

# A Lightweight And Inexpensive In-ear Sensing System For Automatic Whole-night Sleep Stage Monitoring

Anh Nguyen<sup>†</sup>, Raghda Alqurashi<sup>†</sup>, Zohreh Raghebi<sup>††</sup>  
Farnoush Banaei-kashani<sup>††</sup>, Ann C. Halbower<sup>‡</sup>, and Tam Vu<sup>†</sup>  
<sup>†</sup>Mobile and Networked Systems Lab, University of Colorado, Denver  
<sup>††</sup>Big Data Management and Mining Lab, University of Colorado, Denver  
<sup>‡</sup>University of Colorado, School of Medicine  
<sup>†</sup>{anh.t4.nguyen, raghda.alqurashi, tam.vu}@ucdenver.edu,  
<sup>††</sup>{zohreh.raghebi, farnoush.banaei-kashani}@ucdenver.edu,  
<sup>‡</sup>ann.halbower@childrenscolorado.org

## ABSTRACT

This paper introduces *LIBS*, a light-weight and inexpensive wearable sensing system, that can capture electrical activities of human brain, eyes, and facial muscles with two pairs of custom-built flexible electrodes each of which is embedded on an off-the-shelf foam earplug. A supervised non-negative matrix factorization algorithm to adaptively analyze and extract these bioelectrical signals from a single mixed in-ear channel collected by the sensor is also proposed. While *LIBS* can enable a wide class of low-cost self-care, human computer interaction, and health monitoring applications, we demonstrate its medical potential by developing an autonomous whole-night sleep staging system utilizing *LIBS*'s outputs. We constructed a hardware prototype from off-the-shelf electronic components and used it to conduct 38 hours of sleep studies on 8 participants over a period of 30 days. Our evaluation results show that *LIBS* can monitor biosignals representing brain activities, eye movements, and muscle contractions with excellent fidelity such that it can be used for sleep stage classification with an average of more than 95% accuracy.

## CCS Concepts

•Human-centered computing → Mobile devices;  
•Computing methodologies → Supervised learning by classification; Non-negative matrix factorization;

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

*SenSys '16, November 14-16, 2016, Stanford, CA, USA*

© 2016 ACM. ISBN 978-1-4503-4263-6/16/11...\$15.00

DOI: <http://dx.doi.org/10.1145/2994551.2994562>

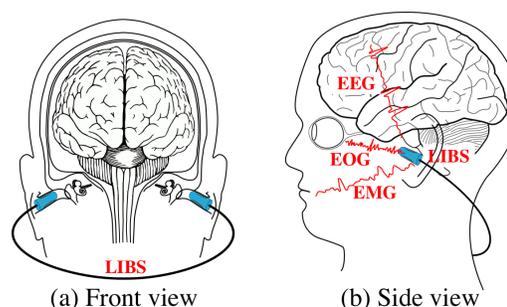


Figure 1: *LIBS* and its relative position to the sources of the three biosignals of interest, EEG, EMG, and EOG.

## Keywords

*LIBS*, In-ear Wearables, Biosignals, Health Care, Sleep Staging

## 1. INTRODUCTION

Sleep occupies nearly a third of human life and acts as a critical daily function to help our body balance and regulate vital systems. Sleep is essential for restorative functions in the brain and is associated with early brain development, learning, memory, and psychological health [55]. Sleep also has the function of stress decrease, hormone balance regulation, cardiovascular stability, and appetite. In other words, losing or skimping on sleep can cause serious harm physically and mentally. In the United States, a number of young adults and adolescents are regularly sleep deprived [40]. Therefore, quantifying sleep quality has significant clinical value in detecting and diagnosing various sleep-related disorders that affects one's health.

Ordinarily, monitoring sleep for clinical reasons requires patients to undergo a sleep study [52] (polysomnography, PSG [45]) in a sleep laboratory in hospitals. During sleep, PSG acquires bioelectrical signals generated by brain activities, eye movements, and muscle contractions using electroencephalography (EEG), electrooculography (EOG), and

electromyography (EMG), respectively, among other auxiliary sensors. While the PSG provides highly reliable sleep study results, it has a number of drawbacks, resulting in decreased utility outside of clinical facilities while remaining very expensive when used in inpatient settings. The issues include (1) the obtrusive attachment of a large number of wired sensors to patient’s head, face, and body, (2) the requirement of travel to a sleep laboratory for expert sensor attachment, (3) the risk of losing sensor contact whenever the patient moves, and (4) the need of a well-trained expert to review a long night sleep staging result.

Recently, with their technological advancement, wearable and mobile devices become a promising hi-tech solution to problems imposed by PSGs. For instance, inertial measurement unit (IMU), as found in many off-the-shelf wearable devices [10, 30, 5, 17], have been utilized for automatic sleep stage tracking through monitoring physical human movement. However, their accuracy remains relatively low since accurate sleep staging requires an access to physiological signals from brain, eyes, and muscles [43]. Addressing such limitations, various head-worn devices such as eye masks [25], headbands [50], and headphones [18] were developed to capture biosignals with a smaller number of electrodes. While the accuracy is improved, such devices still make users uncomfortable when wearing them on forehead, scalp, or face during sleep. As a result, significant effort [11, 23, 13, 21, 28, 29, 48, 4] has been devoted to collect these biosignals at alternative positions of the body. However, none of existing approaches provides a highly comfortable, accurate, and low-cost sleep staging system.

This paper introduces a new low-cost in-ear biosignal sensing system, called *LIBS* as shown in Figure 1, that has the potential to provide vital inputs for a number of healthcare applications. As comfortable as wearing earbuds while listening to music, the *LIBS recorder* designed allows the patient to have very few passive electrodes placed inside the ear for biosignal sensing. Special care was taken to maximize the contact quality between the electrode and user’s outer ear while maintaining a high level of comfort by designing flexible and multi-layered electrodes. Due to the unique location of the ear canal, the signal obtained by our *LIBS* device is a mixture of EEG, EOG, and EMG and unwanted noise. Thus, *LIBS* takes a mixed in-ear signal and adapts a *signal separation model* to extract the three signals of interest without loss of their physiological information. Finally, we apply a *sleep stage classification algorithm* to score every 30-second epoch of sleep data into an appropriate stage using a set of discriminative features extracted from the separated signals. However, not limited to automatic sleep staging, *LIBS* with its three individual biosignal outputs has a potential to become fundamental in divergent healthcare problems including long-term monitoring outside clinical facilities, sleep quality assessment, sleep environment control, brain surgery support, diagnosis of brain-related disease (e.g. sleep disorders, epilepsy), and autonomous audio steering.

Due to an intricate structure of human ear canal as well as unique characteristics of EEG, EOG, and EMG signals, there are three key challenges need to be addressed in this work.

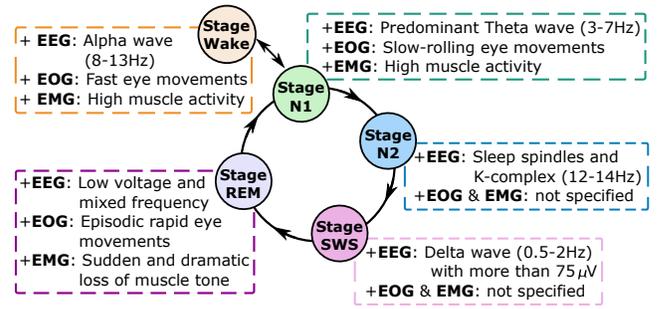


Figure 2: Specification of sleep stages on EEG, EOG, and EMG patterns

First, the anatomy of human head indicates that sources of signals from brain, eye, and facial muscle are not close to the location of the ear canals and generate low amplitude signals. As a result, recording the biosignals from a distance challenges us to develop the *LIBS* sensors capable of not only detecting them but also providing a higher level of comfort while wearing our device during the monitoring period. The second challenge stems from the necessary separation mechanism to extract EEG, EOG, and EMG signals from the single-channel in-ear mixture. It is more challenging because the signals have temporal and spectral characteristics overlapped and their sources can be freely activated at the same time. Finally, the bioelectrical signals acquired vary across people and across recordings due to the displacement of electrodes in different device hookups and the difference of physiological conditions among individuals. Thus, making the separation algorithm robust even with the presence of the variance becomes a significant hurdle.

In addressing the above-mentioned challenges and realizing *LIBS*, we make the following contributions in this work: (1) Augment a low-cost off-the-shelf foam-based earplug-based sensor with a novel design of very sensitive electrodes made of a combination of thin, soft, and highly conductive materials. Thus, *LIBS* itself is a light-weight, low-cost, easily placed device that rests comfortably and safely inside human ears to provide high fidelity and long-term continuous measurement of voltage potential of biosignals.

(2) Derive and implement a single-channel signal separation model, which can decompose EEG, EOG, and EMG signals from the mixed in-ear signals based on their physiologic and electrical properties. In this algorithm, we integrate a learning process to build source-specific prior knowledge for supervising a correct separation and propose a solution to adaptively control the variability of the signals across people and across recordings.

(3) Derive and implement a complete sleep staging system, which takes the three separated biosignals captured by the in-ear device as inputs and then automatically extracts their discriminative features and determines the appropriate sleep stages for the whole sleep duration with the granularity of a 30-second epoch.

## 2. SLEEP STAGING BACKGROUND

There are two states of consciousness: wakefulness and sleep. During sleep, human beings pass through four distinct

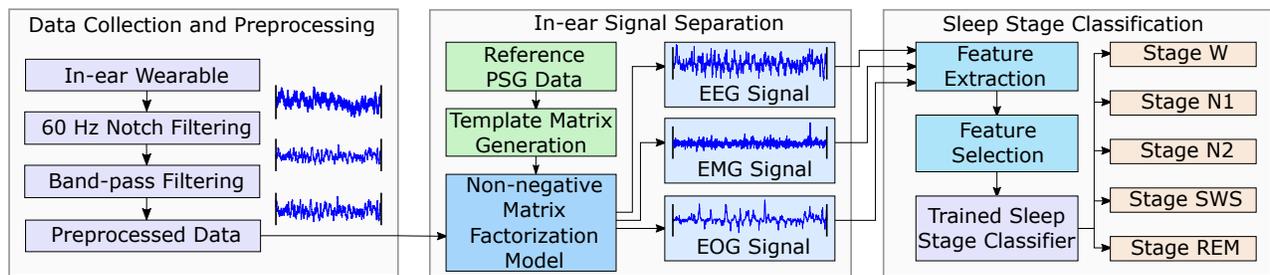


Figure 3: Overall architecture of the proposed automatic whole-night sleep stage monitoring system

stages of sleep: N1, N2, N3, and REM sleep [46, 1], which usually occur in cycles repeatedly. By determining the sleep stages and their distribution during the night, the quantity of sleep can be calculated and the refreshing quality of sleep can be evaluated by examining sleep architecture [51]. Stage N1 in which alpha brain waves begin to disappear provides positive health benefits associated with relaxation and peacefulness during meditation and biofeedback. Stage N3 in which delta brain waves are produced helps the body repair itself by regulating hormones, restoring energy, and restoring emotional health. Finally, the human mind uses REM sleep to stimulate the brain regions used in learning and memorizing, increasing the production of proteins, and affecting certain mental skills such as optimal emotional and social functioning while people are awake [55].

Currently, in clinical sleep study facilities, Polysomnography (PSG) [22] is a conventional monitoring system used as a gold standard to measure a number of body functions during sleep. To study the sleep quantity and quality, the sleep stages are mainly identified by simultaneously evaluating three fundamental measurement modalities including brain activities, eye movements, and muscle contractions [46]. This information is presented as voltage fluctuations derived from ionic current appearing within different areas of the body and measured by surface electrodes placed along suitable positions on the body. Distinguishing between these sleep stages, the EEG measure using electrodes placed around the scalp is essential for interpreting various sleep/wake states of the brain. On the other hand, EMG and EOG using electrodes placed on the skin near the eyes and on the muscles, respectively, are necessary measures in deeply differentiating REM stage from all the other stages and even the wakefulness state. The relationship between the sleep stages and EEG, EOG, and EMG signals is illustrated in Figure 2.

Understanding the importance of EEG, EOG, and EMG in health care, specifically for sleep, and the drawbacks of PSG (cumbersomeness, expensiveness, and easy sensor contact loss), our work outlines an in-ear sensing system *LIBS* that measures and extracts those signals using a light-weight, low cost, precise, and easy to use earplug-based device designed by our team. Its output is then input to an automatic technique that helps classify sleep stages with high accuracy.

### 3. SYSTEM OVERVIEW

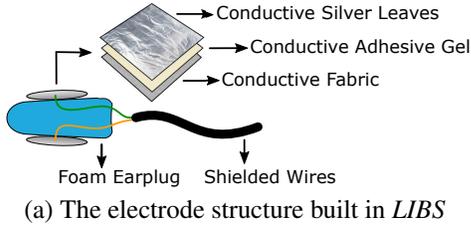
In this section, we generally describe a whole-night sleep stage monitoring system that can capture and separate EEG,

EOG, and EMG from the in-ear biosignal using *LIBS* during sleep and automatically determine different stages of sleep through those signals. The proposed system, as illustrated in Figure 3, consists of three following primary modules.

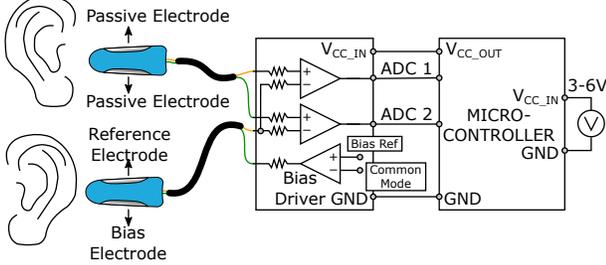
**Data acquisition and preprocessing:** The in-ear wearable recorder in *LIBS* is developed and the in-ear signals captured by *LIBS* is preprocessed to eliminate possible signal interference (e.g. body movement artifact, electrical noise). This module focuses on tackling our hardware challenges that requires (1) an ability to adapt to the small uneven area inside human ear and its easy deformability under the jaw movements (e.g. teeth grinding, chewing, speaking), (2) a potential to acquire the critical biosignals, which are naturally weak (in microVolt amplitude), and (3) a provision of comfortable and harmless wearing to the users. We fulfill these requirements by custom-making a novel design of the wearable deformable earplug using a viscoelastic material with atop sensitive electrodes using thin, soft, and highly conductive materials. To increase signal fidelity, we increase the distance between the electrodes and the reference point. Thus, these solutions allow *LIBS* itself to deform flexibly while still maintaining a good connection between its electrodes and human skin while ensuring its mobility for long-term use. Also, they allow the users to apply *LIBS* by themselves in a similar fashion to using the ear buds without the need of help from others or trained clinical technicians.

**In-ear signal separation:** In this module, a separation algorithm is developed to extract the preprocessed in-ear signal into EEG, EOG, and EMG signals without the loss of their essential nature. This module focuses on overcoming our signal processing challenges that require the ability to deal with (1) overlapping characteristics of three signals in both time and frequency domains, (2) a random activation of the sources generating them, and (3) their variation from person to person and in different recordings. We solve these problems by developing a non-negative matrix factorization (NMF)-based model that can separate the single-channel in-ear mixture into EEG, EOG, and EMG with high similarity to the ground truth given by the gold-standard device. Specifically, our separation algorithm learns prior knowledge of patterns of those biosignals through their individual spectral templates and teaches them to adapt to the variation between people through a deformation step. As a result, the model we build alters itself slightly to return the best fit between the expected biosignals and the given templates.

**Sleep stage classification:** Lastly, this module provides a set of machine learning algorithms to score sleep into differ-



(a) The electrode structure built in *LIBS*



(b) *LIBS* schematic for recording the in-ear signal

Figure 4: A comprehensive design of *LIBS*

ent sleep stages using EEG, EOG, and EMG separated from the recorded signal. Because those signals can have similar characteristics shared in some of stages, this module is challenged by an ability to (1) find features describing all three biosignals in the most informative and discriminative way when they are used together and then (2) construct a model to perform sleep stage classification efficiently. We introduce a two-stage classifier for automatic sleep scoring. Particularly, its first stage is an off-line training stage composing of 3 steps, which are feature extraction, feature selection, and model training. A set of possible features corresponding to each of three separate signals are extracted. Next, a selection process is applied to choose features with a more discriminative process. Using a set of dominant features selected, the sleep stage classifier is trained with a measurement of similarity. Finally, the trained model is utilized in its second stage for on-line sleep stage classification.

#### 4. IN-EAR BIOSIGNAL ACQUISITION

Extensive study of the anatomy of the ear indicates the shape and size changes of ear canals is caused by jaw motion [36, 38, 12], and asymmetry can remarkably occur between the left and right ones within an individual [35, 37]. Furthermore, it is critical to eliminate a gap between the in-ear electrodes and human skin due to the nature of the ion current generated by the biosignals. Hence, the *LIBS* device itself needs to be flexibly adjustable, contacted well with the skin, fitted with different people’s ears and types of muscle contractions without personalization, and comfortable to be worn in long term. Therefore, a commercial earplug with noise-cancelled and flexible wires are offered to form the base of the device. Its viscoelastic material enables *LIBS* to be easily plugged into the canal by squeezing and be quickly reshaped to its original form within the ear canal.

On the other hand, *LIBS* needs to possibly measure low-amplitude biosignals from a distance with high fidelity. Our method integrates several solutions into the hardware design to address this demand. The first solution is to use conduc-

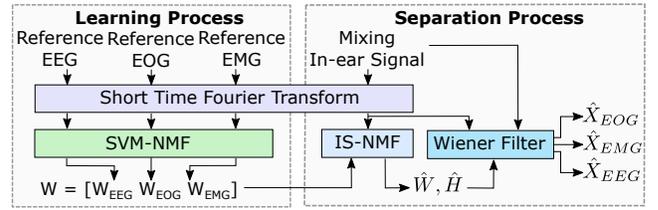


Figure 5: Overview of signal separation process in *LIBS*

tive silver cloth to make sensitive electrodes attached on two opposite sides of the earplug as contact points to the skin. Due to the use of soft cloth, the electrodes neither harm the in-ear skin nor are broken while being squeezed. However, because of the weave pattern of its fibers, we further coat their surface with many layers of thin silver leaves, which gives low and consistent surface resistance for providing reliable signals. Figure 4a shows the comprehensive structure of *LIBS*. Finally, we propose an increase of the distance between the main electrodes (placed in one ear) and the reference electrode (placed in another ear) to improve signal fidelity. To measure the in-ear signal, the sensor is connected to a circuit schemed simply in Figure 4b. An actual implementation of *LIBS* is further discussed in Section 7.

#### 5. SUPERVISED NMF-BASED SIGNAL SEPARATION TECHNIQUE

Due to the limited in-ear space, the biosignal sensed by *LIBS* is inherently a single-channel mixture of at least 4 components including EEG, EOG, EMG, and unwanted noise. Specifically, the mixed signal is seen as a linear combination of multiple signals from a number of individual sources in the spectral domain [15], which is mathematically expressed in Equation 1.

$$X = \sum_{i=1}^3 w_i s_i + \epsilon \quad (1)$$

where  $s_i$  is the power spectrum of the three biosignals with their corresponding weight  $w_i$  and  $\epsilon$  represents noises.

From literature, there exist mainstream techniques [57] such as principal component analysis (PCA), independent component analysis (ICA), empirical mode decomposition (EMD), and non-negative matrix factorization (NMF) built to solve the blind source separation problem. However, to precisely achieve the separation, these techniques, in general, require that (1) the number of collected channels is equal to or larger than the number of source signals (*except NMF*) and (2) the factorized components describing the source signal are selected manually. Based on the requirements, the key challenge in this work stems from the fact that the channel number that *LIBS* has (1 channel) is lower than the number of signals of interest (3 signals). Since NMF works well under the first limitation, we address our challenge by proposing a novel source separation technique that takes advantage of *NMF*. However, as studied in [8], there have existed two potential issues with a *NMF*-based model that might degrade the quality of the decomposed biosignals. They include (1) the inherent non-unique estimation of the original source signals (ill-posed problem) caused by

---

**Algorithm 1** Spectral Template Learning Algorithm

---

```
1: Input:  
2:  $\tilde{X}_{EEG}, \tilde{X}_{EOG}, \tilde{X}_{EMG}$  - One-night PSG signals  
3: Output:  
4:  $W_{ini}$  - Spectral template matrix  
5:  
6:  $X_{EEG}, X_{EOG}, X_{EMG} \leftarrow$   
7:  $\text{ComputePowerSpectrum}(\tilde{X}_{EEG}, \tilde{X}_{EOG}, \tilde{X}_{EMG});$   
8:  $[\hat{W}_{EEG}, \hat{H}_{EEG}] \leftarrow \text{SVM\_NMF}(X_{EEG});$   
9:  $[\hat{W}_{EOG}, \hat{H}_{EOG}] \leftarrow \text{SVM\_NMF}(X_{EOG});$   
10:  $[\hat{W}_{EMG}, \hat{H}_{EMG}] \leftarrow \text{SVM\_NMF}(X_{EMG});$   
11:  $W_{ini} \leftarrow [\hat{W}_{EEG} \hat{W}_{EOG} \hat{W}_{EMG}];$ 
```

---

the non-convex solution space of NMF and (2) the variance of the biosignals on different sleeps. To solve them, our proposed NMF-based model is combined with the source-specific prior knowledge learnt through a training process. We describe next the process of leveraging two different NMF techniques to learn source-specific information and to separate the mixing in-ear signal based on priory training. Figure 5 illustrates the high-level overview of this process.

When a new user uses LIBS, her EEG, EOG, and EMG are acquired using the gold-standard device during a training process. These signals are inputs of the learning process presented in Algorithm 1, which leverages a single-class SVM-based NMF technique (SVM-NMF) [8] to build corresponding template matrices. Then, for any in-ear signal  $\tilde{X}$  recorded by *LIBS*, our model approximately decomposes its power spectrum  $X$  into two lower rank nonnegative matrices

$$X \simeq WH \quad (2)$$

in which  $X \in \mathbb{R}^{m \times n}$  comprises  $m$  frequency bins and  $n$  temporal frames;  $W$  is the spectral template matrix representing basis patterns (components); and  $H$  is the activation matrix expressing time points (positions) when the signal patterns in  $W$  are activated. Finding the best representative of both  $W$  and  $H$  is equivalent to minimizing a cost function defined by the distance between  $X$  and  $WH$ . By that, Equation 2 is achieved through multiplicative update rules for the solution of the following optimization problem:

$$\{\hat{W}, \hat{H}\} = \arg \max_{W, H \geq 0} d(X|WH) \quad (3)$$

While solving Equation 3, the template matrix taken from the learning process is used to initialize  $W$ . As a result,  $W$  is deformed to fit the in-ear signal acquired from that user at different nights. Furthermore, adapting the technique from [7], we use Itakura Saito (IS) divergence  $d_{IS}$  as the cost function for our NMF model (IS-NMF) since it holds a scale-invariant property that helps minimize the variation of the signals acquired from one person in different recordings. The IS divergence is given by:

$$d_{IS}(X|WH) = \frac{X}{WH} - \log \frac{X}{WH} - 1 \quad (4)$$

Hence, Algorithm 2 provides the whole process of separating EEG, EOG, and EMG signals from the single-channel in-ear mixture using a per-user trained template matrix.

---

**Algorithm 2** Signal Separation Algorithm

---

```
1: Input:  
2:  $IS$  - In-ear Signal  
3:  $W_{ini}$  - Spectral Template Matrix  
4:  $ST$  - Segment Time  
5: Output:  
6:  $\hat{X}_{EEG}, \hat{X}_{EOG}, \hat{X}_{EMG}$  - Separated Signals  
7:  
8:  $\tilde{X} \leftarrow \text{PreprocessSignal}(IS);$   
9:  $X \leftarrow \text{ComputePowerSpectrum}(\tilde{X});$   
10:  $Seg \leftarrow \text{SegmentSignal}(IS, ST);$   
11: for  $i = 1 \rightarrow \text{sizeof}(Seg)$  do  
12:  $H_{ini} \leftarrow \text{InitializeMatrixRandomly}();$   
13:  $[\hat{W}, \hat{H}] \leftarrow \text{IS\_NMF}(Seg_i);$   
14:  $V_{EEG}(Seg_i) \leftarrow \hat{W}_{EEG}(Seg_i) * \hat{H}_{EEG}(Seg_i);$   
15:  $V_{EOG}(Seg_i) \leftarrow \hat{W}_{EOG}(Seg_i) * \hat{H}_{EOG}(Seg_i);$   
16:  $V_{EMG}(Seg_i) \leftarrow \hat{W}_{EMG}(Seg_i) * \hat{H}_{EMG}(Seg_i);$   
17:  $\hat{X}_{EEG} \leftarrow \text{resconstructSignal}(X, V_{EEG});$   
18:  $\hat{X}_{EOG} \leftarrow \text{resconstructSignal}(X, V_{EOG});$   
19:  $\hat{X}_{EMG} \leftarrow \text{resconstructSignal}(X, V_{EMG});$ 
```

---

## 6. AUTOMATIC SLEEP STAGING

Human sleep naturally proceeds in a repeated cycle of sleep stages. An expert can visually inspect EEG, EOG, and EMG signals collected from subjects during sleep and label each segment (i.e. a 30-second period) with the corresponding sleep stage based on known visual cues associated with each stage. Below we elaborate on each of aforementioned steps of our data analysis pipeline.

### 6.1 Feature Extraction

The features selected for extraction are from a variety of categories as follows:

*Temporal features:* This category includes typical features used in the literature such as mean, variance, median, skewness, kurtosis, and 75th percentile, which can be derived from the time series. In sleep stage classification, both EOG and EMG signals are often analyzed in the time domain due to their large variation in amplitude and a lack of distinctive frequency patterns. Accordingly, based on our observations about these signals, we include more features that can distinguish N1 from REM, which are often misclassified. In particular, we consider average amplitude that is significantly low for EMG while relatively higher for EOG during the REM stage. Also to capture the variation in EOG during different sleep stages, we consider the variance and entropy for EOG in order to magnify distinctions between Wakefulness, REM, and N1 stages.

*Spectral features:* These features are often extracted to analyze the characteristics of EEG signal because brain waves are normally available in discrete frequency ranges in different stages. By transforming the time series signal into the frequency domain in different frequency bands and computing its power spectrum density, various spectral features can be studied. Here based on our domain knowledge about the EEG patterns in each sleep stage, we identify and leverage spectral edge frequencies to distinguish those stages.

Features	
<i>Temporal features</i>	average amplitude, variance, 75th percentile, skewness, kurtosis
<i>Spectral features</i>	absolute spectral powers relative spectral powers relative spectral ratio spectral edge frequency
<i>Non-linear features</i>	fractal dimension, entropy

Table 1: List of features extracted from the biosignals



(a) Atop silver leaves (b) Fabric (c) Copper

Figure 6: Sensor prototypes with different materials

*Non-linear features:* Bioelectrical signals show various complex behaviors with nonlinear properties. In details, since the chaotic parameters of EEG are dependent on the sleep stages [31], they can be used for sleep stage classification. The discriminant ability of such features is demonstrated through the measures of complexity such as correlation dimension, Lyapunov exponent, entropy, fractal dimension, etc. [54, 9].

For this study, relied on the literature of feature-based EOG, EMG, and EEG classification [31, 20], we consider the features listed in Table 1 from each of the aforementioned categories.

## 6.2 Feature Selection

Although each extracted feature has the ability to partially classify biosignals, the performance of a classification algorithm can degrade when all extracted features are used to determine the sleep stages. Therefore, in order to select a set of relevant features among the extracted ones, we compute the discriminating power of each of them [20, 49] when they are used in combination. However, it is computationally impractical to test all of the possible feature combinations. Therefore, we adopt a procedure called Forward Selection (FSP) [62] to identify the most effective combination of features extracted from our in-ear signal. With FSP, features are selected sequentially until the addition of a new feature results in no performance improvement in prediction. To further improve the efficiency of our selection method, we have considered additional criteria for selecting features. In particular, we assigned a weight to each feature based on its classification capability and relevance to other features. Subsequently, these weight factors are adjusted based on the classification error. Furthermore, a feature is added to the set of selected features if it not only improves the misclassification error but also is less redundant given the features already selected. With this approach, we can efficiently rank discriminant features based on the intrinsic behavior of the EEG, EMG, and EOG signals.

## 6.3 Sleep Stage Classification

Various classification methods are proposed in the literature for similar applications and each has advantages and

### Algorithm 3 Training Algorithm

---

```

1: Input:
2:    $IS$  - In-ear Signal
3:    $SSL$  - Sleep Stage Labels
4:    $ST$  - Segment Time
5: Output:
6:    $RF$  - Trained Random Forest Model
7:    $SF$  - Types of Selected Features
8:
9: procedure A: FEATURE EXTRACTION
10:   $Segs \leftarrow \text{SegmentSignal}(IS, ST)$ ;
11:  for  $i = 1 \rightarrow \text{sizeof}(Segs)$  do
12:     $TF \leftarrow \text{GenerateTemporalFeatures}(Segs_i)$ ;
13:     $FF \leftarrow \text{GenerateSpectralFeatures}(Segs_i)$ ;
14:     $NF \leftarrow \text{GenerateNonlinearFeatures}(Segs_i)$ ;
15:     $FS_i \leftarrow [TF, FF, NF]$ ;
16: procedure B: FEATURE SELECTION
17:  for  $i = 1 \rightarrow \text{sizeof}(FS)$  do
18:     $SF \leftarrow \text{ForwardSelectionProcedure}(FS_i, SSL)$ ;
19:
20:  $TrainingSet \leftarrow \text{SelectSamplesRandomly}[SF, 70, 30]$ ;
21:  $n \leftarrow 25$ ; //trees
22:  $RF \leftarrow \text{TreeBagger}(n, TrainingSet, SSL)$ ;

```

---

disadvantages. Some scholars [47] [31] have chosen the Artificial Neural Network (ANN) classification approach for sleep scoring. In spite of the ANN ability to classify untrained patterns, long training time and complexity for selection of parameters such as network topology. Moreover, since decision tree is easier to implement and interpret as compared to other algorithms, it is widely used for sleep stage classification.

Another classification method used for sleep stage identification is Support Vector Machine (SVM). SVM is a machine learning method based on statistical learning theory. Since SVM can be used for large data sets with high accuracy rates, it has also been widely used by various studies [19] [14] to classify sleep stages. However, this approach suffers from long training time and difficulty to understand the learned function. Based on the existing comparative studies [49] [2], the decision tree (and more generally random forest) classification methods have achieved the highest performance since the tree structure can separate the sleep stages with large variation. As an example, decision tree classifiers are flexible and work well with categorical data. However, overfitting and high dimensionality are the main challenges in decision trees. Therefore, we use an ensemble learning method for classification of in-ear signal. Particularly, we deploy random forest with twenty five decision trees [2, 49] as a suitable classifier for our system. This classifier is able to efficiently handle high dimensional attributes and it also reduces computational cost on large training data sets. The set of features selected through FSP are used to construct a multitude of decision trees at training stage to identify the corresponding sleep stage for every 30-seconds segment of the biosignals in the classification stage. The training procedures is presented in Algorithm 3.

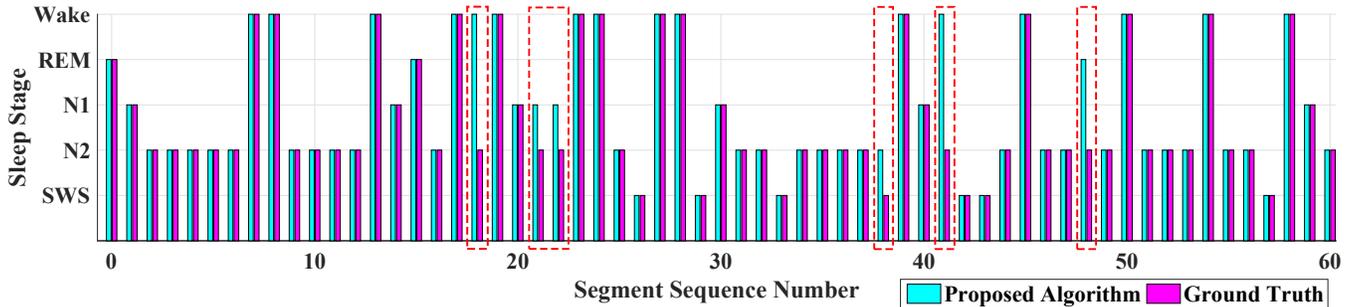


Figure 7: A hypnogram of 30-minute data resulted by *LIBS*. Sleep staging done using *LIBS* (light blue) is compared with the ground truth from PSG (dark pink). The misclassification of our algorithm is marked by red dashed rectangles.

## 7. IMPLEMENTATION

In this section, we discuss the actual construction of *LIBS* prototype using off-the-shelf electrical components.

**Sensor material.** To be able to capture the reliable biosignals, it is important to design *LIBS* that well fits with the user’s ear canal. One possible approach is to personalize a mold. However, this approach entails high cost and time consume. Therefore, we have augmented an over-the-counter sound block foam earplug for its sensor base. The soft elastic material (or memory foam) of the earplug enables the sensor to reshape to its original form shortly after being squeezed or twisted under the strain to insert into the ear. This fundamental property of the foam earplug provides a comfortable and good fit as it allows the sensor to follow the shape of the inner surface in the ear canal. In addition, it not only supplies a stable contact between the electrodes and the in-ear skin but reduces the motion artifact caused by jaw motion as well. Moreover, using the earplug completely eliminate the personalization of the base regarding the canal size. Also, the soft surface and the lightweight property of the earplug make itself more convenient to be worn without much interference during sleep. Finally, as an additional bonus, the foam blocks out noise, hence improves the sleep.

**Electrode construction and placement.** By studying the anatomy of the human ear canal, we first design the electrode with an oval shape (roughly 1 cm length and 0.7 cm wide). To develop robust electrodes to capture the weak in-ear biosignals, we then try different conductive materials as shown in Figure 6. Our experiment leads to that copper is a hard material to be inserted into and placed inside the ear without harm and fabric has very high and non-identical resistivity ( $19\Omega/\text{sq}$ ) on its surface caused by its special weave pattern. Due to the softness of the conductive fabric, we still select this material and further lower as well as stabilize the resistance between the fabric electrode and the outer layer of the skin in the ear canal by coating its surface with three layers of a pure and thin silver leaf. Also, a very small amount of health-grade conductive gel is added. Ultimately, we place the active and reference electrodes in two separate ear canals, hence intensify the potential of the signals by a distance increase. The recorded in-ear signal is finally transferred from the electrode to an amplifier through shielded wires to prevent any external noise.

***LIBS* microcontroller.** In our prototype, *LIBS* uses a brain-computer interface (BCI) board named OpenBCI [39] to sample and digitize the in-ear signal. The board is supplied by a battery source of 6V for safety and configured at a sampling rate of 2kHz and a gain of 24dB. The signal is stored in a mini-SD card on the board while recording and then processed offline in a PC, which include the separation of EEG, EOG, and EMG and the sleep stage classification.

## 8. EVALUATION

In this section, we first present the key results from our main goal of using *LIBS* outputs to perform the automatic sleep stage classification. Next, we illustrate the ability of *LIBS* to capture the usable and reliable biosignal, which contains the mixture of EEG, EOG, and EMG, from inside the ear canal. Based on the outcome of their occurrence, we then show the performance of our proposed separation algorithm for splitting those three signals from the mixing in-ear signal. Lastly, we evaluate user experiences with *LIBS* through a questionnaire reflecting their experience of using *LIBS* for sleep study.

### 8.1 Sleep Study Methodology

We conducted a 38-hours of sleep experiments over 8 graduate students (3 females, 5 males) with an average age of 25 to evaluate the performance of the proposed sleep stage classification system inputting the biosignals returned by *LIBS*. An full board IRB review was conducted and an approval was granted for this study. The participants were asked to sleep in a sleep lab while plugging *LIBS* into their ear canals and have a conventional PSG hook-up around their head simultaneously. We used a portable PSG device named Trackit Mark III supported by LifeLines Neurodiagnostic Systems Inc. company [53] with 14 EEG electrodes placed at the channel Fp1, Fp2, C3, C4, O1, and O2 (in accordance to the International 10-20 system) on the scalp, in proximity to the right and left outer cantus, and over the chin, which were all referenced to two mastoids, to collect the ground truth. This device individually acquired EEG, EOG, and EMG signals at 256Hz sampling rate and pre-filtered them in the range of 0.1-70Hz. After that, the Polysmith program [44] was run to score the ground-truth signals into different sleep stages at every 30-second segment. For all studies, the sleeping environment was set up to be quiet, dark, and cool.

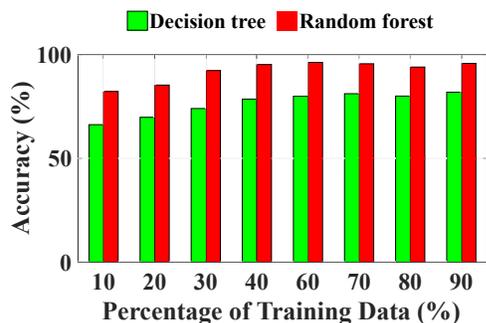


Figure 8: Accuracy comparison for two deployed classifiers

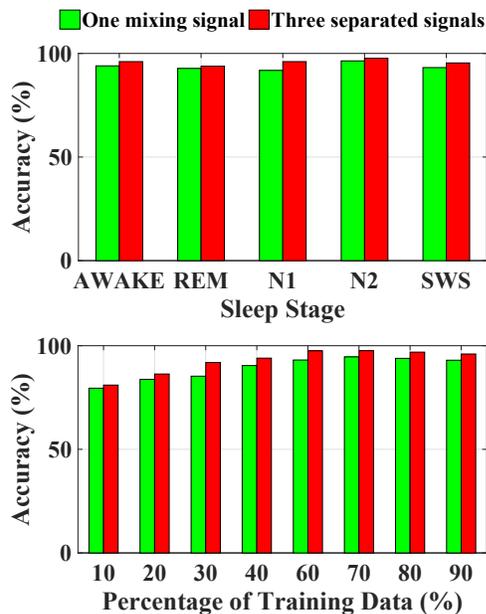


Figure 9: The accuracy of predicting the sleep stage from the sleep data recorded by *LIBS*. The accuracy achieved with the separated signals (dark/red) is compared to the accuracy achieved with the mixed signal (light/green). The results obtained for five sleep stages and for different percentage of training data are shown on the top and in the bottom, respectively.

## 8.2 Sleep Stage Classification Evaluation

To evaluate the performance of our proposed sleep staging method, the features were extracted from 4313 30-second segments from 8 subjects using the original mixed signal as well as three separated signals. Training and test data sets are randomly selected from the same subject pool. We implemented all steps of our sleep staging method in MATLAB. The number of 30-second segments for Wake, REM, N1, N2, and SWS are 886, 242, 490, 1422, and 1273, respectively. The performance of our random forest classifier is determined by computing accuracy, recall, and precision.

Figure 7 displays the results of the sleep stage classification in comparison to the hypnogram of the test data scores out of the gold standard PSG. From this it can be observed that the dynamics of the hypnogram is almost completely maintained in the predicted scores.

After selecting features for the mixed signal and three sep-

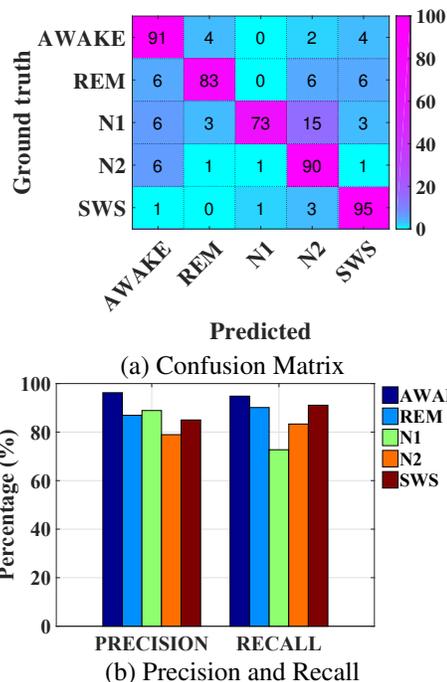


Figure 10: The confusion matrix, precision, and recall of predicting the sleep stage from the separated in-ear EEG, EOG, and EMG signals recorded by *LIBS*

arated signals, we perform the classification. A comparison of the decision tree and random forest classifiers is presented in Figure 8 showing how random forest can overcome disadvantages of decision tree as we explained in Section 6.3. Our results show that our end-to-end system can achieve 95% accuracy in sleep staging on average. Figure 9 (bottom) shows the observed classification accuracy given different proportions of data for training. As illustrated, the classification accuracy is comparable between three separated signals and mixed signal despite the potential noise overhead introduced in the signal separation process. We attribute this to the fact that with the separated signals we are able to leverage on specific features and characteristics of each individual signal for classification. Finally, with 60% of all subject data allocation for training, we achieve the optimal maximal classification accuracy, beyond which our solution will be over-trained without significant improvement.

Figure 10a presents the confusion matrix obtained when the optimal set of features is used. The columns represent the selected sleep stages classified by the random forest classifier and the rows represent the sleep stages as determined by the experts. With the mixed signal and for sleep stage SWS, we achieve maximum accuracy at 96%. With the separated signals on the other hand, N2 is best classified at 89%. On the lower side, classification of the mixed signal and the separated signal result in 78.31% and 87% accuracy for REM and N1, respectively. Figure 10b depicts our system classification precision and recall for the specific sleep stages. We compared the classification sensitivity for N1 and REM using the separated signals with that of the mixed signal and confirmed the superiority of the separated signal

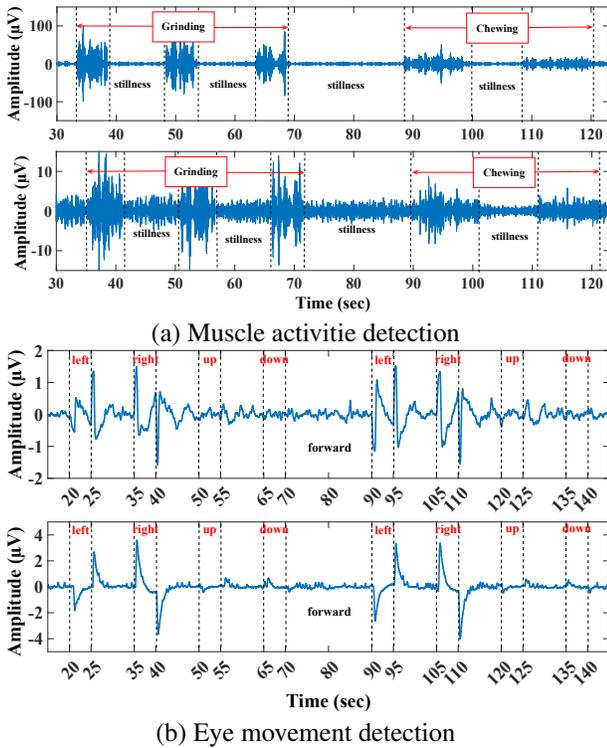


Figure 11: The detection of (a) muscle activities and (b) eye movements from *LIBS* (top) and the gold standard EMG and EOG channels (bottom), respectively

in minimizing misclassification. In particular, with the separated signals we achieve 80% and 83% sensitivity for N1 and REM, versus 73% and 65% sensitivity achieved by the mixed signals for these sleep stages, respectively. We conclude that the selected features demonstrated more discriminating power for the separated signals. It is also important to note that with our experiments, the standard deviation of the accuracy across 8 subjects was about 1.8 that confirms subject-independent quality and robustness of our approach.

### 8.3 Signal Acquisition Validation

In this evaluation, we assess the quality of the mixed in-ear signal acquired by *LIBS* by comparing it with the groundtruth signals acquired from corresponding standard PSG channels when both the devices are worn simultaneously. Particularly, *LIBS*'s capability of recording the three biosignals is illustrated through different experiments.

We first examined *LIBS* to see if it can capture the activities of the facial muscles. To do this, we asked the subject to keep his/her teeth remaining still and then grinding for 5s and chewing for 20s continuously and repeat that four times. From Figure 11a, we noticed that our *LIBS* device could clearly capture those events reflecting the occurrence of the EMG signal.

Similarly, we examine the ability of *LIBS* to capture the events of moving eyes horizontally and vertically that reflect the occurrence of the EOG signal. The subject was asked to remain still and gaze forward for 20s and then move her eyes to pre-specified points set in four directions (i.e. left, right, up, and down) for 5s with 10s of gazing forward between

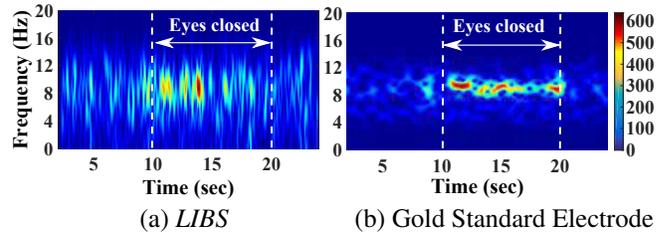


Figure 12: The detection of alpha rhythms from (a) *LIBS* and (b) the gold standard device at Channel C4 on scalp

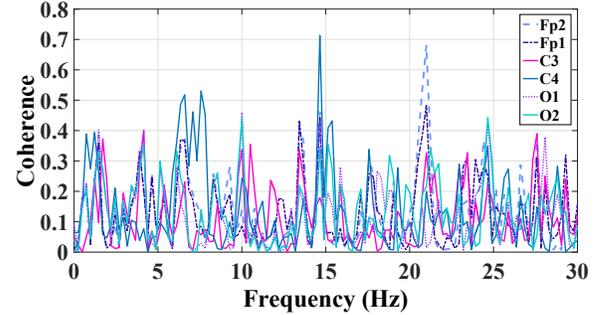


Figure 13: The average spectral coherence results for alpha rhythms recording shown in Figure 12

them. Although the amplitude of the in-ear signal is smaller than the gold-standard one, the in-ear signal still clearly displays the left and right movements of the eyes similar to the EOG signal channeled in the gold-standard device as shown in Figure 11b.

To verify the validity of our *LIBS* device in EEG recordings, we conducted the following standard BCI experiments:

**The alpha attenuation response (AAR):** Alpha waves are brain waves that are specified within a frequency range of 8-13Hz. The brain wave is produced during the N1 stage and is a sign of relaxation and peacefulness [3]. In this experiment of detecting alpha waves, the subject was asked to stay relaxing in a comfortable position and close their eyes for 20s and then open them for 10s five consecutive times. As analyzing the recorded in-ear signal, Figure 12 shows that *LIBS* is able to capture the alpha rhythm from inside the ear. In addition, we computed the magnitude-squared coherence estimation between the in-ear signal and the PSG data using Welch's averaged periodogram method [60]. The highest coherence coefficient of the in-ear signal is 0.72 with the PSG signal at channel C4 in the temporal region of the brain as illustrated in Figure 13. However, the alpha rhythm demonstrated in these figures is not very clear. This can be due to the fact that the alpha waves were produced in frontal lobe that is in a distance from the ear location.

**Auditory Steady-State Response (ASSR):** This EEG paradigm measures the amplitude of EEG responses while modulating auditory stimuli with specific frequency ranges[56]. In our experiment, we applied auditory stimuli in the frequencies of 40Hz and 80Hz in which each stimuli lasted for 30s and was repeated three times with 20s rest between them. The 80Hz and 40Hz ASSR experiments produced a sharp and dominant peak at 80Hz and 40Hz shown in Figure 15 and Figure 14, respectively, as well as higher SNR values of the in-ear signal. Hence, these results demon-

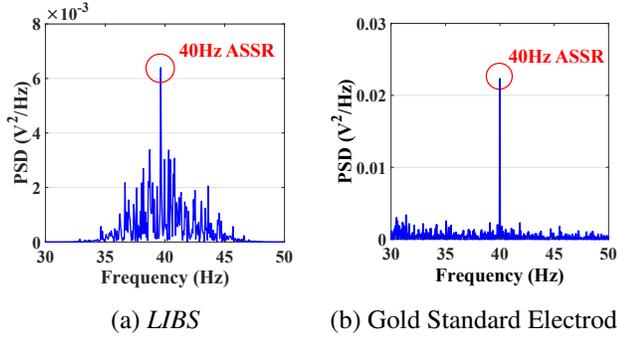


Figure 14: The ASSR for 40 Hz recorded from (a) *LIBS* and (b) the gold standard device at Channel C3 on scalp

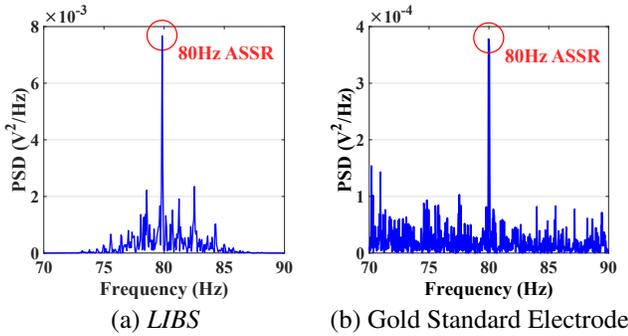


Figure 15: The ASSR for 80 Hz recorded from (a) *LIBS* and (b) the gold standard device at Channel C3 on scalp

strate the ability of *LIBS* to detect those two specific frequencies with high quality signal although the peaks in case of the gold standard electrodes for ASSR were larger than those of the *LIBS* electrode.

**Steady-State Visually Evoked Potential (SSVEP):** Similar to ASSR, SSVEP measures the brain waves responding to visual stimuli with specific frequencies [32]. We stimulated the brain wave responses by blinking stimuli with the frequency of 10Hz played for 20s and repeated three times. Figure 16 presents SSVEP response peak at the 10Hz frequency for *LIBS* and the gold standard on-scalp electrodes.

Generally, as our expectation, the in-ear biosignal captured by our *LIBS* device has its quality reduced in order of magnitude compared to the groundtruth signals because the sources of signals where we placed the *LIBS* and EEG electrodes are different.

**Different Conductive Materials Evaluation.** The same experiments presented above were applied to examine the quality of the recorded in-ear signals using the electrodes made of different conductive materials. In most of the experiments, the copper electrode was not able to record the signals due to the difficulty of securely fitting it inside the ear canal caused by its hardness characteristic. Therefore, we only compared the performance between the fabric electrode and our new prototype that has fine silver leaves on top of the silver fabric. One example of the signals captured by these two types of electrodes from the same person and the same ear canal is shown in Figure 17. It is clearly seen from the figure that the signal captured by the only silver fabric electrode had more artifact impact than the other prototype.

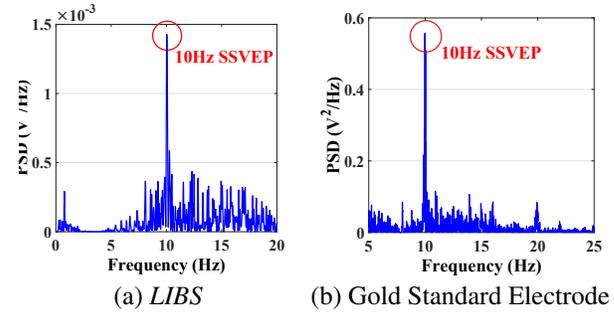


Figure 16: The SSVEP responses recorded from (a) *LIBS* and (b) the gold standard device at Channel O1 on scalp

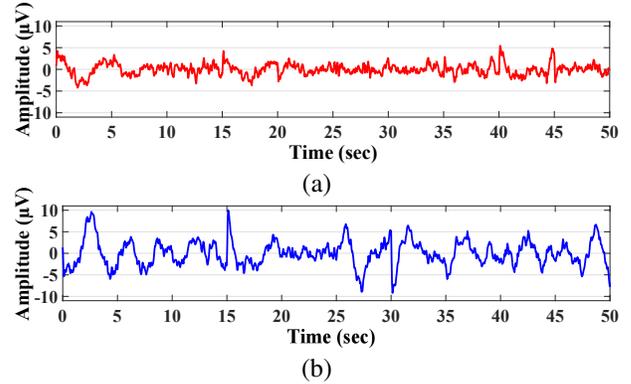


Figure 17: Comparison of the quality of signals obtained from (a) only silver fabric electrodes and (b) silver leaves on top of silver fabric electrodes

Also, the mixed in-ear signal acquired by the new prototype with silver leaves has larger voltage than the signal captured by the fabric electrode.

## 8.4 Signal Separation Validation

From the previous evaluation, we prove that all of the EEG, EOG, and EMG signals are mixed in the in-ear signal and possibly captured by our wearable device. Hence, we now show the result of our proposed NMF-based separation algorithm, which learns the underlying characteristics of gold standard EEG, EOG, and EMG signals individually and adapts its learned knowledge (or the spectral template matrix) to provide the best decomposition from the mixed signal. In this evaluation, because the gold standard device (e.g. PSG device) cannot be hooked up in the ear canal to capture the same signal as our in-ear device does, similarity measures such as mutual information, cross-correlation, etc. cannot be used to provide a numeric comparison between the separated and gold standard signals. We then demonstrate the performance of our proposed model by analyzing the occurrence of special frequencies in the separated EEG biosignal during the sleep study as shown in Figure 18.

In this figure, it is noted that top and bottom panels show the spectrograms and their corresponding time-series signals, respectively. Specifically, Figure 18a provides the spectrogram for a 30-second raw mixed in-ear signal measured by our wearable during the sleep study and labeled as stage SWS by the gold-standard device. In Figure 18b,

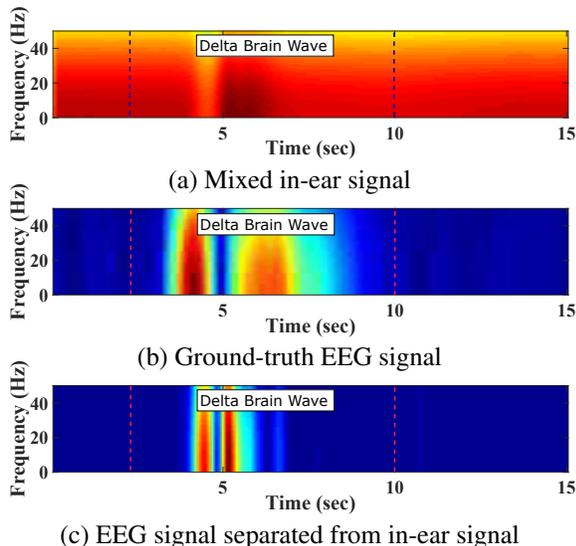


Figure 18: **Signal separation performance obtained by LIBS through a 30-second mixed in-ear signal (a) and compared with the ground-truth EEG signal (b) and its corresponding separated EEG signal (c).**

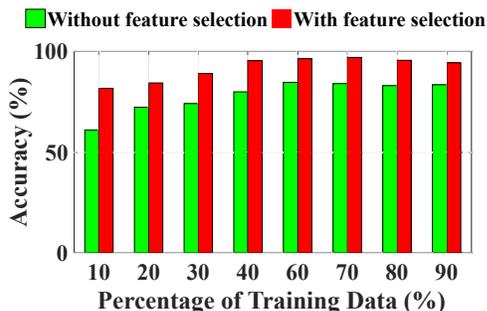


Figure 19: **Classification accuracy with and without feature selection**

the spectrogram for a corresponding 30-second ground-truth EEG signal is presented. By observing these two spectrograms, a delta brain wave in a frequency range lower than 4Hz is correctly found. However, the spectrogram in Figure 18a shows that not only the delta brain wave exists but also other biosignals are added in the raw signal. Finally, Figure 18c exhibits the spectrogram for the EEG signal separated by applying our proposed signal separation algorithm. Analyzing this figure proves that the separation model we propose has a capability of both splitting the signals completely from the mixed one and keeping only the specific characteristics of the separated signal. Thus, the short appearance of the delta brain wave in the decomposed signal can be explained by the fact that the location where our wearable is placed is far from the source of the signal. By that, it is difficult for the signal to be captured when its amplitude is reduced.

## 8.5 Features Selection Evaluation

To demonstrate the effect of feature selection in the performance of classification, performance analysis is done with and without feature selection and results are summarized in Figure 19. As shown in the figure, the classifica-

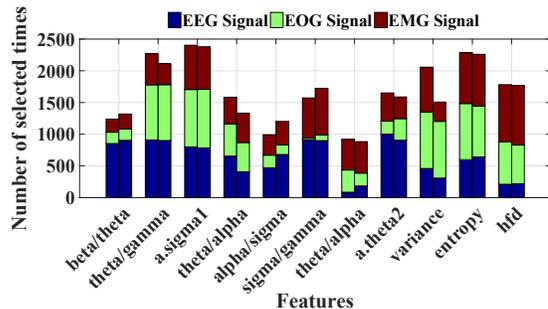


Figure 20: **Stability of the feature selection**

tion accuracy considerably improves when feature selection is applied as it effectively removes irrelevant and noisy features. We also analyzed the stability of the selected feature set across different sessions for a one-subject sleep dataset. In Figure 20, we have illustrated the selection frequency of each feature (in percentage) for various separated signals. This experiment is performed for two different data sets collected from the same object, and the corresponding results are presented in two columns for each feature. As shown in the figure, selection frequency of the features is stable across experiments.

## 8.6 User Experience Survey

No.	Survey Statements	SD	Mean
1	<i>The in-ear device is comfortable to wear during a sleep.</i>	0.58	4.0
2	<i>Wearing this device does not include any harmfulness.</i>	0.76	4.5
3	<i>I would like to use the in-ear device to evaluate my sleep quality.</i>	0.68	4.1
4	<i>Generally, I am satisfied with the use of the in-ear device.</i>	0.47	4.3
5	<i>The in-ear device is more comfortable than the on-scalp electrodes of the PSG device.</i>	0.49	4.4
6	<i>I did not get disturbed during sleep due to the in-ear device.</i>	0.75	4.2
7	<i>I may use the in-ear device every night.</i>	0.98	4.2
8	<i>If the in-ear device is wirelessly and it is available for sale, I would like to buy it to assess my sleep quality.</i>	0.80	4.4

Table 2: **Main statements of user experience survey**

After conducting the sleep study, the participants were asked to answer a survey about their experience of wearing the *LIBS* prototype during their sleeping period. The survey questions concentrate on the comfort and the usability of the device. Table 2 shows main statements used to evaluate *LIBS*.

The results of the questionnaire show an overall agreement of using *LIBS* to assess the sleep stage classification. All the participants were satisfied to use the device and they agreed that wearing *LIBS* did not include any harmfulness. 83.3% of them would like to buy our device to assess their sleep quality if it were wireless. Also, 86.7% of the participants stated that our *LIBS* device is more comfortable than the on-scalp electrodes of the gold-standard PSG device. Also, wearing *LIBS* did not disturb the sleep of 85.8%

of the participants. All these high percentages of agreement show the possibility of our *LIBS* device to be adopted by users and it is promising to be an alternative method for the sleep quality assessment.

## 9. RELATED WORK

Due to their importance and widespread applications in health care [33, 34, 42, 59] and BCI [13, 28], the hardware that drives biosignal collection has been improved dramatically involving both a decrease in size and cost [29] and an increase in comfort in use. Among existing prototypes, researchers have presented reasons and evidences to show that the in-ear position is promising for continuous biosignal recordings [16, 27]. Specifically, physical features of the ear canal allow a tight and fixed electrode placement that is desirable for electrode stability and long-term wearability [26, 58] as well as provide a direct contact with the body in the same time it exhibits a high degree of comfort [6]. Moreover, in contrast to the conventional PSG, *LIBS* can be easily self-applicable by the users themselves.

A number of studies have introduced many wearable devices in the form of a hearing aid that can continuously monitor heart rate and respiratory rate through various measurements such as electrocardiography (ECG), ballistocardiography (BCG), and photoplethysmography (PPG). For examples, He et al. [16] built a MEMS tri-axial accelerometer in the hearing aid to sense BCG. Winokur et al. [61] introduced three techniques embedded in a portable device behind the ear to capture all of them. Vogel et al. [58], Poh et al. [41, 42], and Vollmer et al. [59] developed a microoptic reflective sensor located inside the auditory canal or inserted into ear buds of commercial earphones to measure in-ear PPG. Different from the EEG, EOG, and EMG collected in our work, those types of signals are easy to extract from the in-ear signal due to (1) a fairly unique pattern interpreted in ECG, (2) highly different range of frequency with body movements in BCG, and (3) no effect caused by noise in case of PPG.

Related to the measurement of physiological activities of the brain, eyes, and muscles, many in-ear wearable designs have been made over the past six years. In particular, Sano et al. [48] proposed to attach a sensing system outside the regular ear bud to measure the EMG signal for facial movement classification (i.e. stillness, smile, and chew). On the other hand, Manabe et al. [28] constructed a variety of ear bug prototypes to capture the EOG potentials for eye gaze detection. Furthermore, Merrill et al. [29] bent and placed a 6mm gold cup electrode above a rolled foam earplug or Lee et al. [23] utilized a special material named Carbon Nanotube Polydimethylsiloxane (CNT/PDMS) to produce an earphone cap for collecting the EEG signal. Compared to the *LIBS* we designed, such these designs are generally stiff that can cause a harm on the inside of the ear as well as need personalizing to fit per-user ear canals.

In order to solve their hardness, instability, and the need of personalization, Norton et al. [21] designed a soft and foldable electrode that can capture the EEG from different outer complex surfaces of the ear and the mastoid using

the epidermal electronics with fractal meshes layouts. Also, Goverdovsky et al. [11] suggested a new prototype called Ear-EEG that consists of a viscoelastic substrate memory foam earplug and conductive cloth electrodes to insure conformance with ear canal surface for motion artifacts reduction. However, such these work considers the capture of only one signal among EEG, EOG, and EMG and did not aim at classifying the sleep stages using these three biosignals. Dissimilar from them, the soft electrode we construct is able to provide low and consistent surface-resistance and sensitively captures all EEG, EOG, and EMG signals mixed in one in-ear signal. Moreover, the EEG, EOG, and EMG are extracted from the in-ear mixture to become three independent biosignals.

Finally, our work is distinct from all previous work in determining the sleep stages occurring during sleep using the in-ear EEG, EOG, and EMG. Several standard methods such as artificial neural network (ANN) [47, 31], decision trees [2], random forest [24, 62], and support vector machine (SVM) [19, 14] have been studied for automatic sleep staging. However, most of these techniques require long training time, high complexity for the selection of model parameters, and large training datasets. Additionally, all of them worked on the conventional biosignals captured by many electrodes attached on the head. On the other hand, in this paper, we first try to apply the combination of decision tree and random forest methods into the EEG, EOG, and EMG extracted from the in-ear signal for the sleep staging application.

## 10. CONCLUSION

In this paper, we presented the design, implementation, and evaluation of a wearable sensing system that can sense EEG, EOG, and EMG signals using a small number of electrodes placed inside user's ear. We introduced a set of algorithms for extracting these individual signals which are then used as inputs for our sleep stage classification system. Through our hardware prototype evaluation and 1-month long-term user-study, we show that our in-ear wearable device is comparable to the existing dedicated sleep assessment systems (e.g. PSG) in term of sleep stages classification accuracy, while possessing many desirable properties such as low cost, easy operation, and comfortable wearing during sleep. More than just automatic sleep staging, *LIBS* with its three individual biosignal outputs has a potential to become a fundamental sensing device in divergent healthcare problems including long-term monitoring outside clinical facilities, sleep environment control, brain surgery support, diagnosis of brain related disease (e.g. sleep disorders, epilepsy, etc.), autonomous hearing-aid steering, and a numerous other HCI applications.

## 11. ACKNOWLEDGMENTS

We thank the anonymous ACM SenSys reviewers for their helpful comments. We also thank LifeLines Neurodiagnostic Systems Inc. for their support in providing the gold-standard PSG device and thank Yiming Deng and Titsa Papanтони for their valuable feedback in the early stage of this work.

## 12. REFERENCES

- [1] American Academy of Sleep Medicine. <http://www.aasmnet.org/>. Accessed: August 03, 2016.
- [2] K. A. I. Aboalayon, W. S. Almuhammadi, and M. Faezipour. A comparison of different machine learning algorithms using single channel EEG signal for classifying human sleep stages. In *Proceedings of 2015 IEEE Long Island Systems, Applications and Technology Conference (LISAT)*, pages 1–6, 2015.
- [3] C. Alloway, R. Ogilvie, and C. Shapiro. The alpha attenuation test: assessing excessive daytime sleepiness in narcolepsy-cataplexy. *Sleep*, 20(4):258–266, 1997.
- [4] Aware. <https://goo.gl/eh48HA>. Accessed: August 03, 2016.
- [5] Basis. <http://www.mybasis.com/>. Accessed: August 03, 2016.
- [6] M. Bleichner, M. Lundbeck, M. Selisky, F. Minow, M. Jäger, R. Emkes, S. Debener, and M. D. Vos. Exploring miniaturized EEG electrodes for brain-computer interfaces. An EEG you do not see? *Physiological Reports*, 3(4), 2015.
- [7] C. Damon, A. Liutkus, A. Gramfort, and S. Essid. Non-negative matrix factorization for single-channel EEG artifact rejection. In *Proceedings of 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1177–1181, 2013.
- [8] S. Essid. A single-class SVM based algorithm for computing an identifiable NMF. In *Proceedings of 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2053–2056, 2012.
- [9] J. Fell, J. Roschke, K. Mann, and C. Schaffner. Discrimination of sleep stages: a comparison between spectral and nonlinear EEG measures. *Electroencephalography and Clinical Neurophysiology*, 98(5):401–410, 1996.
- [10] Fitbit. <http://www.fitbit.com/>. Accessed: August 03, 2016.
- [11] V. Goverdovsky, D. Looney, P. Kidmose, and D. Mandic. In-ear EEG from viscoelastic generic earpieces: Robust and unobtrusive 24/7 monitoring. *IEEE Sensors Journal*, 2015.
- [12] M. Grenness. Mapping Morphologic Change in the External Ear. In *Masters dissertation. Hobart, Australia: University of Tasmania*, 1999.
- [13] W. Gu, Z. Yang, L. Shangguan, W. Sun, K. Jin, and Y. Liu. InEar BioFeedController: A Headset For Hands-Free And Eyes-Free Interaction With Mobile Devices. In *CHI*, pages 1293–1298, 2013.
- [14] S. Gudmundsson, T. Runarsson, and S. Sigurdsson. Automatic Sleep Staging Using Support Vector Machines with Posterior Probability Estimates. In *Proceedings of the International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC)*, pages 366–372, 2005.
- [15] H. Hallez, B. Vanrumste, R. Grech, J. Muscat, W. Clercq, A. Vergult, Y. D’Asseler, K. Camilleri, S. Fabri, S. Huffel, and I. Lemahieu. Review on solving the forward problem in EEG source analysis. *Journal of NeuroEngineering and Rehabilitation*, 4(1):1–29, 2007.
- [16] D. He, E. Winokur, and C. Sodini. An ear-worn continuous ballistocardiogram (BCG) sensor for cardiovascular monitoring. In *Proceedings of 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5030–5033, 2012.
- [17] Jawbone. <https://jawbone.com/>. Accessed: August 03, 2016.
- [18] KOKOON. <https://kokoon.io/>. Accessed: August 03, 2016.
- [19] B. Koley and D. Dey. An Ensemble System for Automatic Sleep Stage Classification Using Single Channel EEG Signal. *Comput. Biol. Med.*, 42(12):1186–1195, 2012.
- [20] A. Krakovska and K. Mezeiova. Automatic sleep scoring: A search for an optimal combination of measures. *Artificial Intelligence in Medicine*, 53(1):25–33, 2011.
- [21] A. Kulkarni, P. Rao, S. Natarajan, A. Goldman, V. S. Sabisetti, Y. Khater, N. Korimerla, V. Chandrasekar, R. A. Mashelkar, and S. Sengupta. Soft, curved electrode systems capable of integration on the auricle as a persistent brain-computer interface. In *Proceedings of the National Academy of Sciences of the United States of America*, 2015.
- [22] C. Kushida et al. Practice parameters for the indications for polysomnography and related procedures: an update for 2005. *Sleep*, 28(4):499–521, 2005.
- [23] J. H. Lee, S. M. Lee, H. J. Byeon, J. S. Hong, K. S. Park, and S. Lee. CNT/PDMS-based canal-typed ear electrodes for inconspicuous EEG recording. *Journal of Neural Engineering*, 11(4), 2014.
- [24] S. Liang, C. Kuo, Y. Hu, and Y. Cheng. A rule-based automatic sleep staging method. *Journal of Neuroscience Methods*, 205(1):169–176, 2012.
- [25] S. F. Liang, C. E. Kuo, Y. C. Lee, W. C. Lin, Y. C. Liu, P. Y. Chen, F. Y. Cherng, and F. Z. Shaw. Development of an EOG-Based Automatic Sleep-Monitoring Eye Mask. In *IEEE Transactions on Instrumentation and Measurement*, 2015.
- [26] D. Looney, P. Kidmose, and D. Mandic. *Ear-EEG: User-Centered and Wearable BCI*, volume 6 of *Biosystems & Biorobotics*, pages 41–50. Springer Berlin Heidelberg, 2014.
- [27] D. Looney, P. Kidmose, C. Park, M. Ungstrup, M. Rank, K. Rosenkranz, and D. Mandic. The In-the-Ear Recording Concept: User-Centered and Wearable Brain Monitoring. *IEEE Pulse*, 3(6):32–42,

- 2012.
- [28] H. Manabe, M. Fukumoto, and T. Yagi. Conductive rubber electrodes for earphone-based eye gesture input interface. *Journal of Personal and Ubiquitous Computing*, 19(1):143–154, 2015.
- [29] N. Merrill, M. T. Curran, J. Yang, and J. Chuang. Classifying Mental Gestures with In-Ear EEG. In *Proceedings of 13th International Conference on Wearable and Implantable Body Sensor Network (BSN)*, 2016.
- [30] Misfit. <http://misfit.com/>. Accessed: August 03, 2016.
- [31] S. Motamedi-Fakhr, M. Moshrefi-Torbati, M. Hill, C. Hill, and P. White. Signal processing techniques applied to human sleep EEG signals - A review. *Biomedical Signal Processing and Control*, 10:21–33, 2014.
- [32] A. Nawrocka and K. Holewa. Brain - Computer interface based on Steady - State Visual Evoked Potentials (SSVEP). In *Proceedings of 14th International Carpathian Control Conference (ICCC)*, pages 251–254, 2013.
- [33] A. Nguyen, R. Alqurashi, A. C. Halbower, and T. Vu. mSleepWatcher: Why didn't I sleep well? In *Proceedings of ISSAT International Conference on Modeling of Complex Systems and Environments (MCSE)*, pages 96–103, 2015.
- [34] A. Nguyen, R. Alqurashi, Z. Raghebi, F. Banaei-kashani, A. C. Halbower, T. Dinh, and T. Vu. In-ear Biosignal Recording System: A Wearable for Automatic Whole-night Sleep Staging. In *Proceedings of the 2016 Workshop on Wearable Systems and Applications (ACM MobiSys-WearSys)*, pages 19–24, 2016.
- [35] R. Oliveira. The dynamic ear canal. *Ballachandra B, ed. The Human Ear Canal. San Diego: Singular Publishing Group*, pages 83–111, 1995.
- [36] R. Oliveira, B. Hammer, A. Stillman, J. Holm, C. Jons, and R. Margolis. A look at ear canal changes with jaw motion. *Ear Hear*, 13(6):464–466, 1992.
- [37] R. Oliveira and G. Hoeker. Ear canal anatomy and activity. In *Pirzanski C, ed. Ear Impressions and New Laser Shell Technology. Seminars in Hear*, pages 265–275, 2003.
- [38] R. Oliveira, V. Kolpe, and G. Hoeker. The dynamic human ear canal: Quantitative changes in canal volume with jaw articulation and its relevance to hearing aid use. In *Poster presented at IHCON, 2002*.
- [39] OpenBCI. <http://openbci.com/>. Accessed: August 03, 2016.
- [40] J. Owens. Insufficient Sleep in Adolescents and Young Adults: An Update on Causes and Consequences. *Pediatrics*, 134(3):921–932, 2014.
- [41] M. Poh, K. Kim, A. Goessling, N. Swenson, and R. Picard. Heartphones: Sensor Earphones and Mobile Application for Non-obtrusive Health Monitoring. In *Proceedings of International Symposium on Wearable Computers (ISWC)*, pages 153–154, 2009.
- [42] M. Poh, K. Kim, A. Goessling, N. Swenson, and R. Picard. Cardiovascular Monitoring Using Earphones and a Mobile Device. In *Proceedings of IEEE Pervasive Computing*, pages 18–26, 2012.
- [43] C. P. Pollak, W. W. Tryon, H. Nagaraja, and R. Dzwonczyk. How Accurately Does Wrist Actigraphy Identify the States of Sleep and Wakefulness? *SLEEP*, 24(8), 2001.
- [44] Polysmith - NIHON KOHDEN. <http://www.nihonkohden.de/>. Accessed: August 03, 2016.
- [45] Types of Sleep Studies. <http://www.nhlbi.nih.gov/health/health-topics/topics/slpst/types>. Accessed: August 03, 2016.
- [46] A. Rechtschaffen and A. Kales. A manual of standardized terminology, techniques, and scoring system for sleep stages of human subjects. *Washington: Public Health Service, US Government Printing Office*, 1968.
- [47] M. Ronzhina, O. Janousek, J. Kolarova, M. Novakova, P. Honzik, and I. Provaznik. Sleep scoring using artificial neural networks. *Sleep Medicine Reviews*, 16(3):251–263, 2012.
- [48] A. Sano, T. Tomita, and H. Oba. Applications using Earphone with Biosignal Sensors. In *Human Interface Society Meeting*, volume 12, pages 1–6, 2010.
- [49] B. Sen, M. Peker, A. Cavusoglu, and F. V. Celebi. A Comparative Study on Classification of Sleep Stage Based on EEG Signals Using Feature Selection and Classification Algorithms. *Journal of Medical Systems*, 38(3):1–21, 2014.
- [50] Sleep Shepherd. <http://sleepshepherd.com/>. Accessed: August 03, 2016.
- [51] D. Shrivastava, S. Jung, M. Saadat, R. Sirohi, and K. Crewson. How to interpret the results of a sleep study. *Journal of Community Hospital Internal Medicine Perspectives*, 4(5), 2014.
- [52] What Are Sleep Studies? <http://www.nhlbi.nih.gov/health/health-topics/topics/slpst/>. Accessed: August 03, 2016.
- [53] Trackit Mark III - LifeLines Neurodiagnostic Systems. <https://www.lifelinesneuro.com/>. Accessed: August 03, 2016.
- [54] R. A. U., O. Faust, N. Kannathal, T. Chua, and S. Laxminarayan. Non-linear analysis of EEG signals at various sleep stages. *Computer Methods and Programs in Biomedicine*, 80(1):37–45, 2005.
- [55] Brain Basics: Understanding Sleep. [http://www.ninds.nih.gov/disorders/brain\\_basics/understanding\\_sleep.htm](http://www.ninds.nih.gov/disorders/brain_basics/understanding_sleep.htm). Accessed: August 03, 2016.
- [56] C. van der Reijden, L. Mens, and A. Snik. Signal-to-noise ratios of the auditory steady-state response from fifty-five EEG derivations in adults. *Journal of the American Academy of Audiology*, 15(10):692–701, 2004.
- [57] T. Virtanen, J. F. Gemmeke, B. Raj, and P. Smaragdis.

- Compositional Models for Audio Processing: Uncovering the structure of sound mixtures. *IEEE Signal Processing Magazine*, 32(2):125–144, 2015.
- [58] S. Vogel, M. Hulsbusch, T. Hennig, V. Blazek, and S. Leonhardt. In-Ear Vital Signs Monitoring Using a Novel Microoptic Reflective Sensor. *IEEE Transactions on Information Technology in Biomedicine*, 13(6):882–889, 2009.
- [59] T. Vollmer, P. Schauerte, M. Zink, S. Glüggler, J. Schiefer, M. Schiek, U. Johnen, and S. Leonhardt. Individualized biomonitoring in heart failure - Biomon-HF "Keep an eye on heart failure - especially at night". *Biomedical Engineering (Berl)*, 59(2):103–111, 2014.
- [60] P. D. Welch. The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Trans. on Audio Electroacoustics*, AU-15(2):70–73, 1967.
- [61] E. Winokur, D. He, and C. Sodini. A wearable vital signs monitor at the ear for continuous heart rate and Pulse Transit Time measurements. In *Proceedings of 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 2724–2727, 2012.
- [62] L. Zoubek, S. Charbonnier, S. Lesecq, A. Buguet, and F. Chapotot. Feature selection for sleep/wake stages classification using data driven methods. *Biomedical Signal Processing and Control*, 2(3):171–179, 2007.