LIBS: A Bioelectrical Sensing System from Human Ears for Staging Whole-Night Sleep Study

By Anh Nguyen, Raghda Alqurashi, Zohreh Raghebi, Farnoush Banaei-Kashani, Ann C. Halbower, and Tam Vu

Abstract

Sensing physiological signals from the human head has long been used for medical diagnosis, human-computer interaction, meditation quality monitoring, among others. However, existing sensing techniques are cumbersome and not desirable for long-term studies and impractical for daily use. Due to these limitations, we explore a new form of wearable systems, called *LIBS*, that can continuously record biosignals such as brain wave, eye movements, and facial muscle contractions, with high sensitivity and reliability. Specifically, instead of placing numerous electrodes around the head, *LIBS* uses a minimal number of custom-built electrodes to record the biosignals from human ear canals. This recording is a combination of three signals of interest and unwanted noise. Therefore, we design an algorithm using a supervised Nonnegative Matrix Factorization (NMF) model to split the single-channel mixed signal into three individual signals representing electrical brain activities (EEG), eye movements (EOG), and muscle contractions (EMG). Through prototyping and implementation over a 30 day sleep experiment conducted on eight participants, our results prove the feasibility of concurrently extracting separated brain, eye, and muscle signals for fine-grained sleep staging with more than 95% accuracy. With this ability to separate the three biosignals without loss of their physiological information, *LIBS* has a potential to become a fundamental in-ear biosensing technology solving problems ranging from self-caring health to non-health and enabling a new form of human communication interfaces.

1. INTRODUCTION

Physiological signals generated from human brain, eye, and facial muscle activities can reveal enormous insight into an individual's mental state and bodily functions. For example, acquiring these biosignals is critical to diagnose sleep quality for clinical reasons, among other auxiliary signals. Even though providing highly reliable brain signal Electroencephalography (EEG), eye signal Electrooculography (EOG), and muscle signal Electromyography (EMG), the gold-standard methodology, referred to as Polysomnography (PSG),[9] has many limitations. Specifically, PSG attaches a large number of wired electrodes around human head, requires an expert sensor hookup at a laboratory, and provides a risk of losing sensor contact caused by body movements during sleep. Consequently, this gold-standard approach is uncomfortable, cumbersome to use, and expensive and time-consuming to set up.

As an effort to overcome the inherent limitations of PSG, there exist various wearable solutions developed to acquire the biosignals with high resolution and easy self-applicability. They involve electrode caps, commercial head-worn devices (e.g., EMOTIV, NeuroSky MindWave, MUSE, Kokoon, Neuroon Open, Aware, Naptime, Sleep Shepherd, etc.), and hearing aid-like research devices.[6, 10] However, these solutions are stiff, unstable, and only suitable for either short-term applications or in-hospital use. In other words, they are still inconvenient and less socially acceptable for outdoor, long-term, and daily activities.

To fill in this gap, we propose a Light-weight In-ear BioSensing (LIBS) system that can continuously record the electrical activities of human brain, eyes, and muscles concurrently using a minimum number of passive electrodes placed invisibly in the ear canals. In this work, particularly, the idea of sensing inside human ears has been motivated from the fact that the ear canals are reasonably close to all sources of the three biosignals of interest (i.e., EEG, EOG, and EMG signals) as shown in Figure 1. Furthermore, physical features of the ear canal allow a tight and fixed sensor placement, which is desirable for electrode stability and long-term wearability. Hence, we carefully develop *LIBS* using very flexible, conductive electrodes to maximize the quality of its contact area with the skin in the wearer's ear canals for good signal acquisition while maintaining a high level of comfort.

Figure 1. Conceptual illustration of *LIBS* and its relative position to the sources of EEG, EOG, and EMG signals.
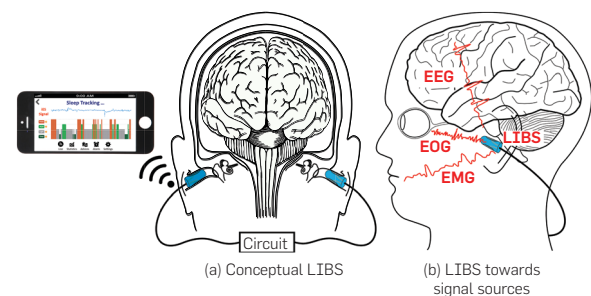
(a) Conceptual LIBS

(b) LIBS towards signal sources

However, as minimizing the number of used electrodes, we can achieve only the single-channel signal, which is a mixture of EEG, EOG, EMG signals, and unwanted noise. We then develop a signal separation model for *LIBS* to extract the three signals of interest from the in-ear mixed signal. To validate the lossless of essential physiological information in the separated signals acquired by *LIBS*, we finally develop a sleep stage classification algorithm to score every 30sec epoch of the separated signals into an appropriate stage using a set of discriminative features obtained from them. Through the hardware prototype and a one-month long user study, we demonstrated that the proposed *LIBS* was comparable to the existing dedicated sleep assessment system (i.e., PSG) in terms of accuracy.

Due to the structural variation across ear canals and overlapped characteristics of the EEG, EOG, and EMG signals, building *LIBS* is difficult because of three following key reasons. (1) The brain signal is quite small in order of micro-Volts (μV). Additionally, the human head anatomy shown in Figure 1(b) indicates that their sources are not too close to the location of *LIBS* placed in the ear canals to be sensed, especially in case of the weak brain source, (2) The characteristics of those three biosignals are overlapped in both time and frequency domains. Moreover, their activation is random and possibly simultaneous during the monitoring period, and (3) The signal quality is easily varied by the displacement of electrodes across device hookups and the variation of physiological body conditions across people. Consequently, our first challenge is to build sensors capable of providing a high level of sensitivity while recording the biosignals from afar and comfort while wearing the device. Our second challenge is then to provide a robust separation mechanism in the presence of multiple variances, which becomes a significant hurdle.

While addressing the above challenges to realizing *LIBS*, we make the following contributions through this work:

1. Developing a light-weight and low-cost earplug-like sensor with highly sensitive and soft electrodes, the whole of which is comfortably and safely placed inside human ears to continuously measure the voltage potential of the biosignals in long term with high fidelity.
2. Deriving and implementing a single-channel signal separation model, which integrates a process of learning source-specific prior knowledge for adapting the extraction of EEG, EOG, and EMG from the mixed in-ear signal to suit the variability of the signals across people and recordings.
3. Developing an end-to-end sleep staging system, which takes the input of three separated biosignals and automatically determines the appropriate sleep stages, as a proof-of-concept of *LIBS*'s potential in reality.
4. Conducting an over 30 day long user studies with eight subjects to confirm the feasibility and learn the usability of LIBS.

2. LIBS'S SYSTEM OVERVIEW

In this section, we present an overall design of *LIBS* in order to achieve the EEG, EOG, and EMG signals individually from the in-ear mixed biosignal. Additionally, we provide a module that automatically determines appropriate sleep stages from *LIBS*'s outputs acquired in sleep studies as its application. Generally, the whole-night sleep staging system, as illustrated in Figure 2, consists of three following primary modules.

2.1. Signal acquisition
Overall, this module focuses on tackling our first challenge 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, and speaking), (2) a potential to acquire the naturally weak biosignals, which have micro-Volt amplitude, and (3) a provision of comfortable and harmless wearing to the users. We fulfill these obstacles by firstly custom-making a deformable earplug-like sensors using a viscoelastic material with atop sensitive electrodes using several layers of thin, soft, and highly conductive materials. To possibly capture the weak biosignals from inside human ears, we then increase the distance between the main electrodes and the reference point to further enhance signal fidelity. Finally, we preprocess the collected signal to eliminate signal interference (e.g., body movement artifact and electrical noise).

2.2. In-ear mixed signal separation
In this module, we form a supervised algorithm to overcome our second challenge for signal separation. This challenge, in detail, is related to (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 supervised Nonnegative Matrix Factorization (NMF)-based model that can separate the preprocessed in-ear mixed signal into EEG, EOG, and EMG with high similarity to the ground truth given by the gold-standard device. Specifically, our separation algorithm initially learns prior knowledge of the biosignals of interest through their individual spectral templates. It then adapts the templates to the variation between people through a deformation step. Hence, the model we build can alter itself slightly to return the best fit between the expected biosignals and the given templates.

2.3. Automatic sleep staging
This last module provides a set of machine learning algorithms to continuously score sleep into appropriate sleep stages using EEG, EOG, and EMG separated from the in-ear mixed signal. Because those signals can have similar

Figure 2. *LIBS* architecture and its sleep staging application.
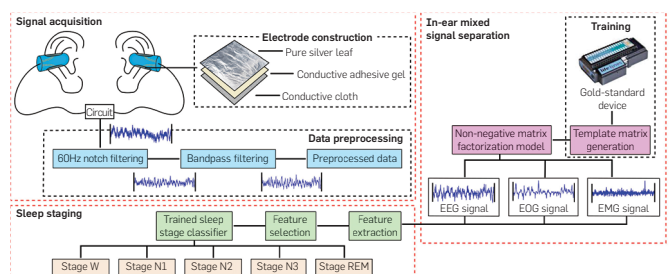
characteristics shared in some of stages, this module is challenging by an ability to (1) find the most informative and discriminative features describing all three biosignals when they are used together and then (2) construct an efficient classifier to perform sleep staging. We introduce a classification model that can automatic score the sleep after well trained. Firstly, we deploy an off-line training stage composing of three steps: feature extraction, feature selection, and model training. Specifically, 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 used in its second stage for on-line sleep stage classification.

3. IN-EAR MIXED SIGNAL ACQUISITION

In this section, we discuss the anatomical structure of human ears that leads to the custom design of *LIBS* sensor as well as its actual prototype using off-the-shelf electrical components.

3.1. Sensor materials

Extensive anatomical study of human ears shows that the form of ear canal is easy to be affected when the jaw moves.[15] More remarkably, a person can have asymmetry between his left and right ears.[14] Beyond those special characteristics, to capture the good signals, it is important to eliminate a gap between the electrodes and human skin due to the nature of the ion current generated by the biosignals. Hence, *LIBS* sensor needs to flexibly reshape itself, well contact to the skin, well fit different ear structures and types of muscle contractions, and comfortably be worn in long term. One possible approach is to personalize a mold. However, this approach entails high cost and time consume. Therefore, a commercial earplug with noise-cancelled and flexible wires are offered to form the sensor prototype. Specifically, we have augmented an over-the-counter sound block foam earplug for its 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. As an additional bonus, the soft surface and the light-weight property of the earplug make itself more convenient to be worn without much interference and to block out noise during sleep for our case study.

3.2. Electrode construction and placement

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. We firstly tried different conductive materials as shown in Figure 3. However, our experiment

Figure 3. Prototypes with different conductive materials.
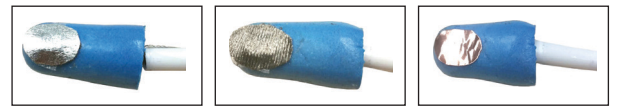
(a) Silver coated fabric electrode **(b)** Fabric electrode **(c)** Copper electrode

resulted that copper is a hard material to be inserted into and placed inside the ear without harm. Oppositely, conductive fabric is a good choice that neither harms the in-ear skin nor is broken while being squeezed. However, because of the weave pattern of its fibers, which cases a non-identical resistivity (19Ω/sq) on the surface, we further coat their surface with many layers of thin pure silver leaves, which gives low and consistent surface resistance for providing reliable signals. Also, a very small amount of health-grade conductive gel is added. In Figure 2, the construction module shows the comprehensive structure of *LIBS* electrodes. Ultimately, we place the active and reference electrodes in two separate ear canals, hence intensify the potential of the signals by a distance increase. Finally, the recorded signal is transferred to an amplifier through shielded wires to prevent any external noise.

3.3. Microcontroller

In this prototype, we use a general brain-computer interface board manufactured by OpenBCI[16] group to sample and digitize the signal. The board is supplied by a battery source of 6V for safety and configured at a 2kHz sampling rate and a 24dB gain. The signal is stored in an on-board mini-SD card while recording and then processed offline in a PC.

4. NMF-BASED SIGNAL SEPARATION

Due to the limited cavity of the ear canal, the biosignal recorded by *LIBS* is inherently a single-channel mixture of at least four components including EEG, EOG, EMG signals, and unwanted noise. We assume that the mixed signal is a linear combination of aforementioned signals generated from a number of individual sources in the spectral domain,[7] which we mathematically express in Equation (1).

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

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

Generally, the problem of separating original signals from their combinations generated by concurrent multi-source activation has long been addressed for different systems. The classical example of this problem is the auditory source separation problem, also called a cocktail party problem, where various algorithms have been developed to extract individual voices of a number of people talking simultaneously in a room. Additionally, the problem of decoding a set of received signals to retrieve the orginal signals transmitted by multiple antennas via Multi-Input and Multi-Output (MIMO)[22] in wireless communication can also be another example. Although there exist mainstream techniques[25] such as Principal Component Analysis (PCA),

Independent Component Analysis (ICA), Empirical Mode Decomposition (EMD), Maximum Likelihood Estimation (MLE), and Nonnegative Matrix Factorization (NMF) built to solve the blind source separation problem, most of them 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 known or selected manually. As a result, it is impossible to directly apply them in our work since their first constraint conflicts with the fact that *LIBS* has only one channel, which is fewer than the number of signals of interest (three signals).

To successfully address this challenge, we propose a novel source separation technique that takes advantage of NMF. However, there have existed two potential issues with a NMF-based model that might degrade the quality of the decomposed signals. They include (1) the inherent non-unique estimation of the original source signals (ill-posed problem) caused by the non-convex solution space of NMF and (2) the variance of the biosignals on different recordings. To solve them, our proposed NMF-based model is combined with source-specific prior knowledge learnt in advance for each user through a training process. Figure 4 demonstrates the high-level overview of this process, which leverages two different NMF techniques to learn source-specific information and to separate the mixing in-ear signal based on priory training.

Particularly, when a new user starts using *LIBS*, his groundtruth EEG, EOG, and EMG are shortly acquired using the gold-standard device (i.e., PSG) and fed into a single-class Support Vector Machine (SVM)-based NMF technique (SVM-NMF)[3] to build a personal spectral template matrix, called W, representing their basis patterns. Then, for any in-ear signal $\tilde{X}$ recorded by *LIBS*, our trained model approximately decomposes its power spectrum X into two lower rank nonnegative matrices

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

in which $X \in \Re^{m \times n}$ comprises m frequency bins and n temporal frames; W is calculated in advance and given; 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 in Equation (3).

$$\left\{ \hat{W}, \hat{H} \right\} = \underset{W, H \geq 0}{\arg\min}\ d(X \| WH) \qquad (3)$$

We solve Equation (3) using multiplicative update rules to achieve a good compromise between the speed and the ease of implementation. While solving this equation, the template matrix taken from the learning process is used to initialize W. Hence, W is deformed to fit the in-ear signal acquired from that user at different nights.

Algorithm 1 Signal Separation Algorithm

```
 1:  Input:
 2:      IS - In-ear Signal
 3:      W_ini - Spectral Template Matrix
 4:      ST - Segment Time
 5:  Output:
 6:      X̂_EEG, X̂_EOG, X̂_EMG - Separated Signals
 7:
 8:  X̃ ← PreprocessSignal(IS);
 9:  X ← ComputePowerSpectrum(X̃);
10:  Seg ← SegmentSignal(IS, ST);
11:  for i = 1 → sizeof(Seg) do
12:      H_ini ← InitializeMatrixRandomly();
13:      [Ŵ, Ĥ] ← IS_NMF(Seg_i);
14:      V_EEG(Seg_i) ← Ŵ_EEG(Seg_i) * Ĥ_EEG(Seg_i);
15:      V_EOG(Seg_i) ← Ŵ_EOG(Seg_i) * Ĥ_EOG(Seg_i);
16:      V_EMG(Seg_i) ← Ŵ_EMG(Seg_i) * Ĥ_EMG(Seg_i);
17:  X̂_EEG ← rescontructSignal(X, V_EEG);
18:  X̂_EOG ← rescontructSignal(X, V_EOG);
19:  X̂_EMG ← rescontructSignal(X, V_EMG);
```

In this work, adapting the technique from Ref. Damon et al.,[2] we specifically select the Itakura-Saito (IS) divergence d_{IS} as a measure to minimize the error between the power spectrum of the original signal and its reconstruction from W and H. The IS divergence, in detail, is a limit case of the β-divergence introduced in Ref. Févotte and Idier,[4] which is defined here

$$d_\beta(x \,|\, y) = \begin{cases} \dfrac{1}{\beta(\beta-1)}(x^\beta + (\beta-1)y^\beta - \beta x y^{\beta-1}) & \beta \in \mathbb{R} \setminus \{0, 1\} \\[2ex] x(\log x - \log y) + (y - x) & \beta = 1 \\[2ex] \dfrac{x}{y} - \log\dfrac{x}{y} - 1 & \beta = 0 \end{cases} \qquad (4)$$

The reason is that a noteworthy property of the β-divergence (in which the IS divergence corresponds to the case $\beta = 0$) is its behavior w.r.t scale. Alternatively, IS divergence holds a scale-invariant property $d_{IS}(\lambda x \,|\, \lambda y) = d_{IS}(x \,|\, y)$ that helps minimize the variation of the signals acquired from one person in different recordings. The IS divergence is given by,

$$d_{IS}\left(X \,|\, WH\right) = \frac{X}{WH} - \log\frac{X}{WH} - 1 \qquad (5)$$

Hence, Algorithm 1 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.

5. SLEEP STAGES CLASSIFICATION

Human sleep naturally proceeds in a repeated cycle of four distinct sleep stages: N1, N2, N3, and REM sleep. To study the sleep quantity and quality, the sleep stages are mainly identified by simultaneously evaluating three fundamental

---

Figure 4. Overview of signal separation process in *LIBS*.
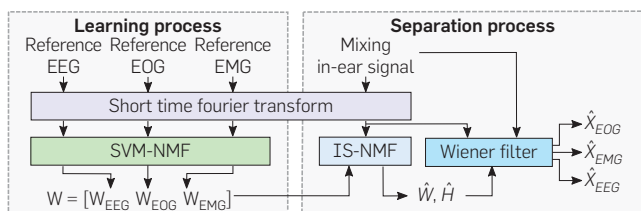
measurement modalities including brain activities, eye movements, and muscle contractions. In hospital, an expert can visually inspect EEG, EOG, and EMG signals collected from subjects during sleep and label each segment (i.e., a 30sec period) with the corresponding sleep stage based on known visual cues associated with each stage. Below we elaborate on each of aforementioned steps of our data analysis pipeline.

5.1. 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.

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,[11] 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.[23]

For this study, relied on the literature of feature-based EOG, EMG, and EEG classification,[11] we consider the features listed in Table 1 from each of the aforementioned categories.

Table 1. List of features extracted from the biosignals.

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

5.2. 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[19] 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 Sequential Forward Selection (SFS)[26] to identify the most effective combination of features extracted from our in-ear signal. With SFS, 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.

5.3. Sleep stage classification
Various classification methods are proposed in the literature for similar applications and each has advantages and disadvantages. Some scholars[11] 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 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[18] 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,[19] 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 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 SFS are used to construct a multitude of decision trees at training stage to identify the corresponding sleep stage for every 30sec segment of the biosignals in the classification stage.

6. EVALUATION

In this section, we first present the key results in proving the feasibility of *LIBS* to capture the usable and reliable bio-signals, in which all EEG, EOG, and EMG is present. From the success of our proof-of-concept, we then show the performance of our proposed separation algorithm for splitting those three signals without loss of information. Finally, we evaluate the usability of *LIBS*'s outputs through the performance of the automatic sleep stage classification.

6.1. Experiment methodology

Beyond our *LIBS* prototype shown in Figure 5, we used a portable PSG device named Trackit Mark III supported by LifeLines Neurodiagnostic Systems Inc. company[21] 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 acquires EEG, EOG, and EMG signals at 256Hz sampling rate and pre-filtered them in the range of 0.1–70Hz.

6.2. Validation of signal presence

In this evaluation, we assess the presence of the signals of interest in the in-ear mixed signal measured by *LIBS* by comparing the recording with the groundtruth signals acquired from the gold-standard PSG channels. While the user wears both devices at the same time, we illustrate the feasibility of *LIBS* to produce the usable and reliable signals through different experiments.

We first examined if *LIBS* can capture the EMG signal by asking a subject to do two different activities for contracting his facial muscles. Specifically, the subject kept his teeth remaining still and then grinding for 5sec and chewing for 20sec continuously. This combination was done for four times. From Figure 6a, we noticed that our *LIBS* device could clearly capture those events reflecting the occurrence of the EMG signal.

Similarly, we asked the subject to look forward for 20sec and then move his eyes to points pre-specified in four directions (i.e., left, right, up, and down) for 5sec. As a result, although the amplitude of the in-ear mixed signal is smaller than the gold-standard one, it still clearly exhibits the left and right movements of the eyes similar to the EOG signal channeled in the gold-standard device. As shown in Figure 6b, *LIBS* also has the ability to capture the horizontal and vertical eye movements as the reflection of EOG occurrence.

Figure 5. Demonstration of a sleep study and the first prototype of *LIBS*.
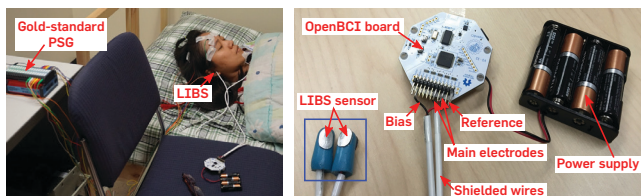
On the other hand, we conducted the following standard Brain-Computer Interaction (BCI) experiments to verify the occurrence of the EEG signal in *LIBS*'s recordings:

Auditory Steady-State Response (ASSR). This EEG paradigm measures the response of human brain while modulating auditory stimuli with specific frequency ranges.[24] In this experiment, we applied auditory stimuli in the frequencies of 40Hz in which each stimuli lasted for 30sec and was repeated three times with 20sec rest between them. Then, by looking at Figure 7, it is easy to recognize a sharp and dominant peak at 40Hz produced during the 40Hz ASSR experiment. Clearly, this result demonstrates the ability of *LIBS* to capture such the specific frequency in the in-ear mixed signal although the peak extracted from the gold standard electrodes was larger than that of *LIBS* electrode.

Steady-State Visually Evoked Potential (SSVEP). Similar to ASSR, SSVEP measures the brain wave responding to a visual

Figure 6. The detection of (a) muscle activities and (b) eye movements from *LIBS* (top) and the gold standard EMG and EOG channels (bottom), respectively.
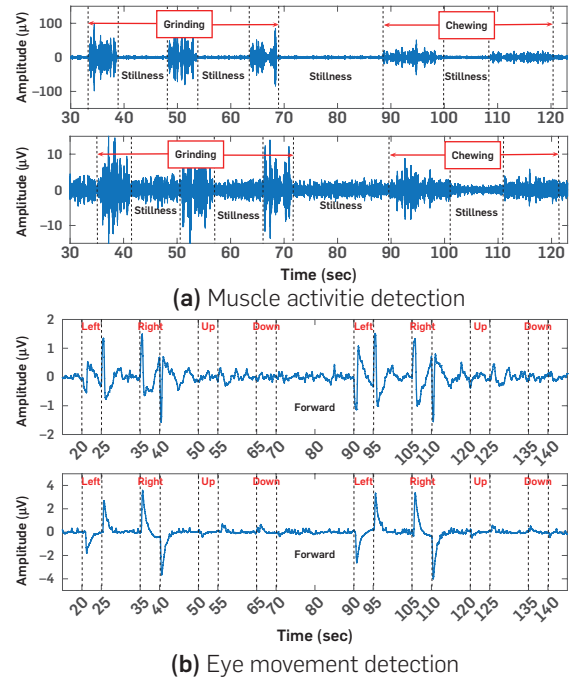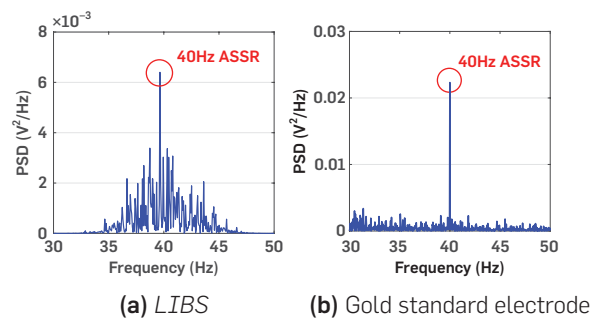

(a) Muscle activitie detection

(b) Eye movement detection

Figure 7. The ASSR for 40Hz recorded from (a) *LIBS* and (b) the gold standard device at Channel C3 on scalp.

(a) *LIBS*   **(b)** Gold standard electrode

stimuli at specific frequencies.[12] Particularly, we created a blinking stimuli at 10Hz and played it for 20sec with three time repetition. Accordingly, the brain response in this SSVEP experiment comprehensibly presented as a dominant peak for *LIBS* and the gold standard on-scalp electrodes in Figure 8.

Alpha Attenuation Response (AAR). Alpha wave is a type of brain waves specified in the range of 8–13Hz. This brain wave is a sign of relaxation and peacefulness.[1] In this experiment, we asked the subject to completely relax his body while closing his eyes for 20sec and then open them for 10sec in five consecutive times. As analyzing the recorded in-ear mixed signal, Figure 9 shows that *LIBS* is able to capture the alpha rhythm from inside the ear. However, the detection of alpha rhythm in case of *LIBS* was 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.

6.3. Signal separation validation
From the previous experiments, we proved that all of the EEG, EOG, and EMG signals appeared in the recordings of *LIBS* and were mixed in the original in-ear signal. 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 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
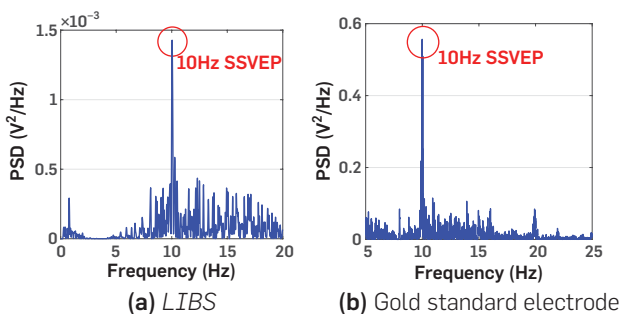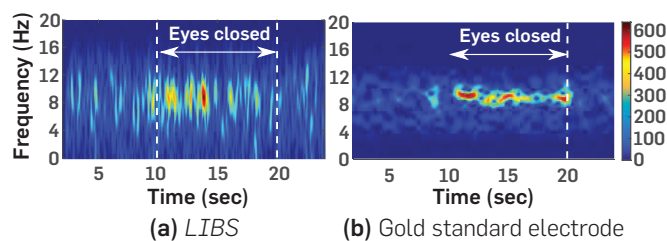
by analyzing the occurrence of special frequencies (i.e., the delta brain wave) in the separated EEG biosignal during the sleep study.

Specifically, Figure 10a provides the spectrogram of a 30sec original in-ear mixed signal captured by *LIBS* during a sleep study and labeled as stage Slow-Wave Sleep (SWS) by the gold-standard device. In Figure 10b, the spectrogram of a corresponding 30sec ground-truth EEG signal is presented. By observing the second spectrogram, a delta brain wave in a frequency range lower than 4Hz is correctly found. However, the spectrogram in Figure 10a cannot show the detection of such the brain wave clearly. Its reason is that not only the delta brain wave exists but also other biosignals are added in this original signal. Finally, Figure 10c exhibits the spectrogram of the EEG signal separated from the original mixed signal by applying our proposed signal separation algorithm. Analyzing this figure proves that the separation model we propose has a capability of not only splitting the signals from the mixed one but keeping only the specific characteristics of the separated signal as well. Otherwise, the short appearance of the delta brain wave in the decomposed signal can be explained by the fact that the location where *LIBS* is placed is far from the source of the signal. By that, the amplitude of the signal is highly reduced.

6.4. Sleep stage classification evaluation
To evaluate the performance of our proposed sleep staging method, we conducted a 38hrs of sleep experiments over eight graduate students (three females and five 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 Institutional Review 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. After that, the Polysmith program[17] was run to score the ground-truth signals into different sleep stages at

**Figure 8. The SSVEP responses recorded from (a) *LIBS* and (b) the gold standard device at Channel O1 on scalp.**

(a) *LIBS*   **(b)** Gold standard electrode

Figure 9. The detection of alpha rhythms from (a) *LIBS* and (b) the gold standard device at Channel C4 on scalp.

(a) *LIBS*   **(b)** Gold standard electrode

Figure 10. Signal separation performance obtained by *LIBS* through a 30sec mixed in-ear signal (a) and compared with the ground-truth EEG signal (b), and its corresponding separated EEG signal (c).

(a) Mixed in-ear signal

(b) Ground-truth EEG signal

(c) EEG signal separated from in-ear signal
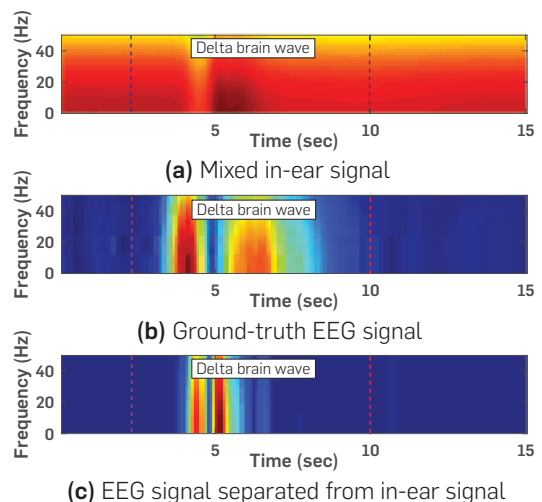
every 30sec segment. For all studies, the sleeping environment was set up to be quiet, dark, and cool.

Statistically, we extracted the features from 4313 30sec segments using the original mixed signal as well as three separated signals. Training and test data sets are randomly selected from the same subject pool. Figure 11 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, we observe that the dynamics of the hypnogram is almost completely maintained in the predicted scores. Moreover, our result show that the end-to-end sleep staging system can achieve 95% accuracy on average.

We refer the readers to Ref. Nguyen et al.[13] for more detailed validations of signal acquisition and separation, their comparison with the signals recorded by the gold-standard device, and our user study.

7. POTENTIALS OF LIBS

We envision *LIBS* to be an enabling platform for not only healthcare applications but also those from other domains. Figure 12 illustrates the eight potential applications including in-home sleep monitoring, autism onset detection, meditation training, eating habit monitoring, autonomous audio steering, distraction and drowsiness detection, child's interest assessment, and human-computer interaction. We discuss these exemplary applications below.

7.1. Healthcare applications

We propose three applications that *LIBS* can be extended to serve in healthcare: autism act-out onset detection, meditation coaching, and eating habit monitoring.

Autism onset detection. Thanks to its ability to capture muscle tension, eye movements, and brain activities, *LIBS* has a potential to be an autism on-set detection and prediction wearable. Particularly, people with autism can have very sensitive sensory (e.g., visual, auditory, and tactile) functions. When any of their sensory functions leads to an overload, their brain signal, facial muscle, and eye movement are expected to change significantly.[5] We hope to explore this phenomenon to detect the relationship between these three signals and the on-set event from which a prediction model can be developed.

Meditation training. Meditation has a potential for improving physical and mental well-being when it is done in a right way. Hence, it is necessary to understand people's mindfulness level during the meditation to be able to provide more efficient instructions. Existing commercial off-the-shelf devices (e.g., MUSE) only capture the brain signal to tell how well users are meditating. Different from them, *LIBS* further looks at the eye and muscle signals to analyze the level of relaxation they have more accurately. As a result, *LIBS* promisingly helps improve the users' meditation performance.

Eating habit monitoring. Eating habits can provide critical evidences for various diseases.[8] As *LIBS* can capture the muscle signal very clearly, such information can be useful to infer how often the users chew, how fast they chew, how much they chew, and what the intensity of their chewing is. From all of that, *LIBS* can then predict what foods they are eating as well as how much they are eating. As a result, *LIBS* can provide users guidance to avoid their bad habits by themselves or to visit a doctor if necessary.

7.2. Non-health applications

LIBS can benefit applications and systems on other domain such as improving hearing aid devices, improving driver's safety, helping parents with orienting their child early on.

Autonomous audio steering. This application helps solve a classical problem in hearing aid, which is called cocktail party problem. As known, state-of-the-art hearing aid devices try to amplify the sounds coming from the area that has large amplitude, which is assumed as human voice, in a party. Consequently, the hearing aid will fail to support the wearers if any group of people behind them is talking very loudly, which is not the right person they want to talk to. Using our technology, using the eye signal *LIBS* can capture, it will possibly detect the area that the users are paying their attention to. Furthermore, combining with their brain signal, *LIBS* can further predict how please the wearers are with the output sound that their hearing aid is producing. With that in mind, *LIBS* can steer the hearing aid and improve its quality of amplification so that the hearing aid can provide the high-quality sounds coming from the right source to the users.

Distraction and drowsiness detection. Distraction and drowsiness are very serious factors in driving. Specifically, if people feel drowsy, their brain signal will be in alpha state, their eyes will be closed, and their chin muscle tone will become relax.[20] Also, it is easy to detect if people are distracted based on the localization of eye positions when we analyze the changes of the eye signal. Hence, *LIBS* with three separated brain, eye, and muscle signals should be able to determine the driver's drowsiness level or distraction to further send an alert for avoiding road accidents.

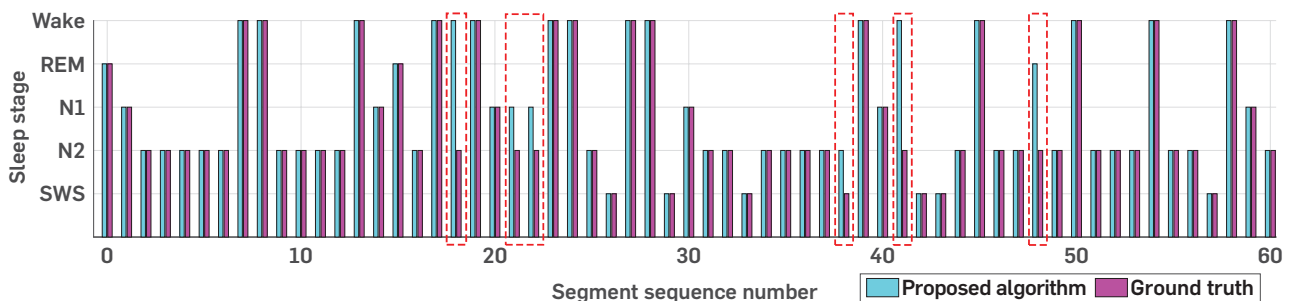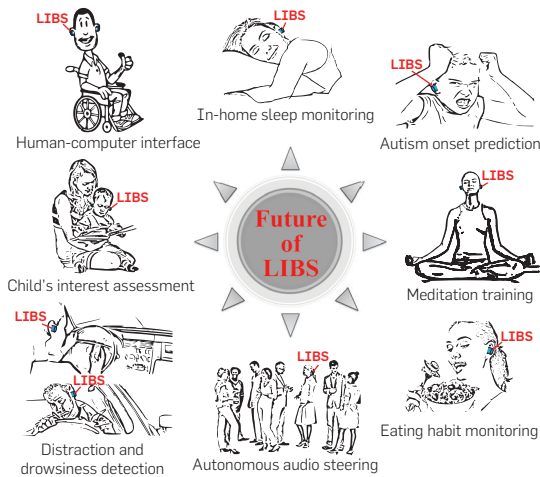**Figure 11. A hypnogram of 30min data resulted by our classification algorithm.**

Figure 12. Potential applications of *LIBS*.

Child's interest assessment. With *LIBS*, child's interest assessment can be done less obtrusively and yield more accurate outcomes. Moreover, from that, the parents will be able to orient them accordingly so that they can learn what they like the most. Clinically, kids from the age of 0–2yrs don't have the ability to express their interest. More precisely, the only way to express their interest is their crying. As a result, the conventional gold-standard device (i.e., PSG) is usually used to read their biosignals, which relatively reflect their interest in what they are allowed to do. However, it is not comfortable for them to wear and do activities during the assessment. Hence, by leveraging *LIBS* to read the signal from their ears and at the same time letting them play different sports or learn different subjects, *LIBS* should be able to infer what the level of their interest is with high comfort.

Human-computer interaction. In a broader context, *LIBS* can be used as a form of Human Computer Interaction (HCI), which can especially benefit users with disability. In stead of using only the brain signal as found in many HCI and brain-to-computer systems today, *LIBS* can combine the information extracted from the three separated signals to enrich commands the user can build to interact with the computer in a more reliable way. This gives users more choices for integration with computing systems in a potentially more precise and convenient manner.

8. CONCLUSION
In this paper, we enabled *LIBS*, a sensing system worn inside human ear canals, that can unobtrusively, comfortably, and continuously monitor the electrical activities of human brain, eyes, and facial muscles. Different from existing hi-tech systems of measuring only one specific type of the signals, *LIBS* deploys a NMF-based signal separation algorithm to feasibly and reliably achieve three individual signals of interest. Through one-month long user study of collecting the in-ear signals during sleep and scoring them into appropriate sleep stages using a prototype, *LIBS* itself demonstrated a promising comparison to the existing dedicated sleep assessment systems in term of

accuracy and usability. Further than an in-ear bio-sensing wearable, we view *LIBS* as a key enabling technology for concealed head-worn devices for healthcare and communication applications, especially for personalized health monitoring, digital assistance, and the introduction of socially-aware human-computer interfaces.

References
1. Alloway, C., et al. The alpha attenuation test: assessing excessive daytime sleepiness in narcolepsy-cataplexy. *Sleep 20*, 4 (1997), 258–266.
2. Damon, C., et al. Non-negative matrix factorization for single-channel EEG artifact rejection. In *ICASSP* (2013), 1177–1181.
3. Essid, S. A single-class SVM based algorithm for computing an identifiable NMF. In *Proceedings of 2012 IEEE ICASSP* (2012), 2053–2056.
4. Févotte, C., Idier, J. Algorithms for nonnegative matrix factorization with the beta-divergence. *Neural. Comput. 23*, 9 (2011), 2421–2456.
5. Gillingham, G. *Autism, Handle with Care!: Understanding and Managing Behavior of Children and Adults with Autism*. Future Education, Texas (1995).
6. Goverdovsky, V., et al. In-ear EEG from viscoelastic generic earpieces: Robust and unobtrusive 24/7 monitoring. *IEEE Sens. J.* (2015).
7. Hallez, H., et al. Review on solving the forward problem in EEG source analysis. *J. Neuroeng. Rehabil. 4*, 1 (2007), 1–29.
8. Kalantarian, H., et al. Monitoring eating habits using a piezoelectric sensor-based necklace. *Comput. Biol. Med. 58*, Supplement C (2015), 46–55.
9. Kushida, C., et al. Practice parameters for the indications for polysomnography and related procedures: an update for 2005. *Sleep 28*, 4 (2005), 499–521.
10. Lee, J.H., et al. CNT/PDMS-based canal-typed ear electrodes for inconspicuous EEG recording. *J. Neural Eng. 11*, 4 (2014).
11. Motamedi-Fakhr, S., et al. Signal processing techniques applied to human sleep EEG signals—A review. *Biomed. Signal Process. Control 10* (2014), 21–33.
12. Nawrocka, A., Holewa, K. Brain—Computer interface based on Steady—State Visual Evoked Potentials (SSVEP). In *Proceedings of 14th ICCC* (2013), 251–254.
13. Nguyen, A., Alqurashi, R., Raghebi, Z., Banaei-kashani, F., Halbower, A.C., Vu, T. A lightweight and inexpensive in-ear sensing system for automatic whole-night sleep stage monitoring. *SenSys '16* (2016), 230–244.
14. Oliveira, R. The dynamic ear canal. *Ballachandra B, ed. The Human Ear Canal. San Diego: Singular Publishing Group* (1995), 83–111.
15. Oliveira, R., et al. A look at ear canal changes with jaw motion. *Ear Hear 13*, 6 (1992), 464–466.
16. OpenBCI. http://openbci.com/.
17. Polysmith–NIHON KOHDEN. http://www.nihonkohden.de/.
18. Ronzhina, M., et al. Sleep scoring using artificial neural networks. *Sleep Med. Rev. 16*, 3 (2012), 251–263.
19. Sen, B., et al. A Comparative Study on Classification of Sleep Stage Based on EEG Signals Using Feature Selection and Classification Algorithms. *J. Med. Syst. 38*, 3 (2014), 1–21.
20. Silber, M.H., et al. The visual scoring of sleep in adults. *J. Clin. Sleep Med. 3*, 02 (2007), 121–131.
21. Trackit Mark III—LifeLines Neurodiagnostic Systems. https://www.lifelinesneuro.com/.
22. Tse, D., Viswanath, P. *Fundamentals of Wireless Communication*. Cambridge University Press, New York (2005).
23. R.A.U., et al. Non-linear analysis of EEG signals at various sleep stages. *Comput. Methods and Programs Biomed. 80*, 1 (2005), 37–45.
24. van der Reijden, C., et al. Signal-to-noise ratios of the auditory steady-state response from fifty-five EEG derivations in adults. *J. Am. Acad. Audiol. 15*, 10 (2004), 692–701.
25. Virtanen, T., et al. Compositional Models for Audio Processing: Uncovering the structure of sound mixtures. *IEEE Signal Processing Mag. 32*, 2 (2015), 125–144.
26. Zoubek, L., et al. Feature selection for sleep/wake stages classification using data driven methods. *Biomed. Signal Process. Control 2*, 3 (2007), 171–179.

Anh Nguyen and Tam Vu, University of Colorado, Boulder, CO, USA.

Raghda Alqurashi, Zohreh Raghebi, and Farnoush Banaei-Kashani, University of Colorado, Denver, CO, USA.

Ann C. Halbower, University of Colorado, School of Medicine, Aurora, CO, USA.