunier J, Arnaud A-S, Rousseau J (2011) Robust video surveillance for fall detection based on human shape deformation. *IEEE Trans Circuits Systems Video Tech J* 21(5):611–622
16. Wu G (2000) Distinguishing fall activities from normal activities by velocity characteristics. *Biomech J* 33(11):1497–1500
17. Lin C-W, Ling Z-H, Chang Y-C (2007) Compressed domain fall incident detection for intelligent homecare. *VLSI Signal Process Systems J* 49(3):393–408
18. Lee T, Mihailidis A (2005) An intelligent emergency response system: preliminary development and testing of automated fall detection. *Telemed Telecare J* 11(4):194–198
19. Sadigh S, Reimers A, Andersson R, Laflamme L (2006) Falls and fall related injuries among the elderly: a survey of residential-care facilities in a Swedish municipality. *J Commun Health* 29(2):129–140
20. Ingram D-A, Fulton R-A, Portal R-W, Aber C-P (1980) Vomiting as a diagnostic aid in acute ischaemic cardiac pain. *Br Med J* 281(6241):636–637
21. Talley N-J (2007) Functional nausea and vomiting. *Aust fam physician J* 36(9):6694–6697
22. Ruseva A (2005) Laboratory diagnosis of acute myocardial infection. *Trakia J Sci* 3(1):8–14
23. Ruppert S-D (1997) Chest pain: a diagnostic dilemma. *Internet J Adv Nurs Pract* 1(1)
24. Blanchard E-B, Andrasik F, Arena J-G, Neff D-F, Jurish S-E, Teders S-J, Saunders N-L, Pallmeyer T-P, Dudek B-C, Rodichok L-D (1984) A bio-psycho-social investigation of headache activity in a chronic headache population. *Headache* 24(2):79–87
25. Juang K-D, Wang S-J, Fuh J-L, Lu S-R, Su T-P (2000) Comorbidity of depressive and anxiety disorders in chronic daily headache and its subtypes. *Headache J* 40(10):818–823

26. Blair P-G (2007) The fainting phenomenon: understanding why people faint and what to do about it. Blackwell Press, Malden
27. Olisky M, Murray J (1990) Dizziness and fainting in the elderly. *Emerg Med Clin North Am* 8(2):295–307
28. Foroughi H, Yazdi H-S, Pourreza H, and Javidi M (2008) An eigenspace-based approach for human fall detection using integrated time motion image and multi-class support vector machine. In: *Proceedings of the 4th international conference on intelligent computer communication processing*, pp 83–90
29. Foroughi H, Aski B-S, and Pourreza H (2008) Intelligent video surveillance for monitoring fall detection of elderly in home environments. In: *Proceedings 11th international conference on computer and information technology*, pp 219–224
30. Vishwakarma V, Mandal C, Sural S (2007) Automatic detection of human fall in video. In: *Proceedings of 2nd international conference on pattern recognition and machine intelligence*, pp 616–623
31. Khan A-M, Lee Y-K, and Lee S-Y (2010) Accelerometer's position free human activity recognition using a hierarchical recognition model. In: *Proceedings of 12th IEEE international conference on e-health networking applications and services*, pp 296–301
32. Mäntyjärvi J, Himberg J, Seppänen T (2001) Recognizing human motion with multiple acceleration sensors. In: *Proceedings IEEE international conference on systems, man, and cybernetics*, pp 747–752
33. Rajpoot N-M, Masood K (2005) Human Gait Recognition with 3-D Wavelets and Kernel based Subspace Projections. In: *Proc. International Workshop on Human Activity Recognition and Modeling, HAREM, Oxford*
34. Chung P-C, Liu C-D (2008) A daily behavior enabled hidden Markov model for human behavior understanding. *Pattern Recognit* 41(5):1572–1580
35. Zhu P, Hu W, Li L, Wei Q (2009) Human activity recognition based on R-transform and fourier mellin transform. In: *Proceedings of 5th international symposium advances in visual computing*, Springer, Berlin, pp 631–640
36. Niu F, Mottaleb M-A (2004) View-invariant human activity recognition based on shape and motion features. In: *Proceedings of IEEE sixth international symposium on multimedia software engineering*, pp 546–556
37. Nguyen N-T, Venkatesh S, Bui H-H (2006) Recognizing behaviors of multiple people with hierarchical probabilistic model and statistical data association. In: *Proceedings of BMVC, Edinburgh*, pp 1239–1248
38. Oliver N, Horvitz E, Garg A (2002) Layered representations for human activity recognition. In: *Proceedings of 4th IEEE international conference on multimodal, interfaces*, Washington, pp 3–8
39. Chang J-Y, Shyu J-J, Cho C-W (2009) Fuzzy rule inference based human activity recognition. In: *Proceedings of IEEE international symposium on intelligent control*, pp 211–215
40. Weinland D, Boyer E, Ronfard R (2007) Action recognition from arbitrary views using 3D exemplars. In: *Proceedings of 11th international conference on computer vision*, Oct 2007, pp 1–7
41. Elgammal A, Harwood D, Davis L (2000) Non-parametric model for background subtraction. In: *Proceedings of 6th European conference computer vision*, pp 751–767
42. Tabbone S, Wendling L, Salmon J-P (2006) A new shape descriptor defined on the Radon transform. *Comput Vision Image Underst* 102(1):42–51
43. Deans S-R (1983) *Applications of the Radon Transform*. Wiley Interscience Press, New York
44. Schölkopf B, Smola A-J (2002) *Learning with Kernels*. MIT Press, Cambridge
45. Schokopf B, Smola A-J, Muller K (1998) Nonlinear component analysis as a kernel eigenvalue problem. *Neural Comput* 10(5):1299–1319
46. Kanungu T, Mount D-M, Netanyahu N, Piatko C, Silverman R, Wu A-Y (2000) The analysis of a simple k-means clustering algorithm. In: *Proceedings of 16th ACM symposium on computational geometry*, pp 101–109
47. Yamato J, Ohya J, Ishii K (1992) Recognizing human action in time-sequential images using hidden markov model. In: *Proceedings of IEEE Comp Soc Conf Vision Pattern Recognition*, pp 379–385
48. Guo Y, Xu G, Tsuji S (1994) Understanding human motion patterns. In: *Proceedings of 12th international conference on pattern recognition, computer vision and image processing*, pp 325–329