Combining Wearable and Environmental Sensing into an Unobtrusive Tool for Long-Term Sleep Studies

Marko Borazio
Embedded Sensing Systems
Hochschulstr. 10, 64289 Darmstadt
TU Darmstadt
borazio@ess.tu-darmstadt.de

Kristof Van Laerhoven
Embedded Sensing Systems
Hochschulstr. 10, 64289 Darmstadt
TU Darmstadt
kristof@cs.tu-darmstadt.de

ABSTRACT

Long-term sleep monitoring of patients has been identified as a useful tool to observe sleep trends manifest themselves over weeks or months for use in behavioral studies. In practice, this has been limited to coarse-grained methods such as actigraphy, for which the levels of activity are logged, and which provide some insight but have simultaneously been found to lack accuracy to be used for studying sleeping disorders [8]. This paper presents a method to automatically detect the user's sleep at home on a long-term basis. Inertial, ambient light, and time data tracked from a wrist-worn sensor, and additional night vision footage is used for later expert inspection. An evaluation on over 4400 hours of data from a focus group of test subjects demonstrates a high recall night segment detection, obtaining an average of 94%. Further, a clustering to visualize reoccurring sleep patterns is presented, and a myoclonic twitch detection is introduced, which exhibits a precision of 74%. The results indicate that long-term sleep pattern detections are feasible.

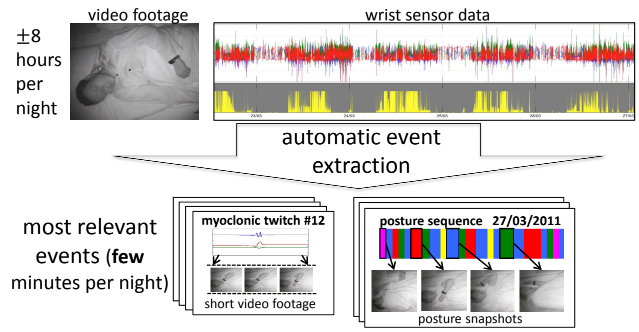

Figure 1: Overview of the detection system: long-term inertial-, time-, and ambient light data from a wrist-worn unit are analyzed for sleep events of interest, and synchronized with camera footage for easy visual inspection.

Categories and Subject Descriptors

H.m [**Information Systems**]: Miscellaneous

General Terms

Design, Hidden Markov Model, Kohonen Self Organizing Map, Support Vector Machine.

Keywords

Activity recognition, sleep detection, long-term monitoring.

1. INTRODUCTION

Spending one third of our life with sleeping, it belongs to one of the prime activities humans pursue. With sleep researchers steadily discovering new ways in which sleep impacts quality of life, it has already been shown that a healthy sleep is at least equally important for our well being as nutrition [7], and that it contributes significantly to regeneration and healing [3]. On the other hand, habits and choices made during the wake state are also known to impact the quality of sleep, such as shifting circadian rhythms, the intake of alcohol [21], or changes in ambient light. There is a reciprocal relationship between sleep and daily life, with shortcomings and problems in the one tending to easily influence the other.

Since general awareness toward the importance of sleep is increasing, personal sleep monitoring applications have recently undergone a surge in variety and commercial success. Many of these units are meant to be worn only during night while being charged on the nightstand during the day, other types of units are recording full circadian rhythm data for multiple, successive days that give more insight in changes in the user's habits and rhythms.

The golden standard for observing sleep/wake patterns is reached by performing a polysomnography (PSG). It incorporates multiple sensing modalities to capture relevant sleep information with typically 20, mostly wired sensors attached to the patient's face, torso and limbs. In addition, the patient is recorded by video for the entire night. As a result, relevant data can be accurately captured allowing in depth analysis. However, PSG is limited for short-term observations, that is, often a few nights are observed only, due to its cumbersome setup. Such monitoring tends to be uncomfortable and less feasible over longer periods.

An alternative approach is performed by actigraphy, which typically captures levels of activity on a minute-by-minute

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
IHI'12, January 28–30, 2012, Miami, Florida, USA.
Copyright 2012 ACM 978-1-4503-0781-9/12/01 ...$10.00.

basis. It is most commonly implemented as a wrist-worn device that can be easily deployed and worn without any additional effort. Used successfully for more than a decade, it is applied for monitoring patients suffering from sleep disorders [12] and psychiatric illnesses.

While PSG captures in-detail data during sleep over a single or a few nights only, actigraphy has been used in somnology as the de facto method for observing long-term trends that become only evident during weeks or months [8]. Given such long-term capabilities, actigraphy then also captures activity levels during the day, resulting in a more holistic view on human activity. However, in order to interpret the data captured, patients are usually required to annotate the sleeping and awake times [15]. Also, obtained data from current actigraphy devices are coarse and limits the interpretability.

This paper focuses on advancing actigraph-like monitoring by increasing the expressiveness of observed data, while keeping its long-term monitoring capability over weeks or months. Our approach is based on three types of observation techniques relevant to sleeping studies, which were derived after discussions with the leading somnologist of the local sleeping lab:

- **Sleep posture changes.** Persons without sleeping disorders tend to change postures between 12 and 20 times during relaxed nights [9], but this can go up significantly in other situations and this can also change drastically over the course of a night. A view on how often the patient changed posture and between which reoccurring postures this happened is therefore helpful, not only to asses quality of sleep [6], but especially for particular sleeping disorders such as obstructive sleep apnea (where certain postures increase apnea [18]) and restless leg syndrome (where frequent posture changes are witnessed).
- **Myoclonic twitches during sleep.** A second feature of interest are involuntary muscular contractions made during sleeping, commonly known as myoclonic twitches, that some people tend to experience when drifting off to sleep, but also during REM phase, or while dreaming. Their detection is not straightforward as these range from subtle and short flexings of (mostly limb) muscles to violent shakes that can last over a few seconds [22, 4].
- **Video feedback.** Real-time and high-frequent video footage is required by expert somnologists to properly assess any posture sequences and myoclonic contractions in sleep study. Even with a more accurate, non-vision motion detection system of changes in sleep posture and muscle contractions, interviews with experts made clear that visual playback of these events is essential in their assessment and in case of doubt.

These techniques were selected as a trade-off between modalities of interest for several sleeping disorders and diagnosis types, but yet to obtain with minimal involvement of the patient. Our overall approach is illustrated in Figure 1. We show that these observations can be performed by non-obstrusive sensors that can be deployed in any household environment and run for an extended period of time. Furthermore, we present an approach to automatically identify the events of sleeping posture changes and myoclonic twitches. Our experimental results indicate that awake and sleep times can be successfully segmented and twitches can be detected with high precision for further analysis. Based on such events, video data is pre-filtered in order to extract and show relevant parts only to the expert and speed up the analysis significantly. Given its low requirements, it is a promising approach for monitoring sleep in home settings.

This paper will first situate its approach among related work of Section 2 in both research and commercial products. Then it introduces in Section 3 the two monitoring systems and in Section 4 the methods used. In Section 5 two night segmentation methods are evaluated, leading to Section 6, where the sleep postures within these segments are clustered. Section 7 focuses on the detection of involuntary muscle contractions during sleep. The final section wraps up with the contributions made in this paper and remaining future work.

2. RELATED WORK

The use of actigraphy for sleep study has been widespread for the last decade, with research and standards of practice committees [8] recommending actigraphs as reliable and valid for normal subjects, and a helpful tool in combination with routine clinical evaluation of insomnia, circadian-rhythm disorders, and excessive sleepiness. Several recent wearable products have been targeting sleep phase detection specifically in order to allow the wearer to wake up at a more convenient sleep stage, or display sleeping trends for the users so that they can keep track of their own circadian rhythms. The most prominent are summarized below:

- The Sleeptracker[1] is a wristwatch-shaped unit that apart from telling the time, also infers whether the user is in deep sleep, light sleep, or awake, using an accelerometer.
- The WakeMate[2] is a wrist-band actigraphy unit designed to be used together with a mobile phone, to which it is wirelessly connected, so that the phone's alarm is triggered only during the lightest phase of sleep within a 20-minute window of the desired wake-up time.
- The aXbo alarm clock[3] is packaged as a stand-alone application in the form of an alarm clock that wirelessly communicates with a wrist-band unit.
- The Zeo[4] is similarly using an alarm clock base unit with a worn sensing device, but the latter is a head-band rather than a wrist-band that measures electrical activity produced by the brain.
- The FitBit[5] is a similar inertial sensor-based device that can be clipped to clothing or an arm strap, and comes with software to extract basic sleep information.

More expensive units can be used that target a specific sleep disorder. For obstructive sleep apnea patients, for instance, a large variety in oximetry and breathing monitoring products are available. They are not straightforward to set

[1] SleepTracker: `http://www.sleeptracker.com`, last access 10/2011
[2] WakeMate: `http://www.WakeMate.com`, last access 10/2011
[3] aXbo: `http://www.axbo.com`, last access 10/2011
[4] myZeo: `http://www.myzeo.com`, last access 10/2011
[5] FitBit: `http://www.fitbit.com`, last access 10/2011

up however, are prescribed for shorter studies only and are not able to detect all types of sleep apnea [10]. A minority of the above products reveals details on how night segments are calculated from the basic actigraphy log, making their detection mechanism hard to reproduce.

Other common techniques include non-wearable solutions that deploy sensors in the home of the patient. Bain et al. [1] describe for instance a pressure mapping technology that could give an extremely detailed view on the total body posture of the patients throughout the night. Recent research in sensor networks, such as [11], have offered similar fine-grained approaches to detect and monitor the patient's body positions and movements by active RFID-based accelerometers (WISPs) placed on the mattress. Camera-only methods have also been explored, with [14] as one example where the sleeper's motion is detected by a nightstand camera's frame-by-frame differences and estimates of body posture.

Several studies are dedicated to the detection of body posture and movements during the sleep, motivated by especially sleep apnea [23] and as a tool to measure for sleep quality [14]. This paper addresses the long-term challenges in particular by integrating methods for night segmentation, posture clustering, and myoclonic twitch detection, in order that these can be applied in behavioral monitoring.

3. MONITORING EQUIPMENT

This paper sets out to investigate first and foremost the detection of posture changes and myoclonic twitches during sleep. The sensor that is focused on to drive these detections is a wrist-worn device that records ambient light, the time, and inertial sensor data for an extended period (at least a week at a time). For visual feedback of the detections, an active infrared night-vision unit is used that is synchronized with the wrist worn sensor. This section will introduce both of these monitoring platforms.

3.1 Wrist-Worn Sensor

It is unfortunately hard to find publicly-available wearable data logging platforms that are both comfortable to wear continuously for several days and nights at a time, and able to store data while providing high-resolution inertial data. Yet these requirements are crucial to this paper's scenario, as it assumes a long-term deployment of at least several weeks, no user interaction with the setup in this period, and minimal discomfort to the person that is monitored. Especially the requirement to capture the myoclonic twitches requires a relatively high sampling of the accelerometer data, which makes long-term acquisition particularly challenging.

The presented data in this paper therefore come from a custom-made prototype: The device that delivers the data for the posture- and muscular twitch detections is a self-contained wrist-worn device that records 3D acceleration at 100Hz, as well as ambient light and time and calendar information. It can record these data on local flash storage (a removable microSD card) for about 2 weeks before its rechargeable battery is depleted. Figure 2 depicts the prototype with and without enclosure and straps.

With the OLED display mostly powered off (the wearer can request the current time by double-tapping the watch), the prototype runs for about 2 weeks on a 600mAh Li-Ion rechargeable battery. When the data is uploaded, the USB connection also provides power to an on-board recharging

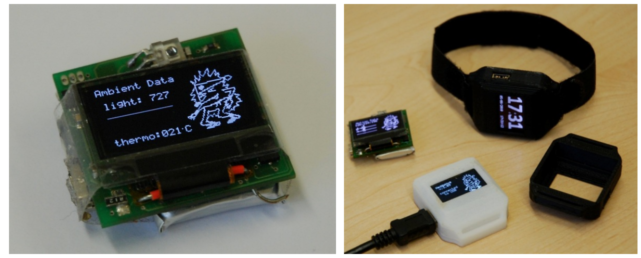

Figure 2: The wrist-worn prototype is intended to be worn continuously and is able to function as a basic wristwatch (via its OLED display and integrated realtime clock). It stores data from a 3D accelerometer at 100Hz, as well as ambient light data, on a local microSD-based flash memory.

Figure 3: The InfraRed camera (shown right) is an off-the-shelf logging device, illuminating the observed area with an array of 7 strong IR LEDs and producing high-quality 640x480 images, tagged with second-resolution timestamps. Left are some examples of test subjects.

circuit connected to the battery. The unit is sealed in a robust resin enclosure to prevent damage from falls and accidental splashes of water (subjects were told to remove the unit when showering or swimming).

With this prototype, continuous study is feasible with 2 week intervals for charging and downloading the data. It is important to note that nothing more is required from the wearer of the sensor after starting and attaching it: apart from its double function as a wristwatch, the user is normally not required to press buttons, annotate data, or fill in questionnaires. The algorithms discussed later will filter out the night segments and sleep events that might be of interest.

3.2 IR Camera

Unlike in the case of the previously discussed wrist-worn sensor prototype, there are already a great deal of commercial night-vision cameras available. The TrendNet TV-IP-422W was chosen in this paper's deployments as it provides an adequate resolution at a frequency of 30 frames per second, and is equipped with an array of infrared Light Emitting Diodes (LEDs) which are powerful enough to sufficiently illuminate an area from up to 5 meters away (see Figure 3). It can be configured to provide the recordings on a network drive via ethernet or the local wireless network, but also on local flash storage via a USB connector. A pan-tilt mo-

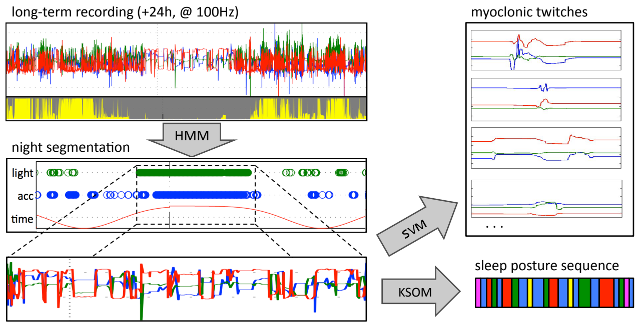

Figure 4: The data flow of our method: A continuous dataset of typically several weeks is taken and analyzed for night sleep segments. Within these, postures are clustered and assigned colors to visually represent entire nights, and myoclonic twitches are detected with their 100Hz 3D signals kept.

tor allows the camera to re-adjust itself to fully focus on the sleeping subject. At deployment, the camera's embedded real-time clock is synchronized to that of the wrist-worn sensor so that data can be merged afterwards.

The IR camera is powered from a wall-socket and as such can be activated for longer periods of time. The camera was scheduled to automatically switch off between 11am and 7pm, and was by default configured to store still pictures on the local flash storage and movies in 10-minute chunks to an ethernet-attached netbook. Since the data produced by the camera for one single day sizes to about 9 Gigabytes on average, this recording setup can remain deployed for longer stretches of time as well without maintenance. The wireless capability of the camera was turned off.

The studies in this paper used one camera as a means to obtain the ground truth for 1) whether the test subject was sleeping, 2) in which posture the test subject was lying, and 3) myoclonic twitches that were spotted in the data. The intended usage scenario in this paper uses the camera for visual inspection by a somnologist: By filtering out everything but the relevant sleeping postures and short movies of possible myoclonic twitches, a sleep expert has the required material for visual inspection on a PC of the raw signals of what the wrist-worn sensor has captured and the subsequent proposed methods have extracted from them.

4. METHOD OVERVIEW

This paper's proposed method to find significant events in sleep based on motion from a highly deployable system consists out of three major steps. First, the continuously recorded information from the wrist-worn sensor is automatically processed and segmented into awake and sleep. After that, non-motion data in the resulting sleep segments that occur and re-occur are automatically clustered into postures in order to visualize general trends over several nights. In a last step, myoclonic twitches are segmented out of the remaining motion segments. Figure 4 illustrates how the original raw data from the wrist-worn sensor is processed.

The envisioned scenario of our method comprises the following steps, from preparation to data analysis:

1. The user obtains the camera and wrist sensor
2. The user places the camera in the bedroom, synchronizes with the wrist sensor, and wears the wrist sensor
3. Throughout the monitoring phase of multiple weeks, both camera and wrist sensor record their data continuously
4. After the monitoring phase, the wrist sensor is synchronized with the camera, and the proposed method provides:
    - the sequence of clustered postures per night
    - the detected myoclonic twitches per night
    - video footage of extracted events

The following section starts with segmenting continuous data into sleep and awake segments.

5. NIGHT SEGMENTATION

We present in this section an adaptive method that extracts nightly sleep segments from the wrist-worn sensor's continuous recordings, using a method that can be bootstrapped from time-use data and personal recordings, combined with measured light intensity and physical motion.

5.1 Time of Day

The time of day is generally a strong clue for estimating the night sleep segment in a day, as most people tend to adhere to a strict circadian rhythm with regular bed- and wake-times. Furthermore, with the help of time-use databases, it is feasible to start off with a prior estimate that is generated from a sizable amount of questionnaire results.

The wrist-worn sensor logs its data with timestamps that are provided by the on-board real-time clock. With an expected deviation of approximately 2 seconds per week, this is sufficient to fuse the recorded data with those from other modalities such as the IR camera. For characterization of the night-time sleep segment, the minute of the day is extracted from the complete timestamp (which holds other information such as year, month, day, and day-of-week as well).

A user-specific model for linking time of day to the night sleep segment is trained from recorded data, as will be done for the other two features. However, in the case of time of day, it is also possible to use existing information that is collected and made available by national statistics agencies or commercial institutions on what activities people tend to do during the day. Inspired by a study by Partridge and Golle [20], which illustrated that it is possible to use these often large-scale datasets to create informed activity classifiers, this was found particularly promising in the special case of sleep, for which an exceptionally large amount of samples are available in time-use study data.

The left plot in Figure 5 shows a minute-by-minute normalized histogram collected from such a database[6]. The right plot of Figure 5 shows for comparison the same information from one of our test subjects.

[6] Obtained from the Federal Statistical Office (Statistisches Bundesamt) in Germany from a total of over 12000 participants, which kept diaries of daily activities for three days each.

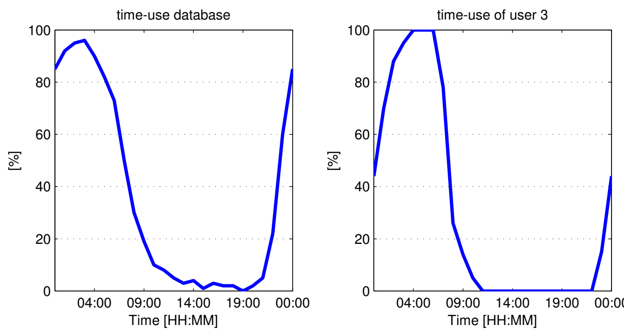

Figure 5: An example of typical sleeping times, from a 12000-subjects time use study over a period of 3 days each (left), and a one-person dataset over 30 days (right). Similarities in both diagrams are visible and are used as prior for 2 different classifiers.

5.2 Ambient Light

A second feature candidate for detecting overnight sleep segments is the absence of light in the environment, as most people tend to sleep in darkened environments. Studies have shown that the gradually dimming light at dusk and increasing light at dawn are used to regulate sleep and wake up times, and that these could be used in sleep therapy [19]. The mean over several light sensor values of a specified time frame is used as a feature.

The wrist-worn sensor has been designed so that the ambient light sensor is directed outward, to avoid as much as possible an occlusion by long sleeves or jackets. Furthermore, the sensor, a TSL250 photodiode, has been chosen and configured for maximal sensitivity to light while providing a low dark (offset) voltage. The prototype reads the voltage at full 12-bit resolution, providing a sensor reading that is capable of detecting particularly small lighting changes under poorly-lit environmental settings.

5.3 Physical Activity Intensity

A large body of research, including many studies using actigraphy, indicates that activity intensity levels tend to be more elevated during the day and fairly low during the night for subjects with normal sleeping behaviors. As such, this can be used as a discriminant feature for the recognition of the night sleep segments. Since the wrist sensor in our studies is worn on the dominant hand, and since the sampling of the on-board accelerometer is set at a frequency of 100Hz in order to capture the myoclonic twitches, the calculation of standard deviation provides a robust value to represent the wearer's activity level.

5.4 Evaluation

This section will discuss the use of the previously described time, ambient light and physical amount of motion features to estimate the start and stop times, as well as duration, for the night sleep segments in continuously recorded long-term data. Two classifiers are used: One baseline algorithm that performs simple histogram-based thresholding on the training data, and one Hidden Markov Model (HMM) based classifier [17, 16] which uses observation sequences of the previously described features.

sub	gen	age	hrs	comments
1	female	33	336	at 5th month of pregnancy
2	male	30	1344	normal sleep
3	male	30	432	normal sleep
4	male	28	624	irregular night segments
5	male	35	360	periodic limb movem. disorder
6	male	35	96	delayed sleep phase syndrome
7	male	61	648	early morning awakening
8	male	26	576	irregular night segments

Table 1: The group of participants used in the evaluation, specifying gender, age, the length of their dataset in hours, and additional factors which are likely to be sleep-relevant. Two of the participants were diagnosed with a sleep disorder, the six others have no known sleep disorders.

For the evaluation of the night segmentation, we used data recordings from a variety of test subjects as summarized in Table 1: These were selected to ensure that several different types of subjects were being included and that the segmentation algorithm can be properly stress-tested with sufficiently contrasting types of sleeping behaviors. Two of the test subjects were included that are diagnosed with a specific sleep disorder, one female subject was monitored being 5 months pregnant, and one elderly test subject was included. All subjects were recruited outside the sleep lab environment to first test the set-up within the research community. A second phase will be set to record patients from a local sleeping lab (see Section 9).

The datasets were split up for the experiment in separate subsets with a duration of 24-hours each, from noon on one day to noon on the next, so that each timespan would contain exactly one night sleep segment. The classification performance of the segment was then measured in both a leave-one-day-out and a leave-one-user-out cross-validation experiment. The purpose of these two experiments is to 1) see how well the algorithm performs on previously trained data from the same user, and 2) how well the classification does on data from a new user.

Classifier 1: Threshold-based segmentation. The first algorithm is a Gaussian model-based classifier that calculates the variance and mean parameters for the light intensity and motion data from the training data, and uses a likelihood per minute of the awake state from the time-use database. New data is used as input to the model and a thresholded vote among all values over a sliding window of several (5, 10, or 15) minutes is then used to classify the night sleep data in a robust way.

Classifier 2: HMM-based segmentation. The second algorithm is based on two two-state discrete Hidden Markov Models (HMMs), which enables capturing changes in sleep habits for new training data efficiently. As features the variance of acceleration, mean of light and time of day (in minutes) are used. The first HMM models the data taken during the awake state, and the other for the sleep state. After training the HMMs, the highest likelihood for new sequences for each classifier of several minutes (5, 10 and 15) from the testing data decides whether the method assigns the latest observation in this sequence as asleep or awake data.

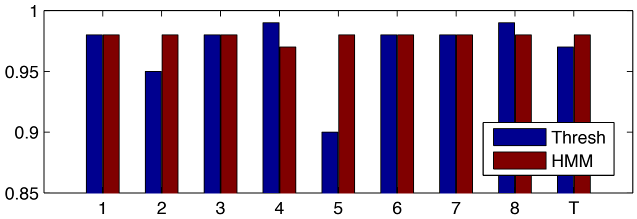

Figure 6: The recall values for the threshold-based (Thres) and HMM-based classifiers for the best performing window size of 10 minutes, obtained through leave-one-night-out cross-validation and displaying the total recall (T).

	per subj			across subj		
subj	prec	rec	acc	prec	rec	acc
1	0.79	0.98	0.91	0.88	0.98	0.95
2	0.92	0.98	0.97	0.92	0.96	0.96
3	0.87	0.98	0.94	0.95	0.98	0.97
4	0.87	0.97	0.94	0.92	0.97	0.96
5	0.82	0.98	0.92	0.95	0.89	0.94
6	0.72	0.98	0.86	0.96	0.97	0.97
7	0.93	0.98	0.97	**0.80**	**0.75**	**0.84**
8	0.89	0.98	0.95	0.96	0.98	0.98
T	0.85	0.98	0.93	0.92	0.94	0.95

Table 2: The HMM precision (prec), recall (rec), and accuracy (acc) results for all subjects under per-subject training (left half of table) and cross-subject training (right half of table). Results are from leave-one-day-out and leave-one-subject-out cross validation respectively.

5.5 Results for night segmentation

Results for recall are summarized in Figure 6, depicting the cross-validation results for person-dependent training. A comparison between the two classifiers resulted in a set of close detection scores, both peaking at a window length of 10 minutes. However, in case of subject 5, where a lot of motion is present during the night, larger differences can be observed in the results for both person-dependent and cross-subject training. Classifier 1 uses as prior for sleep the time-use database, leading to lower recall in contrast to classifier 2.

Table 2 lists the results for the HMM-based method using a 10-minute window, which performed best over all tested parameters. In general, the results perform well for most subjects, with cross-user training performing slightly better for most test-subjects. A significant outlier is the result set from subject 7, where the recognition results are hurt from cross-user training. This is explained by the severe early morning wakeup times (4am on average) and the overall higher amount of motion throughout nights, hurting especially the recall of night segments.

Figure 7 shows an example where the night segments for several weeks of data from one person (subject 2) are classified. Gaps in the data correspond to sections of the dataset between recordings or sections where the device was intentionally turned off. Typical false positives can be seen as small sections in the late evenings, where this test subject was often watching television in a darkened environment.

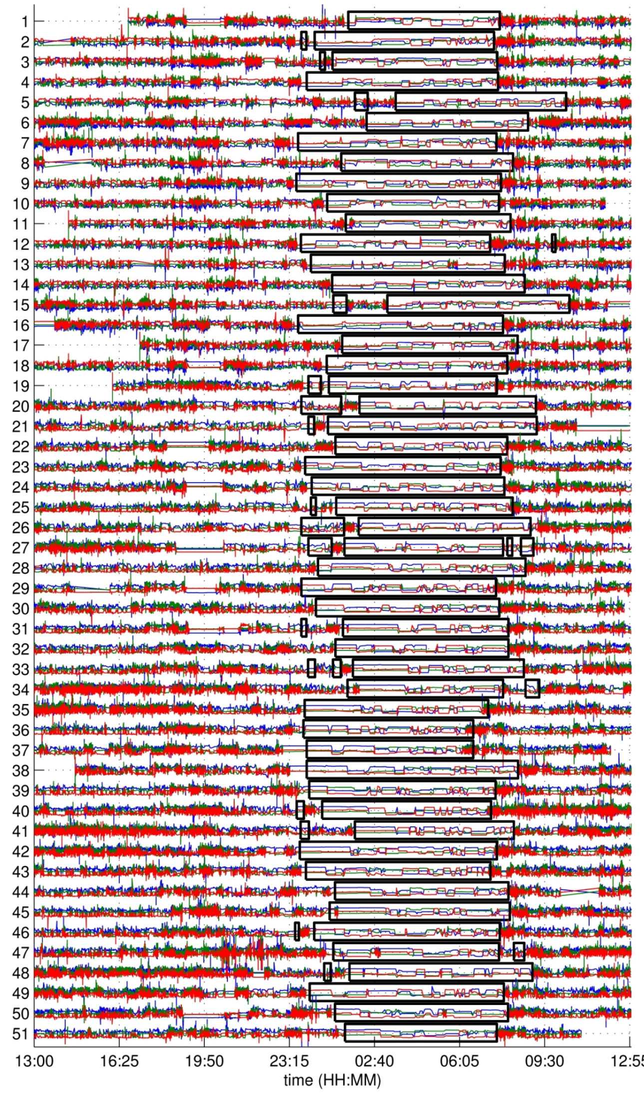

Figure 7: The raw classification output from night segmentation (black rectangles) for the entire dataset from test subject 2 (with each row representing a day's acceleration data) for the HMM-based classifier.

Such false positives were discarded by selecting the largest segment only as the most likely night sleep segment for further analysis.

5.6 Discussion

From the experiments in this night sleep segmentation section follows, that a high recall classifier can extract most of the night sleep segments from continuous activity data. A suitable scenario for our classifier is the detection of advanced or delayed sleep phase syndrom [24], since we train the classifier on personal sleeping habits, which is identified by the classifier in the shift of night segments. This sleeping disease is characterized by a patients habit to go to bed late and wake up late (delayed) or go to bed early and wake up early (advanced). Deriving a prior from such a sleeping habit is feasible with our system and requires a person dependent training for the classifier.

Although all the test subjects were asleep only during the night, the HMM classifier should be able to detect sleep during the day if training data is available. Data from a person napping during the day is used as training data for the HMM, which is able to adapt to this scenario. Further studies need to be conducted to confirm this theory.

The remaining steps taken in our tool split these segments into much smaller segments of either non-motion or motion to detect two types of sleep features of interest.

6. POSTURE CLUSTERING

From the method described in the previous section, our tool obtains a set of candidate night segments, which are focused on to find two features that are relevant to characterize sleep: The first being described in this section will give a coarse-grained view on a subject's postures from static data and their common occurences through all night segments and how long they lasted. The next section will look at the non-static data and detect candidate occurrences of myoclonic twitches. Although the filtering of the night sleep segment for non-motion data is straightforward, the main question that will be answered in this section relates to how these static postures can be visualized most appropriately.

The reoccurring sleep postures are modeled by using a clustering method which facilitates on optimal visualization of the posture sequences later on. Similar to the popular k-means method, the Kohonen Self-Organizing Map (KSOM) [13] is a clustering algorithm that holds a fixed number of cluster centroids, to which new data samples are clustered in an iterative way by selecting and updating the cluster for which the centroid is closest to this new input. 36 cluster centroids are chosen for that, keeping the range of the map small, which are allocated on a semantic map created by the KSOM. Similar values are mapped close together, while dissimilar are mapped apart. The choice for the Kohonen map brings in this case an added value in terms of visualization: By requiring that clusters are structured along a tight topology, neighboring clusters obtain centroids after training that are close in Euclidean space. By then assigning a gradual color map to the Kohonen clusters, the clustering of posture data will result in similar postures being assigned a similar color. Thus, even if not the same cluster is assigned to 2 similar postures, the visual representation for both will look very much alike.

The color-code is obtained by using as training input the night segments from all subjects which leads to a unique posture coloring for each subject. The map grid coordinates are normalized to a unit square, and each coordinate is mapped to a color. As output we receive a grid allocation, which is used as input for a HMM classifier that is trained to show similarities through the dataset per subject, resulting in a typical posture sequence which was used in the following section.

6.1 Evaluation

The method to evaluate the posture representation is to conduct a straight forward survey by multiple users. The users were recruited in the research facilities, but also among friends and family. The survey consists of posture representations of several nights by five subjects and is displayed to 60 users (Figure 9). Then a single night from each of the subjects was shown to the users (Figure 8, a)-e)). The users were asked to assign each of these nights to the night-collection of subjects from Figure 9.

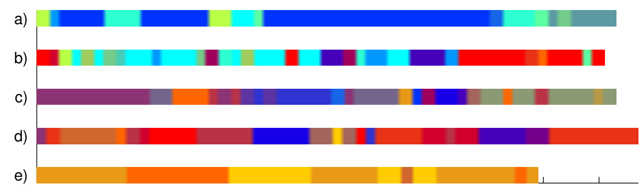

Figure 8: Visualization of the typical nights of the 5 subjects (a-e). The correct combinations to the data sets in Figure 9 are: 1c, 2d, 3b, 4a and 5e.

The results of the survey were obtained as follows: The amount of correct answers per typical night (figure 8 a-e) is divided by the amount of participants, resulting in an accuracy per question. The overall accuracy is then calculated by the sum of each individual accuracy divided by the amount of questions.

6.2 Results

An overall accuracy of 92% is gained for correct allocation of the typical night to the subjects. Figure 9 shows how the test subjects differ in their postures. A suitable scenario of such a representation is to compare a new night of a patient to the previous nights just by the color encoding which leads directly to outliers, which then can be scrutinized further.

While conducting the inquiry, some users had problems allocating the typical night to subjects 1 and 2, since they exhibit similar coloring. The postures of subjects 3 and 4 contain a lot of turquoise-like coloring, but differ in occurrence of red. Assigning the night to subjects 4 and 5 was almost perfect with an accuracy of about 99%, showing explicit postures for each subject.

6.3 Discussion

Interestingly, the instances of nights could be assigned almost perfectly within our study. While we performed a feasibility study that investigates the visual preservation of typical nights using user identification, the promising results are interesting for a different reason. When observing a single user only, the same method can be used to identify non-typical nights which can serve as indicator for somnology analysis.

During the survey we displayed five subjects and their color-coded postures to different users and without explaining for which purpose this inquiry was conducted, participants immediately began answering the questions. This can be explained by the way the information was presented: using colors as representation of the posture characteristics simplifies analysis. However, in a first questionnaire we displayed the postures from all 8 subjects with their typical nights. Again, people were asked to allocate the nights, which proved to be difficult, since too many plots with colors were displayed at once. Therefore, the optimum representation of five subjects was chosen, since the feasibility of assigning postures correctly is the goal of this study.

After the posture data has been removed from the night segments, the data that remains is a mixture of motion data from the subject moving between postures, the subject mov-

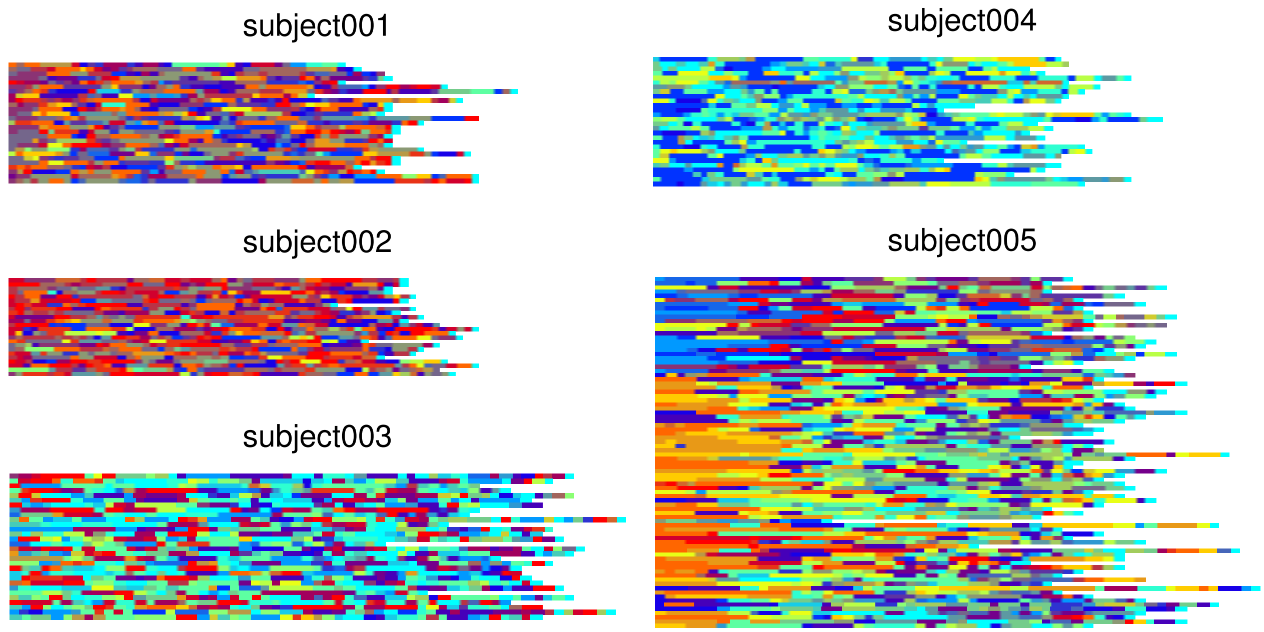

Figure 9: The color coded postures used in the survey for allocating typical nights to subjects. Displayed are all night-collections from subjects 1 to 5.

ing during wakeful periods, and a third type of *involuntary* motion which requires a specific detection step. This step is described in the next section.

7. MYOCLONIC TWITCH DETECTION

Myoclonic twitches describe spontaneous muscle contractions and occur by sporadic limb movements. Note, that it is not the number of twitches that is relevant, but how such twitches are expressed. Therefore, we propose to optimize the precision, leading to a low false positive rate, to identify such events. Given few but correct twitches, the video data can be pre-filtered and facilitate analysis for experts significantly. Our classifier is thus optimized for high precision at the cost of lower recall.

The following two-step process is used to detect myoclonic twitches from the data that was not assigned as posture: First, filtering is done on the duration of the motion in question: since it is known that these twitches are generally not longer than mere seconds, any segment over three seconds is discarded. From the remaining motion data we calculate three different features, which describe a twitch pattern. At first, we calculate the **length of the motion** data, than we determine the **euclidean distance of the start and end values** of the motion segment since a twitch mostly results in almost the same wrist position as before twitching. Lastly, the **distance between the minimum and maximum** of the motion data is computed since non-twitching motion data exhibits larger peaks in contrast to twitches.

These three features are used as input to a support vector machine (SVM) [5], a common linear classifier which is trained on sample data from both myoclonic twitches as positives and posture changes as negatives, obtained by annotations where these events were also clearly visible on the video footage. Figure 10 displays different types of myocloni

night	p 1s	r 1s	p 2s	r 2s	p 3s	r 3s
1	0.58	0.96	**0.62**	0.75	0.55	0.48
2	0.71	0.94	**0.76**	0.71	0.74	0.51
3	0.59	0.70	0.74	0.57	**0.88**	0.41
4	0.77	0.67	**0.81**	0.67	0.80	0.51
5	0.50	0.78	**0.71**	0.59	0.71	0.47
T	0.63	0.81	0.73	0.66	**0.74**	0.48

Table 3: The precision (p) and recall (r) results for a detection window of 1, 2, and 3 seconds on a subset of five nights by one subject for which clear video footage was used to obtain ground truth. Additionally, the total (T) results for p and r are shown.

which occurred in the data set, including multiple ones from subject 5, suffering from periodic limb movement, appearing in short intervals of several seconds.

7.1 Evaluation

Evaluation was performed using a 5-fold leave-one-night-out cross-validation, using video footage to denote ground truth for the twitch detection. The evaluation of the myoclonic twitch classifier was performed on a one-person subset of the dataset only, since participants 4, 5, and 6 displayed obvious myocloni in their data, with only participant 5, suffering from a periodic limb movement disorder, delivering a significant amount of twitches.

7.2 Results

Table 3 shows the precision and recall figures for the SVM classifier based on different initial windows, showing an ideal window size between 2 and 3 seconds. From visual inspection of the false positives, many could be contributed

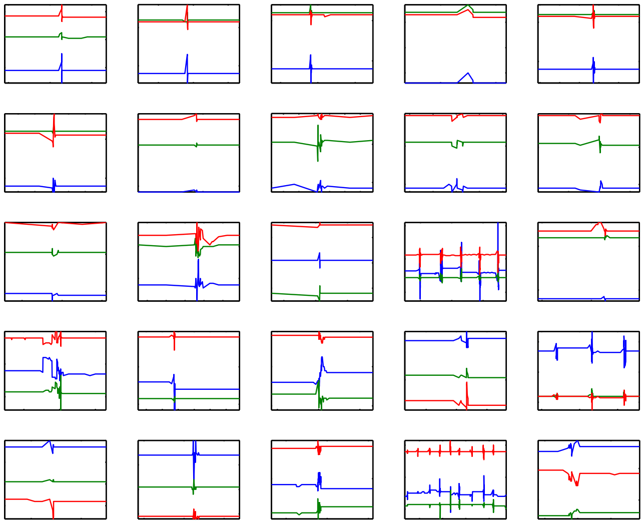

Figure 10: Some examples of the observed myoclonic patterns, including consecutive multiple twitching, which occurred during nights of mostly subject 5 over less than 3 seconds.

to external factors that were neither myocloni nor posture changes, but rather conscious short motions during awake periods. Missed detections largely constitute very slight contractions.

The results are preliminary but demand to be further evaluated within a cooperation at a sleeping lab, where myocloni can be detected accurately and without painstakingly browsing the data for possible twitches, as was done in this study.

8. SYSTEM DISCUSSION

During the process of building and perfecting the prototypes for this system and instructing subjects on its usage, several surprising issues and obstacles had to be solved. Apart from the predictable difficulties that come with designing robust measurement prototypes and deploying ubiquitous technology in domestic environments, especially those worn 24/7, other interesting lessons were learned. The following sections will discuss these in more detail.

8.1 Video Footage

One advantage of the continuous logging that was not specifically focused on so far, is the ability to correlate activity during the day with possible effects for the following nights. This requires an activity recognition component, however, which might be added at a later stage. Similarly, one might expect the video footage to be exploited more and analyzed for steady body postures and sudden motions instead of relying on a body-worn sensor. An initial study investigating this avenue met a lot of obstacles during the first datasets, as illustrated in Figure 11. The same limitations hold also for the visual inspection: especially for the occlusion by blankets, exact body postures and twitches are sometimes hard to verify. In our experimental that was largely taken during wintertime, this amounted for instance to about 18% of all myoclonic twitches observed in the wearable data.

Figure 11: Potential difficult situations for vision-based recognition that were observed during the study: multiple participants in the scene, reflecting metallic objects such as piercings, pets moving into the bed, and thick blankets covering the entire participant.

8.2 Privacy Issues

Recording subjects in their most private room - the bedroom - is a delicate matter and required a privacy consent for the subjects to be signed. For this purpose we used the privacy guidelines from [2]. It was ensured that all the data was anonymized, stored in a secure place and used only by allowed researchers. From our experience with the test subjects, ensuring the privacy with a privacy policy is the minimum requirement for conducting such studies.

8.3 Timing Considerations

As monitoring periods will be extended to the scale of months, one issue that can be expected to become prominent is a possible larger drift between the wrist sensor's timestamps and those embedded in the pictures and movies. For the study data recorded for this paper, an extra 2 seconds were attached before and after the signal, and a variable offset was built in the visualization tool to match the 100Hz accelerometer signal and the recorded movie more perfectly. As drifts due to temperature and humidity fluctuations continue over longer periods, this approach might become less scalable when aiming at year-long logs.

9. CONCLUSIONS AND FUTURE WORK

A tool is presented and evaluated, which combines effortless deployment with capturing of expressive sleep. Given automatically detected events during sleep, data is synchronized with a camera to highlight relevant events and extract them from a large corpus of data which is infeasible to be skimmed by humans. Given these properties it enables video analysis in the wild and additionally uses a power-efficient wrist-worn activity sensor for long-term recording. The sensor data is analyzed with an HMM-based method for occurring night sleep segments, which are in turn analyzed for reoccurring postures and myoclonic twitches. Visual inspection is built in the tool by means of an IR camera, which together with detection techniques that are chosen to deliver a high recall on detections, make the output available for scrutinizing by sleep experts.

Using 'in vivo' datasets from eight test subjects with a high variety of sleeping patterns, evaluation shows that night segmentation with *high recall* (i.e., almost all sleep segment data is retrieved) can be achieved by fusing an HMM-based method with a model obtained by a time-use study. We also found that a common clustering technique is sufficient to accurately capture the most striking sleep postures, which was evaluated by users who assigned typical nights to a subject with an overall accuracy of 92%. The detection of myoclonic twitches, optimized for a *high precision* (to avoid swamping the system with false positives) was able to achieve precision and recall of 73% and 66% respectively.

Previous sleep research involving myoclonic twitches indicates that they are more likely to occur in data from people with an irregular sleep schedule and sleep deprivation. The tool resulting out of this study enables detection of myoclonic twitches over long monitoring periods with an inexpensive setup of devices. Further studies on a variety of sleep disorder patients are planned in local sleep laboratories but also in the patient's home, where this tool's effectiveness is evaluated as a complimentary technique to more expressive features of sleep which allow inference of sleep quality. The variety of people, either with or without sleep disorders will give an insight into features that are suitable for detecting sleep quality.

10. ACKNOWLEDGEMENTS

The authors would like to thank the staff at the collaborating sleep laboratories, as well as the anonymous volunteers that were kind enough to wear our sensor and deploy the camera at their homes.

11. REFERENCES

[1] M. A. Bain D, Ferguson-Pell M. Evaluation of Mattresses Using Interface Pressure Mapping. *J Wound Care*, 12(6):231-5, 2003.

[2] H. Becker, M. Borazio, and K. Van Laerhoven. How to Log Sleeping Trends? A Case Study on the Long-Term Capturing of User Data. *Smart Sensing and Context*, pages 15–27, 2010.

[3] M. H. Bonnet and D. L. Arand. Insomnia, Metabolic Rate and Sleep Restoration. *Internal Medicine*, 254:23–31, 2003.

[4] J. Caviness. Myoclonus. In *Mayo Clinic Proceedings*, volume 71, page 679. Mayo Clinic, 1996.

[5] C. Cortes and V. Vapnik. Support-Vector Networks. *Machine Learning*, 20(3):273–297, 1995.

[6] J. De Koninck, P. Gagnon, and S. Lallier. Sleep Positions in the Young Adult and their Relationship with the Subjective Quality of Sleep. *Sleep*, 6(1):52, 1983.

[7] C. A. Everson, M. Bergmann, and A. Rechtschaffen. Sleep Deprivation in the Rat: III. Total Sleep Deprivation. *Sleep*, 12(1):13–21, 1989.

[8] G.J. Welk and J.A. Schaben and J.A.Jr. Morrow. Reliability of Accelerometry-Based Activity Monitors: A Generalizability Study. In *Medicine & Science in Sports & Exercise Vol.36 No.9*, pages 1637–1645, 2004.

[9] S. Gordon, K. Grimmer, and P. Trott. Self Reported Versus Recorded Sleep Position: An Observational Study. In *Internet Journal of Allied Health Sciences and Practice Vol2No1*, 2004.

[10] S. Gyulay, L. Olson, M. Hensley, M. King, K. Allen, and N. Saunders. A Comparison of Clinical Assessment and Home Oximetry in the Diagnosis of Obstructive Sleep Apnea. *The American review of respiratory disease*, 147(1):50, 1993.

[11] E. Hoque, R. F. Dickerson, and J. A. Stankovic. Monitoring Body Positions and Movements During Sleep Using WISPs. In *Wireless Health 2010*, WH '10, pages 44–53, New York, NY, USA, 2010. ACM.

[12] G. Jean-Louis, F. Zizi, H. Gyzicki, and P. Hauri. Actigraphic Assessment of Sleep in Insomnia: Application of the Actigraph Data Analysis Software (ADAS). In *Physiology & Behavior Vol.65 No.4-5*, pages 659–663, 1999.

[13] T. Kohonen. The Self-Organizing Map. *Proceedings of the IEEE*, 78(9):1464–1480, 1990.

[14] W. Liao and C. Yang. Video-Based Activity and Movement Pattern Analysis in Overnight Sleep Studies. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4. IEEE, 2008.

[15] M. Michael Littner, C. Kushida, D. Dennis Bailey, R. Berry, D. Davila, and M. Hirshkowitz. Practice Parameters for the Role of Actigraphy in the Study of Sleep and Circadian Rhythms: An Update for 2002. *Sleep*, 26(3):337, 2003.

[16] K. Murphy. Hidden Markov Model (HMM) for Matlab. http://www.cs.ubc.ca/ murphyk/Software-/HMM/hmm.html.

[17] K. Murphy. Dynamic Bayesian Networks. *Probabilistic Graphical Models*, 2003.

[18] A. Oksenberg and D. Silverberg. The Effect of Body Posture on Sleep-Related Breathing Disorders: Facts and Therapeutic Implications. *Sleep Medicine Reviews*, 2(3):139–162, 1998.

[19] T. Partonen and A. Magnusson. *Seasonal Affective Disorder: Practice and Research*, volume 66. Oxford University Press Oxford, 2001.

[20] K. Partridge and P. Golle. On Using Existing Time-Use Study Data for Ubiquitous Computing Applications. In *UbiComp'08*, pages 144–153, 2008.

[21] T. Roehrs and T. Roth. Sleep, Sleepiness, and Alcohol Use. *Alcohol Research & Health*, 25(2):101-9, 2001.

[22] E. Roze, P. Bounolleau, D. Ducreux, V. Cochen, S. Leu-Semenescu, Y. Beaugendre, M. Lavallard-Rousseau, A. Blancher, F. Bourdain, P. Dupont, et al. Propriospinal Myoclonus Revisited. *Neurology*, 72(15):1301, 2009.

[23] K. Van Laerhoven, M. Borazio, D. Kilian, and B. Schiele. Sustained Logging and Discrimination of Sleep Postures with Low-Level, Wrist-Worn Sensors. In *Wearable Computers, 2008. ISWC 2008. 12th IEEE International Symposium on*, pages 69–76. IEEE, 2008.

[24] E. Weitzman et al. Delayed Sleep Phase Syndrome: A Chronobiological Disorder with Sleep-Onset Insomnia. *Archives of General Psychiatry*, 38(7):737-46, 1981.