SleepGuard: Capturing Rich Sleep Information Using Smartwatch Sensing Data

LIQIONG CHANG, JIAQI LU, JU WANG, XIAOJIANG CHEN, and DINGYI FANG, ZHANYONG TANG, Northwest University, China
PETTERI NURMI, Lancaster University, United Kingdom, University of Helsinki, Finland
ZHENG WANG, Lancaster University, United Kingdom

Sleep is an important part of our daily routine – we spend about one-third of our time doing it. By tracking sleep-related events and activities, sleep monitoring provides decision support to help us understand sleep quality and 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. However, existing solutions do not take full advantage of the rich sensor data provided by these devices. In this paper, we present the design and development of SleepGuard, a novel approach to track a wide range of sleep-related events using smartwatches. We show that using merely a single smartwatch, it is possible to capture a rich amount of information about sleep events and sleeping context, including body posture and movements, acoustic events, and illumination conditions. We demonstrate that through these events it is possible to estimate sleep quality and identify factors affecting it most. We evaluate our approach by conducting extensive experiments involved 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.

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

Additional Key Words and Phrases: Smartwatch sensing; mobile sensing; sleep events; sleep monitoring

ACM Reference Format:
Liqiong Chang, Jiaqi Lu, Ju Wang, Xiaojiang Chen, Dingyi Fang, Zhanyong Tang, Petteri Nurmi, and Zheng Wang. 2018. SleepGuard: Capturing Rich Sleep Information Using Smartwatch Sensing Data. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 3, Article 98 (September 2018), 34 pages. https://doi.org/10.1145/3264908

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 [6, 19, 43], decrease the level of cognitive performance [3, 5], and affect mood and feelings of personal well-being [57, 58]. 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 the treatment of sleep disorders [46]. Indeed, to highlight the significance of sleep quality, the Centre for

Authors' addresses: Liqiong Chang, Jiaqi Lu, Ju Wang; Xiaojiang Chen; Dingyi Fang, Zhanyong Tang, Northwest University, China, {clq,jqlu,wangju,xjchen,dyf,zytang}@nwu.edu.cn; Petteri Nurmi, Lancaster University, United Kingdom, University of Helsinki, Finland, p.nurmi@lancaster.ac.uk; Zheng Wang, Lancaster University, United Kingdom, z.wang@lancaster.ac.uk.

2474-9567/2018/9-ART98 $15.00
https://doi.org/10.1145/3264908

Disease Prevention (CDC) has declared insufficient sleep as a public health problem in the US [23], and the concern is widely shared amongst other 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), electrooculogram and oxygen saturation [29, 44, 54, 67]. These information sources can then be used to determine sleep stages, sleep efficiency, abnormal breathing, and overall sleep quality. 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, attaching and outfitting the sensing instruments are time-consuming and laborious, 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 where clinical care is required.

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 [68, 73]. By taking advantage of diverse sensors, behaviors and routines associated with sleeping can be captured and modeled. This, in turn, can help users understand their sleep behavior and provide feedback on how to improve 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 [9, 32, 33, 60, 72, 74].

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 [41, 72, 77], 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 actionable recommendations on how to improve sleep quality. This is because current solutions focus on monitoring characteristics of the sleep itself, without considering behaviors 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, the intensity of ambient light [38] and noisiness [51] of the environment can significantly affect sleep quality. Similarly, the user's breathing patterns, postures 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.

This paper presents 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 [68]. By using a smartwatch, the sensors are close to the user during all stages during the sleep, 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

Table 1. Sleep events detected by SleepGuard.

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

orientation relative to the user's body need to be tracked and opportune moments where to capture sensor data need to be detected. SleepGuard addresses these issues by integrating a set of new methods for analyzing 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 [15, 35, 59, 68], 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 15 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 event-detection 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, 6 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 causes of sleep problems for the 4 participants with bad sleep quality, whether it is due to suboptimal hand position, body posture or the 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.

This paper makes the following contributions:

- We present the design and development of SleepGuard, 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 behaviors 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

(Sec. 2.1.1), hand positions (Sec. 2.1.2), body rollovers (Sec. 2.2), micro body movements (Sec. 2.3), and acoustic events (Sec. 2.4) and illumination conditions (Sec. 2.5).
- We extensively evaluate the performance of SleepGuard using measurements collected from two-week monitoring of 15 participants (Sec. 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 (Sec. 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 behaviors and sleeping environment (Sec. 4.3).

2 THE SLEEPGUARD SLEEP MONITORING PLATFORM

SleepGuard is a novel smartwatch-based sleep monitoring system that aims at estimating sleep quality and capturing rich information about behaviors 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 goals, SleepGuard exploits a wide range of sensors that are common on commercial off-the-shelf smartwatches: (i) the accelerometer, the gyroscope, and the orientation sensor are used to collect body and hand movements; (ii) the microphone is used to measure the level of ambient noise and to capture acoustic events; and (iii) the ambient light sensor is used to monitor illumination within the sleep environment. Table 1 summarizes the sleep events that can be detected by SleepGuard using data collected from these sensors.

It is to note that the design choices and algorithm parameters of SleepGuard are empirically determined from our pilot study (Sec. 3.1). The users involved in the pilot study are different from those participating in the experiments for evaluating the developed system (Sec. 3.2).

2.1 Detecting Sleep Postures and Movements

One's sleeping position, 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 analyze effects of sleep disorders [30, 53] while having a good sleep posture has been shown to correlate with subjective assessments of sleep quality [25]. Similarly, the 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 the user's sleep posture and habits. In the following, we detail the techniques we use for capturing the body posture and movements. SleepGuard, currently supports the four basic sleep postures (see Fig. 1); three hand positions (see Fig. 2); six types of body rollovers (see Fig. 3 for an example); and three 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 [2]. Sleep posture also varies across individuals and should fit the personal and physical needs of the individual [42, 69]. 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 unconsciously avoid postures that would be beneficial for health and sleep quality [27]. Having an effective way to detect current posture and track changes in it would thus be essential for estimating overall sleep quality, and avoiding potential harm. SleepGuard captures four basic sleep postures: supine, left lateral, right lateral, and prone. These are illustrated in Fig. 1. Detecting these postures using a single wrist sensor, however, is non-trivial because the sensor cannot accurately track the movement of the entire body. To accomplish posture detection, we observe that the arm position strongly correlates with sleep posture, i.e., the arm is typically located in a specific, stable location for a given posture. This

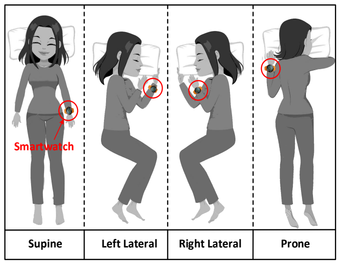


Fig. 1. Four sleep body postures.

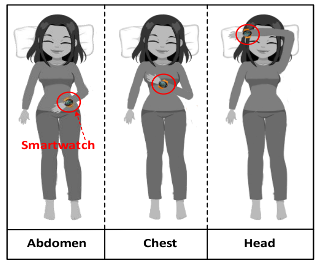


Fig. 2. Three hand positions.

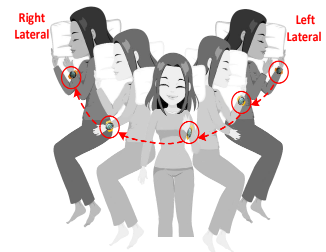
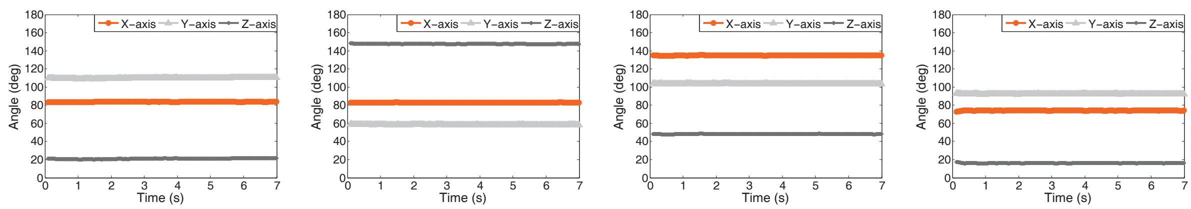


Fig. 3. Body rollover from the left side to the right side.



(a) Supine (b) Left Lateral (c) Right Lateral (d) Prone

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

suggests that we can first identify the user's arm position and the *time the position is approximately stable, which can then be mapped into* a sleep posture. Later in this paper, we show that our approach achieves 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 is placed on the left side of the body, on the abdomen, on the chest or on the head; on the left and right lateral positions we assume the hand to be close to the pillow, on the chest or on the waist; and, 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.

Like SleepMonitor [72], we use three dimensional tilt angles to detect postures. To identify which posture 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. 3.1). 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. We can observe clear differences in the tilt angles of the three axes. The sleeping 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 can be 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.

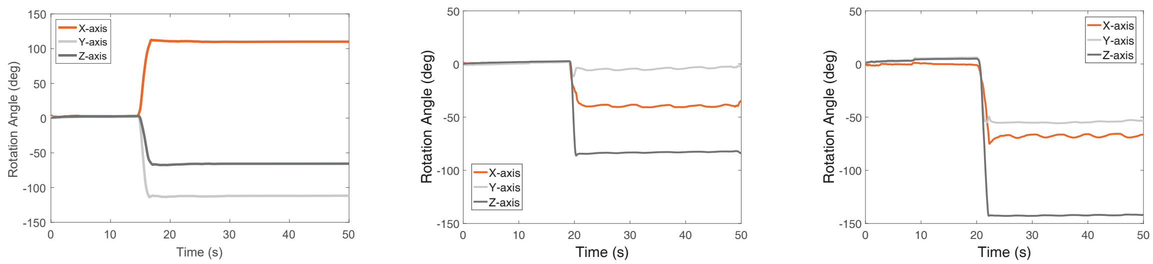
To improve detection accuracy between supine and prone postures, SleepGuard integrates orientation data as an auxiliary feature. This is motivated by the observation that hand directions differ in supine and prone positions. When the result of the previous step is prone or supine and hand is detected 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. When we use the direction sensor, we must limit the pillow orientation remaining unchanged (in the experiment our pillow is placed on the north). In fact, this assumption can be easily satisfied since most people usually have fixed sleep directions.

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 [52]. 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 [52].

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 (Sec. 3.1). 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. When the hand is not on the head, we use respiratory events to identify whether the hand is on the abdomen or chest. We now describe how to detect each of the three positions in more details.

Hands on the head. Fig. 5 shows the change of rotation angle using the gyroscope for one of our pilot study users when his hand was initially placed next to the body and then moved to his head, abdomen, and chest. As can be seen from the figure, when the hand is moved to the head, changes in rotation angles are significantly different from the readings when moving the hand on the abdomen or chest. This is largely due to the palm facing direction - it is upward when the hand is placed on the head but downward when the hand is placed in other positions. SleepGuard exploits this observation to detect if the hand is placed on the head by examining changes of the tilt and rotation angles. We use a hierarchical classifier consisting of two K nearest neighbor 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 facing up. Training data for detecting palm direction (upward or downward) are collected from our pilot study users when their hands are placed on the head, abdomen and chest respectively; see Sec. 3.1) If the first KNN model suggests that the palm was facing upward, the second KNN then uses differences of rotation readings (from the x, y, and z directions) before and after hand movement to determine if the most likely position is on the head or elsewhere. As similarity measure we consider the Euclidean distance from the input data to each of the training samples – consisting of the rotation angle values from the three directions. 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.

Hands on the abdomen or chest. When the aforementioned hierarchical model predicts the hand is not placed on the head, the hand can then be located anywhere, including different parts of the body. However, in our case we are only interested in detecting if the hand is on the chest or abdomen – or at neither of these positions. We build on the intuition that the hand is likely to be affected by breathing whenever it is placed on the chest or abdomen. Specifically, the impact of breathing results in periodical



(a) moving to the head (b) moving to the abdomen (c) moving to the chest

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 the subject's head (a), abdomen (b), and chest (c).

fluctuations on the accelerometer readings. This is because the hand will be pushed up and drop down due to breathing. Our experimental data suggest that this behavior only takes place when the hand is placed on the abdomen or the chest, not when the hand is located elsewhere (such as the shoulder). Therefore, we can separate these two locations from others by examining whether the accelerometer values are impacted by respiration.

To examine if the accelerometer data are affected by the respiration, we calculate the power spectral density (PSD) of the collected accelerometer data. We then check the PSD to see if we can observe any peak that closely matches 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 PSD for one of 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).

Putting things together, we thus first use the PSD detector to identify whether a respiratory event is taking place. When respiratory events are detected, 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 b and c in Fig. 5). On the other hand, if no respiratory peak is detected, we assume the hand to be located at another place on the body that is not supported. Similarly to 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.

2.1.3 Labeling REM/non-REM Stages. We have 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 a prior study showing that when people sleep in the REM stage, their respiratory amplitude is smaller than that in the other stages [56]. Hence,

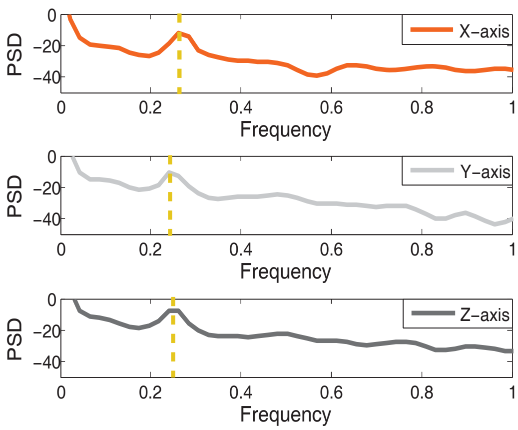


Fig. 6. The power spectral density (PSD) of the accelerometer readings when a user's hand is placed on his chest.

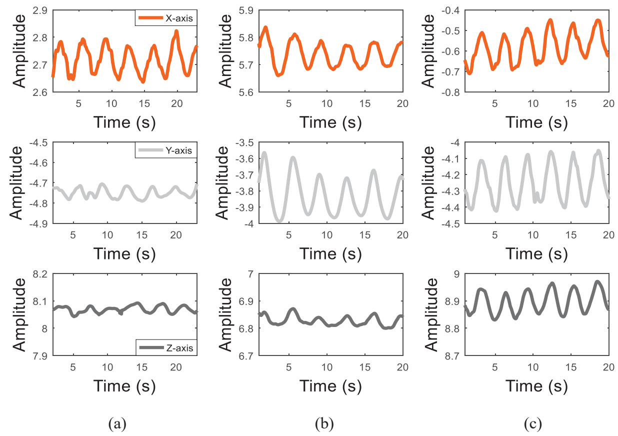


Fig. 7. The periodic change of the acceleration signal. (a) REM–Location 1. (b) REM–Location 2. (c) NREM–Location 1.

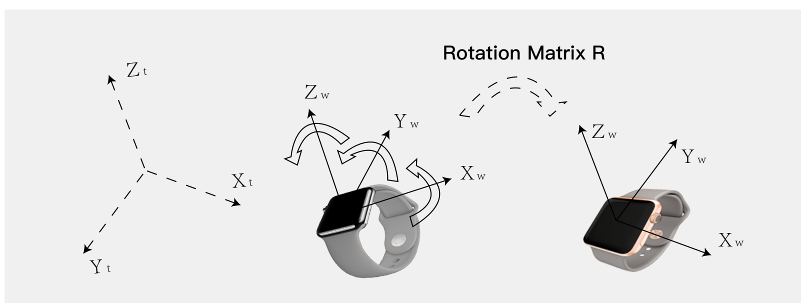


Fig. 8. Coordination system conversion. The left-most figure is the torso coordinate system where Y_t points to the north. The middle figure shows the watch's coordinate system when the watch is placed on an arbitrary position. The right-most figure shows that the resulting watch coordinate system after performing the coordinate system conversion.

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 judgment. However, it serves as an early reference that helps later phases of the sleep stage detection. Normally chest movement amplitude is smaller than abdominal movement amplitude. However, during different sleep stages, respiration amplitudes differ [45]. Therefore, it is likely that there is a situation where chest movement amplitude in the NREM stages is close to the abdominal movement amplitude in the REM sleep stage. As a result, a naïve solution of applying a threshold cannot work satisfactorily but we can combine it with the position of the hand detected earlier. 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 degree of up and down motions, from which we can approximately infer the current sleep stage.

Example. Now we take the case of hands on the abdomen as an example. Even when the hand is placed on the abdomen, there are some minor changes in the exact location of the user's hand, and the intensity of accelerometer fluctuations caused by respiration also varies greatly. 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 the hand is on the abdomen, but the location differs, we cannot directly judge the respiratory amplitude from the amplitude of the three accelerometer axes.

To solve this problem, we convert the acceleration data from the wristwatch coordinate system into the torso coordinate system. As breathing results in chest moving up and down, movements along the z-axis in the torso coordinate system can be used to identify respiratory amplitude. We express the tri-axial 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. Our coordinate alignment aims to find a rotation matrix R that aligns the watch's coordinate system with the torso coordinate system ($[X_t, Y_t, Z_t]$). Matrix R can be obtained from the three-axis direction information of the orientation sensor. Specifically, we have:

$$X_t = (X_w \cos\gamma + Y_w \sin\gamma)\cos\theta + (Y_w \cos\sigma + Y_w \sin\sigma)\sin\theta, \tag{1}$$

$$Y_t = -((Y_w \cos\sigma + Y_w \sin\sigma)\cos\theta - (X_w \cos\gamma + Y_w \sin\gamma)\sin\theta), \tag{2}$$

$$Z_t = (Z_w \cos\gamma - Z_w \sin\gamma)\cos\sigma - (Z_w \cos\gamma - Z_w \sin\gamma)\sin\sigma, \tag{3}$$

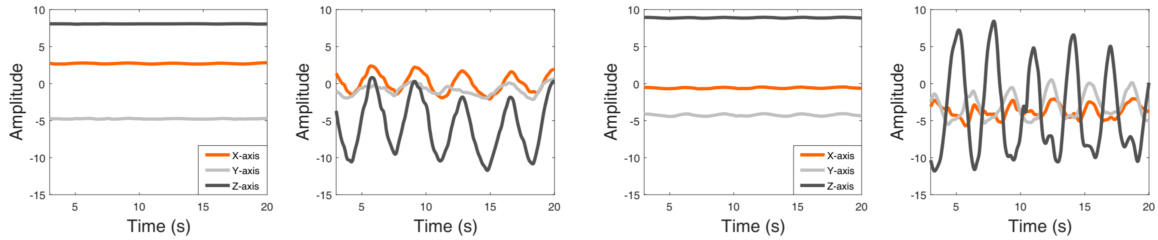
where θ, σ and γ are the x, y and z axis data of the orientation sensor respectively, representing the direction angle, the tilt angle and the roll angle collected from the orientation sensor. After alignment, we can see in Fig. 9 that the z-axis shows a periodic signal with significant fluctuations, while the x- and y-axis data undergo smaller changes around zero, which is consistent with stable sleep influenced only by respiratory patterns.

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

Summary. To summarize, we use respiratory amplitude to detect when the user is in the REM stage. We calculate respiratory amplitude when the hand is found to be placed on the abdomen or the chest. Note that as SleepGuard operates using a wrist-worn device, it can only detect respiratory events when the hand is placed in one of these two positions. As a measure of respiratory amplitude, we use the variance of z-axis acceleration. We then use a KNN classifier to find from our training examples, which training example is most similar to the variance of the acceleration collected from z directions. The similarity is measured by calculating the distance on the feature space. We then use the label (either REM or NREM) associated with the nearest training example as the classification outcome.

2.2 Body Rollover Counts

Normally, people roll their body around 20-45 times during a sleep. The main function of body rollovers is simply to maintain a comfortable sleeping position as maintaining the same position for a prolonged



(a) REM

(b) NREM

Fig. 9. Acceleration data for different sleep stages.

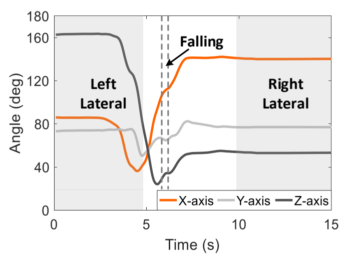


Fig. 10. From left lateral posture to right lateral posture.

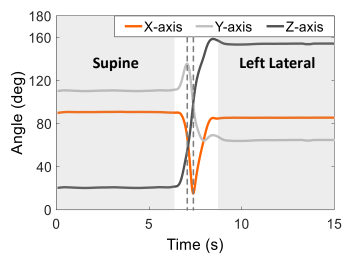


Fig. 11. From supine posture to left lateral posture.

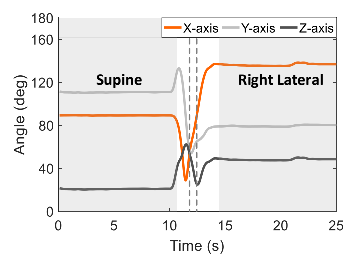


Fig. 12. From supine posture to right lateral posture.

period will result in muscular tension due to the hindered blood supply, which can also lead to local numbness [7]. Therefore, body rollovers are another key indicator of sleep quality. SleepGuard detects the number of body rollovers, which provides cues about sleep quality. Using the body rollover counts, SleepGuard calculates the rollover frequency which is then used to identify the current sleep stage [34].

In general, there are six possible body rollover transitions. These include four posture transitions between the supine and laterals (left and right), and two posture transitions between the left and right laterals. Fig. 3 depicts the case when the body moves from the left to the right side. An intuitive way for detecting body rollover events is to estimate the rotation direction of the arm using the rotation angle data 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 to Fig. 15 show the tilt angle reading changes under different body rollover scenarios. To this end, a naïve 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. Based on this observation, we use a two-step approach for detecting when a body rollover took place. We first detect when the

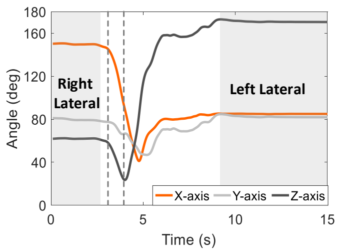


Fig. 13. From right lateral posture to left lateral posture.

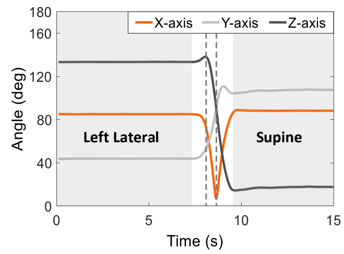


Fig. 14. From left lateral posture to supine posture.

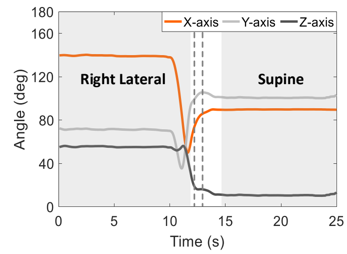


Fig. 15. From right lateral posture to supine posture.

tilt angle value has changed, and mark the time the angle value was changing as when a possible body rollover starting point. Next, we use the sleep posture classification algorithm (described in Sec. 2.1.1) to determine whether the body postures before (within a time window) and after (within a time window) this point were the same. If the postures are considered to be different, our algorithm then assumes a body rollover event is detected; otherwise, it assumes the change of tile angle readings was simply due to a changing arm position. We stress that SleepGuard not only counts the number of rollovers, but also reports the nature of the rollover event (e.g., from which side to which side).

2.3 Detecting Micro-body Movements

Besides major-body movements, 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 behaviors are most likely to occur during the deep sleep stage and the REM stage [8, 39]. 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 micro-body movements including hand moving, arm raising, and body trembling.

Noise canceling. One of the challenges in detecting 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 outliers. To determine the size of the moving window, we apply different parameter settings to our training data. We found that a moving window with a size of 35 data points 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 the Root Sum Square (RSS) value for the data points within a window:

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

$acc_x(t)$, $acc_y(t)$ 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 of two consecutive time stamps, $t-1$ and t:

$$V(t) = Rss(t) - Rss(t-1). \tag{5}$$

We use $V(t)$ to cancel noises in the accelerometer data. Specifically, if $V(t)$ is less than a threshold of 0.03, we consider the change in the RSS values is due to noises and assume there is no micro-body

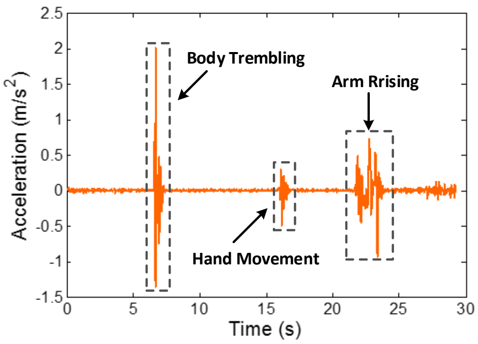


Fig. 16. Accelerometer readings for micro-body movements using samples from one of the users from the pilot study.

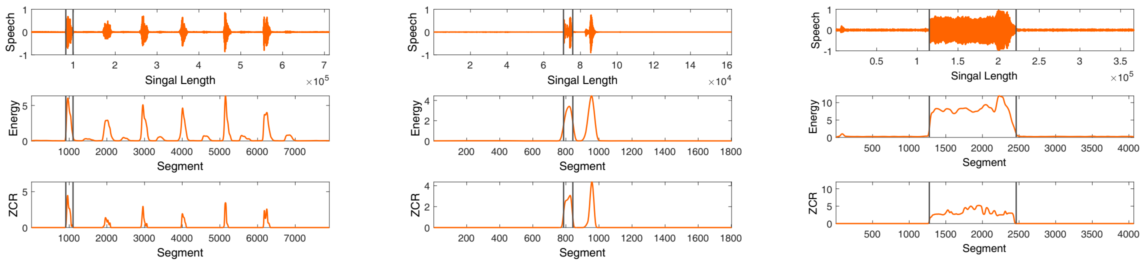
movement between windows $t-1$ and t. This threshold is automatically learned from the training data used in our pilot study (Sec. 3.1).

Differentiating from large-body movements. We observe that the micro-body movement duration is often very short, lasting less than two seconds in our training data. By contrast, the average duration for large-body movements found in our body rollover experiments (Sec. 2.2) lasts for three seconds (as shown in Fig. 10 - Fig. 15). Base on this observation, we divide the body movement events into large- and micro-body movements by measuring the signal duration time. We consider a body movement to be micro is its duration is within two seconds; otherwise, the body movement will be labeled as a large-body movement such as a body rollover.

Distinguishing among micro-body movements. Now we have a way to distinguish between micro- and large-body movements, we need to identify if the detected micro-body movement is a hand movement, an arm raising action or body trembling. To differentiate among those three micro-body movements, we turn again to 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 differentiate arm rising from the other two micro-body movements. We also find that a body trembling event tends to lead to a more drastic change in the accelerometer readings compared to a hand movement. This observation is depicted in Fig. 16 using samples from one of our training users. Based on this observation, we use the change of accelerometer reading to distinguish between the body trembling and the hand movement. Like our noise canceling strategy in Equation 5, 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 seconds, an arm rising is detected. These thresholds are empirically determined from our training data used in the pilot study (Sec. 3.1).

2.4 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, the snore is a possible symptom of cerebral infarction patients; and long-term snoring can also have a serious impact on health and sleep because snoring can cause sleep



(a) Six snore events. (b) Two consecutive coughs. (c) Somniloquy.

Fig. 17. Characteristics of different acoustic events targeted in this work. The thick vertical lines mark the duration of a detected acoustic event.

apnea or narcolepsy, a sleeping disorder [4, 66]. 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 is designed to detect three different acoustic events, including snore, cough and somniloquy. Classical acoustic algorithms use multi-dimensional features to detect acoustic events from a complex environment [33]. By contrast, we tailor our design to the problem domain to derive a simpler yet effective acoustic detection algorithm. This is achieved by exploiting the characteristics of the different events that we target.

Acoustic event characteristics. To understand the characteristics of the acoustic events targeted in this work, we carry out an interesting recognition experiment (as part of our pilot study - see Sec. 3.1) using the microphone built in the 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 wear 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. The figure has six snore events, two consecutive cough and somniloquy events. In the remaining paragraphs of this sub-section, we describe how to exploit these characteristics to detect acoustic events.

Pre-processing. To identify an acoustic event, our acoustic event detection algorithm first preprocesses the recorded audio stream by dividing it into segments of equal durations. In this work, each audio segment consists of 256 samples with a duration of 12 milliseconds.

Observations. Our acoustic event detection algorithm is designed based on the following observations:

- The duration a single acoustic event is different among different types of acoustic events. For example, Fig. 17 suggests that the duration of a snore is shorter than that of a cough or somniloquy. Furthermore, in general, the duration of a cough is shorter than that of a somniloquy signal.
- As shown in Fig. 17, the time intervals between two signals for two different types of acoustic events are significantly different. Specifically, there is a long time interval between two snores, but the time interval between two consecutive coughs is much shorter. In contrast to snores and coughs, the interval between any two consecutive somniloquy signals is irregular and does not exhibit a periodic property.

- The frequencies for different acoustic events are significantly different. Specifically, a snore event has a continuous and periodic signal, while a cough or somniloquy are sudden events. As a result, the number of consecutive occurrences for coughs and somniloquy tends to be small during sleep.

Acoustic event classification. Our acoustic event detection algorithm is a C4.5 decision-tree-based classifier [62]. It takes as input the "duration", "interval" and "frequency" values of an identified acoustic event, and predicts which of the three types acoustic events (snores, coughs, or somniloquy) the event corresponds to. The thresholds of the decision tree are automatically learned from our pilot study training data.

To estimate the *duration* of the acoustic event, we identify the start and the end points of an acoustic event. After identifying the duration, we then perform the peak detection and use the result of peak detection to estimate the *interval* and *frequency* of an acoustic event. Specifically, we use the short-term average energy (Equation 6) to calculate the peak value of an acoustic signal. When the peak exceeds a pre-defined threshold, i.e., 3 dB in this work, we record the position of each peak and calculate the *interval* between two consecutive peaks. We measure the *frequency* of an acoustic event by counting the number of peaks within a time window. Our approach for determining the duration of an acoustic event is described as follows.

Determining an acoustic event duration. In this work, we utilize an existing speech algorithm [61] to detect the start and the end points of an acoustic event. The algorithm works by applying a weight function to the time domain of the signal to detect the start and end points for a signal segment of interests. The weight function calculates two metrics of a signal segment: *the short-time energy* and *the zero-cross rate* (ZCR); it then compares the values against some thresholds to determine the points. The two metrics are defined as follows.

The short-term energy of a signal is computed as:

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

where N is the length of the time window, x is the signal and ω is the impulse response.

The zero-crossing rate is computed as:

$$Z_i = \frac{1}{2}\sum\nolimits_{j=0}^{N}|sgn[x_i(j)] - sgn[x_i(j-1)]|. \tag{7}$$

where Z_i indicates how often (i.e., the number of times) the acoustic signal waveform passes through the horizontal axis (zero level).

While effective, the algorithm presented in [61] has a significant drawback. It uses a fixed set of thresholds which must be obtained by examining a large number of data samples, but the process of data collection can incur significant overhead during deployment. For SleepGuard, we take a different approach to decide the thresholds on a per user per-environment basis; yet, our approach requires significantly fewer data samples compared to [61].

Our heuristics for determining the thresholds work as follows. Since the first and last few audio segments are mostly mute or are background noises, we select the first and last five segments to calculate their short-term energy, which is denoted as E_s and E_e respectively for the first and last segments. The two energy values are then combined to obtain the mean, E_n, as the estimated energy value of noise segments. Let the maximum value of the short-term energy over all segments denoted as $\max(E)$. Then, the average short-term energy, DE, is given as:

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

$$DE = \max(E) - E_n. \tag{9}$$

We can use EH and EL to represent the high and the low thresholds respectively for short-term energy as:

$$EH = \alpha \times DE + E_n, \tag{10}$$

$$EL = \beta \times DE + E_n, \tag{11}$$

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

Here, we need to choose the right values for coefficients 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 night time sound data from our training dataset to determine α and β. Specifically, we tested α with values ranging from 0.1 to 0.5, and β with values ranging from 0.01 to 0.09. We found that setting α to 0.1 and β to 0.06 gives the best overall results in our training dataset.

To minimize the impact of occasionally happened noises, we set the minimum length of the signal segment and measure the duration of the signal when searching for the start and end points. If the length of an acoustic signal is less than the minimum, the signal segment is considered to be a noise segment.

Example. As an example, the two thick vertical lines in Fig. 17 mark a detected acoustic event. The start and end points are used to calculate the duration of an event, as well as the interval between two consecutive acoustic events. To measure the frequency, we count the number of peak points of an acoustic event signal. By feeding the values of the "duration", "interval" and "frequency", our decision-tree-based acoustic detection algorithm can then predict which type of events an acoustic signal corresponds to.

2.5 Tracking Illumination Conditions

Studies have shown that there is a significant interaction between the illuminance level and the mental state of the individual [71]. For example, the bright light can counteract subjective fatigue during the daytime, but at night it will seriously disrupt sleep. 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 a previous study [21], 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. Therefore, the illumination condition in a sleeping environment has a significant impact on sleep.

SleepGuard uses the ambient light sensor to detect the illumination condition during sleep. We visit the bedroom of our ten pilot-study users at night and use 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 the 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. The key to dealing with this practical challenge 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 the 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.

2.6 Sleep Stage and Quality

Sleep is generally considered as a cyclical physiological process composed of three stages: rapid eye movement (REM) stage (see also Sec. 2.1.3), 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 the 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. Moreover, 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. However, 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, the sleep cycle usually repeats four to six times over a sleep. 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, the dependence between two successive sleep stages still exists and different sleep stages have potential conversion probabilities, which also mentioned in Sleep Hunter [33].

To separate between these states, we build a Hidden Markov Model [40] for identifying the current sleep stage of the user. As the input, i.e., the observed states, we use a series of detected sleep events and the sleep stage sequence is modelled as a hidden state, i.e., obs_t=$(NB(t), NB_M(t), BA(t), NA(t))$ represent the feature vector at the 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

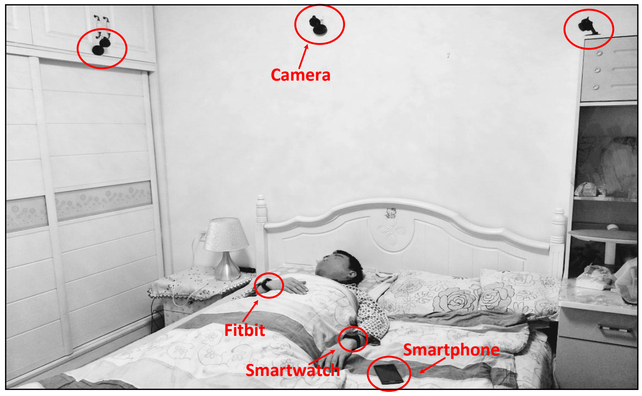


Fig. 18. Experimental setup in one of our participants' home.

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 of the acoustic events. The $states_t$ ={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 use nocturnal sleep data from 10 volunteers who participated in the pilot study. We use the sleep-related events of the 10 volunteers as observation sequences and the corresponding sleep stages detected by Fitbit as hidden state sequences, to generate HMM models. Specifically, we first use the maximum likelihood estimation method for parameter estimation, the state transition matrix and the confusion matrix; we then use the Viterbi algorithm [10] to acquire a series of implicit state sequences corresponding to the observed sequences. 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 sleep quality, we use the Sleep Quality Report model introduced in [33]. Let SQ be the value of the sleep quality, then SQ is given as follow:

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

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

3 EVALUATION METHODOLOGY

3.1 Pilot Study: Training Data

Prior to our main study, we carried out a pilot study that was used to inform our algorithm design, and to provide training data for the algorithms integrated into SleepGuard. Our pilot study consisted of two groups (aged ranged from 15 to 60). One group consisted of randomly selected 100 volunteers to conduct questionnaire surveys to provide the basis for our algorithm design. The other group consisted of 10 users whose data was used to train our models.

To improve the effectiveness of the algorithms integrated into SleepGuard, especially sleep posture and hand position detection, we elicited questionnaires to 100 volunteers to identify their common sleep posture. The main content of the questionnaire was about their common arm position in the four basic sleeping postures. Based on this investigation and previous research [24, 52], we selected the positions to consider in SleepGuard. We also found that these arm positions are representative during the training and testing of our algorithms. In addition, we also survey the extent to which the 100 volunteers are interested in all detected events (see Table 1) in SleepGuard.

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 are different from those taking part in our evaluation (Sec. 3.2).

3.2 Evaluation Setup

Participants. 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 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. 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.

Setup. 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 accurately capture the sleep activity in dark.

Data collection and annotation. The recorded video footage was manually labeled with different sleep activities and the labels were used as ground truth in our evaluation. Specifically, we consider the respiratory amplitude during the NREM stage as large amplitude, and the one during REM stage as normal amplitude. For the acquisition of sleep stage information, we confirm labels when both Fitbit and SleepGuard reach a consensus. To demonstrate the overall benefits of SleepGuard and the events captured by it, we separately collected ground truth information 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 [16]. The results of the user survey are presented at Sec. 4.3. Finally, we collected the sleep stage estimations given by a Fitbit Charge2 and use them as ground truth for sleep stage estimation. While the performance of Fitbit is not comparable to medical grade PSG[1], it has been shown to have a good association in adults [26, 31, 50, 55], especially in estimating REM and light sleep stages.

Competitive schemes. In addition to Fitbit Charge2, we also compare our approach against Sleep Hunter [33], a state-of-the-art mobile-based sleep monitoring approach, and a smartphone-based sleep monitoring app named Sleep as Android [9]. The former app is designed to estimate sleep time and assess sleep by recording the state of motions and the number of body exercises. The latter focuses on

[1]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 a much richer set of sleep information.

estimating sleep stages and evaluate the sleep quality using the tracked sleep-related events. 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.

3.3 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, the light and the orientation sensors. To reduce the energy consumption of the smartwatch, in the experiments we analyze 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 30 milliseconds, which was chosen to balance between information quality and energy consumption. SleepGuard starts tracking sleep events when it detects that the light is off (which can also be triggered by the user during daytime) and there has been nobody 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 EXPERIMENTAL RESULTS

In this section, we first evaluate the performance of SleepGuard in detecting individual sleep-related events (Sec. 4.1). We then compare SleepGuard against alternative schemes for sleep stage detection (Sec. 4.2). Finally, we demonstrate the usefulness of SleepGuard in helping users to understand and improve their sleep through user study (Sec. 4.3).

4.1 Evaluation of Event Detection

In this sub-section, we evaluate the accuracy of SleepGuard on detecting five sleep events: the body posture, the body rollover, the hand position, micro-body movements and acoustic events. Unless otherwise stated, the ground truth of an event is obtained by watching or listening to the video/audio footage.

4.1.1 Sleep Posture Detection. Fig. 19 reports the sleep posture detection accuracy for individual users. SleepGuard achieves a consistent detection accuracy of over 90% across testing users and postures. The confusion matrix in Table 2 shows that SleepGuard only misclassifies a posture on a few occasions, leading to an overall detection precision of over 96% across users. This performance is better than the one reported by SleepMonitor [72]. In particular, SleepGuard gives an improvement of 5% for the prone state and has an overall lower false positive rate. Furthermore, SleepGuard is able to detect more hand positions during sleeping when compared to SleepMonitor. In terms of errors, due to angular characteristics of acceleration being similar between the supine posture where the hand is put on the head and the left-lateral posture, a small amount of the supine postures are misclassified as the left lateral. From the results, we can also observe that the total number of detected prone postures is smaller than that for other postures. This suggests that our testing users are not used to sleep in this position, because it is neither healthy nor comfortable.

4.1.2 Body Rollover Counting. Table 3 reports the performance of body rollover counting. Three of our participants, users 3, 4 and 13, have an unusually high number of rollovers. For users 3 and 4, they have difficulties in falling asleep due to the sleep disorder, while user 13 needs to rollover frequently because of his loudly snoring. As we will demonstrate in Sec. 4.3, these participants also suffered from poor sleep quality and hence indicate how the information extracted by SleepGuard can support the detection of

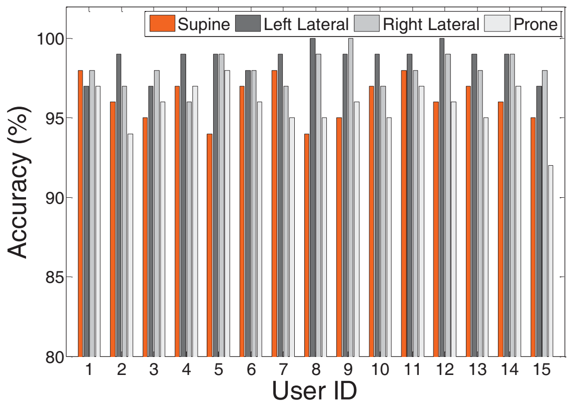


Fig. 19. Detection accuracy of body postures.

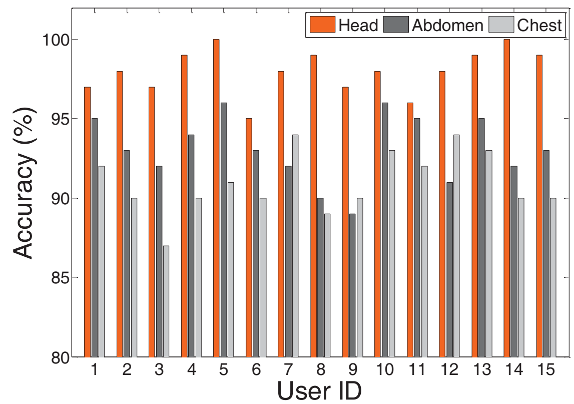


Fig. 20. Identification accuracy of hand positions.

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%	

Table 3. Detection accuracy of body rollovers.

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%

sleep problems. For all the 15 users, the detection accuracies are all high with the lowest accuracy of 87%. Therefore, SleepGuard can accurately distinguish large hand movements from body rollovers. We stress the small detection errors in body rollover events do not have a significant impact on our end result. This is because for sleep stage detection, we consider multiple events in addition to body rollover, including micro-body movement and acoustic events, which can offset the detection errors for body rollover.

4.1.3 Hand Position Recognition. Recall that we combine the tracked hand movement trajectory and periodic signals caused by respiration to identify if the hand is placed on the chest, abdomen or head. In our dataset, 14%, 36% and 22% of the time the hand in the supine posture during sleep were placed on the head, abdomen, and chest respectively. Fig. 20 illustrates the accuracy of hand position across 15 users. As we can see that using just one single user's data for training, our approach achieves an accuracy of over 87% across different testing users. Moreover, we find that at least 4 out of our 15 participants tended to put their hands on their heads (which often lead to bad sleep) and one participate unconsciously

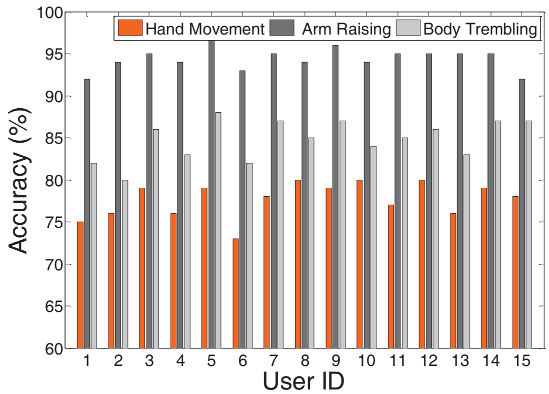


Fig. 21. Micro-body movement detection for each user.

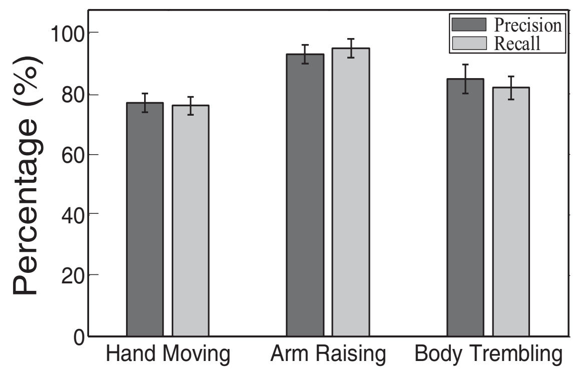


Fig. 22. Average precision and recall for micro-body movement detection. The min-max bars show the range of achieved precision and recall across 15 participants.

put his hand on the chest (which is likely to cause nightmares). By identifying these hand positions, SleepGuard can inform users to change their sleeping behaviors.

4.1.4 Micro-body Movement Detection. In this experiment, in addition to manually labeling the ground truth from the video footage, we also use the accelerometer in the smartphone placed on the bed to record the occurrence of micro-body movements - hand movements, arm raising and body trembling. Using the smartphone data to obtain the ground truth prevents us from missing some movements such as trembling concealed by the duvet (which may not be visible from the video). To gather ground truths from the smartphone data, we first smooth the collected acceleration data along the three axes. Next, we calculate the Root Sum Square (RSS) value to merge the results across axes to obtain the first-order derivative of the merged acceleration data (see Sec. 2.3). Then, we use a simple threshold-based approach to mark body trembling in smartphone data without using the method in Sec. 2.3 to classify these movements, because the body trembling is the easiest to ignore in video. The marked smartphone data is then used together with the manually labeled micro-body movements as the ground truths.

Table 4 lists the total number of the micro-body movements for each participant over the testing period of 14 days, and Fig. 21 reports the accuracy of SleepGuard for detecting these micro-body movements. As can be seen from the table and figure, SleepGuard delivers consistently good precision across our testing users. Figure 22 reports the averaged accuracy across users. The averaged accuracies for arm raising and body trembling are higher, 93% and 84% respectively. While we achieve the lowest accuracy for detecting hand movements, the average procession and recall still exceed 78%. The accuracy for detecting hand movements and body trembling can be further improved by using more training data collected from a longer period of time. It is to note that the purpose of micro-body movement detection is to detect different sleep stages. Since hand movements usually appear in all sleep stages, they offer little information gain for identifying sleep stages compared to other micro-body movements. As a result, the relatively low detection accuracy of hand movements has little impact on sleep stage detection.

4.1.5 Acoustic Events Detection. Table 5 shows the results for acoustic event detection across 15 participants. Overall, SleepGuard is highly accurate in detecting acoustic events, with an average accuracy of over 88%. The precision for detecting coughs is 88.9%, which is slightly lower than for the other three event types. The reason is that different users have distinct cough patterns but the pre-defined parameters

Table 4. The number of micro-body movements per user.

Testing User ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Labeled #hand movement	52	49	67	55	78	65	59	70	61	53	81	55	60	59	63
Labeled #arm raising	48	50	62	53	66	49	57	50	73	45	54	69	57	56	61
Labeled #body trembling	28	32	25	29	34	25	20	30	24	26	27	35	24	22	25

Table 5. The confusion matrix of acoustic events detection.

	Result	Prediction				Recall
		Snore	Cough	Somniloquy	Other	
Ground truth	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%	

Table 6. The confusion matrix of sleep stage detection.

	Result	Prediction			Recall
		REM	Light Sleep	Deep Sleep	
Ground truth	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%	

in the detection model do not include all patterns. For examples, some people have a fast and continuous pattern of coughing, while others have a slower intermittent pattern. Our model for cough detection is trained on 120 sets of nighttime sound data collected from users who are different from our testing users. To further improve accuracy, we can either specialize in our model using data collected from the target user or using a larger and more diverse set of training data. Nonetheless, SleepGuard is able to detect acoustic events with a high accuracy.

4.2 Evaluation of Sleep Stage Detection

In order to prove that SleepGuard is able to identify the sleep stages but not just the user's sleep habits, we use the results given by Fitbit Charge2 as the ground truth. In this evaluation, we randomly choose a total of 50 sets of sleep data for 15 participants across two weeks, yielding at least 3 sets per participant.

SleepGuard uses event-driven methods to detect changes in the sleep stage. To identify sleep stages, SleepGuard uses sleep events detected within a 15-minute window. The window starts from the first detected event. If no event was detected within 15 minutes, it assumes the sleep stage remains unchanged.

Table 7. Effect of respiratory amplitude estimation.

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

Table 8. Performance of sleep stage detection comparison.

	Light Sleep		Deep Sleep	
Stage / System	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	66.7%	66.1%	60.0%	50.7%

4.2.1 Sleep Stage Detection. The averaged precision and recall for sleep stage detection are given in Table 6. Overall, SleepGuard is able to correctly identify over 60% of the sleep stages. While SleepGuard can mispredict between the light sleep and the REM stages, the overall performance is not far from FitBit. Moreover, as we will demonstrate later, the main benefits of SleepGuard are its capability to capture a wide range of sleep events and how do they correlate to sleep quality, but not for accurate sleep stage detection (where medical PSG measurements will be required).

4.2.2 Impact of Respiratory Amplitudes on Sleep Stage Detection. When detecting different sleep stages, SleepGuard also considers the respiratory amplitude when the hand is put on the abdomen or chest. To assess the effectiveness of respiratory amplitude estimation, we evaluate the performance of the sleep stage detection with and without taking the respiration amplitude into account. The performance of sleep stage detection is shown in Table 7. For the three different sleep stages, both precision and recall are improved between 5% to 10% with the help of respiratory amplitude estimation. An alternative to our approach is to use the respiratory frequency as a feature for sleep stage detection. We found that this alternative scheme gives similar performance for sleep stage detection as our approach. This is because the respiratory amplitude is strongly correlated with the respiratory frequency – when the respiratory amplitude is larger, the time taken for one breath will be longer, and consequently, the frequency of breathing will be slower.

4.2.3 Compare to Existing Approaches. We now compare SleepGuard against Sleep As Android and Sleep Hunter [33]. 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 8 shows the detection results. As we can see, SleepGuard significantly outperforms Sleep As Android by improving the accuracy by two folds. Compared to Sleep Hunter, SleepGuard provides a similar accuracy for detecting light sleep stages, but better performance for detecting deep sleep stages. The performance advantage of SleepGuard comes from the incorporation of rich and complicated sleep events.

Porting Smartphone Algorithms to Smartwatches. We also implement the algorithms employed by Sleep Hunter and apply them to the data collected using our smartwatch. This experiment allows us to check if the better performance of SleepGuard is due to the use of a smartwatch instead of a mobile phone. For body movement detection, applying the algorithms employed by Sleep Hunter to our smartwatch data gives a comparable accuracy of around 96% when the system only identifies between drastic and

Table 9. Acoustic event detection on smartwatch data.

System \ Event	Snore		Cough		Somniloquy	
	Precision	Recall	Precision	Recall	Precision	Recall
SleepGuard	96.9%	91.4%	88.9%	90.1%	91.3%	89.4%
Sleep Hunter	71.0%	73.0%	71.0%	63.0%	89.0%	83.5%

Table 10. Sleep stage detection on smartwatch data.

System \ Stage	REM		Light Sleep		Deep Sleep	
	Precision	Recall	Precision	Recall	Precision	Recall
SleepGuard	71.0%	70.0%	66.5%	69.6%	63.3%	59.8%
Sleep Hunter	61.2%	67.6%	61.5%	60.2%	56.8%	47.9%

small body movements. However, the Sleep Hunter algorithms are less effective for detecting finer-grained body movements. SleepGuard outperforms Sleep Hunter by delivering an accuracy of around 90% for detecting body rollovers (see Table 3) and an accuracy for detecting micro-body movements of over 78% (see Fig. 22). For acoustic events, we apply the Sleep Hunter algorithms to detect snore, cough and somniloquy. The results in Table 9 suggest that SleepGuard gives better performance over Sleep Hunter in detecting these acoustic events. Finally, we apply the sleep stage detection model used by Sleep Hunter to combine sleep-related events to identify sleep stages. The results are shown in Table 10. Again, SleepGuard outperforms the Sleep Hunter model with a higher accuracy and recall across different types of sleep stages.

This experiment confirms that the algorithms used by Sleep Hunter for identifying sleep events and stages are not tuned for the smartwatch. Compared to Sleep Hunter, SleepGuard can detect sleep events and stages with a higher accuracy using a set of carefully designed methods to target smartwatches.

4.3 User Study

4.3.1 Methodology. To understand and verify how the additional information captured by SleepGuard supports users, at the end of the experiments the participants are asked to participate in a survey. The survey asks 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 (see Sec. 3.2). We considered two groups of users in our survey. As the main source of information, we consider the 15 participants in 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 the 100 external users participated in our pilot study (see Sec. 3.1). The external participants were asked about their interested 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 neighboring scores into one scale and eventually divide the total scores into four levels, recorded as 0, 1, 2 and 3, representing poor, general,

Table 11. Results of sleep quality assessments. The first three rows show the sleep quality scores of the different systems (mean and standard deviation) for each user across 14 days, whereas the last two rows compare sleep quality labels between subjective assessments and those returned by SleepGuard and FitBit.

User ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
SleepGuard	3 (1.8)	3 (1.5)	0 (1.0)	1 (1.8)	2 (1.2)	2 (1.7)	3 (1.0)	0 (1.3)	2 (1.3)	2 (1.0)	2 (0)	2 (1.5)	1 (1.8)	0 (1.0)	2 (1.3)
Fitbit	3 (1.7)	3 (2.2)	0 (1.3)	0 (1.8)	1 (2.2)	3 (1.8)	2 (1.9)	3 (1.8)	2 (1.7)	2 (1.9)	2 (1.3)	2 (2.0)	2 (2.3)	1 (1.8)	2 (1.7)
User score	3 (0.0)	2 (1.4)	0 (0.0)	0 (1.7)	2 (1.7)	2 (1.0)	3 (1.0)	1 (1.7)	1 (1.4)	2 (1.0)	2 (1.0)	3 (2.0)	0 (1.8)	0 (1.0)	2 (1.0)
SleepGuard	P	O	P	O	P	P	P	O	O	P	P	O	O	P	P
Fitbit	P	O	P	P	O	O	O	**B**	O	P	P	O	**B**	O	P

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.

4.3.2 Results. In Table 11, we show the mean sleep quality score across 14 days given by each user together with the estimations produced by SleepGuard and Fitbit. We also give the standard deviation of the scores across the 14-day period per user. This number is given in the brackets next to the mean score. The last two rows in Table 11 compare the estimation given by SleepGuard and Fitbit against the user self-rating score. In these two specific rows, a label of 'P' indicates an estimated score perfectly matches the user score, a label of 'O' means the estimation error is within one scale point (for example, SleepGuard rates the user sleep to be excellent while the user's self-rating is good), and a label of 'B' indicates the estimation error is greater than a scale point. As can be seen from the table, the estimation given by SleepGuard is more likely to match the user's self-rating compared to Fitbit (as indicated by having more 'P' labels - 9 vs 6) and, unlike Fitbit, the estimation error given by SleepGuard is never greater than one scale point. To further validate this, we calculated the Spearman ρ-correlation [65] between the user scores and each of the two systems, SleepGuard and Fitbit. SleepGuard provides higher correlation coefficient ($\rho = 0.842$) than Fitbit ($\rho = 0.500$). The difference in correlation was found statistically significant using a one-tailed test carried out through a Fisher r-z transformation ($Z = 1.66, p < 0.05$). In summary, SleepGuard thus gives a better sleep quality assessment compared to Fitbit in our evaluation.

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 analyze 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 11). 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 of the 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 the head before sleeping. The 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 the higher risk on nightmares [47], 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 the 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 highlight 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 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 DISCUSSIONS

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. In our experiment, we were unable to directly compare SleepGuard against medical grade solutions for sleep stage detection due to lack of suitable clinical environments and expertise. However, we would expect polysomnography to provide better sleep stage detection performance. This is because polysomnography 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 an 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. Nonetheless, in future work, we strive to perform further studies and compare SleepGuard against a polysomnography-based solution.

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 the 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 an 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 canceled through calibration. This is where we need to measure and consider in our future work. Alternatively, multi-sensor designs, such as a combination of smartwatch 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 smartwatch was 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 the smartwatch, Unlike the 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, e.g., 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 the 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 degrade the performance of sleep posture detection. For the 15 participants, 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

There is an extensive body work on sleep monitoring and tracking [12, 13, 36, 41, 44, 59]. We summarize some of the most relevant work in this section.

6.1 Medical Grade Sleep Monitoring Solutions

Traditionally, the dedicated medical technologies, like EEG, ECG and EMG [67], 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, 44, 54] monitors the brain waves, and then recognizes the sleep stages by leveraging unsupervised learning approaches. Although a high accuracy can be achieved by those technologies, they have two drawbacks. First, those technologies require the dedicated medical devices, which are expensive compared the widely available smartwatch or smartphone. Second, they require the users to be attached many sensors on the human body, which may cause a healthy person hard 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 more cost effective. Second, the smartwatch has little disruption to a user's normal sleep, thus we can monitor the user's sleep quality more precisely.

6.2 Smartphone-Based Approaches

Numerous approaches have been proposed to exploit the use of a smartphone for sleep monitoring. iSleep [36] measures the sleep quality by recording sleep-related acoustic events and evaluates the sleep quality using the Pittsburgh Sleep Quality Index (PSQI) [17]. Bai *et al.* [12] use a wide range of sensor data captured by the smartphone sensors, including the accelerometer, gyroscope, and microphone, to predict a user's sleep quality. The work presented in [41] leverages the smartphone sensors to record the sleep disruptor for a user, while the work presented in [22] explores a series of opportunities to support healthy sleep behaviors. Other approaches predict the sleep quality by leveraging the smartphone to monitor the external factors, such as the daily activity, the sleeping environment and location, and family settings [20, 77]. In addition to the aforementioned approaches, there is a wide range of smartphone base sleep monitoring applications. Examples of such applications include Sleep As Android [9], Sleep Journal [11], and YawnLog [76].

The aforementioned smartphone-based systems, however, require placing the smartphone at a specifical location near to the user, which may not always be feasible. For example, the work presented in [33] requires the smartphone to be placed next to the user's head, and to remain stationary throughout the sleeping process. Such a constraint is hard to satisfy because of body movements during the sleep. Furthermore, prior research also shows that many users do not want to place their mobile phone too close to the body due to health risk concerns [37, 63].

Unlike existing smartphone-based solutions, SleepGuard uses the commodity smartwatch for sleep monitoring. Since many users are willing to wear a smartwatch throughout the sleep, the smartwatch can remain relatively close to the user body. This allows us to collect a wider range of sleep-relevant events with a higher accuracy. This richer set of data thus leads to better sleep monitoring and quality assessment.

6.3 Wearable-Device-Based Approaches

Some of the more recent works have exploited smartwatch or wearable-wrist for sleep monitoring [13, 14, 18, 59]. The Sleeptracker [70] uses the accelerometer data to infer the user's sleep stages. The ubiSleep [59] joints heart rate, accelerometer, and sound signals collected from the smartwatch for sleep

Table 12. Compare the supporting functionalities of mobile based sleep monitoring systems.

	Heart rate	Acoustic events	Postures	Body movements	Hand positions	Sleep stages
SleepGuard		✓	✓	✓	✓	✓
Sleep as Andriod		✓				✓
Sleep Hunter		✓		✓		✓
Sleep Monitor			✓			
Sleeptracker	✓					✓
isleep		✓		✓		
Fitbit	✓					✓
Jawbone						✓
ubiSleep	✓	✓				

monitoring. The Zeo [18] uses a smartwatch together a headband to monitor sleep. The data collected by the headband helps Zeo to achieve good tracking precisions, but wearing a headband can disrupt a user's normal sleep.

The majority of the prior smartwatch-based sleep monitoring systems only gather coarse-grained information but do not provide analysis to help users to understand the correlation between sleep quality and sleep-related events. For examples, many smartwatch Apps, including Jawbone Up [74], FitBit [32], YawnLog [76] and WakeMate [75], only provide the result of sleep quality assessments but there is little information on how the sleep quality is evaluated. Without such information, it is difficult for a user to understand the root cause of poor sleep. Therefore, these systems offer little help in assisting users to improve their sleep. In addition to sleep monitoring, other works use wearable devices to detect specific sleep events, such as roll-over [49], or body acceleration and sleep position changes [64]. However, they only support a limited set of sleep events and thus just scratch the surface of what could be possibly done on smartwatches.

In Table 12, we compare SleepGuard's functionalities with eight other existing sleep monitoring systems for mobile and wearable devices. As can be seen from the table, SleepGuard supports the detection for the largest number of sleep-related events, and it does so without requiring additional hardware. We have shown that this extensive set of sleep events not only leads to more accurate sleep quality assessments, but also enables users to better understand the cause of poor sleep.

6.4 Summary of Prior Work

To summarize, as we can see from Table 13, the advantage of SleepGuard is that it can detect more fine-grained sleep-related events to obtain more abundant sleep information, for which the current commercial or scientific research sleep monitoring system cannot be achieved. Furthermore, 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 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. Compared with some medical grade technologies like PSG, our advantages are inexpensive and easy to deploy at home, so it is suitable for the most general public. Moreover, since it does not need a large number of instruments attached to the user's body, SleepGuard 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. We want to stress that our work 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.

Table 13. Summary of existing solutions.

System	High accuracy	Practicability	Low disruptive	Low cost	Informativeness	Interpretability
PSG	✓				✓	
Smartphone		✓	✓	✓		✓
Wearable		✓	✓	✓		✓
SleepGuard		✓	✓	✓	✓	✓

7 CONCLUSIONS

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 compared to state-of-the-art consumer-grade sleep monitors, while at the same time being able to accurately capture a richer 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 behaviors surrounding sleep. We also compared the sleep quality estimates of SleepGuard against subjective self-assessments, demonstrating a high degree of correspondence.

ACKNOWLEDGMENTS

This work was partially supported by the National Natural Science Foundation of China under grant nos. 61672428, 61772422 and 61602382; and the ShaanXi Science and Technology Innovation Team Support Project under grant agreement 2018TD-O26; and the UK Engineering and Physical Sciences Research Council (EPSRC) under Grant agreements EP/M01567X/1 (SANDeRs) and EP/M015793/1 (DIVIDEND); and the Royal Society International Collaboration Grant (IE161012). Xiaojiang Chen and Zheng Wang are co-corresponding faculty authors.

REFERENCES

[1] 2018. https://en.wikipedia.org/wiki/Respiratory_rate. (2018).

[2] Hanoglu L Agargun M Y, Boysan M. 2004. Sleeping Position, Dream Emotions,and Subjective Sleep Quality. *Sleep & Hypnosis* 6, 6 (2004), 8–13.

[3] Torbjörn Åkerstedt. 2007. Altered sleep/wake patterns and mental performance. *Physiology & behavior* 90, 2-3 (2007), 209–218.

[4] Adriano M. Alencar, Diego Greatti Vaz da Silva, Carolina Beatriz Oliveira, Andr 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.

[5] Paula Alhola and Päivi Polo-Kantola. 2007. Sleep deprivation: Impact on cognitive performance. *Neuropsychiatric Disease and Treatment* 3, 5 (2007), 553567.

[6] Ellemarije Altena et al. 2008. Sleep loss affects vigilance: effects of chronic insomnia and sleep therapy. *Journal of sleep research* (2008).

[7] Amy. 2014. http://www.abc.net.au/science/articles/2014/01/28/3861788.htm. (2014).

[8] 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.

[9] Sleep As Android. 2017. https://play.google.com/store/apps/details?id=com.urbandroid.sleep&hl=zhEN. (2017).

[10] Xavier Anguera, Simon Bozonnet, Nicholas Evans, Corinne Fredouille, Gerald Friedland, and Oriol Vinyals. 2012. Speaker diarization: A review of recent research. *IEEE Transactions on Audio, Speech, and Language Processing* 20, 2 (2012), 356–370.

[11] Sleep Journal App. 2017. http://sleeptrackerapp.blogspot.com/. (2017).

[12] Yin Bai, Bin Xu, Yuanchao Ma, Guodong Sun, and Yu Zhao. 2012. Will you have a good sleep tonight? sleep quality prediction with mobile phone. In *Proceedings of 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] MH Bonnet and DL Arand. 2003. Insomnia, metabolic rate and sleep restoration. *Journal of Internal Medicine* 254, 1 (2003), 23–31.

[15] 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 ACM SIGHIT on International Health Informatics Symposium*. ACM, 71–80.

[16] 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.

[17] 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.

[18] John N Caviness. 1996. Myoclonus. In *Mayo Clinic Proceedings*, Vol. 71. 679–688.

[19] 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).

[20] 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*. 145–152.

[21] 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.

[22] 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 on Human Factors in Computing Systems*. 3053–3062.

[23] CNN. 2015. The great American sleep recession. https://edition.cnn.com/2015/02/18/health/great-sleep-recession/index.html.

[24] Sarah Cummings. 2017. https://www.sleepadvisor.org/sleep-position-and-personality/. (2017).

[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 *Proceedings of the IEEE on Engineering in Medicine and Biology Society*. 1151–1154.

[30] Nathaniel A Eiseman, 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* 8, 06 (2012), 655–666.

[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] 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.

[34] T Matsui H Miwa, S Sasahara. 2007. Roll-over detection and sleep quality measurement using a wearable sensor. (2007), 1507.

[35] Marian Haescher, Denys JC Matthies, and Bodo Urban. 2015. Anomaly Detection with Smartwatches as an Opportunity for Implicit Interaction. In *Proceedings of Human-Computer Interaction with Mobile Devices and Services Adjunct*. ACM, 955–958.

[36] Tian Hao, Guoliang Xing, and Gang Zhou. 2013. iSleep: unobtrusive sleep quality monitoring using smartphones. In *Proceedings of the IEEE on Embedded Networked Sensor Systems*. 4.
[37] Step To Health. 2017. https://steptohealth.com/bad-sleep-near-cell-phone/. (2017).
[38] 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.
[39] 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.
[40] 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.
[41] 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 ACM on Ubiquitous Computing*. 226–234.
[42] DailyBurn Kristen Domonell. 2016. http://edition.cnn.com/2016/03/18/health/sleep-positions-good-bad/index.html. (2016).
[43] Tea Lallukka and Erkki Kronholm. 2016. The contribution of sleep quality and quantity to public health and work ability. (2016).
[44] Martin Langkvist, Lars Karlsson, and Amy Loutfi. 2012. Sleep stage classification using unsupervised feature learning. *Advances in Artificial Neural Systems* 2012 (2012), 107046.
[45] Fonseca P Long X, Foussier J. 2014. Analyzing respiratory effort amplitude for automated sleep stage classification. *Biomedical Signal Processing & Control* 14 (2014), 197–205.
[46] 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).
[47] BIANCA LONDON FOR MAILONLINE. 2016. http://www.dailymail.co.uk/femail/article-3384201/Do-sleep-like-position-lead-nightmares.html?jumpTo=comment-111287510&max=100&offset=181&reply=111287510. (2016).
[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] Matsui T Miwa H, Sasahara S I. 2007. Roll-over Detection and Sleep Quality Measurement using a Wearable Sensor. (2007), 1507–1510.
[50] 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.
[51] Alain Muzet. 2007. Environmental noise, sleep and health. *Sleep medicine reviews* 11, 2 (2007), 135–142.
[52] Neil. 2014. http://www.oddmattress.co.uk/sleeping-positions/. (2014).
[53] 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.
[54] Edgar Oropesa, Hans L Cycon, and Marc Jobert. 1999. Sleep stage classification using wavelet transform and neural network. *International Computer Science Institute* (1999).
[55] Beth Osterbauer, Jeffrey A Koempel, Sally L Davidson Ward, Laurel M Fisher, and Debra M Don. 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.
[56] Pickett C K ouglas N J, White D P. 1982. Respiration during sleep in normal man. *Thorax* 37, 11 (1982), 840.
[57] Tiina Paunio, Tellervo Korhonen, Christer Hublin, Markku Partinen, Mika Kivimäki, Markku Koskenvuo, and Jaakko Kaprio. 2008. Longitudinal study on poor sleep and life dissatisfaction in a nationwide cohort of twins. *American Journal of Epidemiology* 169, 2 (2008), 206–213.
[58] 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.
[59] Nuno Pombo and Nuno M Garcia. 2016. ubiSleep: An ubiquitous sensor system for sleep monitoring. In *Proceedings of the IEEE on Wireless and Mobile Computing, Networking and Communications*. 1–4.
[60] Zeo Sleep Manager Pro. 2017. http://www.digifit.com/Zeo/. (2017).
[61] Bai Y Qingqi L U. 2012. A Speech Endpoint Detection Algorithm Based on Dual-threshold Two Sentences. *Electronic Science & Technology* (2012).
[62] J Ross Quinlan. 2014. *C4. 5: programs for machine learning*. Elsevier.

[63] Quora. 2017. https://www.quora.com/Does-keeping-your-mobile-phone-near-your-head-affect-\your-sleeping-cycle-How. (2017).

[64] Parthasarathy S Razjouyan J, Lee H et al. 2017. Improving Sleep Quality Assessment Using Wearable Sensors by Including Information From Postural/Sleep Position Changes and Body Acceleration: A Comparison of Chest-Worn Sensors, Wrist Actigraphy, and Polysomnography. *Journal of Clinical Sleep Medicine* 13, 11 (2017), 1301–1310.

[65] Alice M Richardson. 2015. Nonparametric Statistics: A Step-by-Step Approach. *International Statistical Review* 83, 1 (2015), 163–164.

[66] Earl Sageron. 2016. http://www.linkedin.com/pulse/effect-chin-strap-snoring-earl-sageron. (2016).

[67] Clifford B Saper, Thomas E Scammell, and Jun Lu. 2005. Hypothalamic regulation of sleep and circadian rhythms. *Nature* 437, 7063 (2005), 1257–1263.

[68] 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.

[69] Jenn Sinrich. 2017. https://www.fitnessmagazine.com/health/sleep/the-best-and-worst-sleeping-positions-for-your-health/. (2017).

[70] Sleeptracker. 2017. http://www.sleeptracker.com/. (2017).

[71] 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.

[72] Xiao Sun, Li Qiu, Yibo Wu, Yeming Tang, and Guohong Cao. 2017. SleepMonitor: Monitoring Respiratory Rate and Body Position During Sleep Using Smartwatch. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* (2017).

[73] 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* (2015).

[74] Jawbone Up. 2017. https://jawbone.com/up. (2017).

[75] WakeMate. 2017. http://www.wakemate.com. (2017).

[76] YawnLog. 2017. http://beta.yawnlog.com. (2017).

[77] 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 *Proceedings of the IEEE on Information Processing in Sensor Networks*. 179–190.

Received February 2018; revised May 2018; accepted July 2018