EarSleep:
In-ear Acoustic-based Physical and Physiological Activity Recognition for Sleep Stage Detection

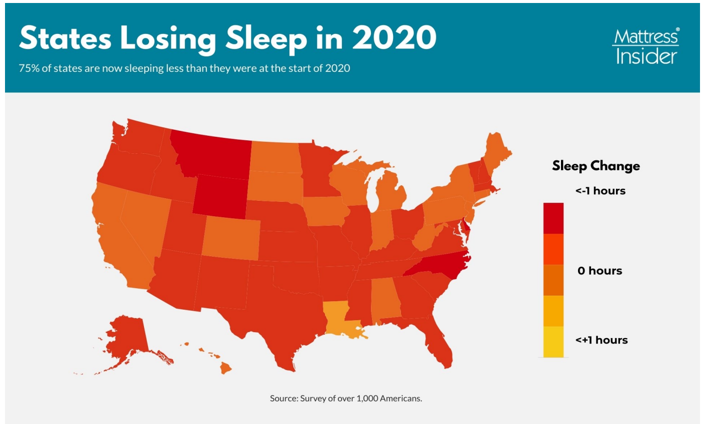
Feiyu Han[1,2], Panlong Yang[2], Yanhao Feng[3], Weiwei Jiang[2], Youwei Zhang[4], and Xiang-Yang Li[1].

[1] University of Science and Technology of China (*USTC*), China
[2] Nanjing University of Information Science and Technology (*NUIST*), China
[3] The Hong Kong Polytechnic University (*PolyU*), China
[4] The University of Electro-Communications (*UEC*), Japan

Presenter: Feiyu Han

Background: Sleep Health Importance

- **Sleep-related diseases** are considered an under-recognized global public health issue and have become **one of the risk factors that seriously threaten public health.**

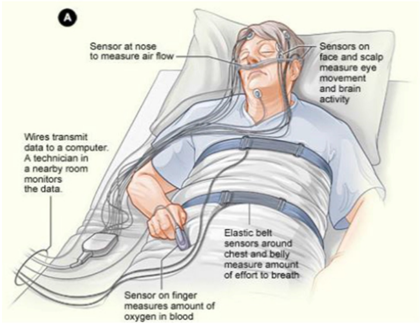


- **50 to 70 million Americans** have some type of sleep disorder[1].
- Nearly **300 million** people in China have poor sleep quality, and **67.24%** of them suffer from insomnia symptoms[2].

[1] https://cfah.org/sleep-statistics/
[2] https://pubmed.ncbi.nlm.nih.gov/38429554/

Sleep Monitoring Technology

- **Traditional approach**



Polysomnography (PSG)

- Fine-grained monitoring in clinical scenarios.
- **Complex** operations and **High** costs

- **Ubiquitous approaches**

Wireless-based

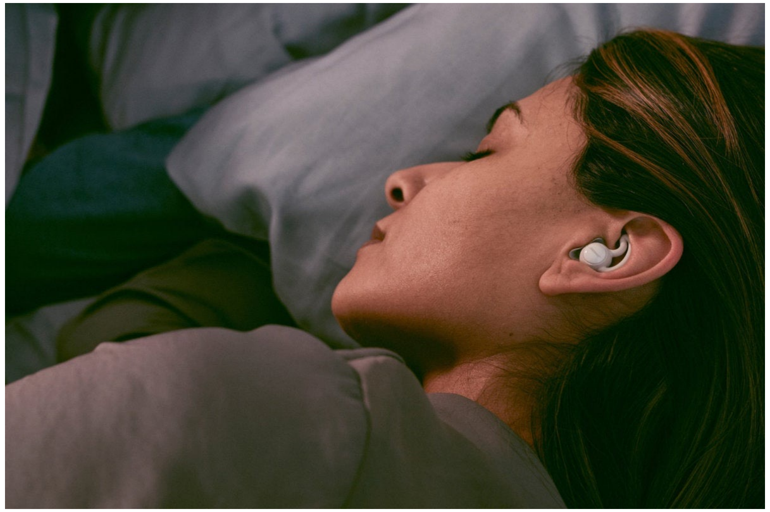
Wearable-based

Mobile-based

- Requiring multi-sensor fusion.
- Only detecting **limited sleep activities**
- Coarse-grained sleep monitoring.
- Wearing discomfort
- ……

Sleep Earbuds

◆ **Earbuds for Sleep**: with the expansion of the sleep economy market, sleep earbuds market is valued at approximately $15 million in 2020



Advantages:

a) Wearing comfort for users:
   - Ergonomic shape design, soft silicone material.

b) Noise isolation:
   - Providing a quiet sleep environment.

c) Ideal positions for measuring physiological parameters.

Becoming the most popular sleep aid tools

- Relying on dedicated biosensors (PPG, ECG), making the cost expensive (above 100 $)

Sleep Earbuds

◆ **Earbuds for Sleep:** with the expansion of the sleep economy market, sleep earbuds market is valued at approximately $15 million in 2020

Advantages:

a) Wearing comfort for users:

b) Noise isolation:

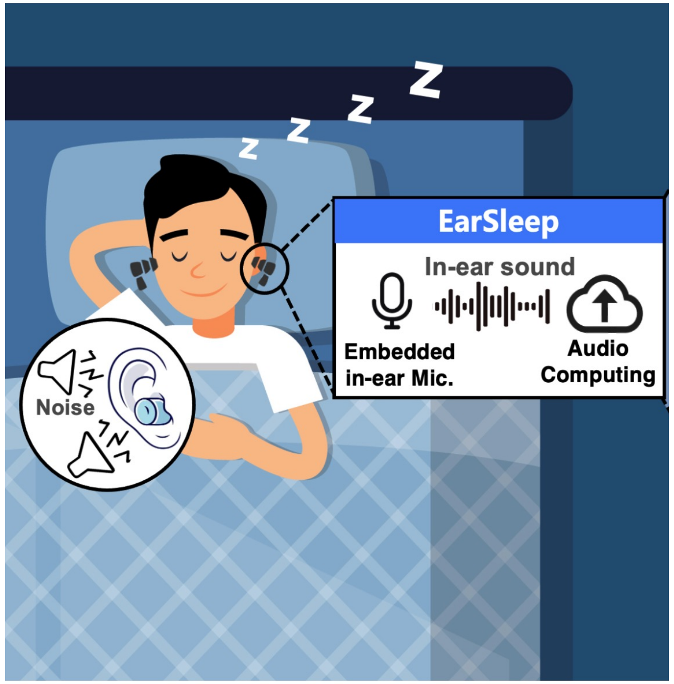
c) Ideal positions for measuring physiological parameters.

Becoming the most popular sleep aid tools

- Relying on dedicated biosensors (PPG, ECG), making the cost expensive (above 100 $)

Recognizing a wide range of sleep activities and achieving fine-grained sleep monitoring in the ubiquitous way

Our Solution: EarSleep



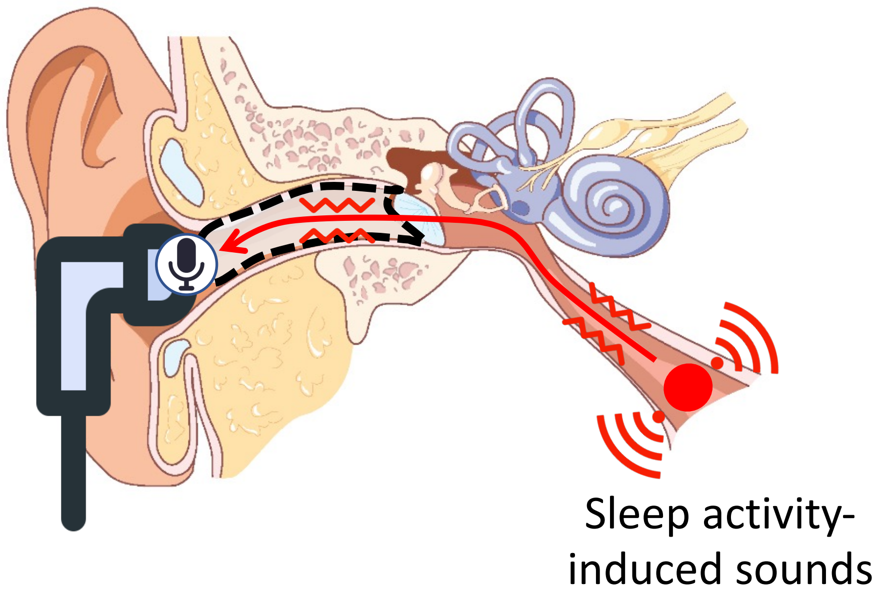
What can EarSleep do!		
Physical Activity	**Body movement**	Turning head, body trembling, limb movement, and body rollover.
	Sound activity	Snore, cough, and somniloquy
Physiological activity		Respiration and heartbeat
Sleep stage		Light, deep, and REM

- EarSleep is built on a pair of **sleep earbuds** with **in-ear microphones**.
- EarSleep can achieve **physical activity recognition** (four-class body movements and three-class sound activities), **physiological activity estimation** (heartbeat and breathing), and **sleep stage detection**.

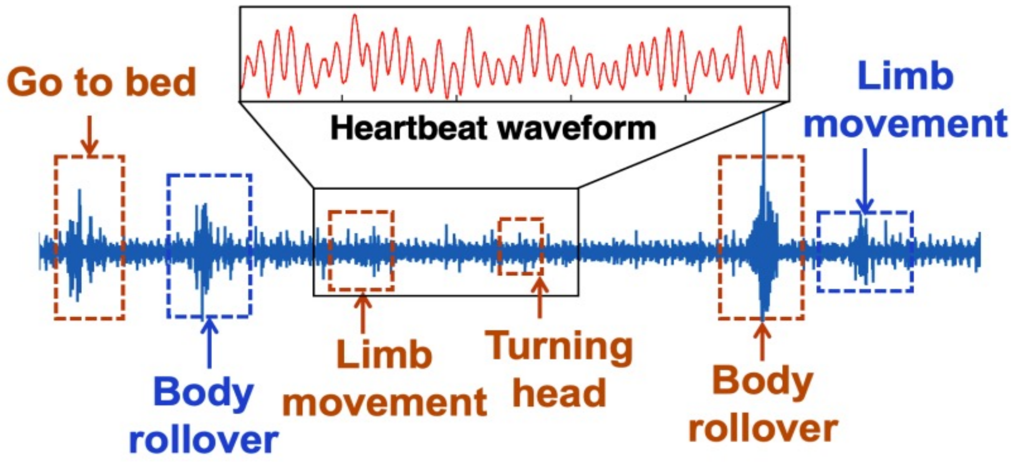
Sensing Principle

- **Body sounds** induced by sleep activities propagate through bone conduction to the ear canal and can be captured by the **in-ear microphone.**

(1) In-ear audio signals generation and propagation

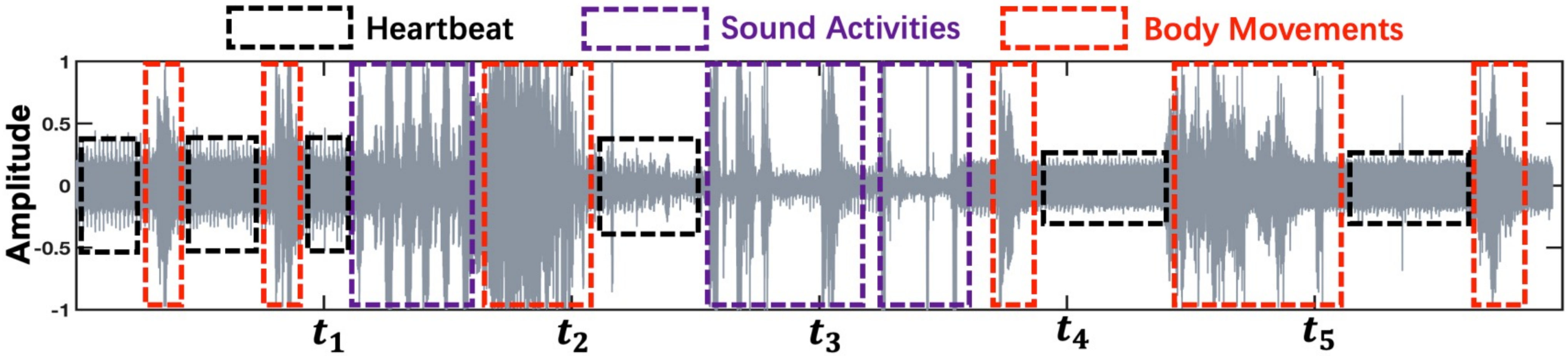


(2) In-ear audio signals captured by in-ear Mic.



Technical Challenge-1

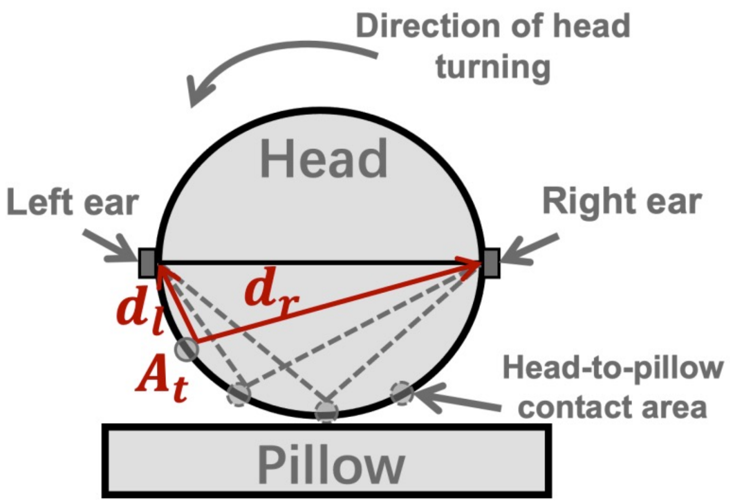
- How to accurately recognize **diverse physical activities** during sleep with only **a single audio modality?**



Sleep is a **continuous and long-term** process and various physical activities that have **behavioral patterns** such as intensity, duration, and periodicity occur.

Unique Acoustic Analysis

- **Taking an example of turning head**



A single channel modeling

$$S_{in}(t,f) = H_{oe}(Vib(t,f) * e^{-(\alpha_r(f)+\alpha_s(f)+\alpha_a(f))*d})$$

Left/right channel joint modeling

$$\frac{S_{inR}(t,f)}{S_{inL}(t,f)} = \frac{H_{oeR}(Vib(t,f) * e^{-(\alpha_r(f)+\alpha_s(f)+\alpha_a(f))*d_R}}{H_{oeL}(Vib(t,f) * e^{-(\alpha_r(f)+\alpha_s(f)+\alpha_a(f))*d_L}}$$

$$\approx H * e^{(\alpha_r(f)+\alpha_s(f)+\alpha_a(f))*(d_L-d_R)}$$

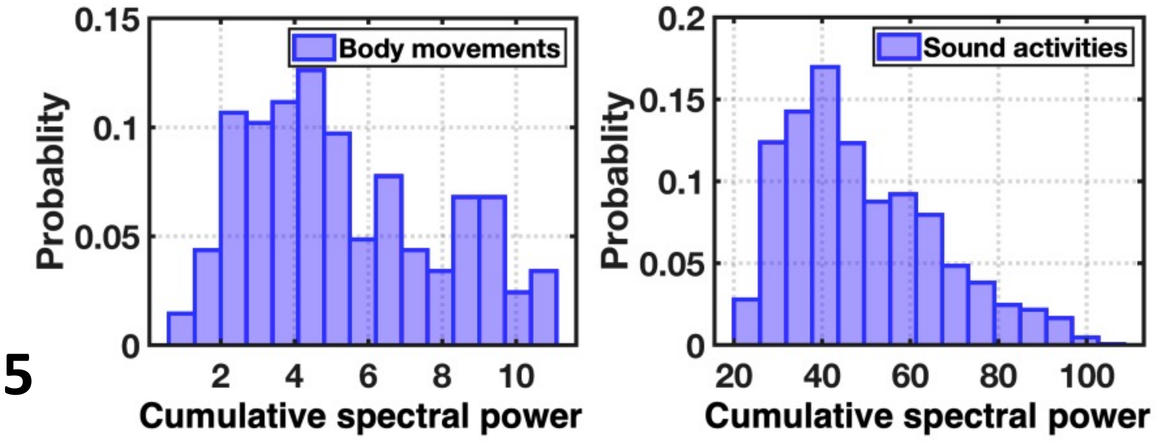
Since the **physical mechanisms** of activities are different, **acoustic attenuation** and **spectral distribution** of physical sleep activities are also distinct, which can inspire us to recognize various physical activities based on a single acoustic modality.

Physical Activity Recognition

① Event Type Identification:

Is Body Movement or Sound Activity?

Threshold of cumulative spectral power : 15



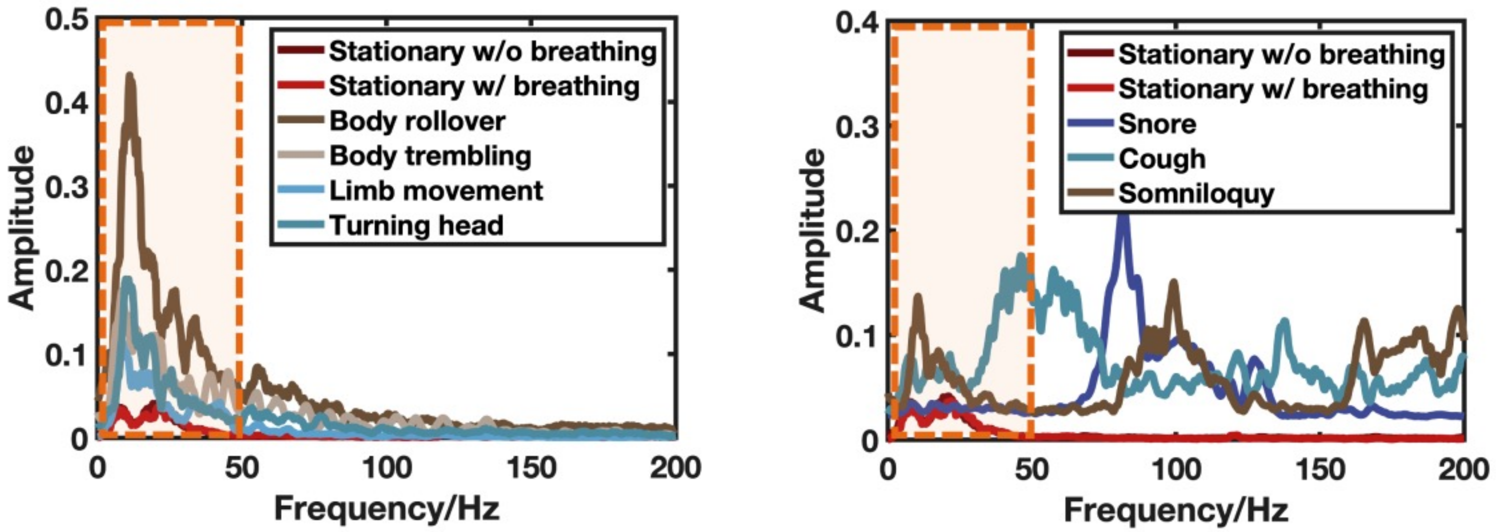
(a) Body movements.
(b) Sound activities.

Physical Sleep Activity

Body movement	• Durations • Zero-crossing rate • Delay profiling • Energy Distribution pattern	Turning head, body trembling, limb movement, and body rollover.
Sound activity	• Autocorrelation • Spectral peaks • Energy distribution	Snore, cough, and somniloquy

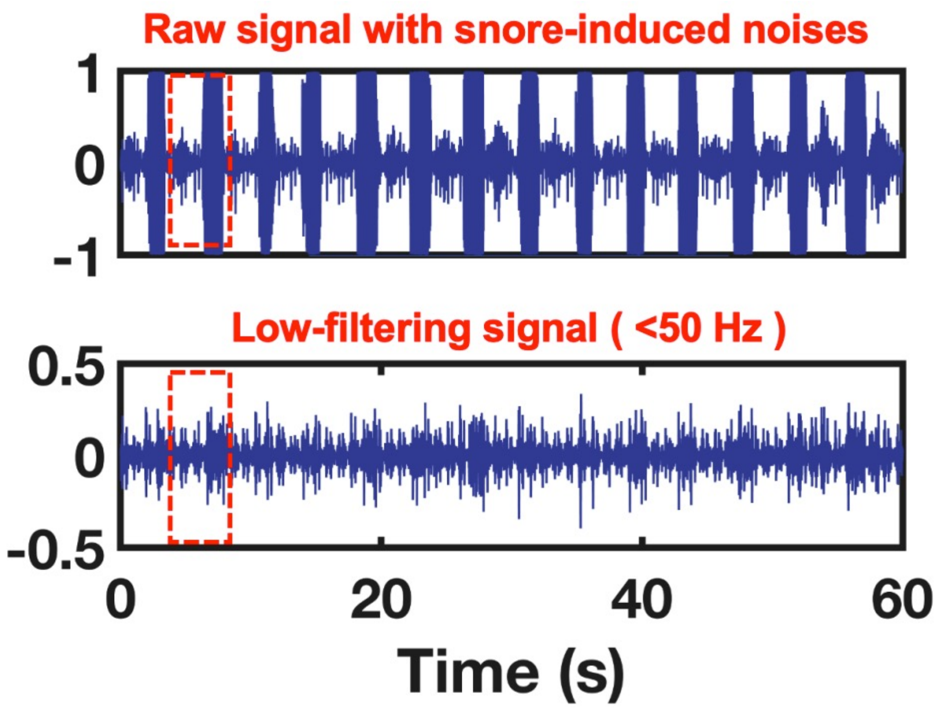
Technical Challenge-2

- **How to obtain accurate physiological activity estimation in the presence of motion artifacts?**



Heartbeat-induced and breathing-induced sounds are heavily disrupted by motion artifacts.

- **Case: raw audio signal with snore noises**



The snoring event overwhelm the original heartbeat waveform

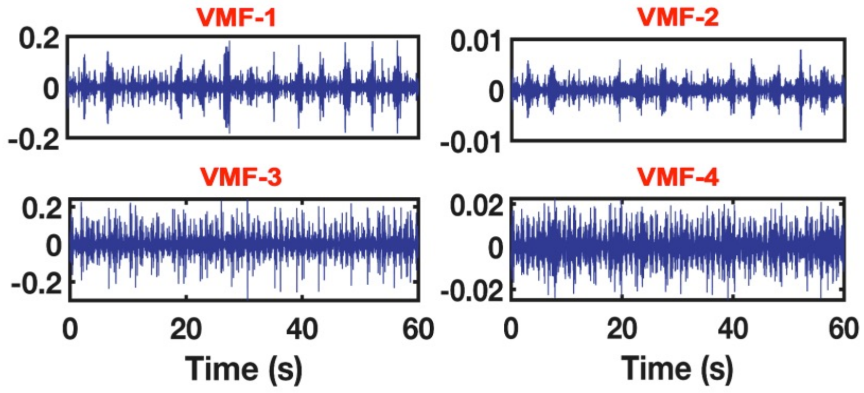
The residual noise still interferes with the original heartbeat waveform

Physiological Activity Estimation

- **Signal Decomposition.**

Decomposing the noisy signal into multiple sub-band signals via Variational Mode Decomposition (VMD)

$$E_r(k) = \frac{\left|\sum_{i=1}^{k} vmf_i\right|^2}{|x|^2} \qquad k = 2, 3, 4, 5, 6.$$

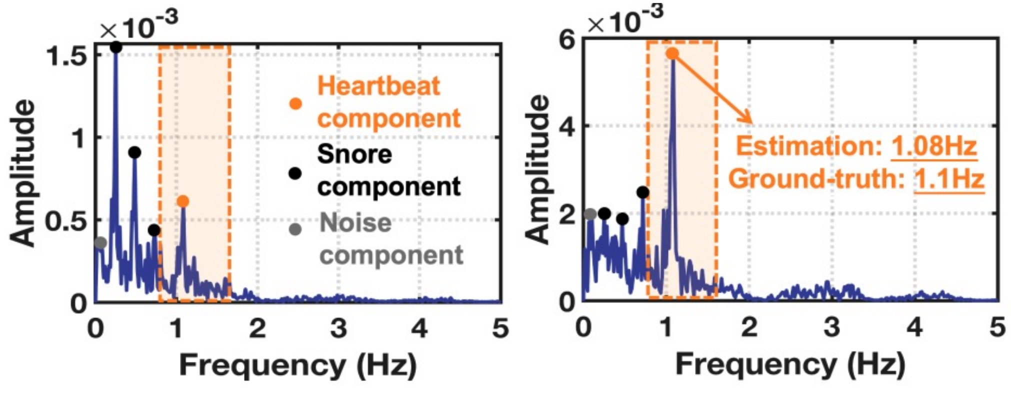


- **Optimal VMF Selection based on HNR and PNR.**

Heartbeat-to-noise Ratio (HNR) and Periodicity-to-noise Ratio (PNR): measuring the the contribution of heartbeat components.

$$HNR = \frac{\sum H}{\sum\{A(i) | 0 < f(i) <= 3\}}$$

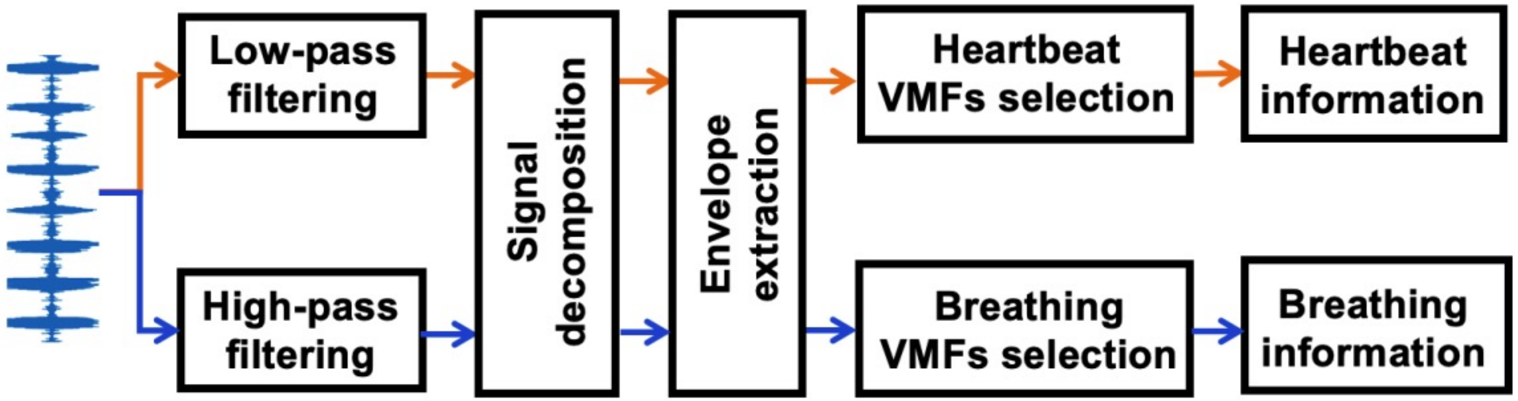
$$PNR = \frac{\sum\{P(i) | P(i) > \mu * P_{max}\}}{\sum H}.$$



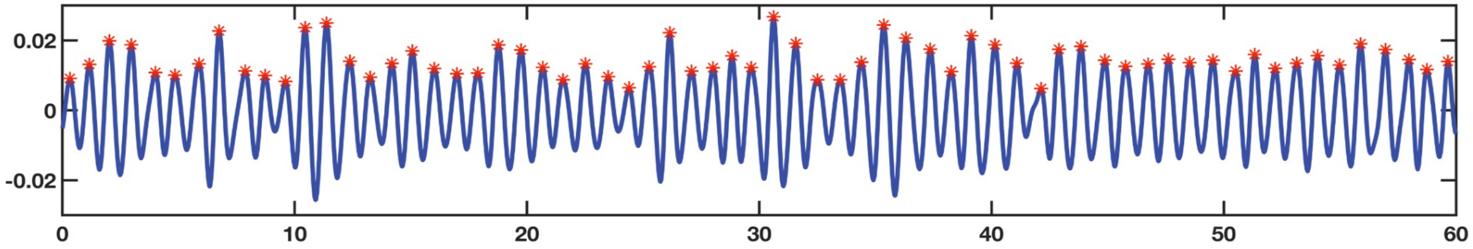
(a) FFT of VMF-2. (b) FFT of VMF-4.

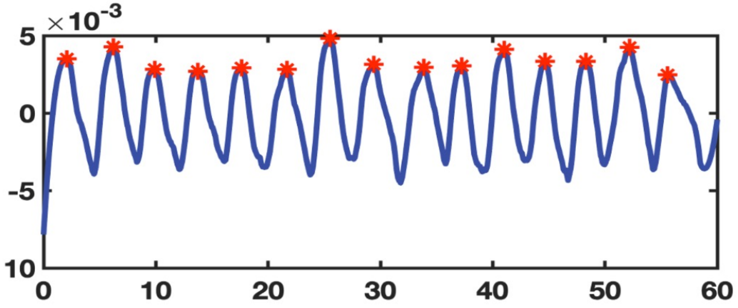
Physiological Activity Estimation

- **Chart flow of physiological activity estimation.**



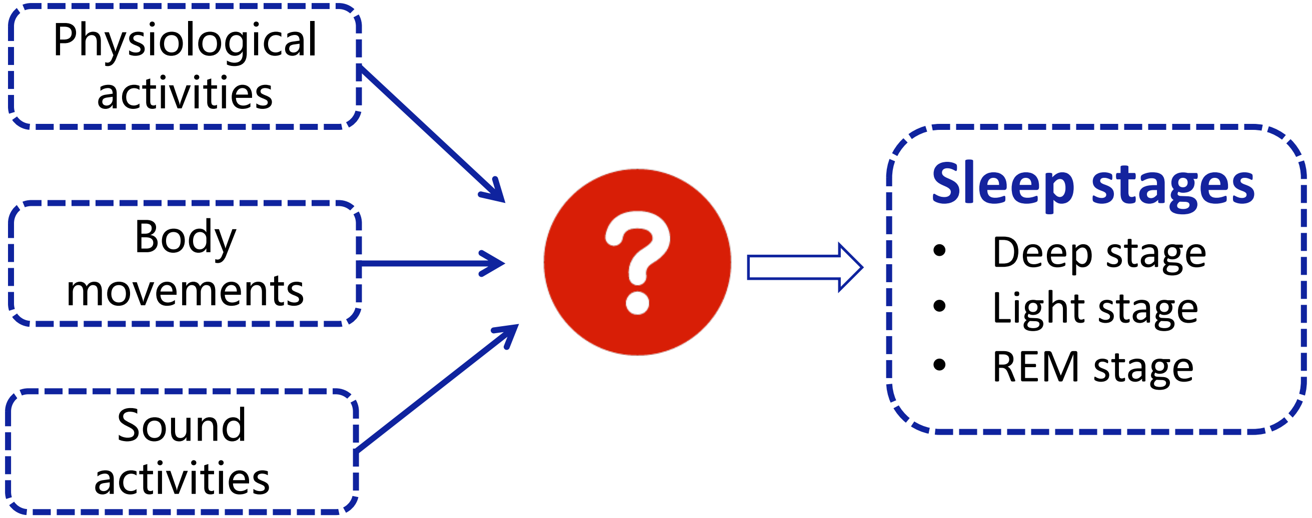
- **Extracted heartbeat waveform (left) and respiration waveform (right).**





Technical Challenge-3

- How to associate various sleep events with sleep stages via representative and interpretable acoustic features?



Sleep-related Acoustic Features Extraction

- Extracting acoustic features from detected sleep activities under the guidance of **sleep medicine knowledge**

Table 1. Variations of sleep events in different sleep stages

Sleep stage	**Characteristics of sleep events**
Light sleep	Large body movements such as body rollover happen. Heartbeat and breathing rates start to slow down.
Deep sleep	Heartbeat, breathing, and body movements become less frequent. Light body movements such as limb and head movement happen. The body is completely relaxed and snoring occurs.
REM sleep	Heartbeat and breathing become more frequent and show long-range correlations. Body movements are concentrated to occur. Some sound activities such as somniloquy and coughing, occur with dreams.

Variation patterns of sleep events in different sleep stages

⇨

(i) Actigraphy Features:

- Occurrence frequency
- Amplitude ratio

(ii) Sound Activity Features:

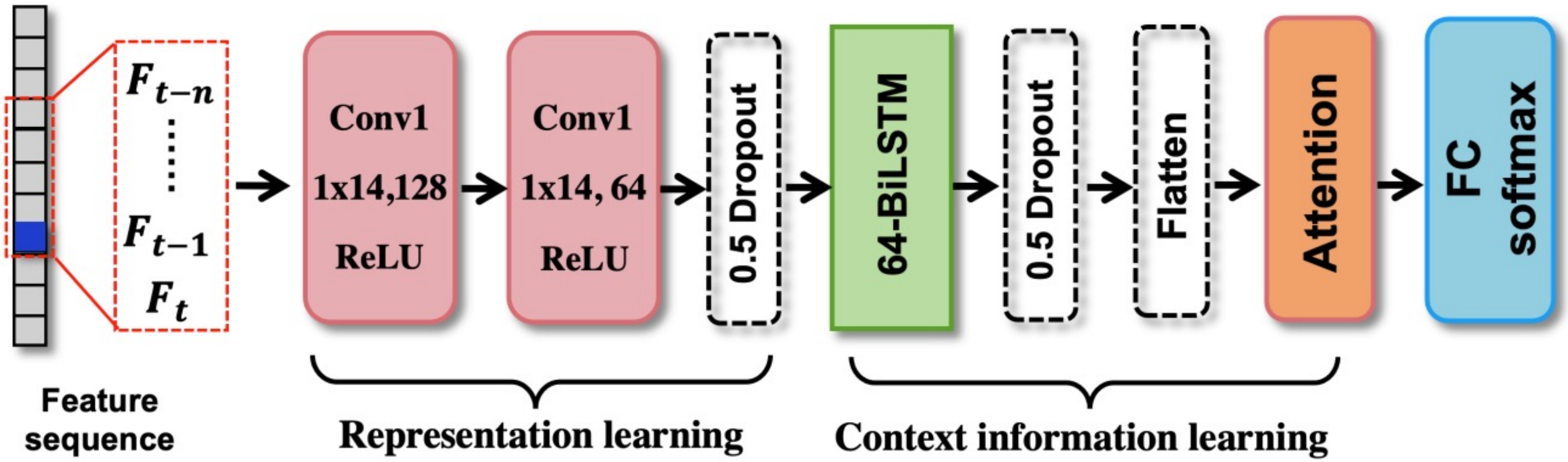
- Occurrence frequency
- Duration ratio

(iii) Physiological Activity Features:

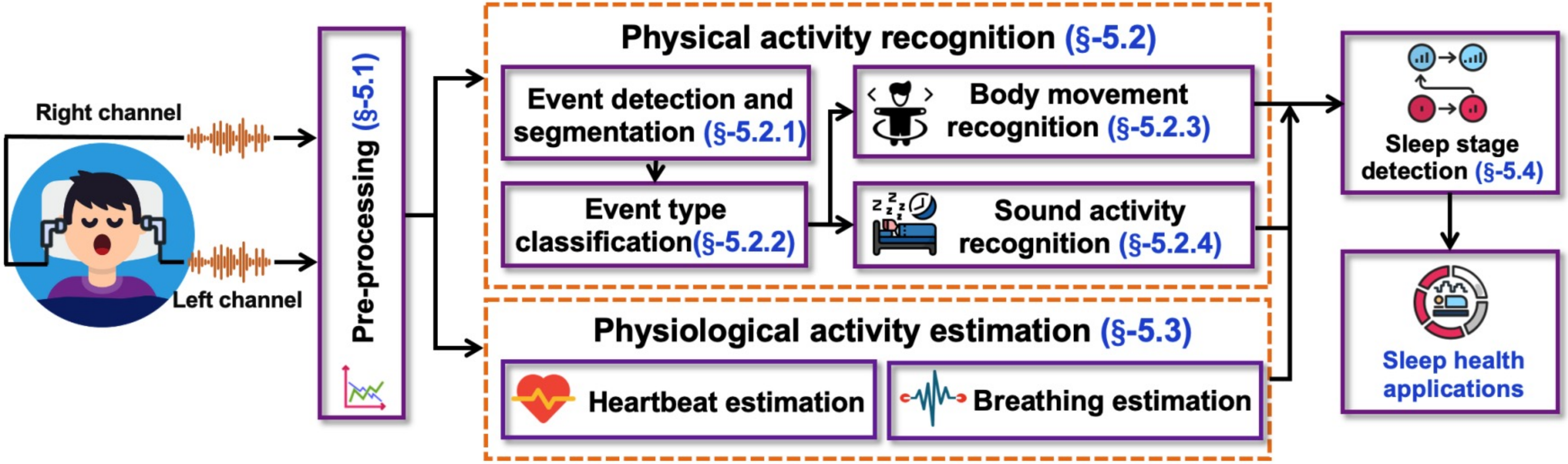
- Waveform Statistical Features.
- Long-time Self-correlation Features.
- Long-time cross-correlation Features.

Attention-based Sleep Stage Detection

- There are **predictable transition patterns** between different sleep states, such as light sleep → deep sleep → light sleep → REM, and there is **context dependency** between different sleep states.



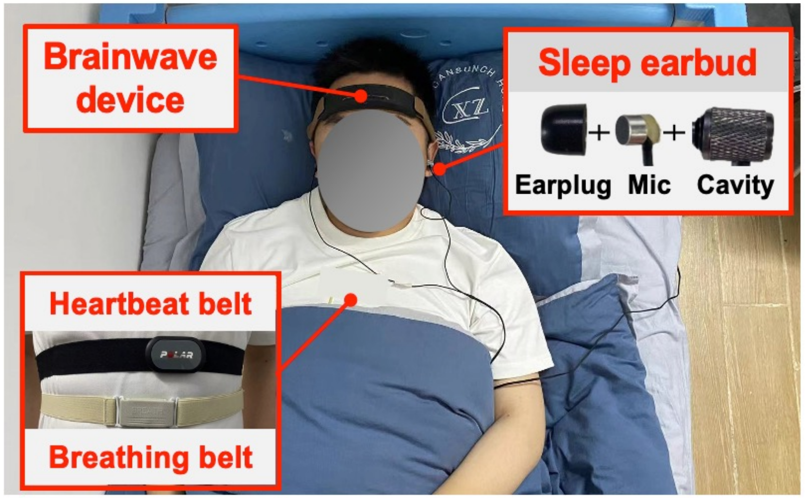
System Overview of EarSleep

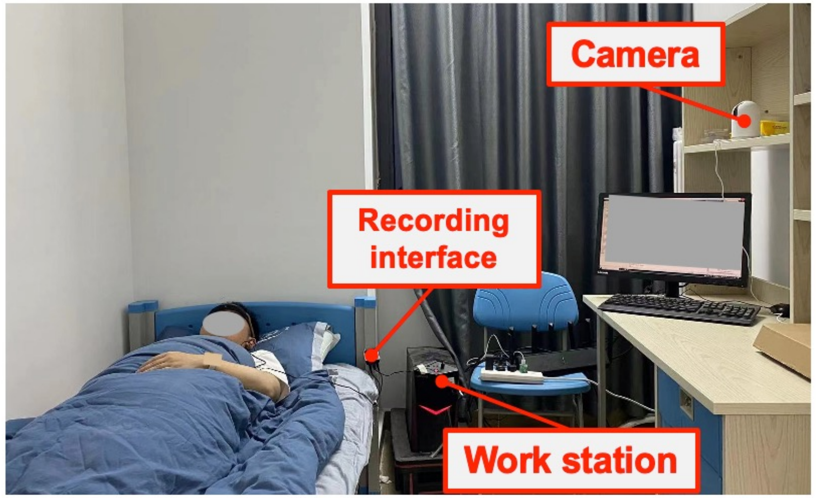


More technical details can be found in our paper.

Performance Evaluation

Experimental Setup



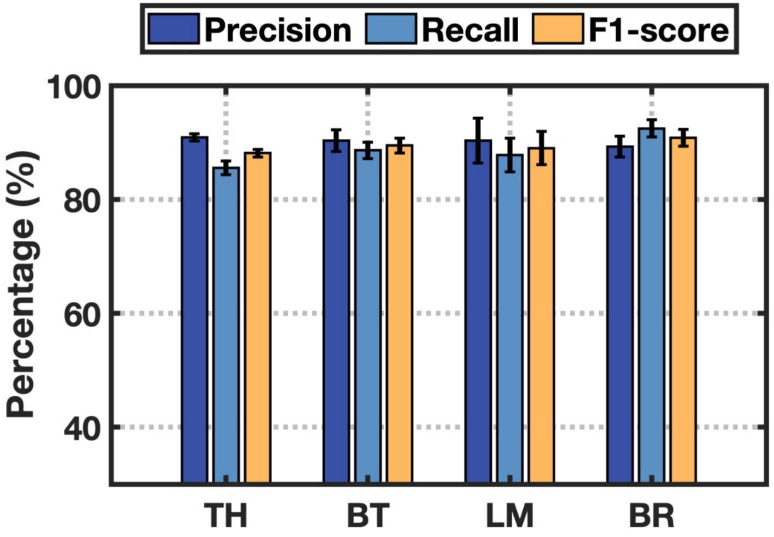


- **18 participants** (12 men and 6 women, 21 to 32 years old) are involved in our evaluation and all participants do not suffer from severe diseases.

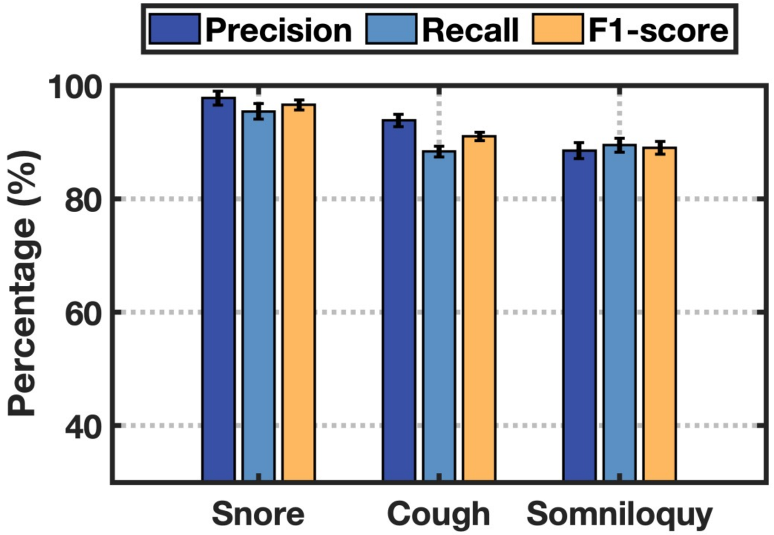
- Each participant sleeps about **6-8 hours** per night during his/her normal sleep schedule. Each participant contributes 2-3 nocturnal sleep data. In total, we collect sleep audio data for **48 nights**.

Performance Evaluation

Evaluation of Physical Activity Recognition



Four-class body movement recognition accuracy of **91.25%**

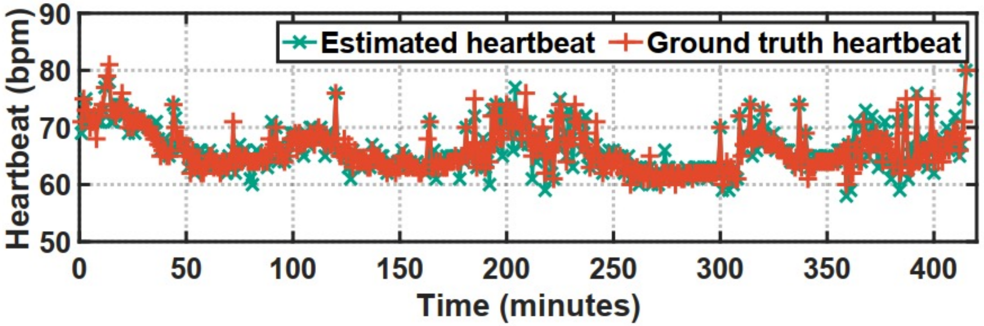


Three-class sound activity recognition accuracy of **97.05%**

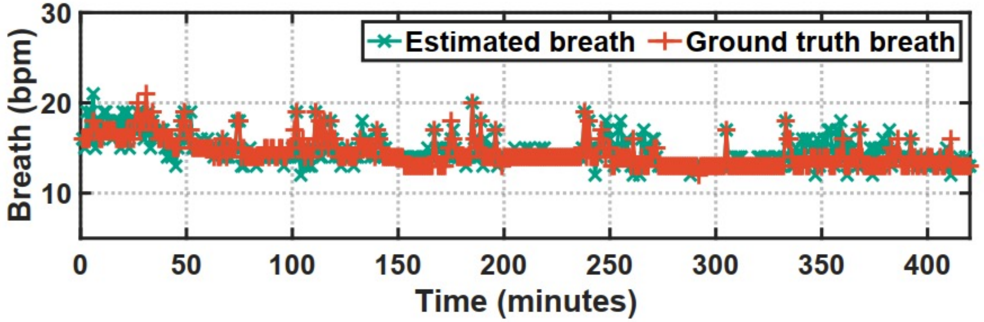
Performance Evaluation

Evaluation of Physiological Activity Estimation

a) Continuous measurement throughout the night of a participant.

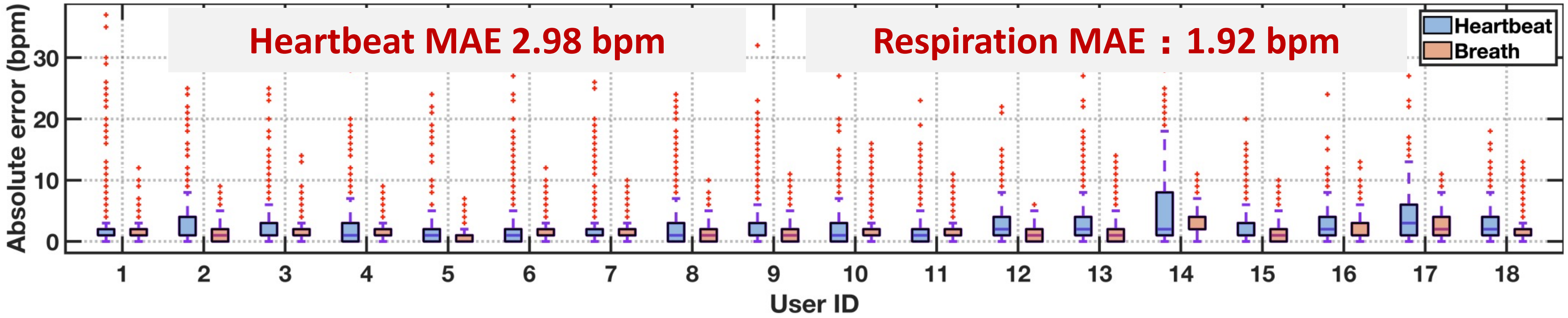


(a) Heartbeat rate estimation.



(b) Breathing rate estimation.

b) Continuous measurement among all participants, including 48 nights.



Performance Evaluation

Evaluation of Sleep Stage Detection

An EEG-based sleep monitoring device provides the ground truth of sleep stage.

- SleepGuard [Ubicomp'17], Sleep Hunter [TMC'15]

	REM sleep			Light sleep			Deep sleep		
	Precision	Recall	F1-score	Precision	Recall	F1-score	Precision	Recall	F1-score
Sleep Hunter	60.91%	57.66%	59.24%	54.12%	50.23%	52.10%	49.48%	45.78%	47.56%
SleepGuard	67.79%	64.27%	65.98%	60.78%	56.42%	58.52%	62.84%	56.32%	59.40%
EarSleep	74.21%	77.37%	75.76%	72.39%	65.32%	68.67%	66.17%	62.29%	64.17%

- **Outperforming with Existing Solutions:**

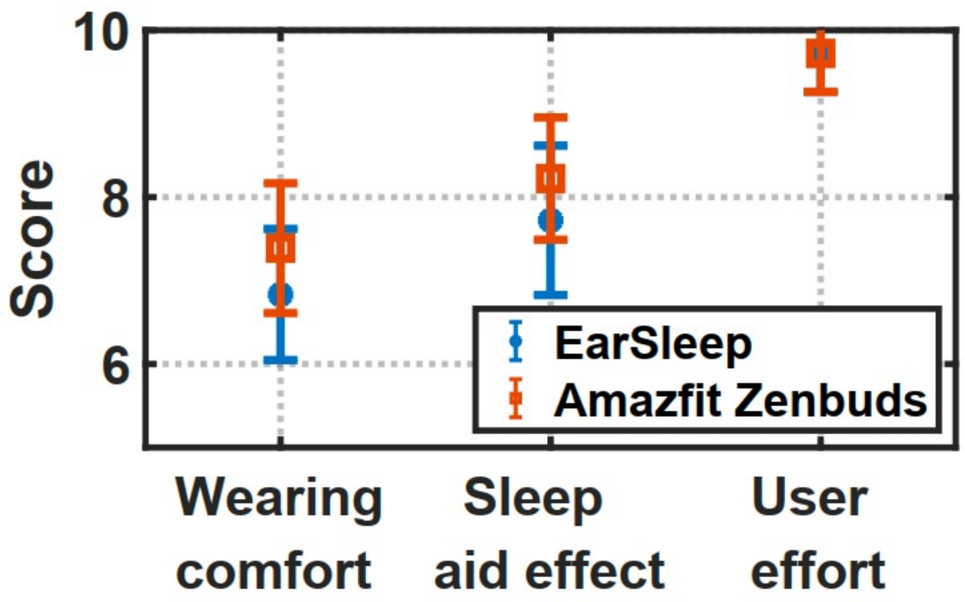
  EarSleep can not only capture the variations in coarse-grained physical activities, but also accurately detect variations in fine-grained physiological activities (breathing and heartbeat).

Performance Evaluation

User Experience Study

Users are required to complete a post-study survey with a 10-point Likert scale (from one to ten) after waking up every day

Amazfit Zenbuds: commercial sleep earbuds



Adopt the softer silicone material and the earbud shape design to improve wearing comfort.

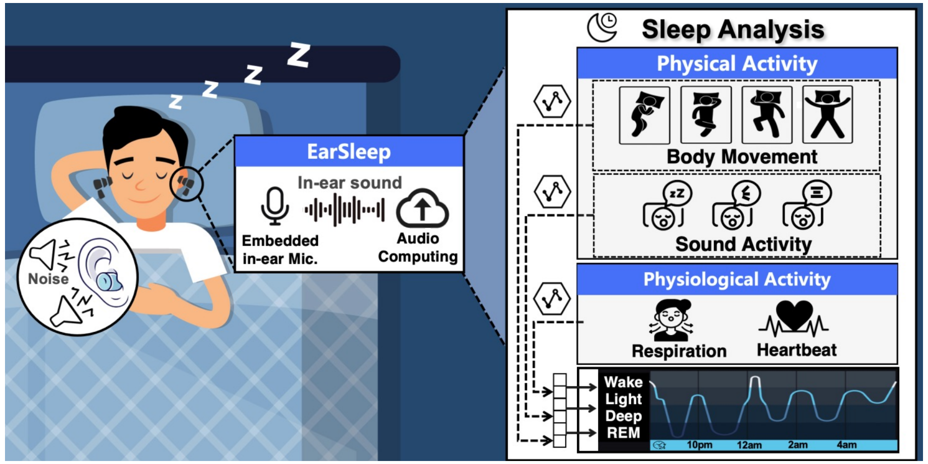
Conclusion

a) **In-ear acoustic-based sleep monitoring approach**

b) **Two low-cost microphones on sleep earbuds**

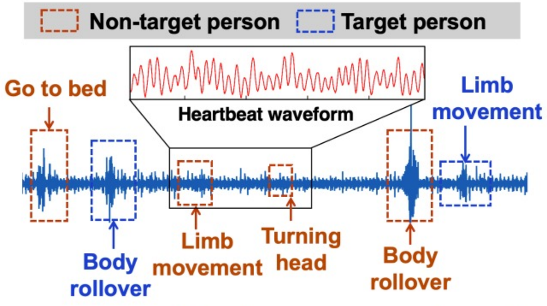
c) **Recognize a wide range of sleep activities**

- Heartbeat and respiration
- Three types of sound activities
- Four types of body movements
- Three types of sleep stages



d) **Discussion:**

- individual Difference
- usage in multi-person scenarios
- ambient noise
- ……



Video: https://www.youtube.com/watch?v=23MpIv_BaVc

Thanks for your listening

EarSleep:

In-ear Acoustic-based Physical and Physiological Activity Recognition for Sleep Stage Detection

Feiyu Han[1,2], Panlong Yang[2], Yanhao Feng[3], Weiwei Jiang[2],
Youwei Zhang[4], and Xiang-Yang Li[1].

Email: fyhan@mail.ustc.edu.cn **Personal website:** https://fyhancs.github.io/

中国科学技术大学 University of Science and Technology of China
南京信息工程大学 Nanjing University of Information Science & Technology
The Hong Kong Polytechnic University 香港理工大學
UEC 国立大学法人 電気通信大学 The University of Electro-Communications