

SLEEPGUARD: Capturing Rich Sleep Information using Smartwatch Sensing Data

Sleep is an important part of our daily routine – we spend about one-third of our time doing it. By tracking the sleep-related events and activities, sleep monitoring provides the decision support to help us understand the sleep quality and the causes of poor sleep. Wearable devices provide a new way for sleep monitoring, allowing us to monitor sleep from the comfort of our own home, but existing solutions do not take full advantage of the rich sensor data provided by these portable devices. In this paper, we develop a novel approach to track a wide range of sleep-related events using smartwatches. We show that it is possible to track, using a single smartwatch, sleep events like body postures and movements, acoustic events, and illumination conditions. From these events, a statistical model can be designed to effectively evaluate a user's sleep quality across various sleep stages. We evaluate our approach by conducting extensive experiments involving fifteen users across a 2-week period. Our experimental results show that our approach can track a richer set of sleep events, provide better decision support for evaluating sleep quality, and help to identify causes for sleep problems compared to prior work. We also show that SLEEPGUARD can help users to improve their sleep quality by helping them to understand causes of sleep problems.

CCS Concepts: • **Human-centered computing** → **Ubiquitous and mobile computing**;

Additional Key Words and Phrases: Smartwatch, Sleep events, Sensing

ACM Reference Format:

. 2018. SLEEPGUARD: Capturing Rich Sleep Information using Smartwatch Sensing Data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 9, 4, Article 39 (February 2018), 30 pages. <https://doi.org/0000001.0000001>

1 INTRODUCTION

Sleep plays a vital role in good health and personal well-being throughout one's life. Lack of sleep or poor quality of sleep can lead to serious, sometimes life-threatening, health problems [7, 20, 44], decrease level of cognitive performance [4, 6], and affect mood and feelings of personal well-being [55, 56]. Besides having an adverse effect on individuals, insufficient or poor quality sleep has a significant economic burden, among others, through decreased productivity, and medical and social costs associated with treatment of sleep disorders [47]. Indeed, to highlight the significance of sleep quality, the Centre for Disease Prevention (CDC) has declared insufficient sleep as a public health problem in the US [24], and the concern is widely shared amongst other industrialized countries.

Traditionally, sleep monitoring is performed in a clinical environment using Polysomnography (PSG). In PSG, medical sensors attached to human body are used to monitor events and information such as respiration, electroencephalogram (EEG), electrocardiogram (ECG), electro-oculogram and oxygen saturation [29, 45, 53, 61]. These information sources can then be used to determine sleep stages, sleep efficiency, abnormal breathing, and overall sleep quality. While PSG is widely considered as the gold standard for sleep monitoring, and while it is extensively used to support clinical treatments of sleep disorders, it has some disadvantages that make it unsuitable for longitudinal and large-scale sleep monitoring. Firstly, the sensing instruments are time-consuming and laborious to put on, and they are prone to disrupting sleeping routines. Secondly, PSG is rather expensive to use and requires a clinical environment and highly trained medical professionals to operate. Due to these disadvantages, PSG is only suitable as a way to support severe disorders that clinical care is required.

Author's address:

ACM acknowledges that this contribution was authored or co-authored by an employee, contractor, or affiliate of the United States government. As such, the United States government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for government purposes only.

© 2018 Association for Computing Machinery.

2474-9567/2018/2-ART39 \$15.00

<https://doi.org/0000001.0000001>

Recently, sleep monitoring based on off-the-shelf mobile and wearable devices has emerged as an alternative way to obtain information about one's sleeping patterns [62, 68]. By taking advantage of diverse sensors, behaviours and routines associated with sleeping can be captured and modelled. This in turn can help users understand their sleep behaviour and provide feedback on how to improve their sleep, for example, by changing routines surrounding sleep activity or improving the sleeping environment. What makes self monitoring particularly attractive is the non-invasive nature of the sensing compared to PSG. Examples of consumer-grade sleep monitors range from apps running on smartphones or tablets to smartwatches and specialized wearable devices [10, 32, 34, 58, 67, 69].

Despite the popularity of consumer-grade sleep monitors, currently the full potential of these devices is not being realized. Indeed, while current consumer-grade sleep monitors can capture and model a wide range of sleep related information, such as estimating overall sleep quality, capturing different stages of sleep, and identifying specific events occurring during sleep [42, 67, 74], they offer little help in understanding the characteristics that surround poor sleep. Thus, these solutions are unable to capture the root cause behind poor sleep or to provide recommendations on how to improve sleep quality. This is because current solutions focus on monitoring characteristics of the sleep itself, without considering behaviours occurring during sleep and the environmental context affecting sleep, e.g., ambient light-level and noise. Indeed, sleep quality has been shown to depend on a wide range of factors. For example, intensity of ambient light [39], temperature [70], and noisiness [50] of the environment can significantly affect sleep quality. Similarly, the user's breathing patterns, posture during sleep, and routines surrounding the bedtime also have a significant impact on sleep quality. Without details of the environment and activities across sleep stages, the root cause of poor sleep cannot be captured and the user informed of how to improve their sleep quality. To unlock the full potential of consumer-grade sleep monitoring, innovative ways to take advantage of the rich sensor data accessible through these devices are required.

The present paper contributes by presenting the design and development of SLEEPGUARD, a *holistic sleep monitoring solution* that captures rich information about sleep events, the sleep environment, and the overall quality of sleep. SLEEPGUARD is the first to solely rely on sensor information available on off-the-shelf smartwatches for capturing a wide range of sleep-related activities (see Table 1). The key insight in SLEEPGUARD is that sleep quality is strongly correlated with characteristics of body movements, health related factors that can be identified from audio information, and characteristics of the sleep environment [62]. By using a smartwatch, the sensors are close to the user during all stages of the night, enabling detailed capture of not only sleep cycles, but body movements and environmental changes taking place during the sleep period. Capturing these sleeping events from sensor data, however, is non-trivial due to changes in sensor measurements caused by hand motions during sleep. To overcome this challenge, changes in sensor orientation relative to user's body need to be tracked and opportune moments where to capture sensor data need to be detected. To address these issues, SLEEPGUARD integrates a set of new methods for analysing and capturing sleep-related information from sensor measurements available on a smartwatch. SLEEPGUARD also incorporates a model that uses the detected events to infer the user's sleep stages and sleep quality. While some prior research has examined the use of smartwatches for sleep monitoring [16, 36, 57, 62], these approaches have only been able to gather coarse-grained information about sleep and often required additional highly-specialized devices, such as pressure mattresses or image acquisition equipment to supplement the measurements available from the smartwatch. In this paper, we demonstrate that, for the first time, using *only a smartwatch*, it is possible to capture an extensive set of sleep-related information – many of which are not presented in prior work. Having a more comprehensive set of sleep-related events and activities available enables users to gain a deeper understanding of their sleep patterns and the causes of poor sleep, and to make recommendations on how to improve one's sleep quality.

We evaluate SLEEPGUARD through rigorous and extensive benchmark experiments conducted on data collected from fifteen participants during a two week monitoring period. The results of our experiments demonstrate that SLEEPGUARD can accurately characterize body motions and movements during sleep, as well as capture different acoustic events. Specifically, the lowest accuracy for SLEEPGUARD in our experiments is 87%, with the best event

detection accuracy reaching up to 98%. We also demonstrate that SLEEPGUARD can accurately detect various sleep stages and help users to better understand their sleep quality. During our experiments, six of the 15 participants suffered from some sleep problems (4 with bad and 2 with general sleep quality), all of whom were correctly identified by SLEEPGUARD. Moreover, we also demonstrate that SLEEPGUARD is able to correctly identify the root cause of sleep problems for the 4 participants with bad sleep quality, whether it is due to suboptimal hand position, body posture or sleeping environment. Compared to state-of-the-art sleep monitoring systems, such as Fitbit and Sleep Hunter, the main advantage of SLEEPGUARD is that can report a wider range of sleep events and provide a better understanding for the causes of sleep problems.

Contributions

This paper makes the following contributions:

- We present the design and development of SLEEPGUARD (Section 2), the first holistic sleep monitoring system to rely solely on sensors available in an off-the-shelf smartwatch to capture a wide range of sleep information that characterizes overall sleep quality, user behaviours during sleep, and the sleep environment.
- We develop novel and lightweight algorithms for capturing sleep-related information on smartwatches taking into consideration changes in orientation and location of the device during different parts of the night. We show how to overcome specific challenges to effectively track events like sleep postures (Section 2.1.1), hand positions (Section 2.1.2), body rollovers(Section 2.2.1), micro body movement(Section 2.2.2), and acoustical (Section 2.3) and lighting conditions (Section 2.4).
- We extensively evaluate the performance of SLEEPGUARD using measurements collected from two-week monitoring of 15 participants (Section 3). Our results demonstrate that SLEEPGUARD can accurately capture a wide range of sleep events, estimate different sleep stages, and produce meaningful information about overall sleep quality (Section 4). We show that SLEEPGUARD successfully reveals the causes of poor sleeps for some of our testing users and subsequently helps them improve their sleep by changing their sleep behaviours and sleeping environment (Section 4.2.4).

2 SLEEPGUARD SLEEP MONITORING PLATFORM

SLEEPGUARD is a novel smartwatch-based sleep monitoring system that aims at estimating sleep quality and capturing rich information about behaviours and events occurring during sleep. By capturing this information, SLEEPGUARD can analyze potential reasons for sleep problems and provide the user with suggestions on how to improve their sleep routine or sleep environment. To achieve its design goal, SLEEPGUARD exploits a wide range of sensors that are common on commercial off-the-shelf smartwatches: (i) accelerometer, gyroscope, and orientation sensor are used to collect body and hand movements; (ii) microphone is used to measure level of ambient noise and to capture acoustic events; and (iii) ambient light sensor is used to monitor illumination within the sleep environment. The different sensors and information extracted from them are summarized in Table 1. In the following we discuss the different subcomponents of SLEEPGUARD in detail.

2.1 Detecting Sleep Postures and Movements

One's sleeping position, also referred to as *sleep posture*, and the extent of body movements are important factors in determining overall sleep quality. Suboptimal posture has been shown to affect the severity of sleep disorders and is widely used in medical diagnoses to analyse effects of sleep disorders [30, 52] while having a good sleep posture has been shown to correlate with subjective assessments of sleep quality [25]. Similarly, high degree of body movements during sleep likely reflects restlessness, which results in poor sleep quality. SLEEPGUARD uses motion sensors (accelerometer, gyroscope, and orientation sensor) to capture user's sleep posture and habits. In the following we detail the techniques we use for capturing the body posture and movements. SLEEPGUARD, currently

Table 1. Sleep events targeted in this work

Event	Type
Sleep postures	Supine, Left lateral, Right lateral, Prone
Hand positions	Head, Chest, Abdomen
Body rollover	Count
Micro body movements	Hand moving, Arm raising, Body trembling
Acoustic events	Snore, Cough, Somniloquy
Illumination condition	Strong, Weak

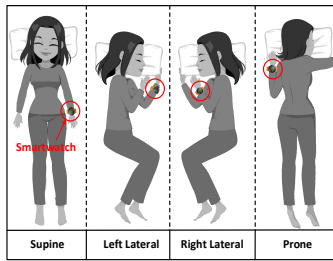


Fig. 1. Four sleep body postures.

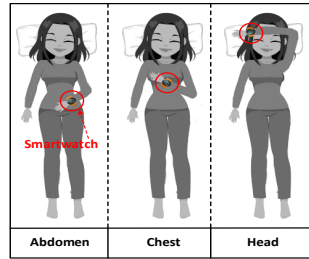


Fig. 2. Three hand positions.

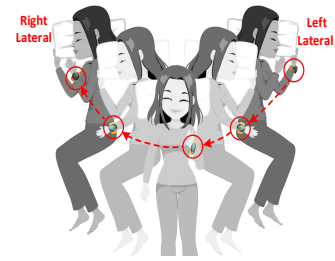


Fig. 3. Body rollover from the left side to the right side.

supports the 4 basic sleep postures (see Fig. 1); 3 hand positions (see Fig. 2); 6 types of body rollovers (see Fig. 3 for an example); and 3 types of body micro movements.

2.1.1 Sleep Posture Detection. Dreaming and sleep quality are associated with underlying brain functions, which in turn are affected by body posture [3]. Sleep posture also varies across individuals and should fit personal and physical needs of the individual [43, 63]. For example, sleeping in a prone position is unsuitable for people with ailments, such as heart disease or high blood pressure. On the other hand, people can consciously avoid postures that would be beneficial for health and sleep quality [27]. Having an effective way to detect the current posture and track changes in it would thus be essential for estimating overall sleep quality, and avoiding potential harm. SLEEPGUARD captures the four basic sleep postures, which are supine, left lateral, right lateral, and prone. This is illustrated in Fig. 1. Detecting these postures using a single wrist sensor, however, is non-trivial because the sensor cannot accurately track movement of the entire body. To accomplish posture detection, we observe that the arm position strongly correlates to the sleep postures, i.e., the arm is typically located in a specific, stable location for a given sleep posture. This observation suggests that we can first identify the user's arm position and the duration of a specific arm location, and then map the information to a sleep posture. Later in this paper, we show that our approach can achieve a high accuracy in identifying sleep postures.

To separate sleep postures, SLEEPGUARD considers a set of feasible hand positions for each posture. In the supine position, we assume the user's hand to be on the left side of the body, on the abdomen, on the chest or on the head; in the left and right lateral positions, we assume the user's hand is close to the pillow, on the chest or on the waist. Finally, in the prone position we assume the user's hand is on the side of the head or above his/her head. These positions were selected based on a pilot carried out in our test environment (see Sec. 3). Fig. 1 shows one possible hand position for each of the postures.

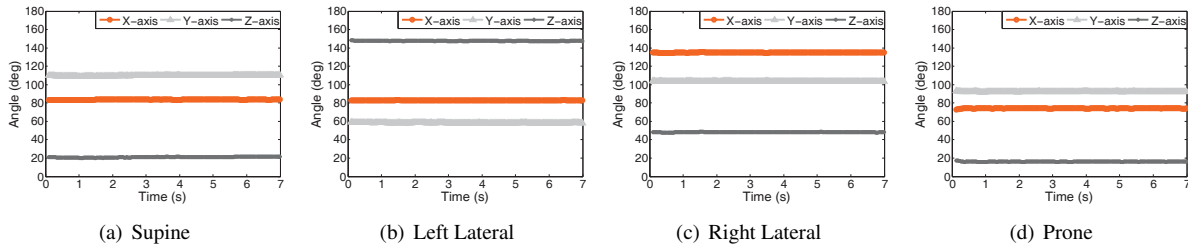


Fig. 4. The tilt angle characteristics of four body postures.

Similarly to SleepMonitor [67], we use tilt angles readings from three dimensions to detect postures. To identify which of the postures the data collected within a time window corresponds to, we average all the calculated tilt angle values of that window in each dimension. We then calculate the Euclidean distance of the input values to a set of posture profiles, which are based on measurements collected in a pilot study that involved 10 users (see Sec. 3.2). We then use the body posture associated with the nearest neighbor as the detection outcome.

Fig. 4 shows the angle values of the four sleep postures targeted in this work. This diagram suggests that the tilt angles of three axes have obvious differences. The sleep posture thus can be inferred based on the position of the smartwatch and the created angle mapping. However, a limitation of this approach is that the hand positions during supine and prone postures are similar when the hand is located on the side of the head (Fig. 4(a) and 4(d)), thus the classification accuracy will be affected.

To improve detection accuracy between supine posture and prone postures, SLEEPGUARD integrates orientation data as auxiliary feature. This is based on the observation that hand directions in the supine and prone positions are different. When the result of the previous step is prone or supine, and the hand is detected to be located next to the body, we combine the tilt angle with three axes data obtained from the direction sensor as a new feature, and classify these postures using a template-based distance matching approach. Specifically, we return the position corresponding to the template with minimum Euclidean distance with current sensor measurements as the user's posture.

2.1.2 Hand Position Recognition. The hand position during sleep can disclose potential health problems, and an improper hand position can even result in health issues [51]. For instance, placing the hand on the abdomen may indicate discomfort whereas placing the hand on the chest can increase the likelihood of nightmares due to long-term pressure on the heart. Similarly, placing the hand on the head can put excess pressure on shoulder nerves and cause arm pain as blood flow is restricted. This can lead to eventual nerve damage, with symptoms including a tingling sensation and numbness [51].

SLEEPGUARD is designed to recognize three common hand positions – if the hand is placed on the abdomen, chest or head when the user is in the supine posture, as shown in Fig. 2. We have chosen these three hand positions because there are found to be the most common and representative positions in our pilot study (Section 3.3). Our hand position recognition algorithm is based on sensor data of rotation angles, tilt angles, and respiratory events. It works by first using the rotation and tile angles to detect if the hand was placed on the head. If the hand was detected to be not put on the head, it then uses the respiratory events to detect if the hand was placed on the abdomen or the chest, but not elsewhere before it utilizes the rotation angles to distinguish the abdomen position from the chest position. We now describe how to detect each of the three positions in more details.

Detect the head position. Fig. 5 show the change of the rotation angle gathered from the x, y, and z directions using the gyroscope for one of our pilot study users when his hand was initially placed next to the body and

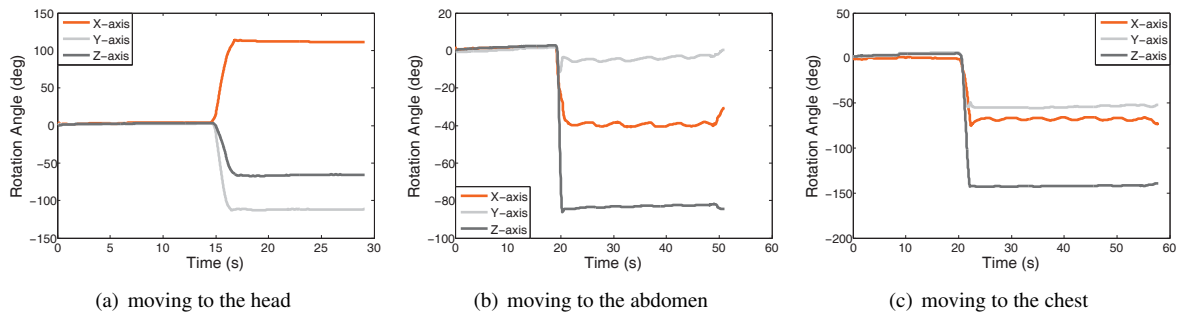


Fig. 5. The differences of the rotation angle when a hand of one of our users, placed next to the subject's body, is moved to his head (a), abdomen (b), and chest (c).

then moved to his head, abdomen, and chest. It is to note that we also observe similar behaviors for our other participants. As can be seen from the figure, when the hand is moved to the head, the changes in the rotation angles are significantly different from the readings given by moving the hand to the abdomen and the chest. This is largely due to the upward facing direction of the palm when the hand is placed on the head compared to the downward palm direction when the hand is placed the other positions. SLEEPGUARD exploits this observation to detect if the hand is placed on the head by examining the changes of the tilt and rotation angles. To this end, we use a hierarchical classifier consisting of two KNN models (with $k = 1$) to predict if the hand is moved to the head based on the tilt and rotation angle readings. Specifically, we use the first KNN model to detect if the input tilt angle reading is closest to one of our training samples where the palm was up. The training data for detecting the palm facing direction are collected from our pilot study users when their hands are placed on the head, the abdomen and the chest respectively (see Sec. 3.2). If the first KNN model suggests that the palm was facing upward, we use the second KNN to take in the differences for the rotation readings (from the x, y, and z directions) before and after the hand movement. The second KNN model then checks if the input data are most similar to a training sample where the hand is placed on the head. The similarity is calculated by measuring the Euclidean distance from the input data to each of the training samples consisting of the rotation angle values from the three directions. Again, the training samples for the second KNN model are the changing rotation angle values obtained from our pilot study users when their hands were moved to the head, the abdomen and the chest respectively. If our hierarchical model predicts that the hand was not placed on the head, we then use the method described in the next paragraph to detect if it was placed on the abdomen or the chest.

Detect the abdomen and the chest positions. If our hierarchical model predicts that the hand was not placed on the head, the hand could be located at anywhere including different parts of the body. Recall that we are interested at detecting if the hand is placed on three positions: the head, the abdomen and the chest. This means in addition to the model for having the hand placed on the head, we still need to find ways to detect if the hand was placed on the abdomen or the chest. We found that the hand is likely to be affected by the breathing if it is placed on these two body locations. Specifically, the impact of breathing leads to periodical fluctuations on the accelerometer readings. This is because the hand will be pushed up and drop down because of breathing when it is placed on the abdomen or the chest. Our experimental data suggest that this is a unique behavior that is only observed when the hand is placed on the abdomen or the chest, but not other part of the body (such as the shoulder).

To examine if the accelerometer data are affected by the respiration because the hand is placed on the abdomen or the chest, we calculate the power spectral density (PSD) of the collected accelerometer data. We then check if

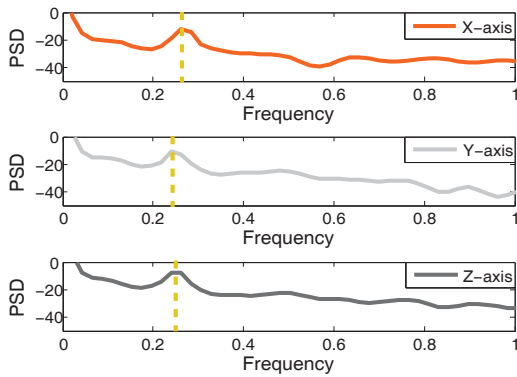


Fig. 6. The power spectral density (PSD) of the accelerometer readings when a user's hand is placed on his chest.

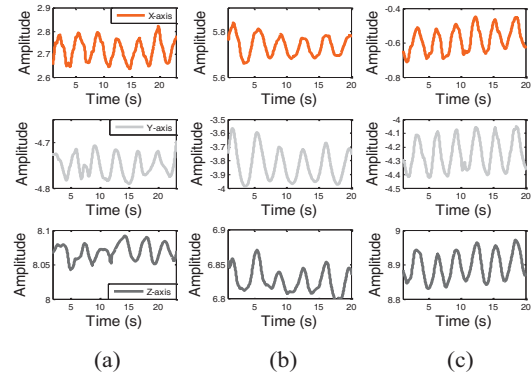


Fig. 7. The periodic change of the acceleration signal. (a) REM-Location 1. (b) REM-Location 2. (c) NREM-Location 1.

there is any peak value matches the human respiratory frequency. A match indicates that the hand is affected by a respiratory event and hence the hand is likely to be placed on either the abdomen or the chest. Fig. 6 provides an empirical evidence to support our design choice. It shows the PSG for one of our pilot study user when his hand was placed on the chest. Here we calculate the PSD for the accelerometer data collected from the x, y and z directions. We can see from the diagram that there is a large peak located at around 0.25Hz (highlighted in the diagrams) when a respiratory event is detected (which was verified by video feed). This peak corresponds to the average respiratory frequency of an adult (0.2Hz to 0.47Hz) [1], suggesting that the PSD reading can be used as a proxy to detect respiratory events. SLEEPGUARD thus exploits the PSD to detect if the hand is placed on the abdomen or the chest by checking if there is any peak value of PSD falls within the range of 0.2Hz (corresponding to 12 breathes per minute) and 0.47Hz (corresponding to 28 breathes per minute).

With the PSD-based method in place (which tells us if the hand is placed on the abdomen or the chest or elsewhere), we then use again a KNN classifier to make a binary decision to determine if the hand is placed on the abdomen or the chest based on the rotation angle readings (see Fig. 5 b and c). Like the head position model, the training samples for the KNN model are also collected from our pilot study users – where each training example includes the rotation angle readings when the hand is either placed on the abdomen or the chest.

2.2 Labeling the Eye Movement Stages

We also found that the extent of body movements can be used to judge the amplitude of respiration, which in turns allows us to detect rapid-eye-movement (REM) and non-rapid-eye-movement (NREM) stages. This is based on the prior study showing that when people sleep in the REM stage, their respiratory amplitude is smaller than that in the other stages [54]. Hence, we can roughly determine the user's current sleep stage based on the respiratory amplitude. Respiratory amplitude is only an indicator of the division of the sleep stage and we can not regard it as a basis for final judgement, but it serves as an early reference that helps later phases of the sleep stage detection. Under normal circumstances, the chest movement amplitude is smaller than abdominal movement amplitude. However, in different sleep stages, the respiration amplitudes are different [46]. It is likely that there is a situation: the chest movement amplitude in the NREM stages is close to the abdominal movement amplitude in the REM sleep stage. As a result, the threshold based method cannot work. At this time, we need to combine it with the position of the hand we detected with the trajectory before. Through the above steps, we have been able to determine whether the hands are on the chest or abdomen, and then we can go further to determine the extent of breathing according to the

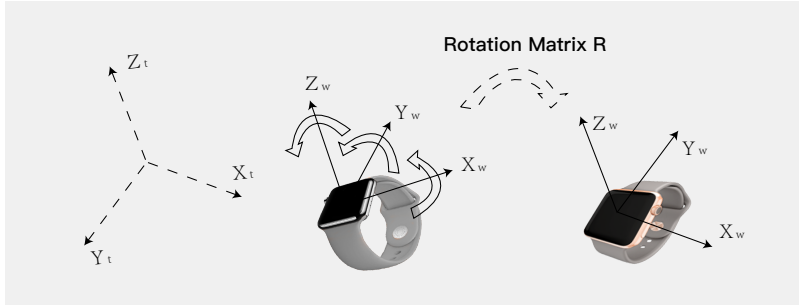


Fig. 8. The first figure on the left is the torso coordinate system, and Y_t points north. The middle figure shows the watch coordinate system when the watch is at an arbitrary position, and the right of the figure shows that the watch coordinate system after we completed the coordinate system conversion.

degree of body ups and downs, and we can roughly infer the sleep stage. Now we take the case of hands on the abdomen as an example.

Even when the hand is placed on the abdomen, due to minor changes in the exact location of the user's hand, the location of intensity of accelerometer fluctuations caused by respiration varies. Hence, we cannot use the amplitude information to determine true respiratory amplitude directly. This problem is illustrated in Fig. 7, where (a) and (b) contain triaxial acceleration measurements at different locations of the abdomen during REM sleep stage, and in (c) which consists of acceleration data for NREM stages at the same approximate position as in (a). We can see that when hand on the abdomen, but the location is different, even in the same location, we can not directly judge the respiration amplitude using only the amplitude of the three axes of the acceleration.

To solve this problem, we convert the acceleration data from the wristwatch coordinate system into the data in the torso coordinate system. Because we can know that when the person is in the supine position, the abdomen and the chest will move up and down due to breathing, that is, move along the z-axis direction of the torso coordinate system. We express the triaxial acceleration data as $Acc_w = [X_w, Y_w, Z_w]$ in the wristwatch coordinate system and $Acc_t = [X_t, Y_t, Z_t]$ in the torso coordinate system, as shown in Fig. 8. And the watch coordinate system ($[X_w, Y_w, Z_w]$) is determined by the position of the watch. Our coordinate alignment aims to find a rotation matrix R to align the watch's coordinate system to the torso coordinate system ($[X_t, Y_t, Z_t]$) and R can be obtained by the three-axis direction information in the orientation sensor. After the coordinate system is aligned, the angle between the y-axis of the wristwatch coordinate system and the y-axis of the torso coordinate system is 180 degrees.

$$X_t = (X_w \cos \gamma + Y_w \sin \gamma) \cos \theta + (Y_w \cos \sigma + Z_w \sin \sigma) \sin \theta, \quad (1)$$

$$Y_t = -((Y_w \cos \sigma + Z_w \sin \sigma) \cos \theta - (X_w \cos \gamma + Y_w \sin \gamma) \sin \theta), \quad (2)$$

$$Z_t = (Z_w \cos \gamma - Z_w \sin \gamma) \cos \sigma - (Z_w \cos \gamma - Z_w \sin \gamma) \sin \sigma, \quad (3)$$

θ , σ and γ are the x, y and z axis data of the orientation sensor respectively, representing the direction angle, the tilt angle and the roll angle collected from the orientation sensor. After the alignment of the coordinate system, we can see from Fig. 9 that the z-axis shows a periodic signal with significant fluctuations, while the x- and y-axis data undergo smaller changes around zero, which is consistent with the actual situation that when the person is in the supine posture with the abdomen up and down caused by respiratory.

The two graphs on the left of Fig. 9 show the same acceleration data as has been used in (a) and (c) of Fig. 7, respectively, whereas the two right-most graphs correspond to the data after coordinate system alignment. We can see that prior to aligning the measurements, we cannot effectively distinguish the respiratory amplitude of REM and NREM stages from the acceleration amplitude. After coordinate alignment, the respiratory amplitudes are

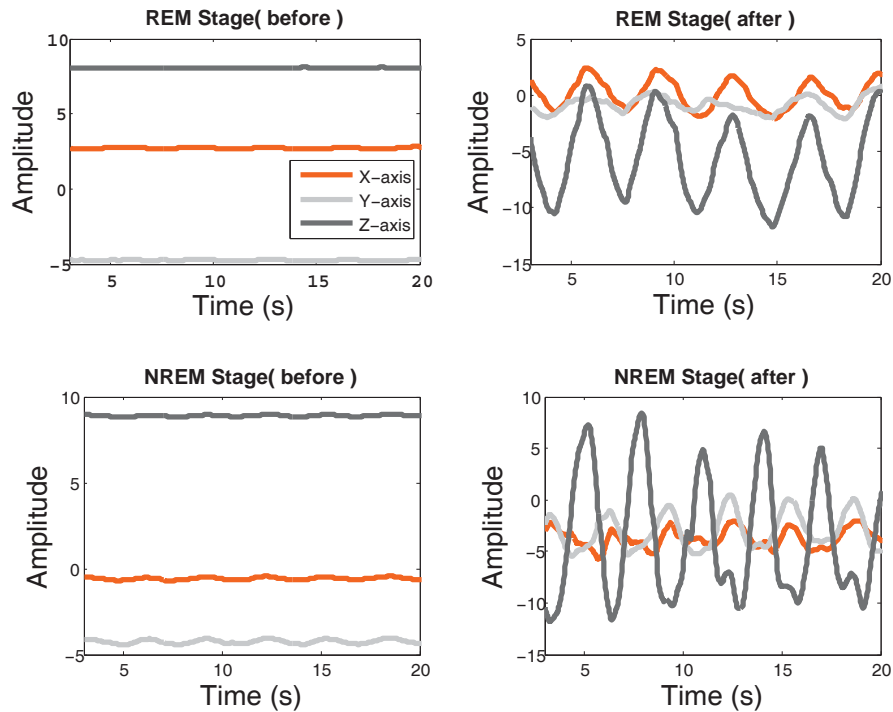


Fig. 9. Acceleration data for different sleep stages.

clearly visible from the z-axis data. We then calculate the variance of z-axis acceleration and use it as a feature to measure the intensity of the fluctuation in a signal, with larger variance corresponding to greater breath amplitude. Note that we cannot use the sum or magnitude of the z-axis as a measure of intensity as the measurements remain affected by gravity.

We use respiratory amplitude to detect whether the user is in the REM stage or not. Specifically, we calculate the respiratory amplitude when the hand is found to be placed on the abdomen or the chest (as SLEEPGUARD operates using a wrist worn device, it can only detect respiratory events when the hand is placed one of these two positions). We then use a KNN classifier to find from our training examples, which training example is most similar to the input respiratory amplitudes collected from the x, y, and z directions. The similarity is measured by calculating the Euclidean distance on the 3-dimensional feature space. We then use the label (either REM or NREM) associated with the nearest training example as the classification outcome.

2.2.1 Body rollover counts. Under normal circumstances, people usually rollover their body around 20-45 times a night. The main function of body rollovers is simply to maintain a comfortable sleeping position as maintaining the same position for prolonged period will result in muscular tension due to hindered blood supply, which can also lead to local numbness [8]. So body rollover is another key indicator about the sleep quality. SLEEPGUARD can detect the number of body rollovers, which can get some reflections of sleep status and the roll-over frequency can also help us to classify the sleep stage [35]. In general, there are six cases: four posture transition cases between the supine posture and lateral (left or right) posture, and two posture transitions between the left lateral posture and right lateral posture. Fig. 3 shows the case when the body moves from the left side to

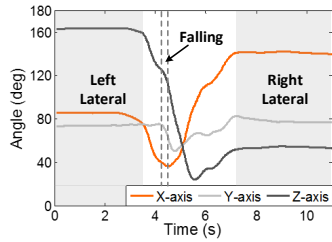


Fig. 10. From left lateral posture to right lateral posture.

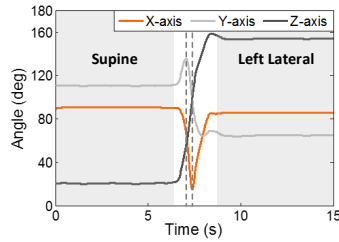


Fig. 11. From supine posture to left lateral posture.

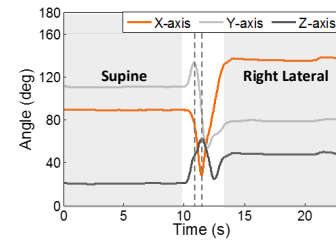


Fig. 12. From supine posture to right lateral posture.

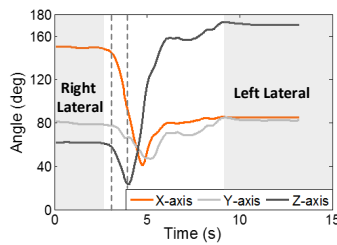


Fig. 13. From right lateral posture to left lateral posture.

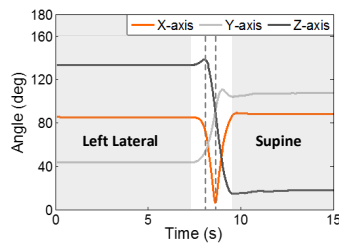


Fig. 14. From left lateral posture to supine posture.

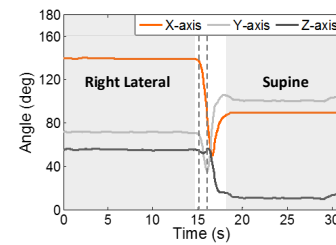


Fig. 15. From right lateral posture to supine posture.

the right side. The most intuitive way for detecting body rollover events is to estimate the rotation direction of the arm based on the rotation angle given by the gyroscope. However, doing so is non-trivial because different users can exhibit drastically different patterns for arm rotations; and the subtle changes in the starting arm position for the same sleeping posture could lead to a misprediction. As an alternative to the rotation angle, we find that the tilt angles to be useful for this task because they strongly correlate to a body rollover event. This correlation thus enables us to effectively translate the change of the tilt angle to a body rollover event to count the occurrence of such events. Specifically, the angle values of three different axes are on the falling edge or rising edge simultaneously during a very short time period. Fig. 10 – Fig. 15 shows the value changes under different body rollover cases. To this end, a naive method to detect rollovers would be to rely on changes in angle measurements. However, this method suffers a very large error since other hand movements will also induce a similar change.

To deal with this challenge, we incorporate the body postures to improve the detection accuracy. As shown in Fig. 3, the body postures are different before and after the rollover. Therefore, after we detect the time when the angle values changes, we term the time as a possible rollover time point. Then we use the sleep posture classification algorithm to determine whether the body postures are the same or not over a period of time before and after this point. If the postures are the same, we exclude this time point, otherwise a body rollover event is found. In other words, a body rollover event is recorded when the posture changes are detected between two time points. Note that SLEEPCARE not only counts the number of rollovers, but also reports the exact nature of the rollover event.

2.2.2 Detecting micro-body movements. Besides major body movement, such as rollovers, there often are involuntary body movements that can affect sleep quality. With the deepening of sleep, limbs are extremely relaxed, and a little stimulus will produce trembling and micro beating. Such behavior is most likely to occur during the deep sleep stage and the REM stage [9, 40]. Therefore, by detecting such micro body movements and

distinguish them from large body movements can help us to further analyze the user's sleep stage. In this paper, we are interested in the sleep-related body movements including hand moving, arm raising, and body trembling.

One of the challenges in detecting the micro-body movements is to cancel the inherent noises brought by the accelerometer. To cancel the noises, we apply a moving window to the collected accelerometer data points to minimize the impact of the outliers. To determine the size of the moving window, we apply different parameter settings to our training data. We found that a moving widow with a size of 35 gives the best average results on our training set. Therefore, we choose to apply a moving window of 35 sample points to the collected user data and then calculate Root Sum Square (RSS) for the data points within a window:

$$Rss(t) = \sqrt{(acc_x(t))^2 + (acc_y(t))^2 + (acc_z(t))^2}, \quad (4)$$

$acc_x(t)$, $acc_y(t)$ and $acc_z(t)$ represent the accelerometer sample value of x-axis, y-axis and z-axis at time stamp t respectively.

We can obtain the first-order derivative of from the RSS values:

$$V(t) = Rss(t) - Rss(t - 1). \quad (5)$$

Eventually, we set the threshold to be 0.03, which can achieve a satisfactory detection performance. We also observe that the micro movement duration is very short, which lasts less than 2 s. However, in our body rollover experiments, we find that the average movement duration is 3 s, as shown in Fig. 10 - Fig. 15. Therefore, we can first divide the body movement events into large movement and micro movement by detecting the signal duration time.

To distinguish among the micro-body movements of hand movement, arm raising and body trembling, we first consider the duration of the movement. We observe from our training data that an arm rising action typically takes around 1.8 second, while a hand movement and a body trembling take around 1 second. Using the duration of the movement, we can distinguish arm rising from the other two movements. We also find that the body trembling tends to lead to a more drastic change in the accelerometer readings compared to the hand movement. This observation is depicted in Fig. fig:micro-move using samples from one of our training users. Based on this observation, we then use the change of accelerometer reading to distinguish between the body trembling and the hand movement. We do so by calculating the first-order derivative of accelerometer data to find out the peak of the acceleration data. If the peak is great than 1.5 and the duration of the movement took between 0.8 and 1.2 second, a body trembling is detected; if the peak is less than 1 and the duration of the movement took between 0.8 and 1.2 second, a hand movement is detected; otherwise, if the duration of the movement took between 1.5 and 2 second, an arm rising is detected. These thresholds are empirically determined from our training data.

2.3 Detecting Acoustic Events

Acoustic events during sleep, such as snore, cough and somniloquy, can reflect and affect user's sleep quality and physical health. For example, snore is one of the possible symptoms of cerebral infarction patients. And long-term snoring can also have a serious impact on health and sleep. It can cause behind sleep apnea or narcolepsy, a sleeping disorder [5, 60]. And when there is a cough, the human cerebral cortex cells are always in an excited state, limiting the depth of sleep, allowing only short sleep between wakefulness, like many other sleep disruptions. SLEEPGUARD can detect these different acoustic events, including snore, cough and somniloquy. Classical acoustic algorithms use multi-dimensional features to detect acoustic events from a complex environment [34]. By contrast, We tailor our design to the problem domain to derive a simpler yet effective acoustic detection algorithm, by exploiting the characteristics of the different events that we target.

Acoustic feature calculation. In order to identify different acoustic events accurately, we select the short-term average energy and the zero-crossing rate as two features. The short-term average energy of acoustic signal is

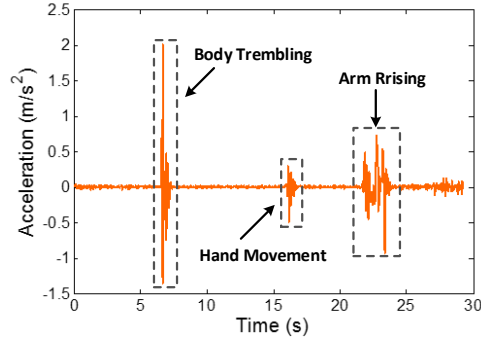


Fig. 16. Accelerometer readings for micro body movements using samples from one of our training users.

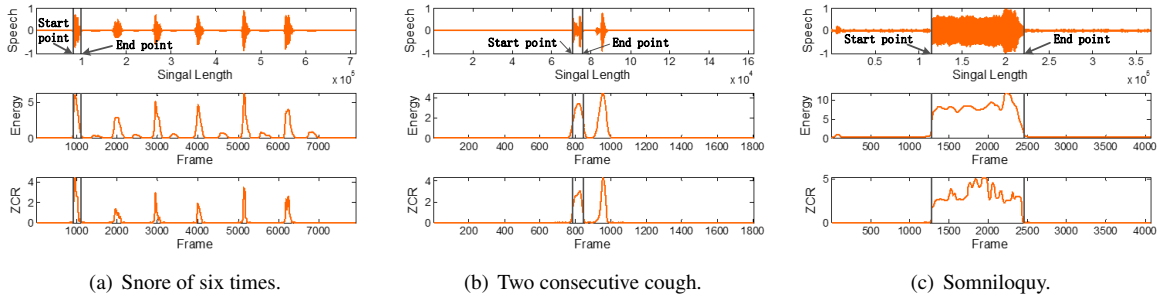


Fig. 17. The characteristics of different acoustic events.

computed as:

$$E_i = \sum_{j=-\infty}^{\infty} [x(j)\omega(i-j)]^2 = \sum_{j=i-(N-1)}^i [x(j)\omega(i-j)]^2, \quad (6)$$

N is the window length, x is the signal and ω is the impulse response. As we can see that the short-term energy is the weighted sum of squared frame sample. The short-term energy can be used to distinguish the segment of unvoiced and voiced sound. It also can be used to differentiate speech segment and noise segment in case of relatively high signal to noise ratio (SNR). The zero-crossing rate is computed as:

$$Z_i = \frac{1}{2} \sum_{j=0}^N |\text{sgn}[x_i(j)] - \text{sgn}[x_i(j-1)]|. \quad (7)$$

It indicates the number of times, which the acoustic signal waveform passes through the horizontal axis (zero level). We carry out an interesting recognition experiment using the microphone built in smartwatch to detect the sound of people during sleep and effectively identify different acoustic events. We focus on three common acoustic events: snore, cough and somniloquy. Ten volunteers worn the smartwatches during sleep to record the acoustic data. We manually label the data with different acoustic events. Fig. 17 shows the acoustic characteristics of three events. There are six times snore event, two consecutive cough event and somniloquy event.

Acoustic event recognition. At beginning, the algorithm divides the audio stream into frames with equal durations. Each frame is composed of 256 samples, and its duration is 12 ms. To identify the three common acoustic events, we introduce an acoustic recognition algorithm based on the following key observations:

- The time interval between two signals for different acoustic events are quite different from each other. As shown in Fig. 17(a), there is a long time interval between two snores. While, the time interval between two coughs is very short. In contrast, the somniloquy signal is irregular and without periodic property.
- The duration of one signal for different acoustic events are different from each other. Fig. 17 shows that the duration of a snore is shorter than the duration of a cough and somniloquy. And in general, the duration of a cough is shorter than the duration of a somniloquy signal.
- The frequencies for different acoustic events are quite different from each other. Snore event has a continuous signal, while the cough and somniloquy are sudden events, thus the number of consecutive occurrences is very small.

In conclusion, the “interval”, “duration” and “frequency” of acoustic events can be used as three unique features. Based on the above three observations, our acoustic event recognition algorithm involves the following two steps. First, in order to estimate the interval and frequency of an acoustic event, we perform the peak detection. We use the short-term average energy to calculate the peak value of the acoustic signal. When the peak exceeds a certain threshold, such as 3 dB in this paper, we record the position of each peak and calculate the interval between two consecutive peaks. Next, we can estimate the frequency of an acoustic event by counting the number of peaks over a time window. Second, to estimate the duration of the acoustic event, we perform the start-point and end-point detection.

Traditional end-point detection algorithm [66], however, uses a fixed double threshold and must be obtained by a large number of data samples, which has two drawbacks. First, the fixed double threshold may cause error detection at the beginning of acoustic event. Second, the requirement of a large number of data samples would lead to large system latency. To address these issues, we develop a detection method that does not require pre-sampling to obtain the optimal thresholds. Instead, our algorithm estimates the thresholds on a per signal basis. This strategy not only reduces the number of data samples needed, but also leads to a higher accuracy.

Specifically, since the first few frames and the last few frames are mostly mute or are the background noise, we select the first five frames and the last five frames to calculate their short-term energy, which are denoted as E_s and E_e , respectively. And then the two are combined to obtain the mean E_n as the estimated energy value of the noise segment. Let the maximum value of the short-term energy over all frames denoted as $\max(E)$. Then, the average short-term energy DE is given as:

$$E_n = \frac{(E_s + E_e)}{2}, \quad (8)$$

$$DE = \max(E) - E_n. \quad (9)$$

So we can use EH and EL to represent the high and low threshold of short-term energy, which are given as:

$$EH = \alpha \times DE + E_n, \quad (10)$$

$$EL = \beta \times DE + E_n, \quad (11)$$

where α and β are the multiplier factors of the energy value DE .

Here, we need to choose the right values for α and β to ensure that we can accurately detect the start and end points of a speech signal. To that end, we use the nighttime sound data from our training dataset to determine α and β . Specifically, we tested α using values ranging between 0.1 and 0.5, and β using values ranging between 0.01 and 0.09. We found that setting α to 0.1 and β to 0.06 gives the best overall results in our training dataset.

Moreover, in order to avoid the interference of sudden noise, we set the minimum length of the signal segment and count the length of the signal during the search for the start and end point. Finally, if the signal length is less than the set minimum, it is considered as a noise segment. The results of the start-point and end-point detection are shown in Fig. 17, we calculate the length of each speech segment and calculate their averages as the duration of the acoustic event. Last, we count the number of peak points to determine whether it meets the third key observation or not.

2.4 Tracking Illumination Conditions

Studies have shown that there is a significant interaction between illuminance level and the mental state of the individual [65]. For example, the bright light can counteract subjective fatigue during daytime, but at night it will seriously affecting the sleep quality. Too much light exposure can shift our biological clock, which makes restful sleep difficult to achieve, it affects our sleep and wake cycle [28]. Besides, we also note that the dim light will affect our sleep too. According to the study [22], it can be learned that the dim artificial light during sleep is significantly associated with the general increase in fatigue, and the proper light can be used to increase the sense of exhaustion and promote sleep. So illumination condition in a sleeping environment has a significant influence on sleep. SLEEPGUARD use the ambient light sensor to detect the illumination condition during sleep. We visit the bedroom of our ten testing users at night and the use of the ambient light sensor to test the lighting conditions in the bedroom. We find that in the absence of lights in the bedroom, the light sensor reading is between 1 Lux to 4 Lux. In some cases, the light of the smartwatch screen can raise the light sensor reading to 4 Lux when the bedroom is dark. In other cases, when the bedroom has weak lights (e.g., when the bedroom is illuminated using a table lamp), the light sensor's average readings are below 10 Lux. Based on these observations, we divide the illumination intensity of the bedroom into two categories. When the bedroom has weak lights (i.e., the light sensor reading is no greater than 10 Lux), and when the bedroom has strong lights (i.e., the light sensor reading is greater than 10 Lux). Using the threshold of 10 Lux, we can then map the light sensor readings to one of these two groups.

However, the light sensor may be obscured, which leads to large errors in measuring the illumination level. For example, a user's smartwatch may be covered under quilt when he/she turned over unconsciously, and thus the illumination readings on the smartwatch may not reflect the real lighting situation. Most of the previous works on smartphone-based light detection have used the proximity sensor to detect whether the light sensor is blocked or not. However, such an approach is not applicable to the off-the-shelf smartwatches because they typically do not have a proximity sensor. To deal with this practical challenge, the key is that the illumination would drop suddenly when the smartwatch is covered by other objects. There are two possibilities for the sudden drop in light intensity. For most smartwatches, the light sensors are usually installed in the front face of it. The first case is the indoor lighting facilities are closed. The second case is the wrist turned so that the back of the hand become downward, thus blocking the light sensor in front of the smartwatch. Such a situation often happens in real life. For examples, when a user changes the sleeping posture to the left side, his/her hand may be close to the pillow with the palm facing up; or the back of the hand may become downward because of a hand movement. To avoid this erroneous illumination condition measurement, we detect whether the user is performing a wrist flip over a period of time during the intensity plummeting. We detect the wrist flip based on two aspects: (i) the rotation angle of smartwatch; (ii) whether the light intensity maintains stable after the sharp drop. If the wrist flips, we use the average of the previous light intensity as the intensity of the time period. It should be noted that, because the illumination condition detection algorithm is relatively simple, it is not explained in the experimental part.

2.5 Sleep stage and quality

Sleep is generally considered as a cyclical physiological process composed of three stages: rapid eye movement (REM) stage, light sleep stage and deep sleep stage. REM is an active period of sleep marked by intense brain activities and dream occurrence. Light sleep stage is a period of relaxation, when the heartbeat, breathing rate and muscle activity slow down. Deep sleep stage triggers hormones to promote body growth, as well as the repair and restoration of energy. The biological characteristics of different sleep stages exhibit distinguishingly. In clinical sleep study, the sleep stages are mainly identified by simultaneously evaluating three fundamental measurement modalities including brain activities, eye movements, and muscle contractions. The EEG measure using electrodes placed around the scalp interpret various sleep/wake states of the brain. And, EMG and EOG using electrodes placed on the skin near the eyes and on the muscles, respectively, measures in deeply differentiating

REM stage from all the other stages. But, apart from the implicit physiological activities, sleepers usually exhibit distinguishable physical activities in different sleep stages. For example, there are somniloquy and body trembles caused by frequent dreams generally appear in REM, large body movements such as body rollovers and arm raising happen in light sleep and micro body movements such as body trembling and snoring occur in deep sleep. In the meanwhile, the breathing amplitude in NREM stage is larger compared with the REM stage. Moreover, sleep cycle usually repeats four to six times over a night. The sleeper usually experiences a transition from light sleep to deep sleep and then enters REM, but sometimes there is also possible a phenomenon of skipping some certain sleep stages occurs during sleep. However, despite this, dependence between two successive sleep stages still exists and different sleep stages have potential conversion probabilities, which also mentioned in Sleep Hunter [34].

To separate between these states, we build a Hidden Markov Model [41] for identifying the current sleep stage of the user. As the input, i.e., the observed states, we use series of detected sleep events and the sleep stage sequence is modelled as a hidden state, i.e., $obs_t = NB(t), NB_M(t), BA(t), NA(t)$ represents the feature vector at detection phase t . The explanation of each item, which is the input of HMM, is listed as follows. $NB(t)$: the number of occurrences of body rollover during the detection phase t . $NB_M(t)$: the number of occurrences of micro body movement. $BA(t)$: the measurement of respiratory amplitude. $NA(t)$: the number of occurrences acoustic events. And $states_t = \{\text{light sleep; deep sleep; REM}\}$ is an output of our model, which represents the sleep stage in the detection phase t . In the training of the HMM model, we used nocturnal sleep data from 10 volunteers who participated in the training parameters of each algorithm and make their sleep-related events as observation sequences and the corresponding sleep stages detected by Fitbit as hidden state sequences, to generate HMM models. Specifically, we first use maximum likelihood estimation method for parameter estimation, the state transition matrix and the confusion matrix, and then use the Viterbi algorithm to acquire a series of implicit state sequences corresponding to observed sequence. As a result, we can estimate the sleep stage during a time window. Finally, we can get the durations of all sleep stages over the whole sleep process.

Further, to quantize the quality of a sleep, we use the Sleep Quality Report model introduced in [34]. Let SQ be the value of the sleep quality, then SQ is given as follow:

$$SQ = \frac{(REM \times 0.5 + Light \times 0.75 + Deep) \times 100}{REM + Light + Deep} \quad (12)$$

where, REM, Deep and Light represents the duration time in a sleep process. The range of SQ is from 50 to 100. A high value of SQ shows a good sleep quality.

3 EVALUATION METHODOLOGY

3.1 Experimental Setup

We evaluate SLEEPGUARD through experiments conducted in 15 single-occupancy homes over a two-week period. The participants include 6 males and 9 females, whose age spans 15 to 60 years. To ensure too little sleep had no effect on the results, each participant was required to sleep at least 6 hours per night during the study period. Two of our participants have been diagnosed with long-term, on-going sleep-related disorders, and one participant has described that his sleep is significantly affected by snoring. The remaining participants reported their sleep quality to go up and down and having poor sleep from time to time. The study was approved by local IRB, and participants were separately asked to consent to release their data for analysis. In total we collected 210 sets of nocturnal sleep data from our participants.

During the study, participants are asked to wear a smartwatch on their wrist. To obtain ground truth of sleep events, three video cameras were placed on the ceiling to monitor the user's sleep activities, as shown in Fig. 18. The cameras have night vision and thus can capture sleep activity with high-quality. The video footage was manually labelled with different sleep activities and the labels were used as ground truth in our evaluation. To demonstrate the overall benefits of SLEEPGUARD and the events captured by it, we separately collected ground truth information

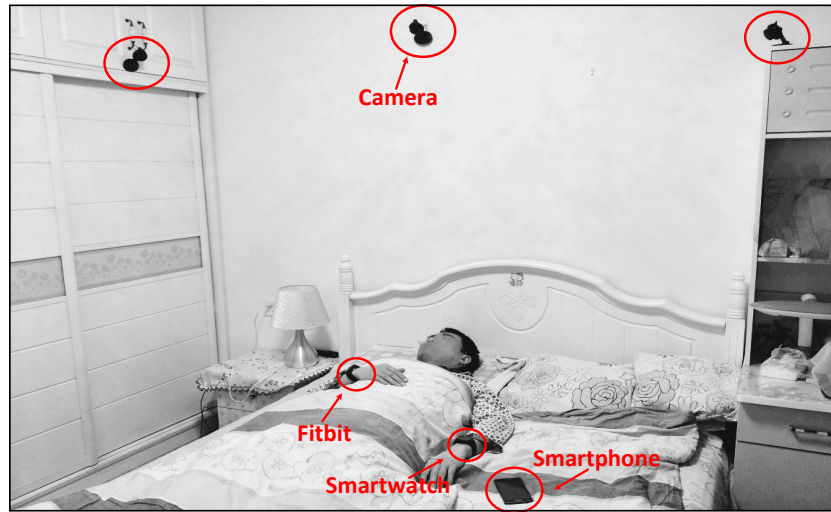


Fig. 18. Experimental setup in one of our participants' home.

about sleep quality using questionnaires which were administrated each morning. The questionnaires were based on the Pittsburgh Sleep Quality Index (PSQI), a widely used and validated questionnaire in sleep quality research [17]. Finally, we collected sleep stage estimates from a Fitbit Charge2 and considered them as ground truth for sleep stage estimation. While the performance of Fitbit is not comparable to medical grade PSG, it has been shown to have a good association in adults [14, 26, 31, 49], especially in estimating REM and light sleep stage. Equipping the participants with PSG was not feasible as it would disrupt their normal sleeping routines and potentially bias and reduce sleep activities, which are the main focus of our work. Moreover, the goal of our experiments is not to demonstrate that SLEEPGUARD is capable of medical grade sleep monitoring, but to demonstrate that it performs comparably to commercial systems in common sleep monitoring tasks, while at the same time being able to capture much richer set of sleep information. We also compare our approach against Sleep Hunter [34], a state-of-the-art mobile-based sleep monitoring approach, and a smartphone-based sleep monitoring app named Sleep as Android [10]. To provide a fair comparison against these baselines, we also place a smartphone next to the user's body on the bed to collect the data for Sleep Hunter and Sleep as Android.

3.2 Pilot Study: Training Data

To train the models used in our system and to determine optimal parameter values, a small-scale pilot study with 10 participants was carried out prior to the main experiment. The training examples used to train our algorithms and to determine the algorithm parameters are collected from 10 users (5 males and 5 females). Our testing users were asked to wear a smartwatch to sleep and collected the sensor data while they were sleeping. Every testing user contributes 10 nocturnal sleep data over a two-week period. These users were different from those took part in our evaluation (Section 3.1).

3.3 Prototype Implementation

We prototype and evaluate SLEEPGUARD on a Huawei Smartwatch 2 wearable device. The smartwatch is equipped with a Quad-core Cortex-A7 processor at 1.1 GHz. It runs the Android Wear 2.0 operating system. We use five

sensors of the smartwatch: the accelerometer, gyroscope, microphone, light sensor and orientation sensor. To reduce energy consumption of the smartwatch, in the experiments we analyse the sensor data on a XiaoMI Note2 Android smartphone to which the smartwatch sends sensor measurements over Bluetooth. The sensors on the smartwatch are sampled every 30ms, which was chosen to balance between information quality and energy consumption. SLEEPGUARD starts tracking sleep events when it detects that the light is off and there has been no body movement for 30 minutes. As part of an initialization process, SLEEPGUARD estimates the initial body posture and hand position. It then uses these as a starting point to monitor sleep events like the body posture, rollovers, hand positions and body movements.

4 RESULTS

In this section, we detail the evaluation results for our system.

4.1 Evaluation of Subcomponents

We focus on the detection accuracy about five events, that are body posture, the body rollover, the hand position, the micro body movement and the acoustic events.

4.1.1 Sleep Posture Classification Performance. We first test the overall classification performance of different body postures. The ground truth of body postures are recorded by the cameras. To avoid biases in evaluation, and to assess the generalization performance of our approach, we consider a cross-validation scheme where all data from a single participant is used for training and data from the remaining 14 participants is used for testing. The motivation for using data from a single user as training data is to highlight the capability of SLEEPGUARD to accurately characterize body posture with very little training data, while at the same time being able to generalize across users. The final performance is then calculated as the averaged accuracy across the 15 folds; as shown in Fig. 19 and Table 2. We can observe that the posture detection accuracy is consistently high across all users, and does not show major variations across users. This good performance benefits from the distinct characteristics of arm position under different sleeping postures. Compared to results reported for SleepMonitor [67], SLEEPGUARD consistently improves performance which is mainly due to the template-based classifier that we use to verify classifications of the prone and supine states. In particular, SLEEPGUARD achieves around 5 percentage units higher performance on the prone state than SleepMonitor and overall has a lower false positive rate. In terms of errors, due to angular characteristics of acceleration being similar between the supine posture with hand putting on the head and the left-lateral posture, a small amount of the supine postures are classified as left lateral. From the results we can also observe that the total amount of the prone posture is smaller than the number of other postures, which suggests that people are not accustomed to sleep in this position, because it is neither healthy nor comfortable.

Table 2. The confusion matrix of body posture classification.

	Result	Prediction				Recall
		Supine	Left Lateral	Right Lateral	Prone	
Groundtruth	Supine	1182	25	4	9	96.7%
	Left Lateral	6	1292	0	0	99.5%
	Right Lateral	7	0	1275	12	98.5%
	Prone	19	2	3	567	95.9%
	Precision	97.3 %	98.0%	99.5%	96.4%	

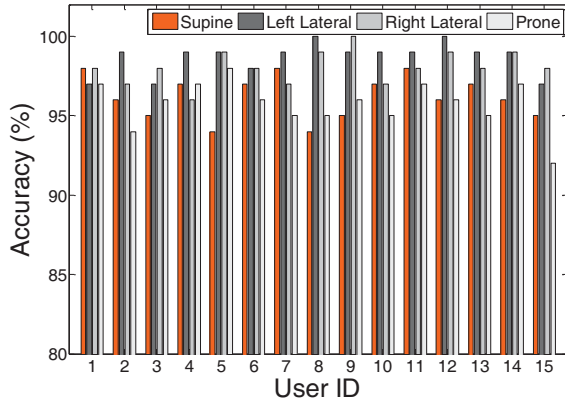


Fig. 19. Detection accuracy of body postures.

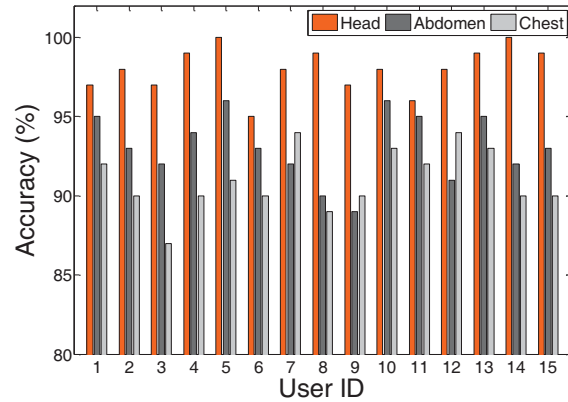


Fig. 20. Identification accuracy of hand positions.

4.1.2 Performance of body rollover counting. To verify the efficiency of body rollover detection algorithm, we compare each user's body rollover events detected by SLEEPGUARD against the data labeled by watching the video. The performance is shown in Table 3. We can see that User 3, User 4 and User 13 have an unusually high number of rollovers. For User 3 and User 4, they have difficulty in falling asleep due to the sleep disorder. User 13 needs to rollover frequently because of his loudly snoring. As we demonstrate in Sec. 4.2.4, these participants also suffered from poor sleep quality and hence indicate how the information extracted by SLEEPGUARD can support the detection of sleep problems. For all the 15 users, the detection accuracies are all very high, and the lowest one is still 87%. Thus SLEEPGUARD can accurately distinguish the large hand movement from the body rollover in bed. Moreover, detecting errors in body rollover events will not have a significant impact on our end result, because the division of sleep stages is a comprehensive consideration of all the detected features in each stage, such as micro body movement and acoustic events.

Table 3. Detection accuracy of body rollover.

Testing User ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Labeled #body rollover	231	204	442	397	198	101	196	164	193	208	131	205	342	149	156
Accuracy	91%	94%	88%	93%	96%	94%	87%	90%	93%	94%	92%	94%	89%	90%	95%

4.1.3 Performance of hand position recognition. To test the recognition performance of different hand positions, we consider the same cross-validation scheme used for body posture detection, i.e., one user's data is used for training the classifier and the remaining 14 users' data as test data. The classifier for detecting the hand movement trajectory is combined with the detection of periodic signals caused by respiration, then the hand position on the chest (or abdomen or head) can be identified. Fig. 20 illustrates the accuracy of hand position across 15 users. As we can see that with just one set of training data, the accuracies for different users are all higher than 87%. Therefore, our system can achieve a good identification accuracy for different hand positions. Moreover, we find that at least four out of fifteen participants tend to put their hands on their heads; one participate unconsciously puts his hand on his chest which makes him have nightmares. Those are all bad habits disrupting a good sleep. SLEEPGUARD can report such key findings to improve the users' sleep qualities.

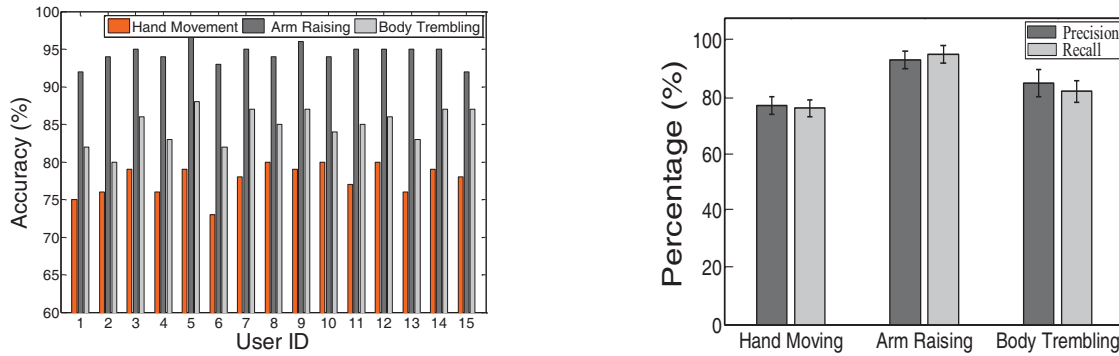


Fig. 21. Detection accuracy of micro body movement. Fig. 22. Performance of micro body movement detection.

4.1.4 Performance of micro body movement detection. To assess the detection accuracy of micro body movements, we manually label the ground truth recorded by the camera during sleep, including hand moving, arm raising, and body trembling. We also use the accelerometer embedded in the smartphone which placed on the bed to record the occurrence of micro body movements, so as to avoid missing some movements such as trembling concealed by the duvet. For the acceleration data collected by smartphone, we first smooth the acceleration along the three axes, calculate Root Sum Square (RSS) to merge them, and obtain the first-order derivative of the merged acceleration. And then we use the threshold detection method to mark the occurrence of motion. Since body trembling is the easiest to be covered, we only focus on such events. So we use smartphone to detect the occurrence of events and the classification of the event is not performed. Fig. 21 illustrates the detection accuracy across 15 users. It shows that the accuracies for all users are very close, that is, there will be no major changes between users. And from Fig. 22, we find that even though the worst classification result belongs to the hand movement, the average precision value and recall value still exceed 75%. The averaged accuracies of arm raising and body trembling are 93% and 84%, respectively. Because the training data volume for the hand movement and body trembling is small, so the performance can be improved by setting each user a threshold by collecting a longer term's sleeping data. In addition, the purpose of micro body movement detection is to detect different sleep stages, and the hand movement usually appears in all sleep stages, thus the poor accuracy of hand moving does not have a significant impact on the final result.

4.1.5 Performance of acoustic events detection. To study the detection accuracies of different acoustic events, we compare the ground truth recorded by the camera with the detected results by our system. Table 4 shows the results across 15 participants. We can see that the precision for cough event is 88.9%, which is slightly lower than for the other three event types. The reason is that different user's cough patterns are different, the pre-defined parameters in the detection model does not include all possible patterns. For example, some people have a fast and continuous pattern of coughing, while others have a slower intermittent pattern. In fact, we train these parameters, namely the "interval", "duration" and "frequency" of acoustic events, with only 120 sets of nighttime sound data. Those data come from 40 (21 males and 19 females) volunteers of different ages (from 15 to 60 years old) who are prone to snoring, coughing, or somniloquy at night. To further improve the detection accuracy, we can train particular parameters for different users. And we can further expand the training data to include more possible patterns, and can also make reasonable estimates of the possible patterns to refine the range of parameters and thus increase the accuracy.

Table 4. The confusion matrix of acoustic events detection.

	Result	Prediction				Recall
		Snore	Cough	Somniloquy	Other	
Groundtruth	Snore	96	0	0	9	91.4%
	Cough	3	64	0	4	90.1%
	Somniloquy	0	3	42	2	89.4%
	Other	0	5	4	325	97.3%
	Precision	96.9%	88.9%	91.3%	95.6%	

4.2 Overall performance

4.2.1 Performance of sleep stage detection. In order to prove that the detected events not only reflect the user's sleep habits, but also effectively identify the sleep stages to assess the sleep quality, we regard the reported results from Fitbit Charge2 as the ground truth. To perform the evaluation, we randomly choose 50 sets of sleep data from the data so that at least 3 sets per participant are chosen for evaluation. For detecting changes in sleep stage, SLEEPGUARD uses event-driven detection. When there is no sleep event detected in 15 minutes, we evaluate the sleep stage. When an event occurs, we immediately evaluates the sleep stage and use this time as the starting point for the next 15 minutes. The averaged precision value and recall value are shown in Table 5. It indicates that though SLEEPGUARD may make misjudgement between the light sleep and REM, overall the performance is satisfying and comparable to current consumer grade monitors. Moreover, as we later demonstrate, the main benefits of SLEEPGUARD result from its capability to estimate a wide range of sleep events and how they relate to sleep quality, not from its performance in sleep stage detection where medical PSG measurements are required for accurate assessment of sleep stages.

Table 5. The confusion matrix of sleep stage detection.

	Result	Prediction			Recall
		REM	Light Sleep	Deep Sleep	
Groundtruth	REM	476	143	61	70.0%
	Light Sleep	131	508	91	69.6%
	Deep Sleep	63	113	262	59.8%
	Precision	71.0%	66.5%	63.3%	

4.2.2 Effect of respiratory amplitude on sleep stage detection. When we detect different sleep stages, we also consider the respiratory amplitude when the hand's position is in the abdomen or chest. To assess the effectiveness of respiration amplitude estimation, we evaluate the performance of the sleep stage detection in two cases, that are with and without taking the respiration amplitude into account. The performance of sleep stage detection is shown in Table 6. For three different sleep stages, both the precision and recall values are improved with the help of respiration amplitude estimation. In fact, respiratory frequency can also be used as a feature to help us to detect sleep stages. But in fact their essence the same. The difference in respiratory amplitude will also affect the difference in respiratory frequency, because when the respiratory amplitude is large, the time taken for one breath will be long, and the frequency of breathing will be slower. In SLEEPGUARD, we choose the respiratory amplitude because the feature is very intuitive.

Table 6. Effect of respiration amplitude estimation.

Features	REM		Light Sleep		Deep Sleep	
	Precision	Recall	Precision	Recall	Precision	Recall
Without Respiration Amplitude	62.9%	63.4%	59.4%	63.9%	57.7%	54.1%
With Respiration Amplitude	71.0%	70.0%	66.5%	69.7%	63.3%	59.8%

Table 7. Performance of sleep stage detection comparison.

System \ Stage	Light Sleep		Deep Sleep	
	Precision	Recall	Precision	Recall
SleepGuard	66.5%	69.6%	63.3%	59.8%
Sleep As Android	27.8%	35.4%	35.7%	50.2%
Sleep Hunter	66.74%	66.11%	60.00%	50.73%

4.2.3 Performance comparison. We compare SLEEPGUARD with two state-of-the-art work, the sleep detection app Sleep As Android and smartphone-based system Sleep Hunter [34]. Considering that Sleep As Android can only detect light sleep stage and deep sleep stage, we only compare the performance of these two stages. Table 7 shows the detection results. As we can see, SLEEPGUARD performs much better than Sleep As Android and slightly better than Sleep Hunter. This good performance comes from the incorporation of rich and complicated sleep events. Therefore, our system helps users understand their sleep more easily and improve sleep quality more effectively.

Further, we compare SLEEPGUARD's functions with 8 other sleep detection products in Table 8, including Sleep As Android, Sleep Hunter, sleepMonitor [67], Sleeptracker [64], Fitbit, isleep [37], Jawbone [69] and ubiSleep [57]. SLEEPGUARD detects a wider range of sleep events and thus provides a better user experience.

Table 8. Functions comparison between different systems.

System	Detected Events					
	Heart Rate	Acoustic Event	Sleep Posture	Body Movement	Hand Position	Sleep Stage
SleepGuard		✓	✓	✓	✓	✓
Sleep as Android		✓				✓
Sleep Hunter		✓		✓		✓
SleepMonitor			✓			
Sleeptracker	✓					✓
isleep		✓		✓		
Fitbit	✓					✓
Jawbone						✓
ubiSleep	✓	✓				

4.2.4 User Survey. To understand and verify how the additional information captured by SLEEPGUARD supports users, at the end of the experiments the participants were administered a survey to the participants asking about their experiences with SLEEPGUARD and their personal sleeping patterns. We combine these results with the subjective sleep quality estimates obtained through the PSQI questionnaires administered during the study. We considered two groups of users in our survey. As the main source of information we consider the 15 participants of our experiments who were asked about their experiences with SLEEPGUARD, their subjective sleep quality assessment, and details of their personal sleep patterns. This set of users was augmented with external participants

who were asked about their interest in the events that SLEEPGUARD is capable of detecting. The questions in our survey include:

- (1) Subjective sleep quality (5-levels, 1 for excellent and 5 for worst),
- (2) Sleep duration,
- (3) Sleep disturbances,
- (4) Daytime dysfunction.

For the above four items, each one is rated on a 1 to 5 scale. These scores are first summed to yield a total score, which ranges from 0 to 20. Then we merge every five neighbouring scores into one scale and eventually divide the total scores into four levels, recorded as 0, 1, 2 and 3, representing poor, general, good and excellent, respectively. This step is necessary to compare the results of the user survey against the sleep quality estimation provided by SLEEPGUARD and Fitbit.

The 14-days average sleep quality scores of the 15 participants are presented in Table 9 together with sleep quality estimates obtained from SLEEPGUARD and Fitbit. For 9 users, SLEEPGUARD's assessment of sleep quality perfectly matches the subjective assessments. For the remaining 6 users, the error is within one scale point, with most of the errors being between 0 and 1 (poor and general) and 2 and 3 (good and excellent). Compared to Fitbit, the results are slightly better overall, and all errors made by SLEEPGUARD are in line with those made by Fitbit. Accordingly, our results demonstrate that SLEEPGUARD is capable of estimating overall sleep quality in line with subjective estimates, as given by the PSQI questionnaire.

While the results above demonstrate that SLEEPGUARD is capable of accurately estimating sleep quality, the main benefit from SLEEPGUARD compared to previous works is not its sleep quality performance but its capability to analyse and capture the *root cause* of sleep issues. To demonstrate this, we carried out a follow up analysis where we examined the events captured by SLEEPGUARD for each of the six users assessing their sleep quality negatively (poor or general subjective quality, i.e., rating 0 or 1 in Table 9). For four of the six users we were able to find clear causes for their poor sleep. For one of the users, SLEEPGUARD indicated difficulties in falling asleep, which was reflected in a high body rollover count. Further analysis of captured events indicated ambient noise and lighting to be most likely reasons for this participant. Another user complained feeling of numbness in the arm after sleep. Events captured by SLEEPGUARD showed that this was likely due to bad hand posture as the person tended to put the hand on top of head before sleeping. Third user complained of frequent nightmares. Analysis of SLEEPGUARD events showed that the person habitually slept on the left side, which has been shown to have higher risk on nightmares [?], and often placed a hand on top of the chest, which creates additional pressure and can lead to nightmares. Finally, one of the users mentioned suffering from long-term snoring problems, which we also were able to detect from the events captured by SLEEPGUARD. The events also highlighted that the person was often sleeping in supine position, which further increases susceptibility to snoring-related problems. Existing systems are only capable of capturing some of these factors influencing sleep quality and thus they are not capable of providing a holistic view of the participant's sleep quality, whereas SLEEPGUARD is capable of providing very detailed information about sleep events. To further demonstrate the benefits of SLEEPGUARD compared to previous works, we asked the 15 participants to make appropriate adjustments according to our recommendations and to conduct a return visit survey three weeks later. It was found that some of the users were able to reduce symptoms and to improve their average quality of sleep based on the suggestions.

As for the user experience, results from the survey highlighted a strong interest in the information captured by SLEEPGUARD. In particular, 80% of participants believe that the detection of sleep posture is very necessary, showing their sleep posture can not only help people to avoid health problems caused by long-term improper sleeping posture, but also help us find out the reasons for the next day's physical discomfort, such as dizziness, muscle soreness may be due to improper sleeping posture. And there are some users are troubled by snoring. This may be due to improper sleeping posture. We map the detected snoring event and sleeping posture to suggest

Table 9. Results of sleep quality assessment.

System \ User ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
SleepGuard	3	3	0	1	2	2	3	0	2	2	2	2	1	0	2
Fitbit	3	3	0	0	1	3	2	3	2	2	2	2	2	1	2
User survey	3	2	0	0	2	2	3	1	1	2	2	3	0	0	2

the user to modify his posture to a suitable posture to reduce the harm caused by long-term snoring. 60% of the participants thought it useful to detect the hand position in supine posture, even one user mentioned that he did often have nightmares and our system found his hand was often placed on his chest, and then SLEEPGUARD could remind him that he should take some measures to avoid such a position and thus reduce the poor sleep quality that nightmare brings. Only 20% of participants found it useful to calculate the number of body rollover. However, detection of rollovers is useful in segmenting sleep stages. Furthermore, body rollover counts could be used to derive additional information to the user, such as how restless or peaceful the sleep has been overall.

5 DISCUSSION

In this paper, we have shown that sensors available on off-the-shelf smartwatches can be used to capture rich information about sleep quality and factors affecting it. The main focus of our work has been to develop the required algorithms for capturing rich sleep related information as accurately as possible. While the recognition performance of our system is very encouraging, there are some issues that would need to be addressed in our system before larger-scale deployment would be feasible. Below we highlight the main issues and briefly discuss possible ways to overcome them.

- **The accuracy of sleep stage detection.** Compared with PSG-based sleep stage detection, the accuracy of SLEEPGUARD is lower and it is rather difficult to achieve comparable performance. This is because PSG monitors and analyzes sleep based on information that directly correlates with sleep such as EEG, EMG, EOG, and oxygen saturation, whereas SLEEPGUARD estimates sleep quality from cues that have an indirect effect on sleep quality. In particular, SLEEPGUARD only combines the body movement, acoustic events, sleep environment and other events during sleep to predict the sleep stage. Therefore, SLEEPGUARD is not a replacement for professional medical equipment for high-precision sleep detection, but serves as a personal technology that provides easy-to-use and non-intrusive way to monitor personal sleep patterns and to obtain feedback about the sleep quality. Moreover, it can trace back to the real causes affecting sleep quality, and guide users to have the direction to improve sleep quality.
- **Battery life.** A critical design requirement for sleep detection is that the monitoring can operate sufficiently long to cover the entire duration of the user's sleep. Battery capacity on smartwatches is rather limited, resource consumption needs to be optimized by considering both the data collection and analysis phases. In our experiments we demonstrated that additional devices in the vicinity of the smartwatch can be taken advantage of, for example, some of the sensing and processing tasks can be offloaded to smartphones or other smart devices located within sufficient proximity. Particularly the acoustic event detection could be offloaded to smartphones that are located on the bedside table or elsewhere in the user's vicinity as smartphones increasingly integrate co-processors for audio processing that allow performing the audio event detection with a small energy footprint. We have also designed our analysis techniques to be as lightweight as possible to minimize energy consumption. Further improvements can be achieved by designing dynamic duty cycling strategies that reduce sampling during periods of regularity, and by using simple triggering mechanisms, such

as a motion intensity detector to reduce processing overhead. Exploring these techniques is part of our future work.

- **Sensor data.** One limitation of SLEEPGUARD is that we have not taken advantage of heart rate when determining the current sleep stage of the user. The main reason for this is programming limitations of the Huawei Smartwatch 2 device used in the experiments. Specifically, the device does not support querying heart rate information, but only provides it through a dedicated application. The output of this application is unfortunately not sufficiently accurate for sleep monitoring purposes, and restricts the rate at which information can be acquired. To compensate for the lack of heart rate data, SLEEPGUARD considers the respiratory amplitude detected from accelerometer instead. As shown in our experiments, the respiratory amplitude detection significantly improves the performance of the sleep stage detection.
- **Single wrist sensor.** In our paper, we only use the sensor data of only one wrist, which loses efficiency for detecting specific events of the other wrist. Even so, it can achieve accurate sleep quality assessment result. Although movement patterns of the left and right wrist can be different during sleep, the technique used for detecting sleep related behaviors is the same. Some sleep related events like the sleep posture, body rollover, acoustic events, illumination conditions, are not affected by different wrists. The reason is that these events are related to the entire body rather than the part of the body. To adapt our approach to these sleep events, the only thing we need to do is adjusting new experimental parameters when the smartwatch is worn on a different wrist. For the hand position detection and micro-body movement detection (including the arm raising and hand movement), wearing the smartwatch on a different wrist does have an impact. This is because the hand movement probability and frequency are different on different hands. However, the degree of impact on our detection performance varies from person to person and can be largely cancelled through calibration. This is where we need to measure and consider in our future work. Alternatively, multi-sensor designs, such as combination of smart watch and intelligent ring, could be used to gather relevant sensor measurements from both wrists. Using an intelligent ring could also help in gathering heart rate information during the sleep. Exploring such multi-sensor designs is an interesting future research direction.
- **Comparing to Fitbit.** In addition, when we use Fitbit as groundtruth, Fitbit is worn on a different wrist from SLEEPGUARD. However, from the analysis of the basic principles of sleep stage detected by SLEEPGUARD and Fitbit, it can be found that this does not have much effect on our assessment of the results. Both Fitbit and SLEEPGUARD have common grounds for detecting sleep stages based on acoustic events, the occurrence and frequency of physical activity, but we go further to conduct more fine-grained detection and classification of these events, and add more consideration about illumination conditions and respiratory amplitude. One thing we can know is that the measurement of events such as acoustic events, body rollover events, body tremble, etc., has little to do with the sensor data collected from the left or right hand. The major difference that may exist is these rich events added in SLEEPGUARD, such as hand position, sleep posture, etc, but these are not detected by Fitbit, so they have no effect when compared. Moreover, we also did a test experiment. The smart watches were worn on the left and right hands respectively and the event detection algorithm in SLEEPGUARD was mainly used to detect those events that are of concern in Fitbit. We can see that the results are not much different. Therefore, in the end, in order to ensure that the user's sleep is as uncomfortable as possible, we choose to make Fitbit and Smartwatch are worn on different hands.
- **Multiple-sleeper scenario.** Currently, SLEEPGUARD considers that the user is sleeping alone, but there are still more complicated situations in reality, such as sleeping with a bed partner, baby, and/or a pet. However, because SLEEPGUARD is based on the detection of smartwatch, Unlike smartphone placed on the bed, it can show more sensitivity to the user's own activities. Therefore, for the detection performance of sleeping posture, body rollover and hand position events has almost no effect, but it may have some influence on the body micro movements and acoustic events. When people around us have relatively large movements, such as body rollover, they may fluctuate to users, making it possible for us to mistakenly detect it as user's micro

movements. For this kind of situation, we can test the change of acceleration data in multi-sleeper situations by popularizing the experiment to adjust the detection threshold of our body's micro movements and achieve better detection performance. This will also be a direction for our future work. As for acoustic events, we can further limit conditions, such as training the different magnitudes of the energy of the sound signals collected by the user's hand at different positions to identify whether it is the user's own acoustic event or the sound of the bed partner. In addition, the related acoustic events of the bed partner can also be considered as a factor affecting the user's sleep.

- **Occurrence of unusual arm's positions.** We detect sleep posture based on arm's positions and focus on three specific positions when detecting the position of the hand. In sleep position detection, we are based on the assumption that between the user's arms position and sleeping postures that the arms have common and (reasonably) stable positions in each posture and we consider as many possible arm positions as possible in four sleeping postures, which are the most common arm's positions for users during sleep. In hand position detection, we chose the three most representative locations that do have an impact on sleep and health. But we know that not all users or a user will not have these common positions all the time. These unusual arm's positions may cause the performance of our sleep posture detection algorithm to degrade. But of the 15 participants we tested, we can see from the video that the unusual arm's positions are present, but these are basically a slight evolution of the common positions, which have little effect on the detection of the sleeping posture. Only a small part of the unusual position will cause us to produce false positives. For this issue, we will expand the test population to further measure the impact of unusual arm's position on our system and consider more hand positions in future work.
- **Actionable feedback.** The current version of SLEEPGUARD has been designed to provide simple recommendations on how users should improve their sleeping environment and habits. These can be linked with additional suggestions that may alleviate the causes. For example, problems in falling asleep can be mitigated by doing some exercise before going to bed or by going to sleep with soft music that can be automatically turned off. Similarly, we can identify poor postures and hand positions and give feedback on what the users should aim to improve to reduce sleep problems. For example, [48] present an anti-supine device mimicking the so-called "tennis ball technique" to control sleep posture, in order to improve OSA hypopnoea syndrome. For some problems, such as persistent snoring or coughing, our system can provide suggestions such as how to improve posture to mitigate these problems, or potentially detect severe cases where medical intervention would be appropriate. Indeed, for long term snoring the medical guidelines suggest undergoing a physical examination so that they can timely discover possible physical diseases that may cause snoring, such as high blood pressure, cardiovascular and cerebrovascular diseases.

6 RELATED WORK

Sleep quality is a crucial issue for each individual, and the poor sleep would lead to numerous diseases, such as endocrine dyscrasia, depression, immunity decline [33, 71]. Thus, a lot of research works have been proposed to monitor the sleep [12, 13, 37, 42, 45, 57]. Below, we summarize the related state-of-the-art research works as the following three categories.

Medical technology based work. Traditionally, the dedicated medical technologies, like EEG, ECG and EMG [61], have been applied for sleep monitoring. Those technologies rely on the certain biomedical signals, such as brain wave, muscle tone and eye movement, to assess the sleep quality. For example, the EEG technology in [29, 45, 53] monitors the brain waves, and then recognizes the sleep stages by leveraging unsupervised learning approaches.

Although a high accuracy achieved by those technologies, they have two drawbacks. First, those technologies require the dedicated medical devices, which are very expensive compared the widely available smart watch/phone.

Second, they require the users to be attached lots of sensors on the human body, which may cause a healthy person had to sleep and result in large errors.

Compared with those medical technologies, our system has two advantages. First, we only need a smartwatch, which is cost effective. Second, the smartwatch does not disturb a user's normal sleep, thus we can monitor the user's sleep quality more accurately.

Smartphone based work. Recently, some researchers use the smartphones for sleep monitoring [12, 23, 37, 42]. iSleep [37] measures the sleep quality by recording the sleep-related acoustic events and evaluates the sleep quality using the Pittsburgh Sleep Quality Index (PSQI) [18]. Bai et al. [12] predicts a user's sleep quality by observing the user's daily activities with the smartphone's sensor data, like the accelerometer, gyroscope, microphone, etc. Work in [42] leverages the smartphone sensors to record the sleep disruptors for a user. The authors in [23] explore a series of opportunities to support healthy sleep behaviors based on the smartphone sensors. Several research works predict the sleep quality by using the smartphone to monitor the external influence factors, such as the daily activity, sleeping environment, location and family settings [21, 74]. Besides those research works, many Smartphone Apps, such as Sleep As Android [10], Sleep Journal [11], YawnLog [73] also have been developed to monitor a user's sleep quality.

Those smartphone based systems, however, require placing the smartphone at a specifically location near to the user, which usually cannot be satisfied in reality. For example, Gu et al. [34] needs the smartphone to be placed next to the user's head, and requires the smartphone keeping stationary throughout the sleeping process. But, many users do not want to place the mobile phone too close to the body due to health risk concerns [38, 59].

Compared with the existing smartphone based systems, our system uses the commodity smartwatch for sleep monitoring. Since many users are willing to wear the smartwatch throughout the night, thus the smartwatch can remain relatively close to the user over the duration of sleep so that more user-related data can be collected. The performance of SLEEPGUARD is better for most sleep events, and comparable to current state-of-the-art solutions for sleep stage estimation. However, the main advantage of SLEEPGUARD results from its capability to capture a wider range of sleep-related events and to use this information to assess sleep quality and potential factors affecting it.

Wearable device based work. With the widely use of wearable devices, more and more researchers and industries try to use the smartwatch or wearable-wrist for sleep monitoring [13, 15, 19, 57]. The Sleeptracker [64] 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 ubiSleep [57] joints heart rate, accelerometer, and sound signals collected into the smartwatch for noninvasive sleep monitoring. The aXbo alarm clock [9] is packaged as a stand-alone application in the form of an alarm clock that wirelessly communicates with a wrist-band unit. The Zeo [19] is similarly using an alarm clock base unit with a worn sensing device, but the latter is a headband rather than a wrist-band that based on electroencephalograph (EEG) to monitor sleep. We know Zeo has good performance in sleep stage detection compared to some wristband sleep monitoring products because products based on some biological signals like EEG are able to get more accurate and more representative sleep data than actigraphy-based [2] sleep monitoring products. However, the vast majority of biosignal-based sleep detection approaches require specialized equipment, which is high cost and complex to operate, while the approaches based on actigraphy are well adapted and user-friendly to accept and understand.

Moreover, these actigraphy-based wristband devices or smartwatch Apps only can gather coarse-grained information and do not design a model for deep understanding the relationship between a user's sleep pattern and the sensor data. For example, many smartwatch Apps, like Jawbone Up [69], FitBit [32], YawnLog [73] and WakeMate [72], do not show how they assess a user's sleep quality based on what kind of sensor data.

Compared with the existing smartwatch or wearable-wrist based systems, our system is a more complete sleep monitoring system, and is based only on sensors in commercial smart watches without additional hardware. It collect an extensive set of sleep-related events, many of which are not supported in prior work. And for sleep

Table 10. Summary of existing solutions.

System	High accuracy	Practicality	Low disruptive	Low cost	Informativeness	Interpretability
PSG	✓				✓	
Smartphone solutions		✓	✓	✓		✓
Wearable solutions		✓	✓	✓		✓
SLEEPGUARD		✓	✓	✓	✓	✓

monitoring in daily life, it is more practical and will not invasion users' normal sleep, and more and more people are willing to accept to wear watches to fall asleep, unlike other wearable devices such as chest-worn sensors, most people are still unwilling to accept to wear her to sleep. Moreover, our original intention and focus are more inclined to enable users to have a deeper and more comprehensive understanding of their sleep, explore the causes of sleep quality, and provide users with more practical advice to point them in a clear direction for improving sleep quality and being healthy.

To summarize, as we can see from Table 10, the advantage of SLEEPGUARD is that it can detect more fine-grained sleep-related events to obtain more abundant sleep information, which is currently on the market for commercial or scientific research sleep monitoring system can not be achieved. And the performance of SLEEPGUARD has been improved to some extent. Our original intention and focus are more inclined to enable users to have a deeper and more comprehensive understanding of their sleep, explore the cause s of sleep quality, and provide users with more practical advice to point them in a clear direction for improving sleep quality and being healthy. And compared with some medical technology like PSG, our advantages are inexpensive and easy to deploy at home, so it is suitable for most general public. And it does not need a large number of instruments attached to the user's body, thus it has less intrusiveness for sleep and does not require professional personnel to operate. Although the accuracy of SLEEPGUARD's ability to detect sleep cannot be compared to medical technology, it is enough for the average family's daily sleep monitoring needs. The most important is that we concentrates on physical activities rather than biomedical signals, so these rich physical activities detected are easily understood by users, and they can be adjusted with improved and improved sleep based on the results of monitoring.

7 SUMMARY AND CONCLUSION

In this paper, we have presented SLEEPGUARD, the first holistic smartwatch-based sleep monitoring solution that can simultaneously estimate sleep quality and provide rich information about sleep events, including body motions, acoustic events related to sleep disorders, and ambient illumination. To capture this information accurately, we have proposed new algorithms for extracting the relevant events from sensor information. We demonstrated the benefits of SLEEPGUARD through rigorous benchmark experiments carried out using measurements collected from a two week trial with 15 participants. The results of our experiments demonstrate that SLEEPGUARD provides comparable sleep quality estimation accuracy than state-of-the-art consumer grade sleep monitors, while at the same time being able to accurately capture a rich set of information about factors influencing sleep quality. This information is particularly important for identifying possible causes of poor quality sleep and can be used to provide the user with suggestions on how to improve their sleep quality, e.g., by improving their sleep environment or behaviours surrounding sleep. We also compared the sleep quality estimates of SLEEPGUARD against subjective self assessments, demonstrating a high degree of correspondence.

REFERENCES

- [1] 2018. https://en.wikipedia.org/wiki/Respiratory_rate. (2018).
- [2] Actigraphy. 2017. <http://en.wikipedia.org/wiki/Actigraphy>. (2017).

- [3] Hanoglu L, Agargun M Y, Boysan M. 2004. Sleeping Position, Dream Emotions, and Subjective Sleep Quality. *Sleep & Hypnosis* 6, 6 (2004), 8–13.
- [4] Torbjörn Åkerstedt. 2007. Altered sleep/wake patterns and mental performance. *Physiology & Behavior* 90, 2 (2007), 209 – 218. <https://doi.org/10.1016/j.physbeh.2006.09.007> Includes a Special Section on Chronobiology Aspects of the Sleep–Wake Cycle and Thermoregulation.
- [5] Adriano M. Alencar, Diego Greatti Vaz da Silva, Carolina Beatriz Oliveira, André P. Vieira, Henrique T. Moriya, and Geraldo Lorenzi-Filho. 2013. Dynamics of snoring sounds and its connection with obstructive sleep apnea. *Physica A: Statistical Mechanics and its Applications* 392, 1 (2013), 271 – 277.
- [6] Paula Alhola and Päivi Polo-Kantola. 2007. Sleep deprivation: Impact on cognitive performance. *Neuropsychiatric Disease and Treatment* 3, 5 (2007), 553–567.
- [7] Ellemarije Altena et al. 2008. Sleep loss affects vigilance: effects of chronic insomnia and sleep therapy. *Journal of sleep research* (2008).
- [8] Amy. 2014. <http://www.abc.net.au/science/articles/2014/01/28/3861788.htm>. (2014).
- [9] Sonia Ancoli-Israel, Roger Cole, Cathy Alessi, Mark Chambers, William Moorcroft, and Charles P Pollak. 2003. The role of actigraphy in the study of sleep and circadian rhythms. *Sleep* 26, 3 (2003), 342–392.
- [10] Sleep As Android. 2017. <https://play.google.com/store/apps/details?id=com.urbandroid.sleep&hl=zhEN>. (2017).
- [11] Sleep Journal App. 2017. <http://sleeptrackerapp.blogspot.com/>. (2017).
- [12] Yin Bai, Bin Xu, Yuanhao Ma, Guodong Sun, and Yu Zhao. 2012. Will you have a good sleep tonight?: sleep quality prediction with mobile phone. In *Proceedings of the 7th International Conference on Body Area Networks*. 124–130.
- [13] D Bain, M Ferguson-Pell, and A McLeod. 2003. Evaluation of mattresses using interface pressure mapping. *Journal of wound care* 12, 6 (2003), 231–235.
- [14] Sally L. Davidson Ward Laure M. Fisher Debra M. Don Beth Osterbauer, Jeffrey A. Koempel. 2016. A Comparison Study of the Fitbit Activity Monitor and PSG for Assessing Sleep Patterns and Movement in Children. *Journal of Otolaryngology Advances* 1, 3 (2016), 24.
- [15] MH Bonnet and DL Arand. 2003. Insomnia, metabolic rate and sleep restoration. *Journal of internal medicine* 254, 1 (2003), 23–31.
- [16] Marko Borazio and Kristof Van Laerhoven. 2012. Combining wearable and environmental sensing into an unobtrusive tool for long-term sleep studies. In *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium*. ACM, 71–80.
- [17] Daniel J Buysse, Charles F Reynolds, Timothy H Monk, Susan R Berman, and David J Kupfer. 1989. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry research* 28, 2 (1989), 193–213.
- [18] Janet S Carpenter and Michael A Andrykowski. 1998. Psychometric evaluation of the Pittsburgh sleep quality index. *Journal of psychosomatic research* 45, 1 (1998), 5–13.
- [19] John N Caviness. 1996. Myoclonus. In *Mayo Clinic Proceedings*, Vol. 71. 679–688.
- [20] Tarani Chandola et al. 2010. The effect of short sleep duration on coronary heart disease risk is greatest among those with sleep disturbance: a prospective study from the Whitehall II cohort. *Sleep* (2010).
- [21] Zhenyu Chen, Mu Lin, Fanglin Chen, Nicholas D Lane, Giuseppe Cardone, Rui Wang, Tianxing Li, Yiqiang Chen, Tanzeem Choudhury, and Andrew T Campbell. 2013. Unobtrusive sleep monitoring using smartphones. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2013 7th International Conference on*. 145–152.
- [22] Ah-reum Cho, Chul-Hyun Cho, Ho-Kyoung Yoon, Joung Ho Moon, Heon-Jeong Lee, and Leen Kim. 2016. Impact of Dim Artificial Light at Night (dALAN) Exposure during Sleep on Human Fatigue. 23 (01 2016), 53–60.
- [23] Eun Kyoung Choe, Sunny Consolvo, Nathaniel F Watson, and Julie A Kientz. 2011. Opportunities for computing technologies to support healthy sleep behaviors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3053–3062.
- [24] CNN. 2015. The great American sleep recession. <https://edition.cnn.com/2015/02/18/health/great-sleep-recession/index.html>. (2015).
- [25] Joseph De Koninck, Pierre Gagnon, and Serge Lallier. 1983. Sleep Positions in the Young Adult and Their Relationship with the Subjective Quality of Sleep. *Sleep* 6, 1 (1983), 52–59. <https://doi.org/10.1093/sleep/6.1.52>
- [26] Claudatos S de Zambotti M, Goldstone A et al. 2017. A validation study of Fitbit Charge 2? compared with polysomnography in adults. *Chronobiology international* (2017), 1–12.
- [27] Gustavo Desouzart, Ernesto Filgueiras, and Rui Matos. 2015. Relationship between Postural Reeducation Technique During Sleep and Relaxation Technique in Sleep Quality. *Procedia Manufacturing* 3 (2015), 6093–6100.
- [28] Dr. Stephen Amira Dr. Lawrence Epstein. 2007. <http://healthysleep.med.harvard.edu/healthy/science/how/external-factors>. (2007).
- [29] Farideh Ebrahimi, Mohammad Mikaeili, Edson Estrada, and Homer Nazeran. 2008. Automatic sleep stage classification based on EEG signals by using neural networks and wavelet packet coefficients. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*. 1151–1154.
- [30] Nathaniel A Eisman, M Brandon Westover, Jeffrey M Ellenbogen, and Matt T Bianchi. 2012. The impact of body posture and sleep stages on sleep apnea severity in adults. *Journal of clinical sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine* 8, 6 (2012), 655.
- [31] Kelly R Evenson, Michelle M Goto, and Robert D Furberg. 2015. Systematic review of the validity and reliability of consumer-wearable activity trackers. *International Journal of Behavioral Nutrition and Physical Activity* 12, 1 (2015), 159.

- [32] FitBit. 2017. <http://www.fitbit.com>. (2017).
- [33] Daniel J Gottlieb, Naresh M Punjabi, A B Newman, Helaine E Resnick, Susan Redline, Carol M Baldwin, and F Javier Nieto. 2005. Association of Sleep Time With Diabetes Mellitus and Impaired Glucose Tolerance. *JAMA Internal Medicine* 165, 8 (2005), 863–868.
- [34] Weixi Gu, Longfei Shangguan, Zheng Yang, and Yunhao Liu. 2016. Sleep hunter: Towards fine grained sleep stage tracking with smartphones. *IEEE Transactions on Mobile Computing* 15, 6 (2016), 1514–1527.
- [35] T Matsui H Miwa, S Sasahara. 2007. Roll-over detection and sleep quality measurement using a wearable sensor. (2007), 1507.
- [36] Marian Haescher, Denys JC Matthies, and Bodo Urban. 2015. Anomaly Detection with Smartwatches as an Opportunity for Implicit Interaction. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*. ACM, 955–958.
- [37] Tian Hao, Guoliang Xing, and Gang Zhou. 2013. iSleep: unobtrusive sleep quality monitoring using smartphones. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*. 4.
- [38] Step To Health. 2017. <https://steptohealth.com/bad-sleep-near-cell-phone/>. (2017).
- [39] Bernadette Hood, Dorothy Bruck, and Gerard Kennedy. 2004. Determinants of sleep quality in the healthy aged: the role of physical, psychological, circadian and naturalistic light variables. *Age and Ageing* 33, 2 (2004), 159–165. <https://doi.org/10.1093/ageing/afh051>
- [40] Girardin Jean-Louis, Daniel F Kripke, Sonia Ancoli-Israel, Melville R Klauber, and Raul S Sepulveda. 2000. Sleep duration, illumination, and activity patterns in a population sample: effects of gender and ethnicity. *Biological Psychiatry* 47, 10 (2000), 921.
- [41] L Steven Johnson, Sean R Eddy, and Elon Portugaly. 2010. Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinformatics* 11, 1 (2010), 431–431.
- [42] Matthew Kay, Eun Kyoung Choe, Jesse Shepherd, Benjamin Greenstein, Nathaniel Watson, Sunny Consolvo, and Julie A Kientz. 2012. Lullaby: a capture & access system for understanding the sleep environment. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. 226–234.
- [43] DailyBurn Kristen Domonell. 2016. <http://edition.cnn.com/2016/03/18/health/sleep-positions-good-bad/index.html>. (2016).
- [44] Tea Lallukka and Erkki Kronholm. 2016. The contribution of sleep quality and quantity to public health and work ability. (2016).
- [45] Martin Langkvist, Lars Karlsson, and Amy Loutfi. 2012. Sleep stage classification using unsupervised feature learning. *Advances in Artificial Neural Systems* 2012 (2012), 107046.
- [46] Fonseca P Long X, Foussier J. 2014. Analyzing respiratory effort amplitude for automated sleep stage classification. *Biomedical Signal Processing & Control* 14 (2014), 197–205.
- [47] Hafner M, Stepanek M, Taylor J, Troxel WM, and van Stolk C. 2017. Why Sleep Matters-The Economic Costs of Insufficient Sleep: A Cross-Country Comparative Analysis. *RAND Health Quarterly* 6, 4 (2017).
- [48] Sue FILSELL Margot A. SKINNER, Ruth N. KINGSHOTT and D. Robin TAYLOR. 2008. Efficacy of the ‘tennis ball technique’ versus nCPAP in the management of position-dependent obstructive sleep apnoea syndrome. *Respirology* 13, 5 (2008), 708–715.
- [49] Bond J A Montgomery-Downs H E, Insana S P. 2012. Movement toward a novel activity monitoring device. *Sleep and Breathing* 16, 3 (2012), 913–917.
- [50] Alain Muzet. 2007. Environmental noise, sleep and health. *Sleep medicine reviews* 11, 2 (2007), 135–142.
- [51] Neil. 2014. <http://www.oddmattress.co.uk/sleeping-positions/>. (2014).
- [52] Arie Oksenberg and Donald S Silverberg. 1998. The effect of body posture on sleep-related breathing disorders: facts and therapeutic implications. *Sleep medicine reviews* 2, 3 (1998), 139–162.
- [53] Edgar Oropesa, Hans L Cycon, and Marc Jobert. 1999. Sleep stage classification using wavelet transform and neural network. *International computer science institute* (1999).
- [54] Pickett C K ouglas N J, White D P. 1982. Respiration during sleep in normal man. *Thorax* 37, 11 (1982), 840.
- [55] Tiina Paunio, Tellervo Korhonen, Christer Hublin, Markku Partinen, Mika Kivimäki, Markku Koskenvuo, and Jaakko Kaprio. 2009. Longitudinal Study on Poor Sleep and Life Dissatisfaction in a Nationwide Cohort of Twins. *American Journal of Epidemiology* 169, 2 (2009), 206–213. <https://doi.org/10.1093/aje/kwn305>
- [56] June J. Pilcher, Douglas R. Ginter, and Brigitte Sadowsky. 1997. Sleep quality versus sleep quantity: Relationships between sleep and measures of health, well-being and sleepiness in college students. *Journal of Psychosomatic Research* 42, 6 (1997), 583 – 596. [https://doi.org/10.1016/S0022-3999\(97\)00004-4](https://doi.org/10.1016/S0022-3999(97)00004-4) Nocturnal Penile Tumescence: Measurement and Research.
- [57] Nuno Pombo and Nuno M Garcia. 2016. ubiSleep: An ubiquitous sensor system for sleep monitoring. In *Wireless and Mobile Computing, Networking and Communications (WiMob), 2016 IEEE 12th International Conference on*. IEEE, 1–4.
- [58] Zeo Sleep Manager Pro. 2017. <http://www.digifit.com/Zeo/>. (2017).
- [59] Quora. 2017. <https://www.quora.com/Does-keeping-your-mobile-phone-near-your-head-affect-your-sleeping-cycle-How>. (2017).
- [60] Earl Sageron. 2016. <http://www.linkedin.com/pulse/effect-chin-strap-snoring-earl-sageron>. (2016).
- [61] Clifford B Saper, Thomas E Scammell, and Jun Lu. 2005. Hypothalamic regulation of sleep and circadian rhythms. *Nature* 437, 7063 (2005), 1257–1263.
- [62] Anita Valanju Shelgikar, Patricia F Anderson, and Marc R Stephens. 2016. Sleep Tracking, Wearable Technology, and Opportunities for Research and Clinical Care. *CHEST Journal* 150, 3 (2016), 732–743.

- [63] Jenn Sinrich. 2017. <https://www.fitnessmagazine.com/health/sleep/the-best-and-worst-sleeping-positions-for-your-health/>. (2017).
- [64] Sleeptracker. 2017. <http://www.sleeptracker.com/>. (2017).
- [65] Kort Y A W D Smolders K C H J. 2014. Bright light and mental fatigue: Effects on alertness, vitality, performance and physiological arousal. *Journal of Environmental Psychology* 39, 4 (2014), 77–91.
- [66] Dan Stowell, Dimitrios Giannoulis, Emmanouil Benetos, Mathieu Lagrange, and Mark D Plumbley. 2015. Detection and Classification of Acoustic Scenes and Events. *IEEE Transactions on Multimedia* 17, 10 (2015), 1733–1746.
- [67] Xiao Sun, Li Qiu, Yibo Wu, Yeming Tang, and Guohong Cao. 2017. SleepMonitor: Monitoring Respiratory Rate and Body Position During Sleep Using Smartwatch. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* (2017).
- [68] Ko Ping-Ru T, Julie A Kientz, Eun Kyoung Choe, Matthew Kay, Carol A Landis, and Nathaniel F Watson. 2015. Consumer Sleep Technologies: A Review of the Landscape. *Journal of Clinical Sleep Medicine (JCSM)* (2015).
- [69] Jawbone Up. 2017. <https://jawbone.com/up>. (2017).
- [70] Helka Urponen, Ilkka Vuori, Joel Hasan, and Markku Partinen. 1988. Self-evaluations of factors promoting and disturbing sleep: An epidemiological survey in Finland. *Social Science & Medicine* 26, 4 (1988), 443 – 450. [https://doi.org/10.1016/0277-9536\(88\)90313-9](https://doi.org/10.1016/0277-9536(88)90313-9)
- [71] Alexandros N Vgontzas, Duanping Liao, Slobodanka Pejovic, Susan L Calhoun, Maria Karataraki, and Edward O Bixler. 2009. Insomnia with Objective Short Sleep Duration is Associated with Type 2 Diabetes: A Population-based Study. *Diabetes Care* 32, 11 (2009), 1980–1985.
- [72] WakeMate. 2017. <http://www.wakemate.com>. (2017).
- [73] YawnLog. 2017. <http://beta.yawnlog.com>. (2017).
- [74] Jin Zhang, Qian Zhang, Yuanpeng Wang, and Chen Qiu. 2013. A real-time auto-adjustable smart pillow system for sleep apnea detection and treatment. In *Information Processing in Sensor Networks (IPSN), 2013 ACM/IEEE International Conference on*. 179–190.