Journal of Mobile Multimedia, Vol. 10, No. 3&4 (2014) 327–342 **PREPRINT**

SLEEP DETECTION USING DE-IDENTIFIED DEPTH DATA

BJÖRN KRÜGER, ANNA VÖGELE, MAROUANE LASSIRI
LUKAS HERWARTZ, THOMAS TERKATZ, ANDREAS WEBER
Institute for Computer Science II, University of Bonn
Friedrich-Ebert-Allee 144, 53113 Bonn, Germany
kruegerb@cs.uni-bonn.de

CARMEN GARCIA, INGO FIETZE, THOMAS PENZEL
Charite - Universitätsmedizin Berlin
Berlin, Germany

The work at hand presents a method to assess the quality of human sleep within a non-laboratory environment. The monitoring of patients is performed by means of a Kinect device. This results in a non-invasive method which is independent of immediate physical contact to subjects. The results of a study which was carried out as proof of concept are discussed and compared with the polysomnography-based gold standard of sleep analysis. When medical data are concerned, confidentiality is always an issue. This is no less important when monitoring people in their own homes, especially when they are in a situation as vulnerable as sleep. To meet the upcoming challenge of protecting people's privacy while still offering analyses of their data we introduce a blurring method to the acquired data and evaluate the use of our sleep detection test on such de-identified data sets.

Keywords: sleep, self-monitoring, polysomnography, privacy protection

1 Introduction

In recent years, the market for wearable technology designed to assist consumers in health care and fitness has virtually exploded. Since the so-called 'quantified self movement' has taken off, various companies have launched smaller and more powerful sensors and consumer electronics devices which allow for tracking of fitness, physical activity, health, diet and other personal metrics.

Although sleeping is a significant component of human activity it involves being relatively inactive. Nevertheless, the quality of sleep is a vital factor in the overall quality of human performance during daytime. Usually, the analysis of serious sleep disorders requires invasive lab-based monitoring. However, such measurements may also be partly replaced by actigraphy as a basic means to measure sleep quality. An overview of the consumer electronics devices that analyze frequency of wrist movement over time or similar information is found in [1].

While wearable devices like the above-mentioned are popular tools to monitor sports and everyday motions during waking hours, reservations against wearing such devices during sleep are a common phenomenon due to their causing impairments on the quality of sleep. In this work we focus on measuring the sleep phases by means of consumer electronics hardware but without immediate physical contact to the subjects. To demonstrate this technology achieves competitive accuracy, we compare our evaluations of Kinect recordings with the results from a professional sleep laboratory.

Naturally, there is an increasing awareness of privacy issues when personal data are involved. This is important when monitoring people in their own homes, particularly when they not conscious of their surroundings. In order to tackle the problem of protecting people's privacy we introduce a blurring method to our video and depth data. Our evaluations show that this method obscures the identity of subjects in the data but does not impair our sleep detection method to a significant level.

2 Previous Work

The standard approach to studying sleep and the detection of specific sleep disorders is a polysomnographic monitoring of patients in a clinical sleep laboratory. This multi-parametric test results in a comprehensive recording of the biophysiological changes that occur during sleep. A survey on standardized specifications and techniques for characterization of natural sleep is given in the scoring manual of Iber et al. [2]. A revision of existing methods for scoring sleep stages as well as new techniques to gather information from the polysomnographic readings are presented by Silber et al. [3] as part of a visual sleep scoring initiative. Standards of practice and clinical diagnostics by means of cardiorespiratory polysomnography are found in the contribution of Penzel [4].

As alternatives to study sleep in the clinical laboratory alone, there is also a wide range of options to record data with professional equipment but within the the patient's domestic environment. A study to compare home polysomnography with the in-lab approach in the diagnosis of sleeping disorders was done by Porter et al. [5]. Their evaluation showed that unattended polysomnography was still problematic in a number of cases since diagnosis depends on the quality of data obtained under unattended conditions. However, since there is an ongoing trend towards monitoring sleep in an environment comfortable for patients, the necessity to meet the obstructions of self-monitoring sleep quality is vital. The increasing variety of mobile devices for self-recording personal metrics can help alleviate by offering more consumer friendly options of data acquisition.

Metsis et al. [6] propose as system which uses several non-invasive sensors to monitor sleep patterns. They pursue a data driven machine learning approach to process and analyze data thus obtained. As opposed to this multi-modal data-driven learning technique the method we introduce uses of a single Kinect device to achieve results of comparable quality as classical polysomnography.

Kay et al. [7] describe how exterior conditions influence the sleep quality in human subjects and introduce a system *Lullaby* which helps patients self-monitor their sleep. This system is designed as a mobile sleep lab which helps the patients track down the negative influence such as light or noise their sleep might suffer from. The tracking is done by a Fitbit sensor along with an infrared camera. Sadeh and Acebo [8, 9] describe the important role of actigraphy, i. e. the monitoring of rest and activities by means of a wearable actimetry sensor, in sleep medicine. They summarize how sensitive actigraphy is in detecting unique sleep patterns corresponding to certain sleep disorders and other phenomena like medical problems or drug use. An overview of systems which allow for consumers to monitor their personal metrics is presented in [10].

Algorithms for scoring sleep stages from readings of actigraph sensors are described in the works of Kripke et al. [11], [12]. The authors develop a method to identify sleep-stages from wrist activity readings and critically evaluate the performance of these systems compared to standard polysomnographic scoring with slight disadvantages to the cell phone accelerometers.

Smartphone applications naturally play an important part as a means of self-assessing sleep quality. Natale et al. compare sleep estimation for healthy adults with a smartphone accelerometer to that of an actigraph accelerometer [13]. They found that the overall agreement between the scorings of the two types of devices was satisfactory. For an overview of current trends in personal sensor informatics based on emerging technologies associated with smartphones which enable the provision of pervasive health informatics services, refer to [14].

The well established methods for face detection and recognition allow systems to automatically link images to known personal information and background. These techniques being readily available makes protecting privacy an urgent necessity. Face detection is a means of localizing and extracting face regions in images from the respective background. A critical survey on existing face detection algorithms is given by Hjelmas and Low in [15]. Similarly, an overview of techniques of face recognition, i. e. determining the identity of the person in the input image, can be found in the works of Jafri and Arabnia in [16].

As a natural consequence of the availability of automated recognition methods, the task of avoiding face recognition to protect privacy has been tackled before. One method to de-identify data in order to avoid unintentional and unwarranted invasion of the privacy of individuals caught in videos is presented by Agrawal and Narayanan in [17]. The authors present automated methods to obscure the identity of subjects in such data exploring exponential space-time blur and line integral convolution. Neustaedter et al. [18] critically evaluate video blurring, an image masking method used to secure privacy for telecommuters and employees who have to attend video conferences from home.

3 Materials and Methods

3.1 Polysomnography

The basic tool and the reference method for sleep medicine diagnosis in the sleep laboratory is attended cardiorespiratory polysomnography. Cardiorespiratory polysomnography records physiological signals to quantify sleep and sleep disorders to achieve a diagnosis according to the International classification of sleep disorders revision 2 [19]. These multi-sensor readings are evaluated by experts in the field, that is, sleep assessment is done according to a specific set of rules which were defined in a manual for the recording and evaluation of sleep [20]. This manual specifies the recording and evaluation of the electroencephalogram (EEG), the electrooculogram (EOG), and the submental electromyogram (EMG). A few years ago the manual had been expanded and revised to include the recording of respiration, cardio-circulatory functions, and movement parameters. Age-specific aspects such as pediatric sleep and respiration in children were included in this revision too [2]. The new manual is based on a systematic literature review and a evidence evaluation [3, 21, 22, 23].

In our experiments, we use the results of polysomnographic sleep detection as gold standard for evaluation of our own results.

3.2 Kinect-Based Setup

For our Kinect based setup, we mounted the system to panel that could be placed under the ceiling and directly over the bed. Thus, the bed and the surrounding area was within the visible range of the sensor. Especially at the edges of observed objects, the depth information, i. e. images of resolution 640×480 per frame (see [24, 25] for further specifications) obtained by the Kinect sensor is very noisy. We computed two measures to quantify the amount of relevant motion in the data - as opposed to noise - and use these to evaluate our methods for noise reduction on a five minute test sequence where no motion was present whatsoever.

First, we computed visual motion summaries (see Figure 2). These show the average motion $\mathcal{D}$ for every pixel $p_{i,j}$ over a fixed time period of T frames.

$$\mathcal{D}(p_{i,j}) = \frac{1}{T} \sum_{t=1}^{T-1} |p_{i,j}(t) - p_{i,j}(t+1)| \tag{1}$$

The results for every pixel of the input image sequence are color coded, starting from black ($\mathcal{D} = 0$) to white ($\mathcal{D} = \max(\mathcal{D}(p_{i,j}))$ for all indices i, j in the image, i.e. $i = 1, \ldots, n$ and $j = 1, \ldots, m$).

Second, we computed the average change $\mathcal{N}$ per frame (see Figure 1):

$$\mathcal{N}(t) = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} |p_{i,j}(t) - p_{i,j}(t+1)| \tag{2}$$

While $\mathcal{D}$ gives an overview on the spatial distribution of motion in the input image sequence, $\mathcal{N}$ gives this information on the temporal domain.

Considering the five minute test sequence, the results of the two measures are shown in Figure 1 and Figure 2 for $\mathcal{N}$ and $\mathcal{D}$, respectively. While the blue curve in Figure 1 illustrates the high amount of noise, it gets clear that most of the noise is actually caused by the edges of the bed and other furniture as can be seen in Figure 2 a). Another source of noise in the data is the floor of the room. The reason for this is the high distance from the floor to the Kinect device as compared to the distance of the bed to the device. To deal with the noise-induced disturbances in the data we combine two strategies: On the one hand, we smooth the data by filtering. This is achieved by applying a simple, temporal box filter to the time series of each pixel of the depth map. The width of the filter window was 3 frames in our experiments. On the other hand, we only consider the relevant part of the image for further analysis by cropping the image to contain the bed itself only.

The benefit of each of these steps as well as their combination is also illustrated by the two above-mentioned Figures. In the absence of any motion in the scene, the average noise per frame is in the range between 15 and 20 millimeters. By applying the described box filter the noise is reduced to 2.5 millimeters per pixel. The restriction of the image to the bed alone reduces the noise from the original 15 or 20 millimeters to 5 millimeters. Finally, the combination of cropping to the relevant area (bed) and filtering reduces noise to a total of 1 to 2 millimeters per pixel. The motion summaries for the modified data are also shown in Figure 2 b)-d).

For all experiments described in the following sections we used the image data thus cropped and filtered. The amount of present motion per frame was measured using the function $\mathcal{N}$,

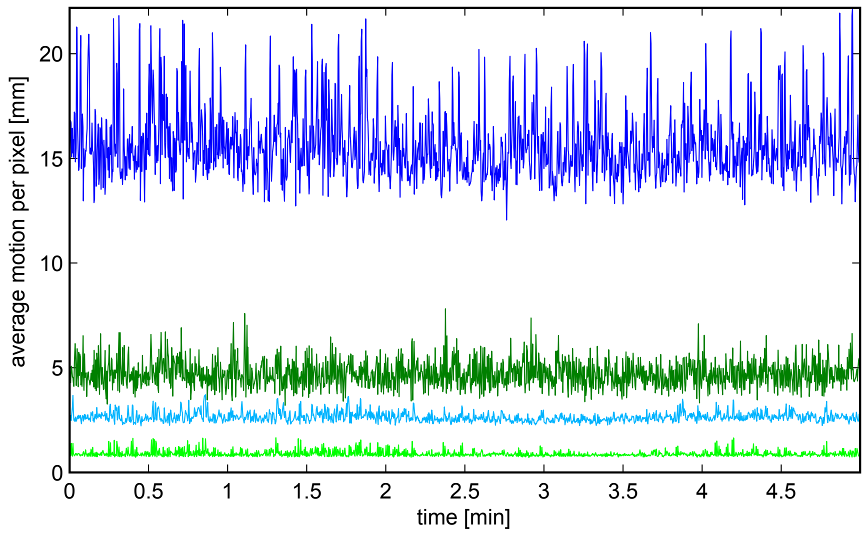

Fig. 1. Illustration of the Kinect sensors noise of a five minute test sequence. Average noise from the original depth image (blue). Average noise if the analysed region is restricted to the bed (green). Noise of the filtered depth sensor sequences light green and light blue respectively.

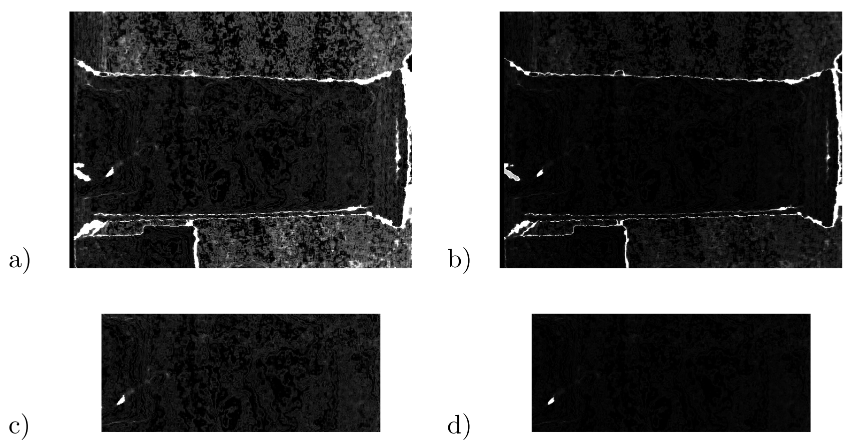

Fig. 2. Visual Motion Summary. a) raw data b) filtered data c) cropped area d) cropped, filtered area.

while the sampling rate of the Kinect sensor was set to a low temporal resolution of 5 frames per second. Since we are interested in the overall amount of motion of the subjects, this sampling rate has proven to be sufficient. Note also that the temporal uncertainty caused by the delay of at most 3 frames introduced by fixing the width of the filter window does not have an impact on our results since a delay of just over half a second is insignificant compared to several hours of testing evaluated as 30 second slots (see Section 3.5).

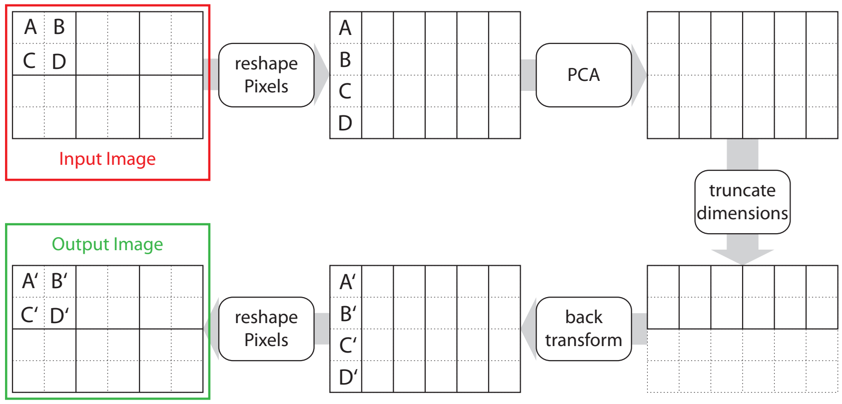

Fig. 3. Overview of our image simplification. The image is split into 2×2 squares that are reshaped to vectors. All resulting vectors are regarded as data matrix for the PCA. By truncating dimensions from the image representation in PCA space and the according back projection we get a simplified image.

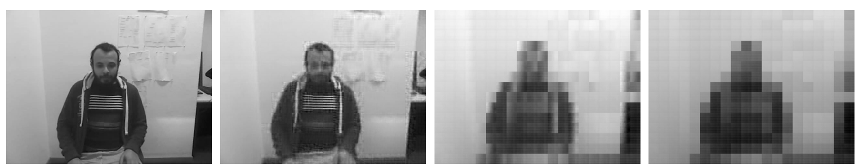

Fig. 4. Example for simplified images. Original image, 16 PCs, 2 PCs 1 PC. 32×32 block size

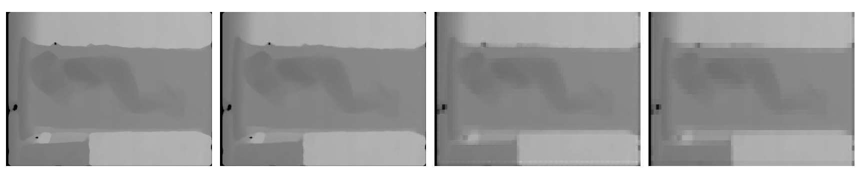

Fig. 5. Example for simplified depth images. Original data, 16 PCs, 2 PCs 1 PC. 16×16 block size

3.3 Image Simplification Using PCA

In this section we describe our image simplification technique, that we use to make persons unrecognizeable while still keeping track of the motion that is performed. To this end we employ a simple PCA based approach. First, we split the image into N quadratic blocks of size $n \times n$. Second, Each of this blocks is transformed to a vector of size $n^2 \times 1$. Thus, the image is now transformed to an $n^2 \times N$ data matrix. Third, we apply a standard principal component analysis (PCA) to this data. Fourth, we truncate the data in PCA space to the first M principal components (with $M \leq n^2$). Fifth, we obtain our simplified image by transforming the truncated data back to the data space and reshaping the data back to the original pixel positions. An overview is given on a simple example in Figure 3, where a 4×6 image is cut into 2×2 blocks, while only the 2 first principal components are left for the reconstruction of the image.

Figure 4 shows a sequence results of gray scale images. We show the original (left) image and the modified versions based on 16, 2 and 1 principal components. Similar we show a sequence based on depth images in Figure 5.

3.4 Sleep Detection by Kripke

In the experiments of Kripke et al. [11], subjects were monitored by an actimetry sensor called actigraph. The actigraph is a device similar to a wrist-watch and can be worn by a patient to measure gross motor activity, e. g. during sleep. In the following, we will briefly describe the algorithm introduced in [11].

Actigraphic activity counts are performed once every 30 seconds. The score $x_i \in \{0, 1\}$ given at each time $t = j$, $x_j = 0$ for inactive, $x_j = 1$ for active. The scores per epoch are collected and evaluated by

$$B(t) = \mathrm{s} \sum_{\hat{t}=t_s}^{t_e} b_{\hat{t}} x_{\hat{t}}. \tag{3}$$

In our case we fixed $t_s = -10$ and $t_e = 3$, and s was an overall scaling parameter for normalization. Refer to Table 1 for the parameter list of the b_ts used in both Kripke's and our algorithm. When the overall result is $B(t) \geq 1$, the epoch $t = j$ is scored wake, otherwise sleep. Note: Kripke et al. [11] did 49 training recordings to optimize for the parameters on the list. In the end, the findings are summarized by a secondary algorithm, scoring 'sleep' only when there have been at least 10 epochs of sleep before epoch $t = j$. Note: this was not done for the final set of $|t_e - t_s + 1|$ epochs of the study.

3.5 Adaptations of Kripke's Algorithm

We modified the original method by using different weights for the input signals:

$$H(t) = \sum_{\hat{t}=t_s}^{t_e} h_{\hat{t}} \cdot x_{\hat{t}} \tag{4}$$

Thus, the input signal from the Kinect which was recorded and preprocessed as described in Section 3.2 was classified. The algorithm to detect the status of the subjects is a binary classification of 'sleep' (positive) or 'wake' (negative). The algorithm used for this is based on the one introduced by Cole and Kripke [12] (see equation 3) where $D \leq \alpha$ means the subject is awake, $D \geq \alpha$ and the weights $b_{\hat{t}}$ are given in Table 1. In addition to these two parameter groups, there was one test using constant weights:

$$C(t) = \sum_{\hat{t}=-4}^{4} c \cdot x_{\hat{t}}. \tag{5}$$

Table 1. Parameters for both algorithms

Kripke													
b_{-10}	b_{-9}	b_{-8}	b_{-7}	b_{-6}	b_{-5}	b_{-4}	b_{-3}	b_{-2}	b_{-1}	b_0	b_1	b_2	b_3
.0064	.0074	.0112	.0112	.0118	.0118	.0128	.0188	.0280	.0664	.0300	.0112	.0100	.0000
Gaussian													
h_{-10}	h_{-9}	h_{-8}	h_{-7}	h_{-6}	h_{-5}	h_{-4}	h_{-3}	h_{-2}	h_{-1}	h_0	h_1	h_2	h_3
.0044	.0104	.0224	.0440	.0790	.1295	.1942	.2661	.3332	.3814	.3989	.1295	.0044	.0000

3.6 Derived Characteristics

The following quantities were computed to describe the nights:

- $\mathrm{T}_{\mathrm{night}}$ [min]: The total duration of the recorded interval in minutes. Note that, in the results tables in Figure 7 and Figure 8, this value differs for each scoring method with respect to the ground truth. This is due to the fact that the Kinect recordings were started before the ground truth measurements. This has no influence on the tested sleep efficiency or any other result.
- $\mathrm{T}_{\mathrm{awake}}$ [min]: The number of minutes the patient was identified as awake during the recorded time interval.
- $\mathrm{T}_{\mathrm{sleep}}$ [min]: The number of minutes the patient was identified as asleep during the recorded time interval.
- T_0 [min]: The time the patient fell asleep for the first time with respect to the time $\mathrm{T}_{\mathrm{night}}$.
- efficiency [%]: The percentage of minutes per night which the patient actually spend sleeping.

3.7 Binary Classification

The recorded raw data were scored according to the algorithms discussed in Section 3.4 and Section 3.5. The scoring results, x_θ, were classified with respect to the ground truth. Note the following distinction:

- A result x_θ at time θ is classified a *true positive*, TP, if the subject was asleep at time θ and was identified as asleep by the algorithm.
- A result x_θ is classified a *false positive*, FP, if the subject was awake at time θ but was scored asleep by the algorithm.
- x_θ is called a *true negative*, TN, if the subject was awake at time θ and this state was scored correctly.
- x_θ is a *false negative*, FN, if the subject was awake at time θ but was predicted asleep by the algorithm.

With the classifications in hand, the evaluation of the algorithm's performance is done according to the accuracy measure:

$$\text{accuracy} = \frac{\sum TP + \sum TN}{\sum TP + \sum TN + \sum FP + \sum FN} \tag{6}$$

The results of the evaluation are discussed in the following sections.

3.8 Collected Data

To evaluate our approach we measured two nights in a sleep laboratory environment. To this end, we recorded the subject with our setup and in parallel with the full equipment of the sleep laboratory. Annotations performed by the lab assistants on the basis of their data are used as ground truth reference. Overall, we captured more than 16 hours of data during the two nights.

a) b)

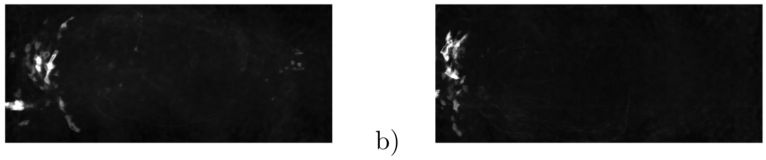

Fig. 6. Visual Motion Summary. a) first night b) second night.

Table 2. Classification results for the first night.

Classifier	T_{night} [min]	T_{awake} [min]	T_{sleep} [min]	T_0 [min]	efficiency [%]
ground truth	408	96.5	311.5	69	76.4
Kripke	463.5	88	375.5	12.5	81
Gauss	464	79.5	384.5	12	82.9
constant	466.5	79	387.5	11	83.1

4 Results Sleep Detection

4.1 Motion Summaries

In our discussion of the results we shall distinguish between spatial and temporal sleep detection respectively motion detection. The spatial motion summary illustrates in which regions of the bed the patient showed the most activity, e. g. if he moved the head or the feet. The temporal motion summary is a curve illustrating, at which times the subject displayed significant instances of activity. This results in a scoring whether the patient was asleep or awake.

Figure 6 is a summary of the accumulated visual findings of activity in the Kinect recordings made of the subject's bed during the two nights which were investigated in the lab (see also Section 3.2). The regions of the bed where the subject's head - which is placed at the left-hand side - hands and feet - which are at the right-hand side - were located are clearly visible: The bright areas in the pattern indicated that the subject moved the upper part of his body, especially the head and wrists most. Note that, although the subject slept under a duvet, the Kinect recordings sufficiently display a significant amount of movement of the subject's feet during the night.

The temporal hypnograms displayed in Figures 7 and 8 show the different phases of the first respectively second night's rest recorded in the experiments. Note that the hypnogram of the first night displays a significant amount of movement, even periods of restlessness. This is due to the subject wearing a complex clinical sensor setup which was a cause of discomfort. On the one hand, this was inevitable because of the necessity of sensor data acquisition for comparison with the Kinect's sleep detection performance. On the other hand, the influence of the setup disturbs the natural sleep patterns which were subject to the test in the first place. The hypnogram of the second night is more reliable because the subject was able to actually sleep normally.

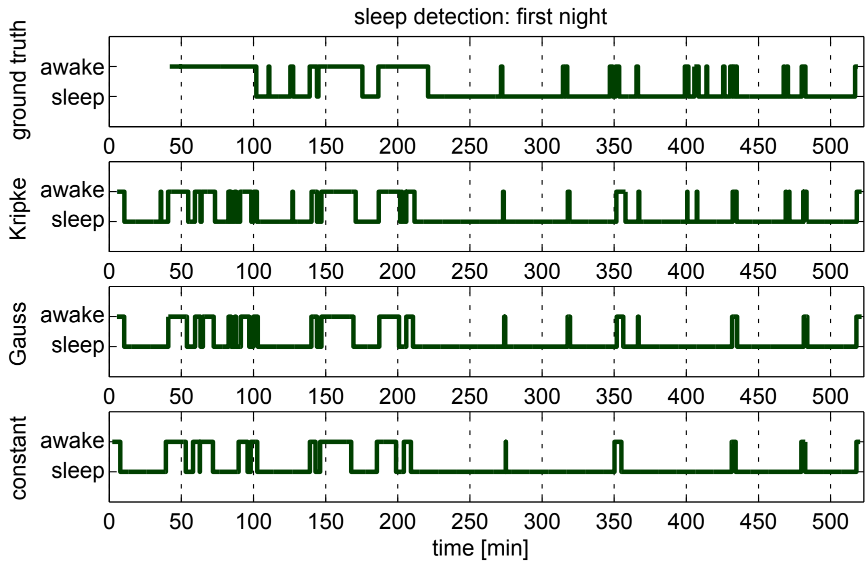

Fig. 7. Temporal hypnogram of the first night.

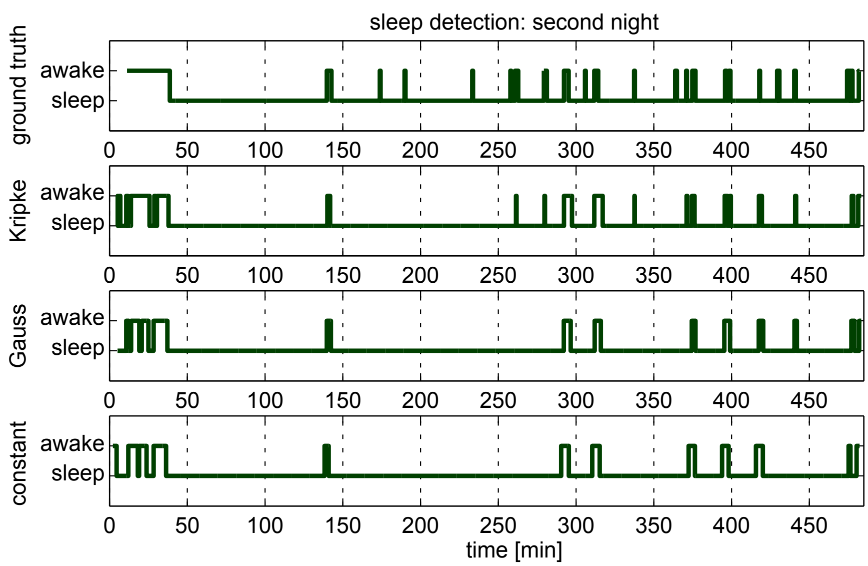

Fig. 8. Temporal hypnogram of the second night.

Table 3. Classification results for the second night.

Classifier	T_{night} [min]	T_{awake} [min]	T_{sleep} [min]	T_0 [min]	efficiency [%]
ground truth	443.5	26	417.5	27.5	94.1
Kripke	470	42	428	0	91.1
Gauss	469.5	42.5	427	1	91
constant	469	45.5	423.5	0	90.3

Table 4. Evaluation of the classifications for the first night.

Classifier	TP	TN	FP	FN	accuracy [%]
Kripke	817	191	131	22	86.8
Gauss	815	169	153	24	84.8
constant	824	170	152	15	85.6

Table 5. Evaluation of the classifications for the second night.

Classifier	TP	TN	FP	FN	accuracy [%]
Kripke	1025	71	41	27	94.2
Gauss	1025	69	43	27	94.3
Konstant	1009	63	49	43	92.1

4.2 Sleep Efficiency

The efficiency of the subject's sleep during the night was tested by computing the ratio between the cumulative length of the intervals of restlessness and the total duration of the rest phase,

$$\text{efficiency} = \frac{\mathrm{T}_{\text{awake}}}{\mathrm{T}_{\text{night}}}. \tag{7}$$

This value was computed only for the time interval starting from the point were the subject fell asleep for the first time. Thus, possible troubles falling asleep are not included in this particular efficiency analysis. Comparing the computations for the ground truth data as well as each of the classifiers shows the values of efficiency were similar for the different methods. For the first night, the results differ stronger, especially with respect to the ground truth. In fact, the reason for this is the following: initially there had been some additions to the sensor setup which, unfortunately brought about a significant amount of noise to the recording. After the measuring assembly was was reduced to the described setup, the data output was a lot more stable, while their was also substantially less nuisance for the subject by the sensors.

4.3 Classification Evaluation

We classified the data from the recordings according to different weights such as in Table 5 and compared the results with the data from the sleep laboratory. Refer to Tables 4 and 5 for details.

5 Results of the Face Recognition on Modified Data

In order to test the effectiveness of the blurring procedure, we had a number of different experiments. The first series of experiments dealt with the effect of the blurring algorithm on image data of test subjects that were made unrecognizable by the PCA method described in Section 3.3. The image organizer and viewer software *Picasa* is generally able to detect human faces in images of adequate quality and also to link those faces to known other pictures of the same subject. Thus, we employed *Picasa* as a state of the art facial recognition system in this work.

Our experiment setup was as follows. From a group of six male test subjects each individual was asked to sit down in front of a camera. Each subject was required to tilt and turn their heads in a standard procedure, such that their expression could be seen from a variety of angles. All video frames were converted to a total number of 648 images from which a set of training and test images were created. The training data remained unchanged, while the images used for testing where blurred by the method described in Section 3.3 using varying block sizes and numbers of principal components for reconstruction. Note that the choice of this training set and test set was done extremely for the benefit of the classifier. This was done against the backdrop of wanting the classifier to have a maximal chance for success in the test runs, while consequently excluding the possibility of unwanted recognition in the practice runs. The blurring of $n \times n$ matrix blocks was tested in practice for three parameters ($n \in \{8, 16, 32\}$). Since the procedure is parameterized also by the number of principal components used in the reconstruction, there were three types of results which were further divided into groups according to this parameter M (compare to Section 3.3). The value of M ranges from 1 to n^2.

For the face recognition by *Picasa* there are two additional parameters which need to be set. One is the *face recognition threshold* and the other is the *cluster threshold*. The face recognition threshold defines the required precision for the recognition of an individual subject. The number of candidates is naturally influenced by this definition. The *cluster threshold* influences the process of grouping subjects. Both these thresholds were fixed to the standard value of 85%.

A preliminary training of the software on the training set resulted in a recognition of 584 out of 648 images. Note that the rest of the images did not have any human faces in them at all. For the actual test, a training set was created in *Picasa* for each data sequence. All blurred images from each sequence were processed by *Picasa*. The number of recognized faces is a function in the number of principal components and in the parameter n employed in the blurred reconstruction of the images from the video sequences.

Even though the choice of the training and test set was fortunate for the classifier it failed to detect or recognize the displayed subjects when using a certain number of principal components in the reconstruction. Specifically, we distinguished two notions of success for the recognition tests with *Picasa*. A stronger notion of success for the classifier means it can not only detect faces where there are human individuals in the image but also recognize the identity of at least one of the faces. A weaker notion of success is given by only generally detecting the present faces as human expressions. For the parameter $n = 8$, we found there is always at least a weak success regardless of the paramer M: Even for $M = 1$ principal component, *Picasa* is able to still detect the presence of human faces in the images, while for $M > 3$ the software begins to recognize the individual faces for the first time. But since any success classifying should be avoided in practice to make for a real de-identification, our evaluations will, in the following, only be performed on the two other parameter options $n \in \{16, 32\}$. For the blurring parameter $n = 16$, only the number of $M = 1$ principal components left all faces undetected, while for $M \leq 5$ at least the recognition failed for all faces. The results of this particular recognition tests are summarized in Figure 11. A complete de-identification took place for $M = 1$. For $n = 32$, the choice of parameters $M \geq 4$ would result in the detection of a number of faces but not in any recognition. The recognition task was only successful for $M > 16$. A total de-identification was achieved for $M \in \{1, 2, 3\}$.

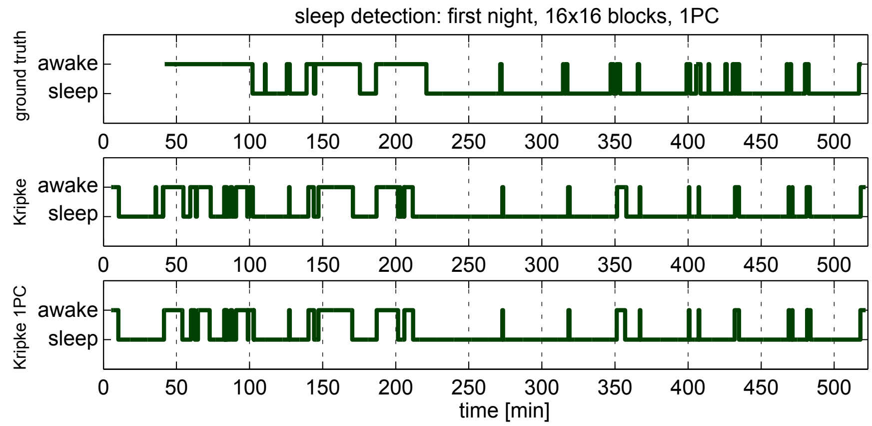

Fig. 9. Temporal hypnogram of the first night. The de-identified data were processed by our sleep detection method (cf. Figure 7)

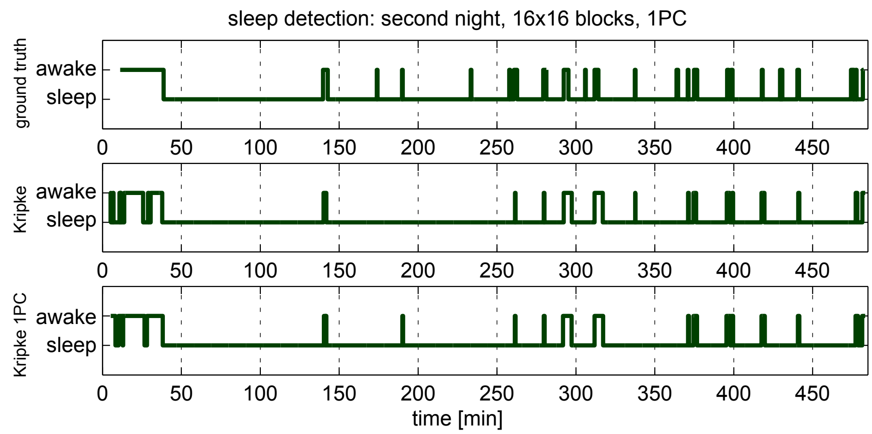

Fig. 10. Temporal hypnogram of the second night. The de-identified data were processed by our sleep detection method (cf. Figure 8)

6 Results of the Sleep Detection from Modified Depth Data

The final objective is testing the sleep detection method also on the blurred depth data. This means applying the blurring procedure to the depth data employed for the sleep detection (see Sections 3 and 4) and investigating how the results differ from the ones presented earlier. The evaluation of the classification after blurring is compared to the evaluation before blurring. Typical resulting hypnograms can be seen in Figures 9 and 10. The examples show the results of a PCA blurring with parameters $n = 16$ and $M = 1$ for the data of both nights. The accordance with the sleep detection results on unchanged data is very high. In fact, it is less than $1,5\%$ worse. To be precise, for the first night, $86,8\%$ on the original vs. 86% on the blurred data, for the second night, $91,1\%$ vs. $89,7\%$.

7 Conclusion and Future Work

In this study we investigated the possibility of monitoring human sleep without the professional equipment and expense factors involved in the clinical analysis scenario. Our tests

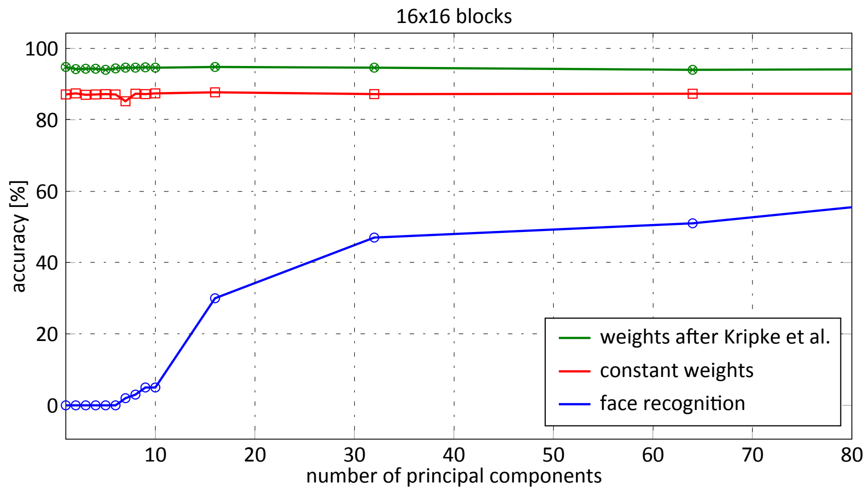

Fig. 11. For the blurring parameter $n = 16$, a complete de-identification of the data took place with our method for $M = 1$. The number of the recognized faces as a discrete function in the number of principal components is shown in blue, the other two sets of values, namely the number of recognized faces according to the choice of weights are green (Kripke) and red (constant).

already achieved fairly good results and proved the effectiveness of unattended sleep detection in the subjects' home by means of a Kinect or similar device. While our study, above all, serves as prove of concept, future work should focus on a number of convincing tests and the consideration of a reasonably large group of test subjects.

Future work will focus on comparing the presented approach against different data-driven analysis techniques to identify sleep and wake stages from sensor readings. To this end, we plan to employ segmentation techniques [26] to segment the stream of data in the temporal domain. On the basis of short motion segments the search for similar, annotated sequences in a large database, can be performed efficiently [27]. This facilitates a more unsupervised assessing of sleep patterns with the help of annotations and analysis based on the evaluation of existing data by experts in the field. Another strand of future research could be the exploration and visualization of motion summaries [28] that provide users as well as experts with condensed information on the activities performed. Another idea worth pursuing is to discuss variations in sensor placement for activity data. The comparison of the Kinect findings with the respective results could offer more scientific insight on the quality of contact vs. non-contact sensor sleep assessment.

Particularly for the study of clinical patients and other experimental subjects, data must be made anonymous. Finding means to represent data in a more abstract way is crucial to help preserve each subject's anonymity and dignity. This especially requires developing techniques to reduce the quality of data in the sense that individuals cannot be recognized but without losing other information.

Being a binary classification, the approach discussed in the paper at hand is a rather simplified method. However, there are various stages of sleep such as sleep characterized by Rapid - or Non-rapid Eye Movement and also different waking stages which, for instance, are easily documented by classical electroencephalography but have yet been neglected by the binary scoring. Finding ways to include these phenomena in the sleep analysis will help asses variations of natural sleep patterns and also detect certain abnormalities. New developments such as the Kinect2 might help improve such techniques further, as could other novel devices which also track eye activity.

References

1. Vandrico Inc. List of wearable devices. http://vandrico.com/database. Accessed: 2014-09-22.
2. C. Iber, S. Ancoli-Israel, A. Chesson, and S.F. Quan for the American Academy of Sleep Medicine. The aasm manual for the scoring of sleep and associated events: Rules, terminology and technical specifications, 1st ed. American Academy of Sleep Medicine, 2007.
3. Michael H Silber, Sonia Ancoli-Israel, Michael H Bonnet, Sudhansu Chokroverty, Madeleine M Grigg-Damberger, Max Hirshkowitz, Sheldon Kapen, Sharon A Keenan, Meir H Kryger, Thomas Penzel, et al. The visual scoring of sleep in adults. *J Clin Sleep Med*, 3(2):121–131, 2007.
4. Thomas Penzel. *Sleep Laboratory.* Wiley Encyclopedia of Biomedical Engineering, 2006.
5. Florence Portier, Adriana Portmann, Pierre Czernichow, Lionel Vascaut, Etienne Devin, Benhamou, Antoine Cuvelier, and Jean Franois Muir. Evaluation of a portable device based on peripheral arterial tone for unattended home sleep studies. *American Journal of Respiratory and Critical Care Medicine*, 162(3):814–818, 2003.
6. Vangelis Metsis, Dimitrios Kosmopoulos, Vassilis Athitsos, and Fillia Makedon. Non-invasive analysis of sleep patterns via multimodal sensor input. *Personal Ubiquitous Comput.*, 18(1):19–26, January 2014.
7. Matthew Kay, Eun Kyoung Choe, Jesse Shepherd, Benjamin Greenstein, Nathaniel Watson, Sunny Consolvo, and Julie A. Kientz. Lullaby: A capture & access system for understanding the sleep environment. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 226–234, 2012.
8. Avi Sadeh. The role and validity of actigraphy in sleep medicine: An update. *Sleep medicine reviews*, 15(4):259–267, 2011.
9. Avi Sadeh and Christine Acebo. The role of actigraphy in sleep medicine. *Sleep medicine reviews*, 6(2):113–124, 2002.
10. Christopher Trickler. An overview of self-monitoring systems. *SAIS 2013 Proceedings. Paper*, 37, 2013.
11. Daniel F Kripke, Elizabeth K Hahn, Alexandra P Grizas, Kep H Wadiak, Richard T Loving, J Steven Poceta, Farhad F Shadan, John W Cronin, and Lawrence E Kline. Wrist actigraphic scoring for sleep laboratory patients: algorithm development. *Journal of sleep research*, 19(4):612–619, 2010.
12. Roger J. Cole, Daniel F. Kripke, William Gruen, Daniel J. Mullaney, and J.Christian Gillin. Technical note automatic sleep/wake identification from wrist activity. *Sleep*, 15(5):461–469, 1992.
13. Vincenzo Natale, Maciek Drejak, Alex Erbacci, Lorenzo Tonetti, Marco Fabbri, and Monica Martoni. Monitoring sleep with a smartphone accelerometer. *Sleep and Biological Rhythms*, 10(4):287–292, 2012.
14. Katarzyna Wac. Smartphone as a personal, pervasive health informatics services platform: literature review. *arXiv preprint arXiv:1310.7965*, 2013.
15. Erik Hjelmas and Boon Kee Low. Face detection: a survey. *Computer Vision and Image Understanding*, 83(3):236 – 274, 2001.
16. Rabia Jafri and Hamid R. Arabnia. A survey of face recognition techniques. *JIPS*, 5(2):41–68, 2009.
17. Prachi Agrawal and P. J. Narayanan. Person de-identification in videos. In *Proceedings of the 9th Asian Conference on Computer Vision - Volume Part III*, ACCV'09, pages 266–276, Berlin, Heidelberg, 2010. Springer-Verlag.
18. Carman Neustaedter, Saul Greenberg, and Michael Boyle. Blur filtration fails to preserve privacy for home-based video conferencing. *ACM Trans. Comput.-Hum. Interact.*, 13(1):1–36, March 2006.
19. American Academy of Sleep Medicine. Icsd - international classification of sleep disorders, revised:

Diagnostic and coding manual. Library of Congress Catalog No. 97-71405, ISBN: 0-9657220-1-5, 2001.

20. A. Rechtschaffen and A. Kales. A manual of standardized terminology, techniques, and scoring system for sleep stages of human subjects. US Department of Health, Education, and Welfare Public Health Service National Institute of Health, 1968.
21. Thomas Penzel, Max Hirshkowitz, John Harsh, Ron D Chervin, Nic Butkov, Meir Kryger, Beth Malow, Michael V Vitiello, Michael H Silber, Clete A Kushida, et al. Digital analysis and technical specifications. *J Clin Sleep Med*, 3(2):109–120, 2007.
22. S Redline, R Budhiraja, V Kapur, CL Marcus, JH Mateika, R Mehra, S Parthasarthy, VK Somers, KP Strohl, LG Sulit, et al. The scoring of respiratory events in sleep: reliability and validity. *Journal of clinical sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine*, 3(2):169–200, 2007.
23. Sean M Caples, Carol L Rosen, Win K Shen, Apoor S Gami, William Cotts, Michael Adams, Parvin Dorostkar, Kalyanan Shivkumar, Virend K Somers, Timothy I Morgenthaler, et al. The scoring of cardiac events during sleep. *Journal of clinical sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine*, 3(2):147, 2007.
24. MicrosoftCorporation. Kinect for windows sensor. http://msdn.microsoft.com/en-us/library/hh855355.aspx.
25. Zhengyou Zhang. Microsoft kinect sensor and its effect. *IEEE Multimedia*, 19(2):4–10, 2012.
26. Anna Vögele, Björn Krüger, and Reinhard Klein. Efficient unsupervised temporal segmentation of human motion. In *2014 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, July 2014.
27. Björn Krüger, Jochen Tautges, Andreas Weber, and Arno Zinke. Fast local and global similarity searches in large motion capture databases. In *2010 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pages 1–10, July 2010.
28. Jürgen Bernard, Nils Wilhelm, Björn Krüger, Thorsten May, Tobias Schreck, and Jörn Kohlhammer. Motionexplorer: Exploratory search in human motion capture data based on hierarchical aggregation. *IEEE Transactions on Visualization and Computer Graphics (Proc. VAST)*, December 2013.