Non-invasive analysis of sleep patterns via multimodal sensor input

Article *in* Personal and Ubiquitous Computing · January 2014

DOI: 10.1007/s00779-012-0623-1

Vangelis Metsis
Texas State University

Dimitris Kosmopoulos
University of Patras

Fillia Makedon
University of Texas at Arlington

Pers Ubiquit Comput
DOI 10.1007/s00779-012-0623-1

ORIGINAL ARTICLE

Non-invasive analysis of sleep patterns via multimodal sensor input

**Vangelis Metsis · Dimitrios Kosmopoulos ·
Vassilis Athitsos · Fillia Makedon**

Received: 30 September 2011 / Accepted: 23 June 2012

Abstract The monitoring of sleep patterns is of major importance for various reasons such as the detection and treatment of sleep disorders, the assessment of the effect of different medical conditions or medications on the sleep quality, and the assessment of mortality risks associated with sleeping patterns in adults and children. Sleep monitoring by itself is a difficult problem due to both privacy and technical considerations. The proposed system uses a combination of non-invasive sensors to assess and report sleep patterns: a contact-based pressure mattress and a non-contact 3D image acquisition device, which can complement each other. To evaluate our system, we used real data collected in Heracleia Lab's assistive living apartment. Our system uses Machine Learning techniques to automatically analyze the collected data and recognize sleep patterns. It is non-invasive, as it does not disrupt the user's usual sleeping behavior and it can be used both at the clinic and at home with minimal cost.

Keywords Sleep disorders · Sleep patterns · Machine Learning · Motion recognition

V. Metsis (✉) · D. Kosmopoulos · F. Makedon
Heracleia Human-Centered Computing Lab,
Department of Computer Science and Engineering,
University of Texas at Arlington, Arlington, TX, USA
e-mail: vmetsis@uta.edu

D. Kosmopoulos
e-mail: kosmopo@uta.edu

F. Makedon
e-mail: makedon@uta.edu

V. Athitsos
Department of Computer Science and Engineering,
University of Texas at Arlington, Arlington, TX, USA
e-mail: athitsos@uta.edu

1 Introduction

According to the American Academy of Sleep Medicine, there are 81 official sleep disorders, presented in [1]. Seventy million people in the USA have a sleep disorder, the vast majority of which remain undiagnosed and untreated. It is estimated that sleep-related problems incur $15.9 billion to national healthcare budget. There is then great need for automatic non-intrusive methods for sleep disorder recognition that patients can use in their homes. This would not only help decrease healthcare costs but also increase the number of diagnosed patients.

Another reason why sleep disorder detection is important is the fact that it is related to other potentially more serious medical conditions. According to [8], results of their study involving 1,506 participants (out of which 83 % reported some medical condition) show that sleep disorders are related to comorbidities rather than age. This is most likely because major comorbidities such as stroke, heart disease, osteoporosis, or arthritis impact the patients' quality of sleep. Detection of sleep disorders could therefore be an indication of another important disorder.

Vandeputte and de Weerd [19] studied 917 patients from a wide range of ages and suggest that patients with chronic sleep disorders are more likely to have depression and in fact about 1 in 4 patients who went to a sleep disorder clinic admitted to be experiencing depression, although only 3.5 % were found with moderate to severe depression.

In [18], it is mentioned that especially in older adults, there are three sleep disorders frequently seen: sleep disordered breathing (SDB), restless legs syndrome (RLS)/periodic limb movements in sleep (PLMS), and REM sleep behavior disorder (RBD). Adults with SDB may experience insomnia, nocturnal confusion, and daytime cognitive impairment including difficulty with concentration,

Published online: 01 November 2012

attention, and short-term memory loss. Patients with SDB are also at greater risk for cardiovascular consequences such as hypertension, cardiac arrhythmias, congestive heart failure, stroke, and myocardial infarction. The RLS is characterized by dysesthesia in the legs which is usually described as pins and needles or a creepy and crawly sensation in the legs that is only relieved with movement. Chronic RBD has been associated with narcolepsy and other idiopathic neurodegenerative disorders such as Lewy body dementia, multiple system atrophy, and Parkinson's disease.

In the past, some methods using Electroencephalograms (EEG) or Electromyograms (EMG) have been proposed for sleep disorder monitoring (e.g., [7, 9]). However, these methods are very inconvenient for the patients due to the cumbersome wiring that is required for the biosignal acquisition. On the contrary to those methods, here, we propose a non-invasive system that is able to analyze and recognize sleep patterns which can be further utilized to detect various types of sleep disorders. The first sensor that we employ is a bed pressure mat (product of Vista Medical Ltd.[1]) where the patient sleeps. The second sensor is the Kinect 3D image acquisition device by Microsoft [11]. Our approach is strongly motivated by the fact that by combining the information acquired by the two sensors, it is possible to attain better results than from a single one, due to the complementarity of the acquired information. Indeed, the pressure mattress is very reliable in capturing the information about the users' body parts that are in contact with it, but cannot provide any information about the rest of the body. On the other hand, the Kinect cannot see the body parts touching the mattress, but can provide rich data about the rest of the body parts that are visible. To the best of our knowledge, this is the first such multimodal approach for non-invasive recognition of sleep patterns.

We analyzed the acquired data using Supervised Machine Learning techniques, and the system classified the sleep patterns of the user in one or more predefined categories regarding both *posture* and *motion*. In this work, we experimented with data collected from seven individuals. The different patterns included periods of normal sleep and periods of abnormal sleep such as restlessness and frequent changes of body position. Preliminary results show that our system is able to successfully recognize sleep patterns and classify them among a predefined set of categories.

The remainder of this article is organized as follows. Section 2 presents related previous work in sleep pattern and sleep disorder detection. Section 3 elaborates on our methodology and experimental results in sleep pattern detection. Finally, Sect. 4 gives the conclusions of our findings.

[1] http://www.pressuremapping.com/.

2 Related work

Related research has focused on detecting various parameters of sleep for humans and animals as well as sleep quality and body posture recognition. More specifically, studies on rodents focus mainly on detecting if the animal is asleep or awake using piezoelectric films, used as a filtering stage for traditional classifiers using EEG and EMG [7]. The authors use EEG signals, preprocessed using Fast Fourier Transform (FFT), Principal Components Analysis (PCA) for feature selection, and classified using the k-Nearest Neighbor (k-NN) algorithm. Jansen and Cheng [9] also uses EEG and other signals and Markov modeling techniques to classify normal and abnormal human sleeping patterns. These types of signals require traditional Digital Signal Processing techniques such as Discrete Fourier Transform (DFT) and PCA for extracting meaningful features and k-NN or Artificial Neural Networks for the recognition step. Nevertheless, these methods require sensors or cables attached to the skin of the subject which is not acceptable for assistive pervasive applications. Other researchers use additional types of data, such as oximetry information to detect respiratory abnormalities [14]. The authors evaluate classification results using spectral and nonlinear analysis for feature extraction and Quadratic Discriminant Analysis (QDA), Linear Discriminant Analysis (LDA), k-NN and Linear Regression (LR) for classification. In [13], the authors try to assess sleep quality using near-infrared video only. The authors apply a homomorphic filtering technique to tackle the problem of over exposure in the center, common in near-infrared cameras. The authors also learn a threshold to differentiate noise from actual motion, since this type of cameras have very low signal-to-noise ratio (SNR). They then use the Motion History Image (MHI) technique that provides direction of movement to identify motion.

As human bodies and motions are in essence three-dimensional, the information loss in the depth channel could cause significant degradation of the representation and discriminating capability for camera-based feature representations. The recent emergence of cheap depth sensors (e.g., Microsoft Kinect) has made it feasible and economically sound to capture in real-time depth maps with appropriate resolution (e.g., 640×480 in pixel) and accuracy (e.g., =1 cm). A work using RGB + Depth sensor for human activity recognition is [12], where a bag of 3D points are sampled from the depth map and Gaussian mixture models are used to model the human postures. In [17], skeleton motion data are extracted from Kinect SDK for activity representation, while methods to extract skeletal data are presented in [3]. The very detailed human models used in the aforementioned works show the high utility of the depth data; however, in our application, the human limbs are not easily recognizable due to self occlusions or due to the use of blankets.

Pressure has also been used to infer if the subject is asleep or awake by detecting movements and respiration of rodents. There exists one previous approach to our knowledge that recognizes sleeping posture of humans using pressure sensors. More specifically, 32 pressure sensors were used to record the pressure pattern of the subject at a particular pose and Naive Bayes as well as Random Forests were used for classification and compared to each other [15]. In [2], the authors use a pressure mat to identify sleeping postures of babies possibly assisting prevention of Sudden Infant Death Syndrome. The authors collected the data from a 1-year-old baby freely moving on the pressure mat, and after a feature selection stage, they classified each posture using majority vote of k-NN, SVM, linear and quadratic classifiers and then applied a sliding window algorithm to eliminate possible misclassifications.

In our literature survey, we did not find any other non-invasive method that would be able to combine the benefits of a contact-based sensor such as a pressure mattress with the merits of non-contact sensors such as 2D or 3D cameras. Furthermore, the related work is rather limited to posture identification and does not cover motion patterns, which may be of importance. In this work, we aim to cover this gap.

3 Multimodal sleep pattern analysis

3.1 Description of datasets

For the needs of our experiments, we collected data from 7 different individuals simulating their sleep habits. Each individual lied on the bed for a period of time and performed the actions that they would normally perform if they went to bed. That involved getting in bed, staying still for periods of time in different postures, changing body postures, moving parts of the body like the arms or the legs, and getting out of the bed. The different actions performed during that period of time were recorded using 2 different sensors. The first one was a bed pressure mat (see Sect. 3.1.1) that we put under the sheets, and the second one was a Microsoft Kinect sensor (see Sect. 3.1.2) that we mounted on the ceiling. The recorded data were then manually annotated according to the various classes of interest, such body posture, motion occurrence, etc.

3.1.1 Data collected from FSA bed pressure mat

The FSA bed mat system produced by *Vista Medical Ltd.* provides a 1,920 mm $\times$ 762 mm sensing area which contains an array of 32 $\times$ 32 pressure sensors. Each of the sensors can capture a measurement in the range 0–100 mmHg (1.93 PSI) with a scan frequency of up to 5 Hz. The measurements can be recorded over a period of time and can be exported as a set of time stamped vectors containing the values of each of the 1,024 pressure sensors for each time stamp. To make visualization easier, we can consider each of these vectors as a frame of a video. Each of the sensors can be considered as pixel of a gray-scale image with an intensity ranging from 1 to 100. Thus, each frame can be considered as a 32 $\times$ 32 pixel image. Figure 1 illustrates a visualization example of the pressure values captured in one frame. The color coding is just a convention to facilitate visualization.

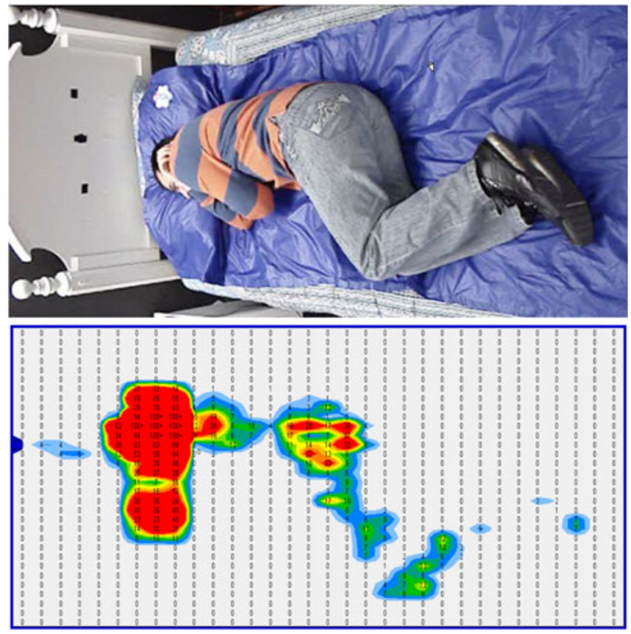


Fig. 1 An example of a subject lying on his side on the pressure mat (*top*) and the measurement values obtained (*bottom*)

3.1.2 Data collected from Kinect

Kinect is a motion sensing input device designed by Microsoft for the Xbox 360 video game console [11]. Kinect outputs 3 different data streams, RGB video stream, depth sensing video stream and audio. The video output frame rate is 30 Hz. The RGB video stream uses 8-bit VGA resolution (640 $\times$ 480 pixels), while the monochrome depth sensing video stream is in VGA resolution (640 $\times$ 480 pixels) with 11-bit depth, which provides 2,048 levels of sensitivity. In our experiments, we used only the depth sensing video stream. The depth sensor consists of an infrared laser projector combined with a monochrome CMOS sensor, which captures video data in 3D under any ambient light conditions. That feature makes the Kinect usable even in very low lighting conditions,

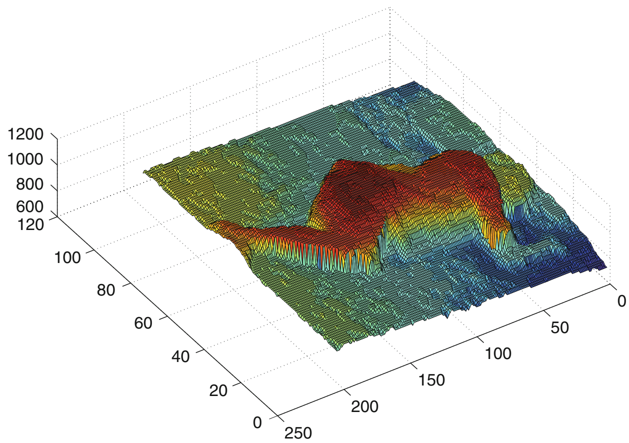


Fig. 2 A 3D representation of the input obtained by the Kinect depth sensor

which is usually the case during the night sleep. Furthermore, the 3D input that we get regarding the subject's body posture is more informative compared to the 2D information that we could get from the RGB video. The value of each pixel in a depth video stream frame is the distance, in millimeters, of the corresponding surface part of the object from the sensor (Fig. 2).

3.2 Data analysis and classification

The detection/recognition of sleep disorders usually boils down to the recognition of a set of symptoms that are related to a specific sleep problem. Such symptoms are as follows: how long it takes for the person to fall asleep, how many times (if any) they wake up during the night, how often do they move during their sleep time, how many hours on average do they sleep, etc. These indicators are difficult to monitor at home. Our immediate goal is to create a system that can recognize these indicators and make them easily accessible to the physicians. The long-term goal is to create a system that will be able to automatically detect specific sleep disorders based on training data from previous known cases.

To achieve our goal, we break our problem into a set of classes and we employ a combination of rule-based and supervised learning methods to classify the various instances into one of those classes. To evaluate the classification accuracy, we perform leave-one-out cross-validation experiments where every time we test the classification accuracy on the data collected from one user, by training it on data collected from the other users.

In more detail, we are attempting to recognize the following situations: (1) if the person is in bed or not, (2) when does motion occur while in bed, (3) what type of motion is that, and (4) while the person does not move what is their body posture in bed. Being able to detect and recognize the above situations and then combining them together can be a very rich information source with regard to the symptoms that we want to identify. In the following sub-sections, we will describe how we approach each of the above situations and how efficient our system is in terms of recognition accuracy.

3.2.1 Detecting if the person is in bed or not

The first case of interest in our experiments would be to detect whether the person is in bed or not. This is useful in cases, for example, where we want to know how many hours in total does the person spend in bed and how often do they get up during their sleep time. It turns our that this is a very easy problem to solve by just using the bed pressure mat. All we had to do is just define a threshold of the total amount of pressure that we get in the pressure mat. If the total pressure exceeds that threshold, it means that the person is in bed. Using this approach, we got 100 % accuracy in detecting whether the person is in bed or not in our experiments. Note that we did not consider cases where somebody puts something heavy on the bed that might confuse our system, since we assume that participants are willing to be examined and they are not willing to mislead the system.

3.2.2 Motion detection

Another case of interest is to detect when motion occurs while the person lies on bed. The detection of motion can be related to various sleep disorder symptoms. For example, it can be an indication of how long does the person take to fall asleep after they go to bed or how often do they wake up during the night.

To detect motion, we used the standard computer vision technique of frame differencing. That means that we compared consecutive frames by subtracting the frame n from the frame $n+i$, where $i \geq 1$ depending on the frame rate, and summing up the absolute differences. The value of that sum S is a very good indicator of the existence of motion in the time slot between the two frames.

For example, by using the only bed pressure mat, this can be achieved by calculating the sum of absolute differences of the values of each of the 1,024 pressure sensors between consecutive frames represented as vectors. Assuming a frame vector $X_k = \{x_1, x_2, \ldots, x_n\}$, where $n = 1,\ldots,1{,}024$, at each time point k, this sum S can be can be calculated as follows:

$$S = \sum_{i=1}^{n} |x_{k+1,i} - x_{k,i}| \tag{1}$$

It turns out that motion can be easily detected by specifying a threshold T on the value of S. If S becomes greater than T,

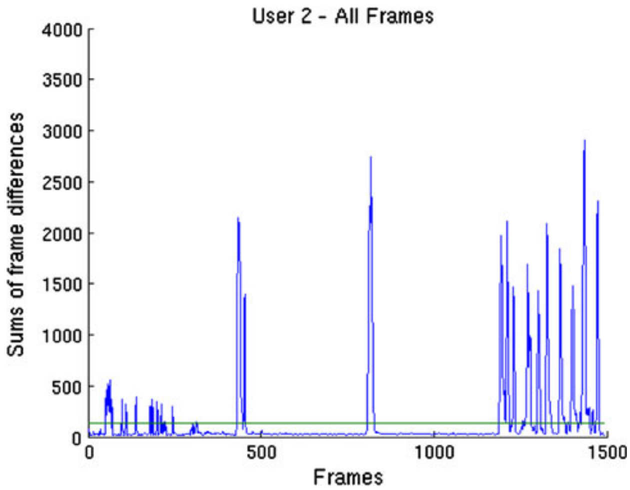


Fig. 3 Detection of motion using the sum of absolute frame differences (S) and a threshold $T = 130$

the subject is moving. The optimal value of T can be calculated from the training dataset, and it is almost constant among subjects of similar weights. Figure 3 shows a graph of the values of S over a period of about 1,500 frames obtained from one of the subjects. The green horizontal line defines the threshold.

Exactly the same approach can be used on the data collected from the Kinect sensor. The only differences compared to the pressure mat frames are the frame rate, the frame resolution, and the pixel value range. However, the formulas to calculate the sums S and the optimal threshold are exactly the same.

Using this approach, we classified each frame in the stream as containing motion or not. We tested our system's accuracy against the manually annotated data where human had specified the time points where motion occurred. We experimented using the pressure mat only and the combination of pressure mat and Kinect. To combine the two different data sources, we re-sampled the Kinect data to meet the pressure mat frame rate (3 Hz), and then, we aligned the frames using their timestamps. Each frame was classified as containing motion, if the value of S in either of the two data sources exceeded the predefined threshold.

Using the pressure data only, we achieved an average motion detection accuracy of 96.83 %, whereas adding the Kinect data, the accuracy increased to 97.57 %. The increase in accuracy can be attributed to cases where a motion (e.g., hand movement) is not strong enough to be detected by the pressure mat but it can still be detected by Kinect. The majority of the misclassified frames were spotted either at the beginning or at the end of movement of the subject where the levels of motion are very low. Hence, some of those might have actually been misannotated during the manual annotation process. In any case, the results of motion detection accuracy can be considered satisfactory.

3.2.3 Recognition of motion types and body postures

After detecting motion, our next step was to recognize the motion type, when motion occurred, and the subject's body posture, when there was not motion. To do that, we first used our motion detection method to segment the data steams into sequences of frames which are part of a motion and sequences of frames where there is no motion. Then, we classified each of those sequences into one of the motion classes or body posture classes.

The basic motion classes that we defined were the following:

1. Changing body posture.
2. Moving arms or legs.
3. Getting in bed or out of bed.
4. Making bed.

The first class refers to the case where the subject is changing sides, for example, they are sleeping on their back, and then, they turn their left. The second class refers to more subtle motion types where the subject moves a part of their body, usually a limb, but they do not completely change their body position. The third class occurs when the person gets in or out of the bed. This motion type differs from the previous two considerably. The last class refers to the case where the person is not actually in bed but there is still some type of motion detected by the pressure mat or the Kinect. This is usually the case when someone makes their bed.

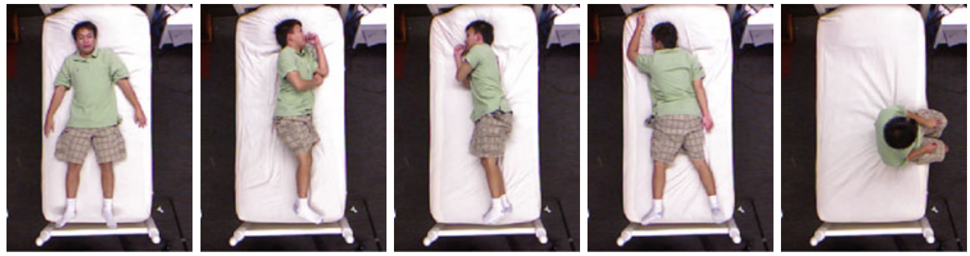
Regarding the body postures we defined the following classes:

1. Back
2. Left side
3. Right side
4. Stomach
5. Sitting on bed

The first four classes cover the basic usual sleeping postures, whereas the fifth class occurs when the subject is on the bed but they are not actually lying on it. Such cases usually occur when the subject is about to get in or out of the bed, but there could also be cases where they do not feel good and the temporarily get up for a few seconds. Figure 4 gives an overview of these 5 postures.

To recognize the body postures, we experimented with two different techniques. The first one is a Computer Vision-based technique, called Template Matching (TM), which has been used in face detection [10] and other similar applications. The idea behind this technique is that for each posture, we pick a representative frame to use it as

Fig. 4 The 5 different body postures



(a) Back **(b)** Left Side **(c)** Right Side **(d)** Stomach **(e)** Sitting

a template, after possibly cropping it appropriately, and then for every other frame to be classified, we compare it with all the templates and see which one matches better according to some distance criterion. In our case, we used the simple frame difference as a distance criterion. That means we calculated the sum of absolute differences of each pixel of the template subtracted from the corresponding pixel in the frame to be classified. To accommodate for cases where the subject lied in a different position of the bed compared to the template or they were taller/shorter compared to the subject used in the template, we tried different scales and different centering positions of the template.

The second technique that we used was based on supervised learning. In order to perform supervised learning, we converted each frame to a feature vector where each pixel represented a feature. To remove redundant features and reduce noise before classification, we performed a PCA [6] transformation on the data. The PCA eigenvectors were calculated on the training dataset matrix every time, and then, the testing dataset was projected to them. The optimal number of principal components to keep at each experiment was determined using an incremental search on the training dataset. The number of components that gave the highest classification accuracy on the training dataset was used for the final experiment with each classification method, and it usually varied from 7 to 40 components depending on the method. In addition, we calculated the Central Image Moments of the original frames, and we added those as features to the feature vector the resulted from PCA. An image moment is a certain particular weighted average (moment) of the image pixels' intensities. The advantage of central moments is that they are translation invariant which could be useful in cases where the subject is lying in an unusual position of the bed. For a digital gray-scale image with pixel intensities $I(x, y)$, the raw image moments M_{ij} are calculated by

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y) \qquad (2)$$

The central moments can be calculated using the following equation:

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \qquad (3)$$

where $\bar{x} = \frac{M_{10}}{M_{00}}$ and $\bar{y} = \frac{M_{01}}{M_{00}}$ are the components of the centroid. We used central moments of order up to 2, which yields 8 different moments. For the data coming from the Kinect, we assumed a near-constant background and we defined the region of interest to be the area that covers the dimensions of the bed.

In our experiments, we evaluated our methods using each of the data sources separately and in combination. To combine features from the two different data sources, we aligned the pressure and the depth sensing frames using their timestamps and we created a composite feature vector which included the top Principal Components and the Central Moments of each pair of frames. To deal with the difference in the frame rate, we re-sampled the depth sensing stream to reduce its frame rate. The classification methods that we used to classify each single frame were based on the well-known KNN [4] and Linear Kernel SVM [5] algorithms. For KNN, we used $K = 10$ nearest neighbors. To determine the optimal number of neighbors to use, we did an incremental search and used the number that gave the highest accuracy across the training datasets of all users.

In order to recognize the body posture in the sequence of frames, in both the cases of template matching and supervised learning, we first classified each of the frames in the sequence to one of the predefined classes, and then, we used majority voting to decide the final posture class. The voting approach was chosen in order to improve robustness in case the segmentation process into static and motion frames did not perfectly determine the boundary between the two states. For example, if a user was lying on their back and then turned to their left, the accuracy in detecting the motion would be virtually perfect; however, some of the frames at the beginning and the end of the motion could be misclassified as static (since motion levels at these boundaries are relatively low). To prevent these boundary frames from having an effect to the final recognition of the posture, we decide the final class to be the class of the majority of the frames since the overall number of frames in a segment is much bigger than the boundary frames. Another approach would be to ignore the boundary frames, but then, we would have to determine how many frames to ignore.

At each round, we trained our system using data coming from 6 out of the 7 users and classified the motion sequences of the 7th user. This ensures that if the system is to be used in real life, it can be trained off-line in advance and it does not require any re-training for the specific user.

For the classification of the sequences of motion frames into one of the 4 classes, we used Hidden Markov Model (HMM). A HMM is a statistical model of a system having hidden states and operating under the Markovian assumption. HMMs have been proven to model effectively temporal sequences as well as other forms of sequential data. The models are trained using the Baum–Welch algorithm that calculates their parameters. As for the recognition step, it is done using the Forward–Backward algorithm [16].

For the KNN classifier, we found that the combination of the top 40 principal components from each data source plus the central moments for each frame gave us the best classification accuracy. Similarly, for SVM, we used the top 30 principal components plus the central image moments. For the classification of motion using HMM, we used the top 7 principal components plus the central moments from the pressure sensing datasets and the top 14 principal components plus the central moments from the depth sensing datasets.

Table 1 presents the classification accuracy results for each user and the weighted average accuracy, where the weight represents the number of instances per dataset. In the different columns of the table, we present the result for body posture recognition and motion type recognition, separated by the classification algorithm that was used and also by the type of data source that was used to perform the training and testing. At each experiment, we evaluated our system using the pressure sensing (*P*) data only, the depth sensing (*D*) data only, and their combination (*C*).

As one can notice, combining the two different data sources by fusing their features gives the best classification accuracy in most cases. Also, with the exception of Template Matching (TM), using the pressure sensing data alone to recognize body postures and motion types gives better accuracy compared to using the depth sensing data alone. The supervised learning methods (KNN and SVM) outperform the template matching classification in the majority of the cases. In our experiments, we constructed the templates from one user (User 1) and we applied those same templates to all the other users. That is the reason why template matching works particularly well on "User 1". Since the template construction only requires the capturing of one frame for each posture and some cropping to match the body dimension, it would not be unreasonable to construct new templates for each new user in real life.

4 Conclusion and future work

In this paper, we presented our work on analysis of sleep patterns using non-invasive sensors and applying a combination a rule-based and Machine Learning methods. Our experimental results on real user datasets show that the task of analyzing sleep patterns with the intent to detect symptoms related to sleep disorders can be successfully achieved. Although the available dataset was relatively small, the classification accuracy results are promising and show that the proposed tools and methods could be used in the future for the detection of sleep disorders and other

Table 1 Classification accuracy results for body posture and motion type recognition

	Posture recognition									Motion recognition		
	TM			KNN			SVM			HMM		
	P	*D*	*C*	*P*	*D*	*C*	*P*	*D*	*C*	*P*	*D*	*C*
User 1	87.75	**<u>91.83</u>**	89.79	**83.67**	57.14	**83.67**	89.79	87.75	**<u>91.83</u>**	80.39	74.51	**92.15**
User 2	47.72	56.81	**77.27**	**<u>90.90</u>**	77.27	88.63	81.81	77.27	**84.09**	95.74	76.59	**97.87**
User 3	31.91	57.44	**65.95**	95.74	**<u>97.87</u>**	95.74	91.48	89.36	**93.61**	94.23	78.84	**96.15**
User 4	52.17	63.04	**84.78**	**89.13**	67.39	86.95	89.13	91.30	**<u>93.47</u>**	75.51	63.26	**79.59**
User 5	30.43	08.69	**47.82**	69.56	**73.91**	69.56	78.26	56.52	**<u>86.95</u>**	90.90	27.27	**95.45**
User 6	30.76	33.33	**<u>66.66</u>**	**56.41**	41.02	53.84	51.28	**64.10**	56.41	76.08	65.21	**78.26**
User 7	57.14	52.38	**73.81**	73.81	**<u>92.85</u>**	76.19	90.47	76.19	**<u>92.85</u>**	95.45	54.54	**86.36**
Average	53.10	64.82	**83.79**	**81.38**	72.76	80.69	82.76	79.66	**<u>86.21</u>**	86.50	66.24	**89.07**

"*P*" as a column title denotes that only pressure sensing date were used, "*D*" denotes that only depth sensing data were used, and "*C*" denotes that a combination of pressure and depth sensing data was used. For the posture recognition, the best accuracy per data source is in boldface and the best accuracy across the different classification methods is underlined. For the motion recognition, the best accuracy per data source is in boldface related diseases affecting sleep quality. To this end, further experimentation with bigger datasets and improved fusion methodology would be of high interest.

In the future, we plan to apply the system to large-scale clinical tests and we believe that it will be possible to associate our findings with pathological cases such as SDB, RBD RLS/PLMS, as well as depression. The big challenge is the diagnosis of diseases by recognizing the sleep patterns, which may lead to more focused medical treatments. The more focused treatments are expected to enhance the quality of life for millions of patients suffering from sleep disorders.

References

1. American Academy of Sleep Medicine (2001) International classification of sleep disorders, revised: diagnostic and coding manual. American Academy of Sleep Medicine, Illinois
2. Boughorbel S, Bruekers F, Breebaart J (2010) Baby-posture classification from pressure-sensor data. In: ICPR, pp 556–559
3. Charles J, Everingham M (2011) Learning shape models for monocular human pose estimation from the microsoft xbox kinect. In: IEEE international conference on computer vision workshops (ICCV workshops), pp 1202–1208
4. Cover T, Hart P (1967) Nearest neighbor pattern classification. IEEE Trans Inf Theory 13(1):21–27
5. Cristianini N, Shawe-Taylor J (2000) An introduction to support Vector Machines: and other kernel-based learning methods. Cambridge University Press, Cambridge
6. Dunteman GH (1989) Principal components analysis. Sage Publications Inc., Thousand Oaks
7. Flores AE, Flores JE, Deshpande H, Picazo JA, Xie X, Franken R, Heller HC, Grahn DA, O'Hara BF (2007) Pattern recognition of sleep in rodents using piezoelectric signals generated by gross body movements. IEEE Trans Bio-med Eng 54(2):225
8. Foley D, Ancoli-Israel S, Britz P, Walsh J (2004) Sleep disturbances and chronic disease in older adults: results of the 2003 national sleep foundation sleep in america survey. J Psychosom Res 56(5):497–502. doi:10.1016/j.jpsychores.2004.02.010. http://www.sciencedirect.com/science/article/B6T8V-4CG6MR1-5/2/e6f128d71f0fcc69c302ec316c194a62
9. Jansen BH, Cheng WK (1987) Classification of sleep patterns by means of markov modeling and correspondence analysis. IEEE Trans Pattern Anal Mach Intell PAMI-9(5):707–710. doi:10.1109/TPAMI.1987.4767968
10. Jin Z, Lou Z, Yang J, Sun Q (2007) Face detection using template matching and skin-color information. Neurocomputing 70(4–6): 794–800
11. Kinect-xbox.com. http://www.xbox.com/en-US/kinect. Accessed 26 Oct 2012
12. Li W, Zhang Z, Liu Z (2010) Action recognition based on a bag of 3d points. In: CVPR4HB10, pp 9–14
13. Liao W, Yang C (2008) Video-based activity and movement pattern analysis in overnight sleep studies. Proceedings 3(Conf 19):1774–1777
14. Marcos JV, Hornero R, Alvarez D, del Campo F, Zamarron C (2009) Assessment of four statistical pattern recognition techniques to assist in obstructive sleep apnoea diagnosis from nocturnal oximetry. Med Eng Phys 31(8):971
15. Ni H, Abdulrazak B, Zhang D, Wu S, Yu Z, Zhou X, Wang S (2010) 7th Conference, I.: Towards non-intrusive sleep pattern recognition in elder assistive environment
16. Rabiner L (1989) A tutorial on hidden markov models and selected applications in speech recognition. Proc IEEE 77(2):257–286. doi:10.1109/5.18626
17. Sung J, Ponce C, Selman B, Saxena A (2012) Unstructured human activity detection from RGBD images. In: Proceedings of the International Conference on Robotics and Automation (ICRA). IEEE, pp 842–849
18. Susan R, Sonia AI (2010) Sleep disorders in the elderly. Indian J Med Res 131:302–310
19. Vandeputte M, de Weerd A (2003) Sleep disorders and depressive feelings: a global survey with the beck depression scale. Sleep Med 4(4):343–345. doi:10.1016/S1389-9457(03)00059-5. http://www.sciencedirect.com/science/article/B6W6N-4949GT9-G/2/08444f7069f4981ee4e60c13d836b96f