UHead: Driver Attention Monitoring System Using UWB Radar

CHONGZHI XU, Beijing University of Posts and Telecommunications, China
XIAOLONG ZHENG, Beijing University of Posts and Telecommunications, China
ZHIYUAN REN, Beijing University of Posts and Telecommunications, China
LIANG LIU, Beijing University of Posts and Telecommunications, China
HUADONG MA, Beijing University of Posts and Telecommunications, China

The focus of Advanced driver-assistance systems (ADAS) is extending from the vehicle and road conditions to the driver because the driver's attention is critical to driving safety. Although existing sensor and camera based methods can monitor driver attention, they rely on specialised hardware and environmental conditions. In this paper, we aim to develop an effective and easy-to-use driver attention monitoring system based on UWB radar. We exploit the strong association between head motions and driver attention and propose UHead that infers driver attention by monitoring the direction and angle of the driver's head rotation. The core idea is to extract rotational time-frequency representation from reflected signals and to estimate head rotation angles from complex head reflections. To eliminate the dynamic noise generated by other body parts, UHead leverages the large magnitude and high velocity of head rotation to extract head motion information from the dynamically coupled information. UHead uses a bilinear joint time-frequency representation to avoid the loss of time and frequency resolution caused by windowing of traditional methods. We also design a head structure-based rotation angle estimation algorithm to accurately estimate the rotation angle from the time-varying rotation information of multiple reflection points in the head. Experimental results show that we achieve 12.96° median error of 3D head rotation angle estimation in real vehicle scenes.

CCS Concepts: • **Human-centered computing** → **Ubiquitous and mobile computing systems and tools**.

Additional Key Words and Phrases: Ultra-Wideband, Wireless Sensing, Driver Attention Monitoring, Head Rotation

1 INTRODUCTION

Advanced driver-assistance systems (ADAS), which make driving more automated and safe, are attracting increasing attention. According to a Canalys 2021 research report [11], around 33% of the new cars sold in the US, Europe, Japan and China have ADAS capabilities, which is predicted to increase to 50% by 2030. As technology evolves, the goal of ADAS is gradually extending from vehicle and road conditions [35] to the most critical element when driving – the driver [14]. Driver's attention is closely related to driving safety. Researches have shown that 90% of road accidents are closely related to distracted driving [51]. In Europe, legislation has been passed

Authors' addresses: Chongzhi Xu, xuchongzhi@bupt.edu.cn, Beijing University of Posts and Telecommunications, Beijing, China; Xiaolong Zheng, zhengxiaolong@bupt.edu.cn, Beijing University of Posts and Telecommunications, Beijing, China; Zhiyuan Ren, renzhiyuan@bupt. edu.cn, Beijing University of Posts and Telecommunications, Beijing, China; Liang Liu, liangliu@bupt.edu.cn, Beijing University of Posts and Telecommunications, Beijing, China; Huadong Ma, mhd@bupt.edu.cn, Beijing University of Posts and Telecommunications, Beijing, China.

to prohibit behaviours that take a driver's attention off the road [2]. Besides safety concern, monitoring driver attention allows for applications such as predicted vehicle control that greatly enriches the driving experience. Therefore, there is an urgent need for a reliable and easy-to-use driver attention monitoring system to enhance driving comfort and safety.

To enhance the driving experience, car companies have created various systems to monitor driver attention, including the Pre-Crash Safety (PCS) system by Lexus [1], ATTENTION ASSIST by Mercedes-Benz [5] and Driver Alert Control (DAC) by Volvo [14]. Meanwhile, researchers have proposed various ways to monitor driver attention. One straightforward way is to use wearable sensors to monitor various physiological indicators of the driver, such as EEG [65], ECG [20] and EMG [24]. These methods can infer driver fatigue based on physiological indicators. However, wearing those sensors leads to physiological discomfort. Another way is using a camera to obtain pictures or videos of the driver's face, and analyse them to get the driver's facial motions, such as blinking [29], and head posture [34]. However, the effectiveness of vision-based systems is difficult to guarantee due to the diversity of environmental conditions such as weather and lighting while driving. In addition, there are serious privacy issues with cameras.

The development of wireless sensing opens up new possibilities for monitoring driver attention. Wireless signals such as acoustic and RF can be reflected from a person and carry human vital signs [27, 41, 44, 62] or motion information [21, 42, 54]. Such a system allows for contactless sensing of people and is not susceptible to interference from ambient light. Since wireless signals are virtually incapable of acquiring any information other than the target's motions and physical reflective properties, these systems typically do not invade the user's privacy. Acoustic signals enable the extraction of a driver's breathing waveform [62] and blink detection [44]. In [42], risky driving behaviours are identified based on the unique characteristics of the acoustic signals. However, acoustic signals are easily contaminated by background noise from driving. RF signal can be a good way to avoid these problems. In particular, WiFi is used to recognise the driver's in-vehicle activity [22, 50]. mmWave can monitor the driver's respiratory rate [55] and heart rate [64]. However, these motions and physiological indicators only hint at possible changes in the driver's attention and do not accurately monitor where the driver's attention is focused. This makes these systems difficult to apply in practice. Moreover, these systems usually require the deployment of additional RF equipment, such as multiple antennas, which is difficult to implement in assembly-line produced cars.

The ultra-wideband (UWB) technology is being used in more and more consumer electronics products due to its advantages of low energy consumption, low multipath interference, and high distance resolution. Apple has been carrying UWB chipsets since iphone 11 [9]. Smartphone products from Samsung [10], Xiaomi [7] and Google [8] also support UWB functionality. If the UWB chips in smartphones can be fully utilised, a simple data interaction between the phone and the car's control system is all that is needed to achieve easy and widely applicable driver attention monitoring that can be placed on a phone holder in the car. It avoids the problem of the difficulty of deploying multi-antenna devices. However, leveraging the UWB radar on smartphone to accurately monitor the driver's attention is still a challenging problem.

Based on our measurement study and existing studies [43, 53], we find that head direction and driving behaviour are highly correlated, because drivers rely almost exclusively on visual information to observe road conditions. Research [30] shows that head motions can reliably predict lane change intentions. In addition, study [52] suggests that distracted driving conditions such as fatigue are highly correlated with head motions. Moreover, since the driver's head usually faces the road ahead, when the driver's head rotates at an excessive angle, his attention is inevitably taken away from the road, resulting in a shift of attention away from the road. Therefore, we focus on estimating the driver's head rotation angle to determine where the driver's attention is focused.

However, there are several challenges to achieving reliable estimation of the driver's head rotation angle on in-vehicle phone holder. Firstly, head rotation usually occurs simultaneously with some body motions (e.g., tiny body swings). The coupling of the tiny dynamic information and the head rotation information in the

reflected signal makes it challenging to analyse the head rotation accurately. Secondly, head rotation is associated with a frequency shift in the signal, but the rotational property is not constant in time. Hence, estimating fine-grained head rotation angles requires precise time and frequency information. Traditional time-frequency representation calculation methods, such as the Short Time Fourier Transform (STFT), calculate the frequency information of the time-series signal within a certain window. As the window size increases, the frequency resolution improves while the time resolution degrades. When the window size decreases, the time resolution raises while the frequency resolution declines. It results in a trade-off between time and frequency resolution, leading to the dilemma of obtaining both precise time and frequency information. Finally, the angular velocity of head rotation is hard to estimate because of the coupled signals from multiple reflective points on the head. The head surface morphology is so complex that the rotation of a large number of reflective points on the head will affect the signal simultaneously. How to extract accurate rotation velocity from time-varying head rotation features remains a challenge to be solved.

In this paper, we present UHead, a UWB radar-based system placed on a phone holder in car to monitor the driver's attention by detecting the head rotation of the driver. Head rotation in different direction leads to positive or negative frequency shift of UWB signal, whereas frequency shift due to body motions is unpredictable. Compared to body motions while driving, head rotations have greater amplitude and higher velocity, and thus cause larger frequency shifts to the signal. We use Complex Variational Mode Decomposition (CVMD) to decompose the signal at positive and negative frequencies, respectively, from which we extract high-frequency head rotation information. To address the trade-off between time and frequency resolution of conventional methods, UHead utilises a bilinear joint time-frequency representation, Pseudo Wigner-Ville Distribution (PWVD) to avoid the inevitable loss of resolution when using single linear approach. Different from methods such as Fourier transform that perform windowing in only time dimension, PWVD performs a Fourier transform of the instantaneous autocorrelation function of the signal, which yields an instantaneous time-frequency representation of the signal and therefore avoid the loss of resolution caused by windowing in time dimension. Finally, we design the head structure-based rotation angle estimation algorithm to accurately estimate the head rotation angle from the complex head reflection and time-varying rotation information. Since all the reflection points of the head share the common angular velocity at one moment, we can know that the reflection point with the largest rotation radius leads to the largest signal frequency shift. Accordingly, we extract the rotational features among all moments of the same reflection point to estimate the head rotation angle.

The main contributions of this paper can be summarized as follows.

- We propose UHead, a driver attention monitoring system that uses UWB radar placed on a phone holder in vehicle to detect the head rotation angle to infer the driver attention, which can enable smart in-cabin applications such as distracted driving monitoring and human-vehicle interaction prediction.
- We extract fine-grained accurate head rotation information by removing signal dynamic noise and calculating joint time-frequency representations with high resolution in both time and frequency. We develop a novel head structure-based rotation angle estimation algorithm, which calculates the head rotation angle from time-varying rotation information of a certain reflection point.
- We conduct extensive experiments to evaluate the performance of the UHead and the effectiveness of the components. The results show that the UHead is effective in estimating the 3D head rotation angle.

The rest of the paper is organized as follows. Sec. 2 describes the driver's head rotation pattern. Sec. 3 presents the system design of the UHead. Sec. 4 shows the system evaluation. Sec. 5 discusses the limitations and applications of our work. Sec. 6 reviews related work. Sec. 7 draws conclusions.

2 DRIVER HEAD ROTATION PATTERNS

In this section, we first introduce the head motion patterns of driving, and then analyze the Doppler profiles caused by head rotation.
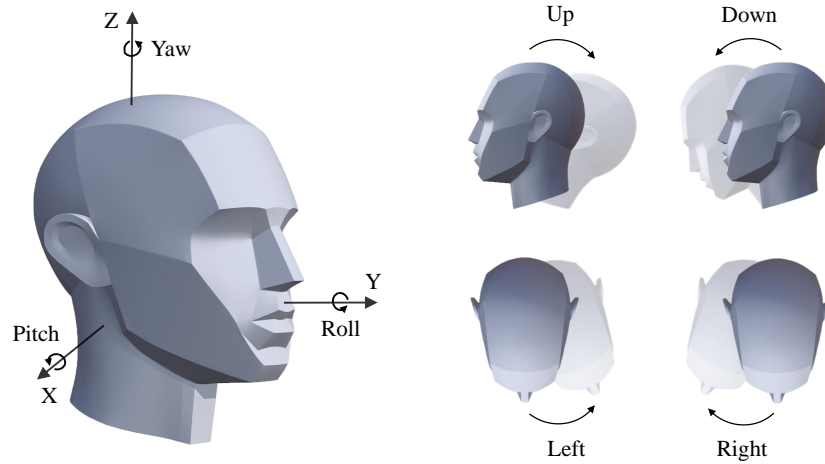
2.1 Patterns of Driver Head Motions

Fig. 1. Head rotations described by the Euler angles and the directions of head rotations that would cause the driver's attention to be diverted: up, down, left and right.

As shown in Fig. 1, the Euler angle is commonly used to describe head motions [16]. The angle around X, Y and Z axes is pitch, roll and yaw, respectively. Only pitch and yaw changes will cause the driver's line of sight to move, which includes four directions of head rotation:

- **Up.** Rotating the head up will increase pitch angle while causing the driver's view to be focus on the sky rather than the road. For example, when a driver is sleepy, his head will unconsciously raise to yawn.
- **Down.** Rotating the head down will decrease pitch angle as well as leading the driver's eyes to focus on the steering wheel. When the driver is nodding due to fatigue, his head will inevitably rotate downward.
- **Left.** Rotating the head left will increase yaw angle and make the driver ignore the road, such as looking in the left rear-view mirror.
- **Right.** Rotating the head right will decrease yaw angle while leading the driver's line of sight to focus on the right. It occurs when the driver is attracted to the scenery on the right side of the road.

Although the human head can rotate to various orientations in 3D space, the driver's head will usually rotate around one axis only while driving. There is almost no rotation in the other two axes at the same time. Therefore, we can concentrate on analysing the head rotation in one axis instead of analysing all three axes at the same time. To validate this, we ask a volunteer driver to wear the VR headset on his forehead and repeat the head rotation around the X and Z directions. The IMU sensor in the headset measures the rotation angles of the driver's head in all three axes, as shown in Fig. 2. The results show that when the driver's head rotates around one axis, there is little rotation around other two axes. Therefore, we can focus on analyzing the head rotation in one axis for each time.
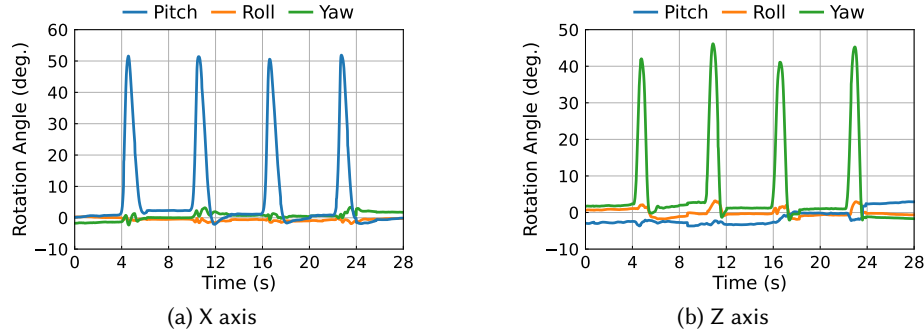
Fig. 2. Rotation angle around X (a) and Z (b) axes. The head is rotated four times on each axis. Rotation angle is measured by the IMU.

2.2 Head Rotation Patterns via Doppler Shift

The Doppler shift is a change in frequency of the signal as the target moves relative to the source. For a rotating target, the Doppler shift is related to the angular velocity [26]. As the target rotates, each reflected point on the target produces such a Doppler shift, which is

$$f_D = \frac{2f}{c}\omega r \tag{1}$$

where f is carrier frequency, c is the speed of light, ω is the angular velocity of the target, and r is the rotation radius of the reflection point.

In UHead, we use a mono-static UWB radar with the transmitter and receiver in the same location. Each reflection point in the head corresponds to a signal transmission and reception path. As the target rotates, reflection points have the same angular velocity, but different radius of rotation. These reflection points result in different Doppler shifts during once head rotation. However, the Doppler profiles of the signal consists of the Doppler shifts caused by all reflection points mixed together, which makes it difficult to extract stable rotation information from the Doppler profiles. How to extract the head rotation information and estimate the accurate rotation angle needs elaborate designs.

3 SYSTEM DESIGN

In this section, we introduce the design details of UHead, our driver attention monitoring system, which leverages the Doppler shift of UWB signals to estimate the head rotation angle to monitor driver attention .

3.1 System Overview

Fig. 3 shows the overview of UHead, which consists of three main components, *direction recognition*, *dynamic noise reduction* and *rotation angle estimation*. Firstly, we perform head rotation direction recognition. The complex static environment in the car causing the head rotation information to be immersed in static noise, and we use a loopback filter to remove the static noise from the original signal and get the head signal that contains all the dynamic information. The rotation of the head in various directions affects the signal differently, resulting in different signal features. Accordingly, we identify the direction of head rotation in three dimensions. The driver's random body motion leads to dynamic noise in the Doppler profiles that is coupled to the head rotation information and has different distributions in the positive and negative frequencies. To reduce dynamic noise, we use Complex Variational Mode Decomposition (CVMD) to separate the positive and negative frequency portions
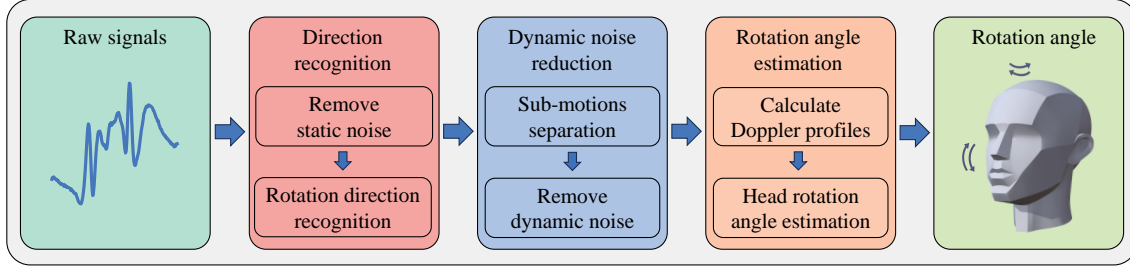
Fig. 3. Overview of UHead.

of the Doppler profiles, and then decompose the coupled head rotation information and body motion information to extract pure head rotation features. Head rotation features are reflected in the Doppler profiles of the signal. Conventional single linear time-frequency representation results in a trade-off between time and frequency resolution due to windowing. To obtain refined head rotation Doppler profiles, we use PWVD to acquire bilinear joint time-frequency representation with the highest resolution in both time and frequency. Finally, we design the head structure-based rotation angle estimation algorithm, which extracts the time-sequential rotation information of a certain reflection point from the time-varying head rotation Doppler profiles of multiple reflection points based on the correlation between the Doppler profiles and the structural characteristics of the head, so as to accurately estimate the head rotation angle.

3.2 Modeling UWB Channel

In this section, we introduce the basic theory of UWB radar. The UWB transmitter generates a Gaussian pulse $g_k(t)$, which is modulated to the carrier frequency f_c to obtain the transmitted signal in k-th frame, $s_k(t)$:

$$s_k(t) = g(t - kT_s) \cdot cos(2\pi f_c(t - kT_s)) \tag{2}$$

where T_s is the frame duration. Since all the pulses in different frame are the same, Eq. (2) can be written as $s_k(t) = g(t) \cdot cos(2\pi f_c t)$. The transmitted signal is shown in Fig. 4. The signal propagates inside the vehicle. The impulse response v_k reflecting the in-vehicle environment can be expressed as:

$$v_k(t) = \sum_{p=1}^{P} A_p \delta \left(t - t_p - t_p^R (kT_s) - t_p^D (kT_s) \right) \tag{3}$$

where A_p is the channel gain of the p-th reflection path signal in the vehicle, t_p is the time delay of the p-th path, $t_p^R(kT_s)$ and $t_p^D(kT_s)$ are time delays caused by Doppler shifts of the p-th path due to head rotation and other parts of the body, respectively. For the mono-static UWB radar, $t_p = \frac{2R_p}{c}$, $t_p^R(kT_s) = \frac{2k\omega_p rT_s}{c}$ and $t_p^D(kT_s) = \frac{2kv_pT_s}{c}$, where R_p is the distance between the target to be detected and the UWB radar.

The received signal can be obtained by the convolution operation of the transmitted signal and the impulse response as $v_k(t) * s_k(t)$. The received baseband signal $r_k^b(t)$ after applying IQ downconversion and low-pass filtering can be obtained by:

$$r_k^b(t) = \left| [v_k(t) * s_k(t)] \cdot e^{j2\pi f_c t} \right|_{LPF} = \sum_{p=1}^{P} A_p e^{2\pi f_c \left(t_p + t_p^R(kT_s) + t_p^D(kT_s) \right)} \cdot s \left(t - t_p - t_p^R (kT_s) - t_p^D (kT_s) \right) + n(t) \tag{4}$$

where $n(t)$ is Gaussian channel noise, and the symbol $*$ denotes convolutional operation.

In a radar system, the time between two pulses is called Pulse Repetition Interval (PRI), which equals to the reciprocal of the sampling rate F_s. PRI denotes the ranging accuracy of a radar, that is, the error between the measured value of the target distance and its true value. Range resolution of a radar is the minimum distance at which two targets can be separated in the radial direction of the radar. The range resolution of a UWB radar is proportional to its bandwidth B, which is equal to $\frac{c}{2B}$.
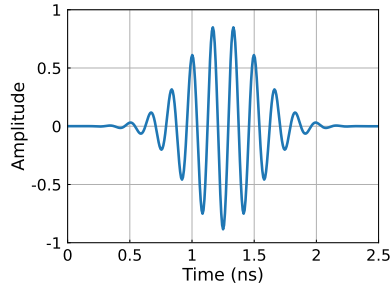

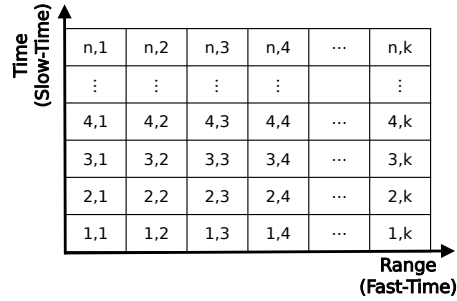Fig. 4. Transmitted signal $s_k(t)$.

Fig. 5. Fast-Time and Slow-Time.

In order to facilitate the analysis and processing of radar signals, we put the reflected signals of the same pulse in the same row, and the different pulses are arranged in different rows according to the time sequence. We define the dimension of the time slots that compose a single PRI, which is the row dimension, as the "fast-time" dimension, and the dimension of the columns that are updated by each PRI as the "slow-time" dimension. As shown in Fig. 5. The fast-time axis represents the time delay at different ranges and the slow-time axis is used to estimate the Doppler shift through a long time observation. In the fast-time axis, each unit is called a range bin, which corresponds to the distance between a target and the radar.

3.3 Rotation Direction Classification

3.3.1 Driver Identification. The general approach usually takes the signal being perturbed as an indication of the target signal [70]. However the driver is usually not alone when driving. This makes it very likely that passenger-caused signal changes will be mistakenly attributed to the driver, and thus the method of target identification by dynamics is no longer applicable. In most cases the reflected signal from the driver has the maximum strength, but the absence of significant movement by the driver as well as the redundant physical environment and the presence of passengers in the vehicle result in the reflected signals from other targets sometimes being stronger than the driver's reflected signal. Therefore, it is difficult to achieve stable identification of the driver by the strength of the reflected signal.

We analyse driver behaviour and find that the driver's position usually remains constant and the displacement of his head is usually within limited 10 cm. We make full use of high ranging accuracy of UWB to determine the reflected signal of the driver at a distance. This makes it easy to distinguish head reflection signals from those of other targets. As the head diameter is typically greater than the range of one range bin, the reflected signal of the driver's head will exist in several neighbouring range bins at the same time. In the limited space inside the vehicle, before and after a slight displacement of the driver's position (e.g. seat adjustment), the driver's head rotation can still be detected in several co-occurring bins. As shown in Fig. 6, before the head displacement, the range bins of the head is [3, 8], which changes to [5, 10] after head displacement. Then there are still four co-occurring range bins (depicted as red ones in Fig. 6), [5, 8], can be used for head position detection. This makes the detection of the driver's head position by distance robust. The effectiveness of this approach is validated in Sec. 4.2.3.
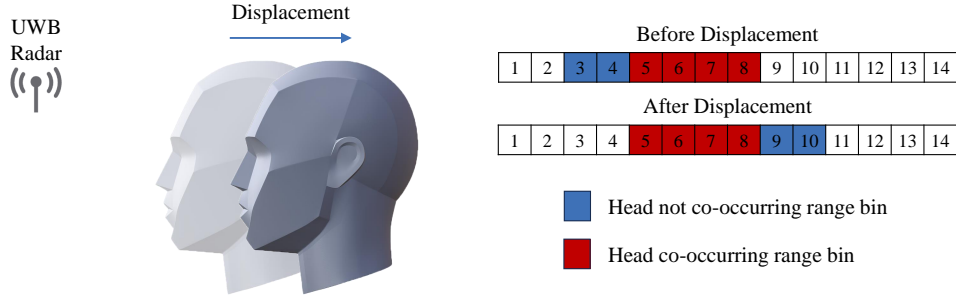
Fig. 6. Range bins of head presence before and after driver displacement.

3.3.2 *Static Noise Reduction.* There are many static objects in the vehicle cockpit, such as seats and doors. These static objects, which are close to the driver, strongly reflect the signal, causing the head rotation information to be immersed in the static background, as shown in Fig. 7(a). It is necessary to remove the static noise and retain the dynamic information in the signal for further processing.

(a) Signal before static noise reduction.
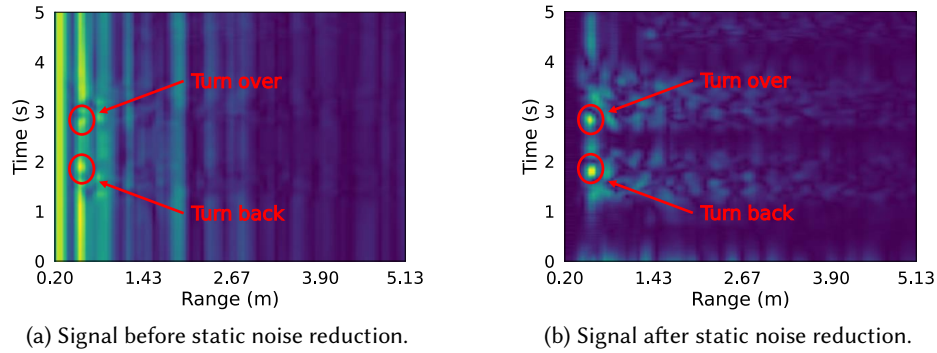
(b) Signal after static noise reduction.

Fig. 7. The signal without static noise reduction (a) and with static noise reduction (b).

We use a loopback filter [17] to remove static background. The static noise of the system can be described as:

$$c_n(t) = \beta c_{n-1}(t) + (1 - \beta)r_n(t) \tag{5}$$

and the background subtracted signal can be represented as

$$r_n^-(t) = r_n(t) - c_n(t) \tag{6}$$

where $r_n(t)$ denotes the n-th frame and the weight β is empirically set to 0.9.

The signal removed the static noise is shown in Fig. 7(b). The signal changes caused by the head rotation can be seen distinctly.

3.3.3 *Rotation Direction Recognition.* The Doppler shift is related to the direction of relative motion of the target and the radar. When the target moves toward the radar, the Doppler shift is positive. Conversely, it is negative. It can reflect the head rotation direction to some extent. As shown in Fig. 8, for the sake of description, we represent the position of the head by a certain reflection point on the head, i.e. the red dot in the figure. For a typical position of mobile phone holder placement, on the right side of the driver, a down head motion causes the head

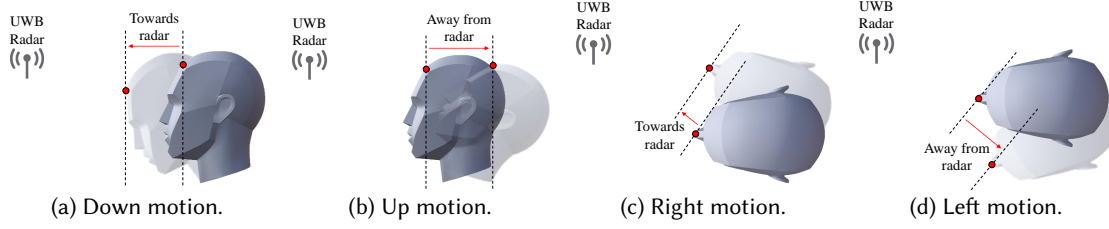(a) Down motion. (b) Up motion. (c) Right motion. (d) Left motion.

Fig. 8. Head rotation relative UWB radar when the radar is put on the right side of the driver. (a) Down motion causes head moves towards radar. (b) Up motion causes head moves away from radar. (c) Right motion causes head moves towards radar. (d) Left motion causes head moves away from radar.

to move towards radar, which will result in a positive Doppler shift of the signal, as shown in Fig. 8(a). An up motion of the head results in the head moving away from radar, as in Fig. 8(b), which will appear as negative values in the Doppler profiles of the signal. For the case of Fig. 8(c), the head moves to the right, the direction of head motion is the same as the relative direction of the radar to the driver, causing the head to move close to the radar. In the case of Fig. 8(d), the head moves to the left, which results in the direction of head motion being opposite to the relative direction of the radar from the driver, and therefore the head moves away from the radar. However, it is not sufficient for distinguishing head rotations in all four directions. For manoeuvres in which the relative position of the head to the radar changes in the same manner, e.g. down and right motions, as well as up and left motions, it is not feasible to differentiate them merely by the relative movement in the radial direction of the head to the radar.
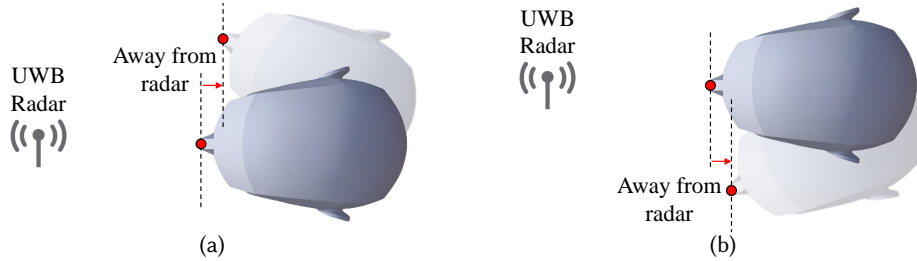

(a) (b)

Fig. 9. The device just needs to be placed in the usual $(0°, 90°)$, rather than a very few special positions, head rotates in different directions can have different patterns.

We then analyse the head 3D rotation pattern in more detail. Due to the ellipsoidal shape of the human head [45], the radius of head rotation around the X axis is larger than the radius around the Z axis. When the driver's head rotates up and down around the X axis, the relative distance to the radar changes significantly and causes a larger Doppler shift. When the head rotates left and right around the Z axis, the relative distance to the radar changes less and the Doppler profiles are smaller. Furthermore, due to differences in muscle traction, head rotation around different axes produces different velocity trends. As a result, although sometimes head rotations around different axes shows the same positive or negative frequency shift, they will show different magnitudes of values and trends in the Doppler profiles. Despite the significant differences in the Doppler profiles of different rotations, the differences are difficult to quantify. Hence, we use the classification ability of neural networks to identify the head rotation direction. Considering that the head rotation feature is time dependent, we use a simple two-layer LSTM network. Due to the structural properties of the human head, it achieves good classification performance

even for user data that the model has never seen before. Then we can use a pre-trained model to recognize the rotation direction first and avoid the training cost for different users. For a very few device placements, where the radar is placed at an angle of 0° or 90° to the driver's head, head rotations in different directions around one axis may produce the same features. These cases are rare as phone holders in the car are hardly ever placed at these angles. The rotation of the driver's head in different directions produces fully distinguishable features when the radar placement is at an angle of (0°, 90°) to the driver's head, as shown in Fig. 9. This is also usually the placement of phone holder. In practice, the user can determine the location of the device based on positive or negative frequency shift by simply checking the left rear view mirror in advance.
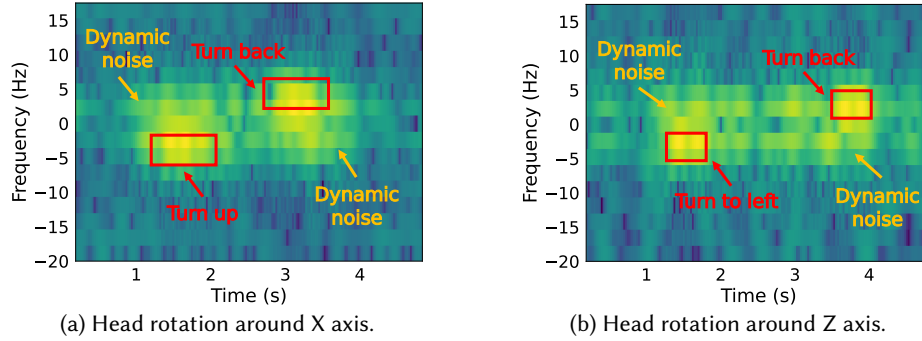

(a) Head rotation around X axis. (b) Head rotation around Z axis.

Fig. 10. Doppler profiles of head rotation around X axis (a) and Z axis (b) by STFT.

3.4 CVMD-based Head Rotation Doppler Profiles Extraction

Driver's head rotation while driving is usually in two consecutive parts: turning over and turning back, which produces two parts of the Doppler profile. As shown in Fig. 10, when the UWB radar is placed on the driver's right side, negative Doppler shifts are generated by head turning up and turning to left, while positive Doppler shifts are generated when the head are turning back, respectively. At the same time, small motions such as tiny body swings occur together. These dynamic noises caused by small body motions are arbitrarily distributed over $[-5Hz, 5Hz]$, while typical head rotation frequency shifts range in $[-18Hz, -2Hz] \cup [2Hz, 18Hz]$. Dynamic noise and head rotation information overlap in frequency and are therefore difficult to separate by a simple bandpass filter.

Algorithms such as VMD [31] are commonly used to separate coupled signals with frequency overlapping between components. These methods are designed for real signals, which allow the decomposition of positive and negative frequencies after stacking them uniformly in the positive frequency. However, the randomness of the body swing makes the dynamic noise have different distributions in the positive and negative frequencies, so stacking the signal frequencies to positive before processing results in the loss of head rotation information in a certain direction. To separate the information about the rotation of the head in both directions effectively, we utilize Complex Variational Mode Decomposition (CVMD) [57], which decomposes the positive and negative frequency parts of the signal into multiple components respectively, and extracts the components related to the head rotation therein. Further details of CVMD can be found in Appendix A.

For a complex signal $s(t)$ received by UWB radar, we first extract the corresponding positive and negative frequency components by an ideal band-pass filter, thus obtaining two signals containing different directions of head motions, namely two sub-motions of head rotation. As shown in Fig. 10, the Doppler profiles contain dynamic noise caused by small motions in addition to the Doppler shift resulted from head rotation. Dynamic

noise and head rotation usually appear at the same time and overlap in frequency. Fortunately, the tiny motions that cause the noise are usually small in amplitude and velocity, thus concentrating on the low-frequency portion of the Doppler profiles. We then decompose the positive as well as the negative frequencies of the signal into three separate components. The one closest to zero frequency is regarded as the dynamic noise. The remaining two components are signal changes caused by head rotation. This is due to the fact that there are a very large number of reflection points in the head and these reflection points usually have different radius of rotation, thus resulting in a larger range of Doppler shift. Therefore, more components are required to accurately contain all the head rotation information.
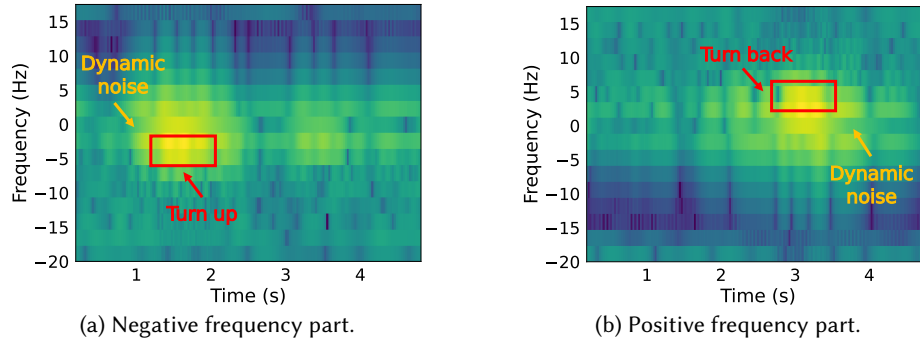

Fig. 11. Doppler profiles of negative frequency part (a) and positive frequency part (b) by STFT.

Take the Doppler profiles in Fig. 10(a) as an example. After the separation of the sub-motions, we obtain the Doppler shifts in both directions separately, as shown in Fig. 11(a) and 11(b). Among them, the low-frequency parts of the positive and negative Doppler profiles have some dynamic noises unrelated to the head rotation, which last much longer than the head rotation duration, and even lead to both positive and negative Doppler shifts at some times. Moreover, these dynamic noises are coupled in frequency with the Doppler shifts due to head rotation, without a clear boundary. After removing the low-frequency dynamic noise components, we get clear contoured head rotation information, which contains temporal rotation features of a large number of head reflection points.

3.5 Obtain High Resolution Time-Frequency Distribution

3.5.1 Disadvantage of Short-Time Fourier Transform. As shown in Fig. 12, the user's head rotation speed usually varies with time, rather than rotating at a constant angular velocity. Besides, head rotation time is difficult to estimate. This makes it difficult to estimate the head rotation angle which needs accurate integration in time and frequency domain.

Since the Doppler shift is related to the angular velocity, an intuitive idea is to use Short-Time Fourier Transform (STFT) to obtain the frequency-time Doppler profiles from the Fourier transform by windowing the sequence signal, from which the angular velocity at different times will be derived. However, according to the Heisenberg's uncertainty principle [49], the temporal resolution and frequency resolution of the signal are mutually constrained, making it hard to estimate rotation angle.

We elaborate on this limitation by the following analysis. For the receveid signal $s(t)$, its STFT is

$$STFT(t,\omega) = \int s(\tau)\gamma_{t,\omega}^*(\tau)d\tau = \int s(\tau)\gamma^*(\tau-t)e^{-j\omega\tau}d\tau = \left\langle s(\tau), \gamma(\tau-t)e^{j\omega\tau} \right\rangle \tag{7}$$
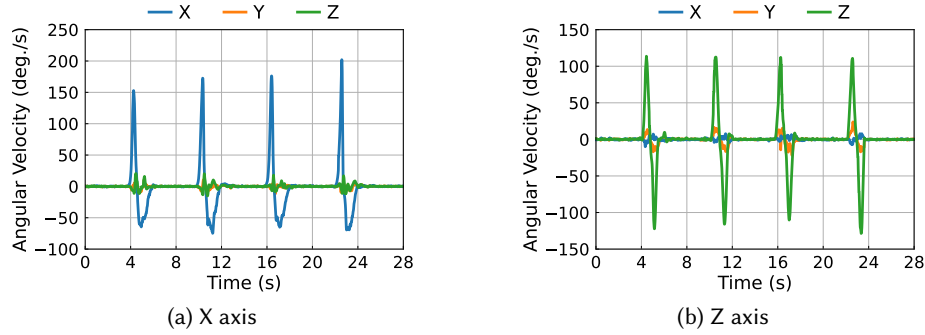
Fig. 12. Angular velocity around X (a) and Z (b) axes. The head is rotated four times on each axis. Angular velocity is measured by the IMU.

where $\gamma(t)$ is window function, which shares the time variable t with the received signal $s(t)$. STFT is to intercept the signal $s(t)$ in the time domain with the window function $\gamma(t)$, and then do the Fourier transform on the intercepted signal. STFT transform makes an inner product of the signal with a sequence of elementary functions $\gamma(\tau - t)e^{j\omega\tau}$ that are finitely supported in the time-frequency domain. If the function $\gamma(t)$ is concentrated at $t = 0$ and its Fourier transform $\Gamma(\omega)$ is concentrated at $\omega = 0$, then the time center $\tau_0 = t$ of the sequence of basis functions $\gamma_{t,\omega}(\tau)$ (note that t is the shift variable) has the time width:

$$\Delta_t^2 = \int (\tau - t)^2 \left|\gamma_{t,\omega}(\tau)\right|^2 d\tau = \int \tau^2 \left|\gamma(\tau)\right|^2 d\tau \tag{8}$$

It means that the time center of $\gamma_{t,\omega}(\tau)$ is determined by t, but the time width is independent of t. Similarly, the Fourier transform $\Gamma_{t,\omega}(v)$ of $\gamma_{t,\omega}(\tau)$ has frequency center $v_0 = \omega$, while the frequency width

$$\Delta_\omega^2 = \frac{1}{2\pi} \int (v - \omega)^2 \left|\Gamma_{t,\omega}(v)\right|^2 dv = \frac{1}{2\pi} \int_{-\infty}^{\infty} v^2 \left|\Gamma(v)\right|^2 dv \tag{9}$$

is also independent of the center frequency ω. Then $STFT(t, \omega)$ in Eq. (7) represents the characteristics of the signal in the neighborhood of $[t - \Delta_t, t + \Delta_t] \times [\omega - \Delta_\omega, \omega + \Delta_\omega]$.

According to Heisenberg's uncertainty principle, if, when $|t| \to \infty$, there is

$$\sqrt{t}s(t) \to 0 \tag{10}$$

then

$$\Delta_t \Delta_\omega \geq \frac{1}{2} \tag{11}$$

where Δ_t, Δ_ω are the time width and frequency width, respectively. For the equal sign to hold, it is required that $s(t)$ is a Gaussian function, i.e.

$$s(t) = Ae^{-\alpha t^2} \tag{12}$$

It can be seen that the performance of STFT in the time and frequency domains are in conflict with each other. If $\gamma(t)$ is chosen to make the time domain have a higher resolution (Δ_t becomes smaller), then the frequency resolution deteriorates (Δ_ω becomes larger); and vice versa. Therefore, the linear joint time-frequency representations, such as STFT, has to trade-off between time-domain and frequency-domain resolution.

3.5.2 Pseudo Wigner-Ville Distribution (PWVD). Pseudo Wigner-Ville Distribution (PWVD) does not use a single linear window for the Fourier transform and is thus not constrained by the Heisenberg's uncertainty principle. PWVD performs Fourier transform of the bilinear instantaneous autocorrelation function of the signal for each instant of time, enabling the highest resolution to be achieved in both time and frequency. This provides the possibility to obtain the exact angular velocity of the head rotation at each moment.

PWVD equation is as follows:

$$PWVD_s(t, \omega) = \int_{-\infty}^{\infty} h(\tau)s(t + \frac{\tau}{2})s^*(t - \frac{\tau}{2})e^{-j\omega\tau}d\tau \tag{13}$$

where $s(t + \frac{\tau}{2})s^*(t - \frac{\tau}{2})$ is the instantaneous autocorrelation function of the signal $s(t)$, τ is the integral variable, and $h(\tau)$ is smoothing function, which is used to reduce the cross terms interference problem [33]. For a signal with multiple components, e.g., $s_n(t) = s_1(t) + s_2(t) = A_1 e^{j\omega_1(t)} + A_2 e^{j\omega_2(t)}$, applying the Fourier transform directly to its instantaneous autocorrelation function produces cross terms:

$$\begin{aligned} W_n(t, \omega) &= \int_{-\infty}^{\infty} s_n(t + \frac{\tau}{2})s_n^*(t - \frac{\tau}{2})e^{-j\omega\tau}d\tau \\ &= \int_{-\infty}^{+\infty} \left[s_1\left(t + \frac{\tau}{2}\right) + s_2\left(t + \frac{\tau}{2}\right) \right] \cdot \left[s_1\left(t - \frac{\tau}{2}\right) + s_2\left(t - \frac{\tau}{2}\right) \right]^* e^{-j\omega\tau}d\tau \\ &= W_{s_1}(t, \omega) + W_{s_2}(t, \omega) + W_{s_1 s_2}(\omega) + W_{s_2 s_1}(t, \omega) \end{aligned} \tag{14}$$

where $W_{s_1}(t, \omega)$ and $W_{s_2}(t, \omega)$ are the Fourier transforms of the instantaneous autocorrelation functions of signals $s_1(t)$ and $s_2(t)$, respectively, and $W_{s_1 s_2}(\omega)$ and $W_{s_2 s_1}(t, \omega)$ are cross terms caused by the interaction of the two signal components. After processing by CVMD, we separate the positive and negative frequency components of the signal and remove the dynamic noise component. This eliminates the cross terms to some extent. The signal contains only the head rotation component at this time. However, due to the complexity of head reflections, multiple reflection points produce multiple effective frequency components, and these frequency components result in cross terms that cause frequency confusion in the time-frequency representation. We extracted an accurate time-frequency representation of the head rotation by suppressing these cross terms using a smoothing function $h(\tau)$. Here we use the Hamming window as the smoothing function $h(\tau)$. Unlike the STFT, the smoothing function $h(\tau)$ is only intended to smooth the result of the Fourier transform of the instantaneous autocorrelation function of the signal at each instant. It does not cause any loss of time and frequency resolution.

In order to compare the effects of STFT and PWVD, we calculate the joint time-frequency representation using each of the two methods for UWB radar data with 40 samples per second for 5 seconds. We set the window size of STFT to 16 and the step size to 1. In Fig. 13, STFT shows fuzzy Doppler profiles. With this setting, the frequency resolution of the STFT is $\frac{40}{16} = 2.5Hz$. The STFT slides the window to perform the Fourier transform, and its time resolution is difficult to compute but numerically greater than $\frac{5}{185} \approx 0.027s$. PWVD can present a fine-grained joint time-frequency representation, as shown in Fig. 14 which achieves the frequency resolution of $\frac{40}{200} = 0.2Hz$ and the time resolution of $\frac{5}{200} = 0.025s$.

3.6 Head Structure-Based Rotation Angle Estimation Algorithm

The high resolution joint time-frequency representation derived from PWVD contains the Doppler profiles caused by the head rotation. However, due to the complexity of the geometrical features of the head [61], the Doppler profiles at one moment contains rotation information for a large number of reflection points. The different properties of these reflection points lead to different Doppler shifts. In order to estimate the head rotation angle from fine-grained massive reflection point Doppler profiles, we designed head structure-based rotation angle estimation algorithm. The specific flow of the algorithm is as follows.
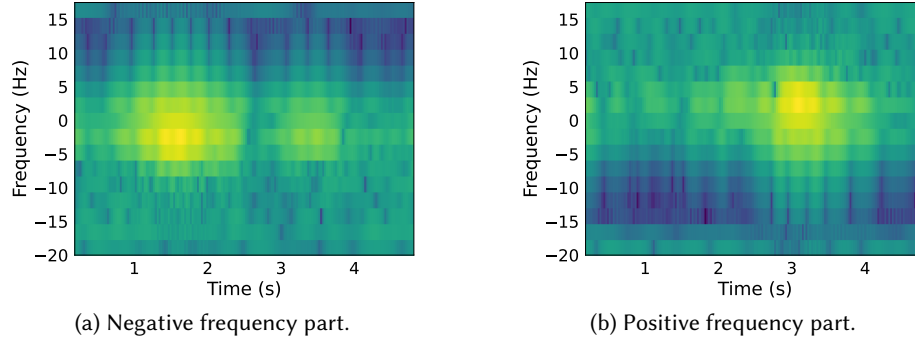
(a) Negative frequency part.

(b) Positive frequency part.

Fig. 13. Joint time-frequency representation by STFT.

(a) Negative frequency part.
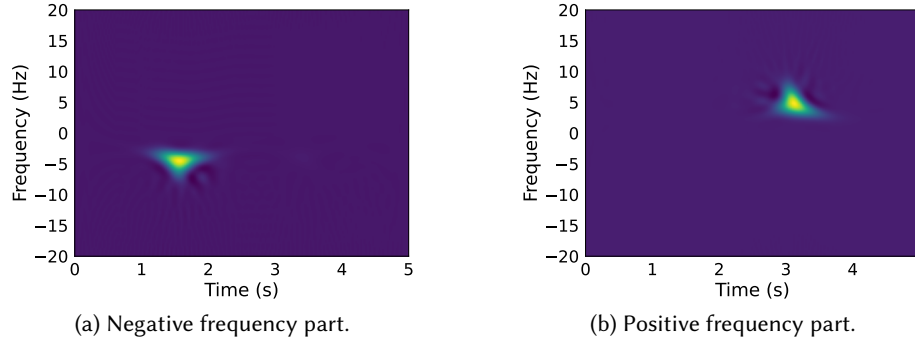
(b) Positive frequency part.

Fig. 14. Joint time-frequency representation by PWVD.

3.6.1 Time-Frequency Representation Binarization. After the previous processing, there are still some noises in the time-frequency representation, such as tiny cross terms that are still present. These noises blur the edges of the Doppler profiles, making it difficult to extract accurate Doppler frequency shift points. Since the energy of these noises is less than the head rotation information, we convert the time-frequency representation to image and then binarise it to remove these noises and extract the head rotation Doppler profiles. Specifically, we use adaptive threshold to take the maximum energy of the time-frequency representation as the benchmark, and points with energy less than a certain scale of the benchmark are considered as noise to be removed.

3.6.2 Reflection Point Extraction. The binarization provides the possibility to extract the angular velocity of the head rotation at each instant of time. All the reflection points have the same angular velocity at the same moment, but different rotation radius. It results in a larger Doppler shift for reflection points with larger rotation radius. Therefore, We extract the maximum frequency shift points in the Doppler profiles at each instant. These points correspond to the reflection points with the largest rotation radius. This ensures that the points selected at all instant correspond to the same reflection point.

3.6.3 Head Rotation Angle Estimation. In order to calculate the head rotation angle, we need to get the Doppler shift at each instant. So we construct the coordinate system and map the reflection points to coordinates. Specifically, a two-dimensional coordinate system is constructed with the zero Doppler shift at time zero as the origin. The direction of increasing time is taken as the positive direction of the time axis, and the positive

direction of the frequency axis is aligned with the direction of positive Doppler shift. The instant of time and Doppler shift corresponding to each extracted point are used as its coordinates. Then integrate it on the time axis to obtain the area S_D of the Doppler profiles. According to Eq. (1) it is obtained that

$$S_D = \int \frac{2f}{c} \omega(t) r dt = \frac{2f}{c} \theta r \qquad (15)$$

So the head rotation angle can be calculated by $\theta = \frac{cS_D}{2fr}$. The θ is in radians and the radius is in metres. The conversion to angle is $\theta = \frac{90cS_D}{\pi fr}$. The key to accurately estimating the angle of rotation is the calculation of the precise Doppler area, which plays a decisive role in the calculation of θ. Since the radius of the human head does not vary much, small variations within a reasonable range have little effect on the results of the angle estimation.

Here, we summarise the processing of UHead for Doppler profiles by taking the case of Fig. 10(a) as an example. First, we use CVMD to split the Doppler profiles into the positive frequency part and the negative frequency part, as shown in Fig. 15(a) and 15(e), which correspond to the two sub-motions of turn over and turn back. The positive frequency part and the negative frequency part are then divided into three components respectively. Among the three components, the one closest to zero frequency is the dynamic noise. So we remove it and retain the remaining two components, which correspond to the head rotation features, as shown in Fig. 15(b) and 15(f). It can be seen that we remove the dynamic noise properly and extract the accurate head rotation features. Then we use the novel head structure-based rotation angle estimation algorithm to estimate the rotation angle from the head rotation features. Specifically, we binarise the joint time-frequency representation of head rotation after converting it to a picture, and extract the head rotation Doppler profiles, as shown in Fig. 15(c) and 15(g). Next, we extract the Doppler shifts for each moment of the head reflection point with the largest radius of rotation and map them to coordinates by establishing coordinate axes, as shown in Fig. 15(d) and 15(h). Finally, we integrate it to estimate the head rotation angle.

(a) Negative frequency part.   (b) Removing dynamic noise of the negative part.   (c) Binarization of the negative part.   (d) Reflection point extraction of the negative part.

(e) Positive frequency part.   (f) Removing dynamic noise of the positive part.   (g) Binarization of the positive part.   (h) Reflection point extraction of the positive part.
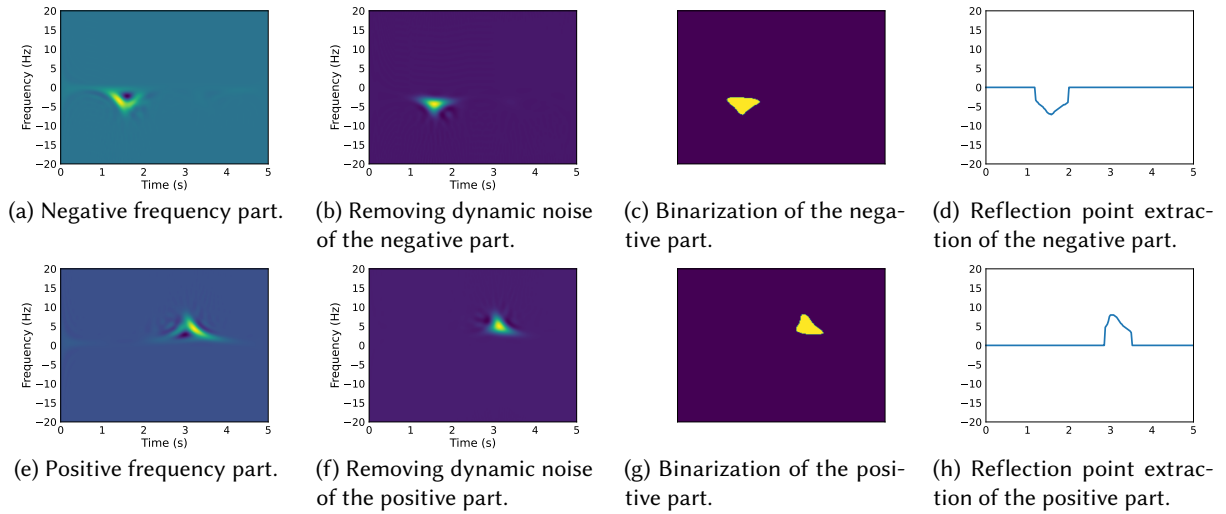
Fig. 15. The rotation angles estimation process of the negative part (a), (b) and the positive part (c), (d). The joint time-frequency representations are obtained by PWVD.

4 EVALUATION

We develop a prototype of UHead using off-the-shelf an IR-UWB radar. In this section, we evaluate UHead's performance by conducting driving scenario tests.
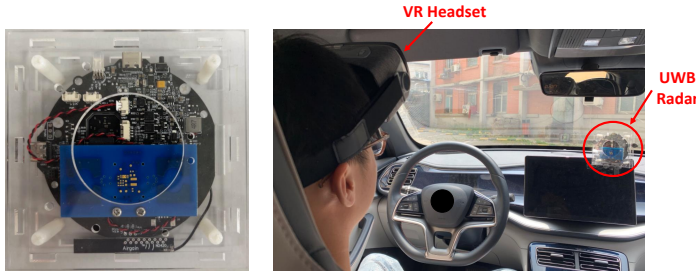
4.1 Experiment Setup and Metrics

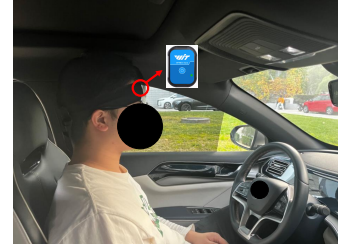Fig. 16. UWB radar used in UHead and setup inside the vehicle.

Fig. 17. The setup of a hair tie worn on the head.

Experiment settings. UHead uses UWB signals for head rotation angle estimation. The core component is a compact, low-cost Novelda X4M05 [4] UWB transceiver. The UWB radar with a high-speed ADC module has a sampling rate of $23.328GS/s$, which makes its ranging accuracy as high as $c/2F_s = 6mm$. The radar transmits pulses with a width of $0.4ns$. The bandwidth of the radar is $1.4GHz$, and it has a distance resolution of about $10cm$. The center frequency of the radio is $8.02GHz$, and the number of frames per second is 40. The distance between neighbouring range bins in a frame of the UWB radar is $5.14cm$. The radar is placed on a phone holder as shown in Fig. 16. We connect the UWB radar to a laptop and use a python program to control data acquisition of radar.

Data collection. We recruit a total of 13 volunteers to systematically evaluate the performance of UHead. In the experiment, we instruct volunteers to perform head rotations in different directions in the driver's seat of vehicle. Specifically, these involve head rotation in four directions: up, down, left and right. These motions consists of different head rotation angles (from $20°$ to $80°$). As with driving, the volunteers rotate the head each time and then turned it back. We performed all experiments in real vehicle scenarios and real driving scenarios. The experimental scenarios involve opening and closing the windows or doors, and turning on and off the air conditioning. The experimental weather included sunny, cloudy, rainy and windy. Our study has received IRB approval.

Ground truth. To obtain the ground truth of head rotation angle, our implementation uses a VR headset with a smart phone inside, which uses the builtin IMU sensor to get the head orientation. To avoid blocking the view, the VR headset is worn on the forehead. The VR headset is merely for collecting ground truth, which is not necessary when UHead is running. In comparison experiment, we use a tiny IMU in hair tie to collect the ground truth, as shown in Fig. 17.

Evaluation metrics. Angle estimation error is used for the evaluation. Considering the effect of positive and negative sign, the error is defined as the absolute value of the difference between absolute value of the estimated angle $|A_C|$ and absolute value of the actual rotation angle $|A_R|$, i.e., $||A_C| - |A_R||$.

4.2 Overall Performance

4.2.1 Overall Performance. We evaluate UHead by calculating the cumulative distribution function (CDF) of the angular estimation error. First, we split each head rotation into two parts, turning over and turning back. Here,

we set the rotation radius of the head up and down around the X-axis to 0.16m , and left and right around the Z-axis to 0.15m. We calculate the angle estimation error for each of the two parts for all the data, as shown in Fig. 18(a). We can see that UHead achieves a median error of 12.59° for turning over and a median error of 12.96° for turning back, which is more than enough to monitor driver attention. We then evaluate further for head rotation in each direction. The result is shown in Fig. 18(b). The median errors for up, down, left, and right are 8.63°, 9.48°, 12.91°, and 18.24°, respectively. All angle estimation errors do not exceed 53°. Since most people have a central field of view of about 60° [12], these errors will not affect the monitoring of the driver's attention too much. For the right motion of the driver's head, as shown in Fig. 19, the windscreen results in a wider observable field of view to the right of the driver's head because the cockpit is on the left. It is therefore able to tolerate marginally larger errors than head rotation in other directions. When monitoring the driver's attention in the car, the relative position of the driver's head orientation to the car component is of concern. The right angle estimate error does not tend to cause a misjudgement of which vehicle component the driver's head is facing.

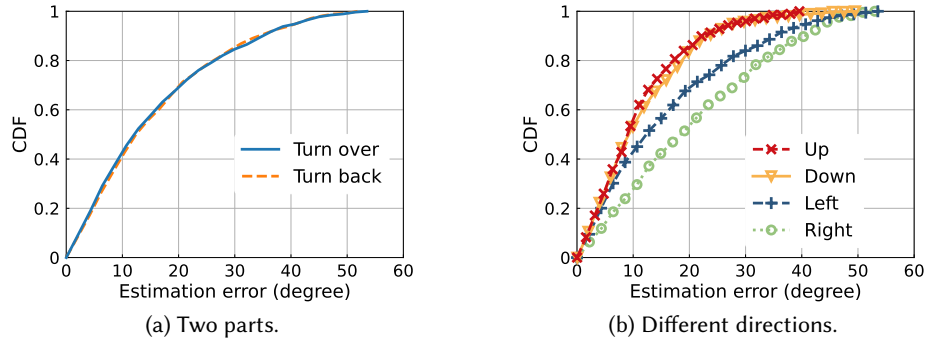|	
(a) Two parts.	(b) Different directions.

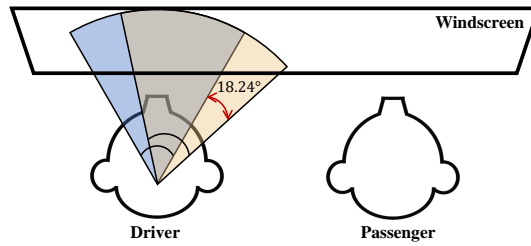Fig. 18. CDF of two parts and different directions of driver head rotation.

Fig. 19. Errors in the estimation of the right rotation angle of the driver's head do not cause much effect on the monitoring of the driver's attention to the right.

4.2.2 Direction Recognition Accuracy. We recognize the direction of driver's head rotation. We train and test the direction recognition module on a 9 person dataset. To verify the generalization ability of the model, we also test it on a completely untrained 3 person dataset. Fig. 20(a) plots the confusion matrix for the 9-person data test set. The average accuracy is 92.1%. The accuracy of the cross person data is shown in Fig. 20(b). We can see that the average accuracy achieves 88.54%. It indicates that the model achieves good generalization ability for cross person data due to the stability of the head rotation features.
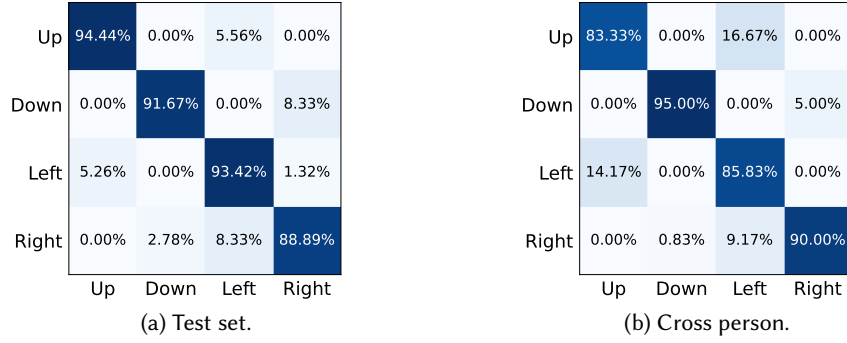
(a) Test set.

(b) Cross person.

Fig. 20. Confusion matrix for test set and cross person.

4.2.3 Range Bins of Head Presence. In order to verify the robustness of identifying the driver's head by distance, we ask 12 volunteers (height from $150cm$ to $185cm$) to adjust driving seat and sitting postures according to their habits, and the head rotation angles are calculated in the bins with ranges from the radar of $71.4cm$, $76.54cm$, $81.86cm$, and $86.82cm$, respectively. The results are shown in Fig. 21. It can be seen that UHead achieves almost the same rotation angle estimation results in all four bins. This shows that, benefiting from the high ranging accuracy of UWB radar, the reflection signals co-occur in range bins, even if the position of the driver changes. The driver's head rotation information can be obtained by selecting any one of the bins.
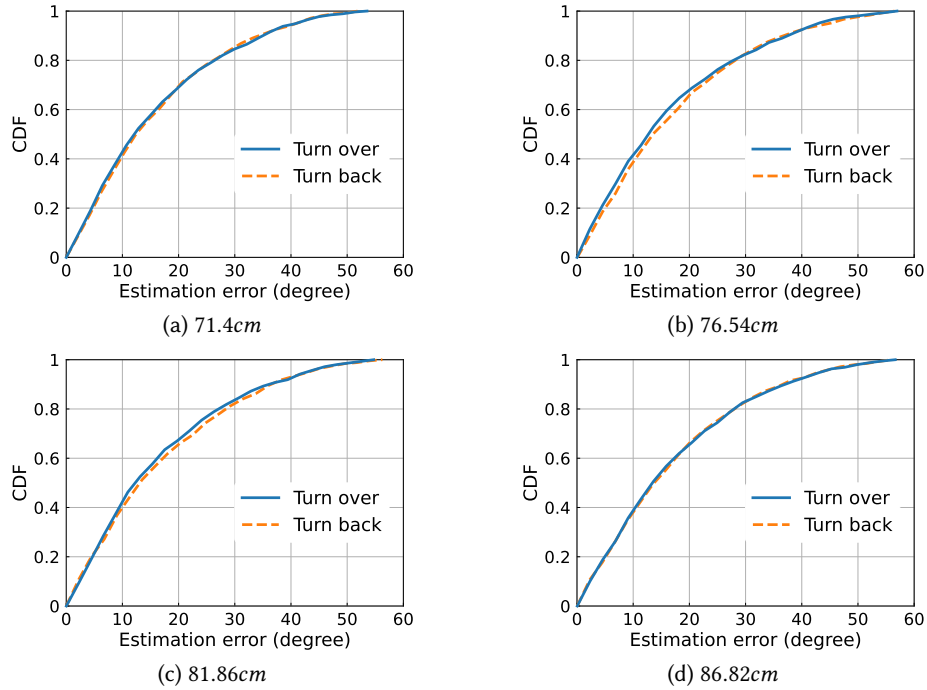

(a) $71.4cm$

(b) $76.54cm$

(c) $81.86cm$

(d) $86.82cm$

Fig. 21. Evaluation results of different range bins.

4.2.4 Effectiveness of Dynamic Noise Reduction. To evaluate the effectiveness of dynamic noise reduction, we use the data containing dynamic noise to estimate the head rotation angle. Fig. 22 plots the angle estimation error with and without dynamic noise removal. The median error without dynamic noise removal is $15.54°$, which is 20.5% higher than after dynamic noise reduction. Moreover, the angle estimation of the data containing dynamic noise is consistently worse than the data after removing the dynamic noise. This is due to the fact that dynamic noise is pervasive and it is always intermixed with the head rotation information in the Doppler profiles, resulting in failing to accurately analyse the head rotation information. The experiment results show that we successfully eliminate the effect of dynamic information besides head rotation and effectively extract the head rotation information.
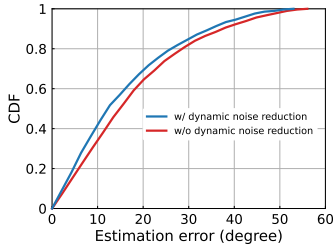

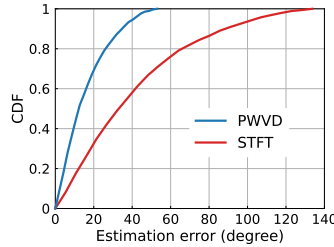Fig. 22. Effectiveness of dynamic noise reduction.

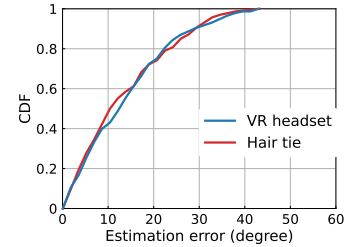Fig. 23. Effectiveness of high resolution time-frequency representation.

Fig. 24. Evaluation results of head worn devices.

4.2.5 Effectiveness of Pseudo Wigner-Ville Distribution. Accurately estimating the head rotation angle relies on high-resolution time-frequency representation. To verify the effect of the resolution of the time-frequency representation, we estimate the head rotation angle using Doppler profiles obtained from both PWVD and STFT. The result is shown in Fig. 23. The median errors using PWVD and STFT are $12.89°$ and $35.19°$, respectively. Moreover, the errors of STFT are always much larger than those of PWVD, and the maximum error of STFT is more than $130°$. The reason for this large error is that each unit in the time-frequency representation obtained by the STFT represents whether or not a certain frequency occurs within a small time window, rather than representing an instantaneous frequency as in the PWVD. Estimating fine-grained head rotation angles requires precise time-frequency information, and thus PWVD achieves far better results than STFT.

4.3 Evaluation of System Robustness

4.3.1 Impact of Head Worn Devices. We test the effect of different size devices worn on the driver's head. We ask volunteer to wear the VR headset shown in Fig. 16 and the hair tie shown in Fig. 17 for the experiment with all other experimental setups being the same. As shown in Fig. 24, almost identical results are achieved when the head wears devices of different sizes. Both the maximum and ninety percent errors are comparable. It indicates that UHead is stable for different sizes of head worn devices.

4.3.2 Impact of Different Relative Locations of Radar and Driver. For real driving, there is a variety of positions of the driver relative to the phone holder. Phone holder usually has limited space on the left side of the driver's seat and more space on the right side. We test the effectiveness of UHead in different locations by placing the UWB radar at the 30 degree angle to the left and the 30, 45, and 60 degree angles to the right of the head orientation when the driver faces straight ahead. The results are shown in Fig. 25. We can see that the UHead achieves similar results for all four locations, and the errors for turning over and turning back are essentially the same. The maximum error of 30 degree angle on the left side is slightly higher than in other cases. This may be caused

by the too narrow space on the left side of the pilot's seat. Even so, the overall error distribution of this case is still very similar to the other cases. It indicates that the UHead is robust to different locations.
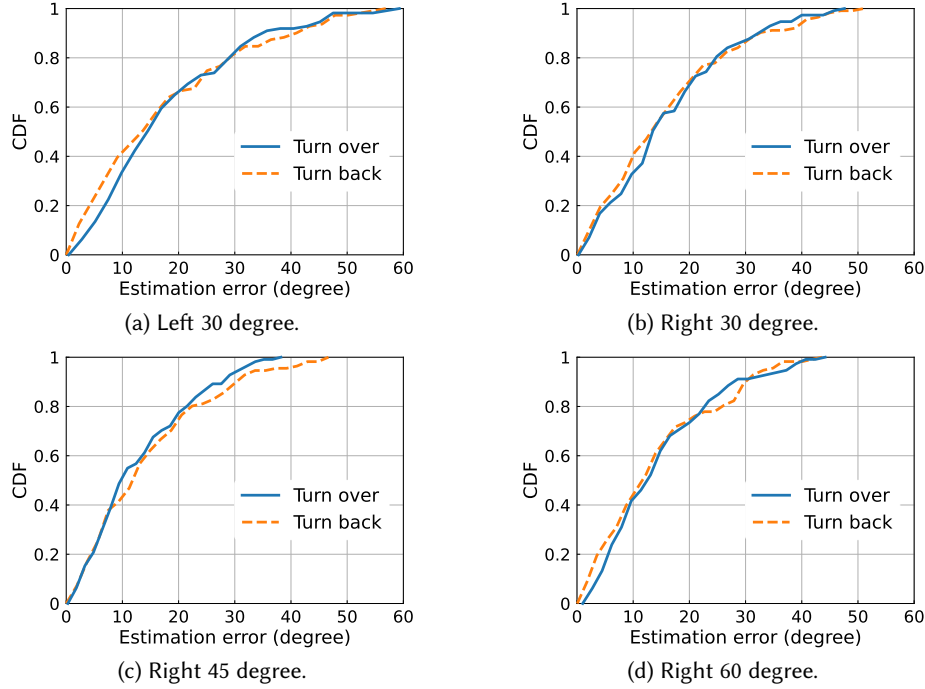

(a) Left 30 degree.

(b) Right 30 degree.

(c) Right 45 degree.

(d) Right 60 degree.

Fig. 25. Evaluation results of different relative locations of radar and driver.

4.3.3 Impact of Presence of Passengers. Drivers are sometimes not alone in the vehicle. Benefiting from the high range resolution of UWB, UHead can focus well on the driver even though he/she is usually very close to the passenger. We test two situations where the passenger sit in the copilot's seat and in the rear. The volunteers involve acted like normal passengers, sitting normally and sometimes suddenly with their head down to look at their phones. The experiment results are shown in Fig. 26. It can be seen that there is no effect when the passenger sits in the rear with the median error being less than 13 degrees. There is a slight increase in the angle estimation error when the passenger is sitting in the copilot's seat, with the median error being about 16 degrees. This may be due to the occasional large motions of the passenger affecting the bin where the driver's head is located, but the effect is very small. Overall, the UHead still achieves good results even with a passenger in the car.

4.3.4 Impact of Interfering Motions. We test the impact on the UHead of common interfering motions while driving, including talking on a phone and steering. Drivers sometimes talk on the phone while driving. Since talking on a mobile phone in the hand while driving is usually an offence to the law [15], we evaluate the situation of talking without phone in hand while the head rotates. To exclude the influence of other factors, we conduct separate experiments with and without talking during head rotation under the same experimental setup. The experimental results are shown in Fig. 27. The results show that both sets of experiments have similar error distributions. The results indicate that talking doesn't influence our method.

(a) Passenger sits in the copilot's seat.
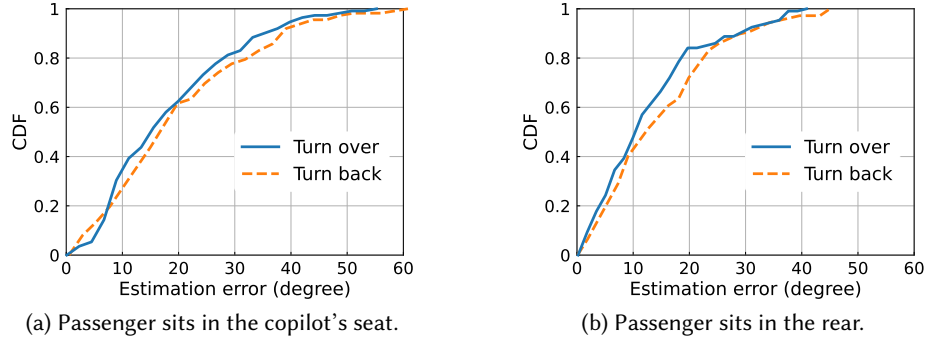
(b) Passenger sits in the rear.

Fig. 26. Evaluation results of presence of passengers.

The experimental results of head rotation while steering are shown in Fig. 28. The overall median error is 15.82°. Compared to the results without steering in Fig. 18(a), the median error for turn over and turn back is 1.9° and 4.11° larger. This may be due to the fact that the steering influence will be in many range bins at the same time, and the occasional shoulder movement during steering affects the range bin where the head is. But the error is still within reasonable levels.

4.3.5 Impact of Angular Velocity. We find in our experiments that the maximum angular velocity of head rotation may be different for different users. In order to verify the influence of the maximum angular velocity on UHead, we analyze the data of user A, whose head rotation angular velocity are in the range of 0-159°/s, and the data of user B, whose angular velocity are in the range of 0-242°/s. The evaluation results are shown in Fig. 29, where different maximum angular velocities have almost identical error distributions. This indicates that the UHead is robust to higher angular velocities.
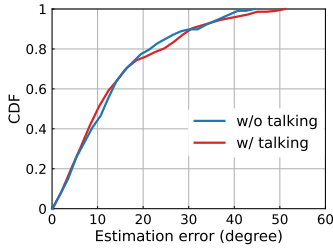
Fig. 27. Evaluation results of head rotation with and without talking.
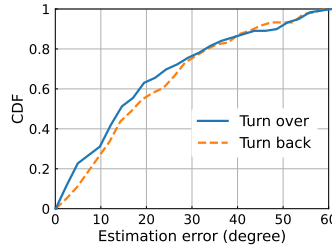
Fig. 28. Evaluation results of head rotation when there is steering activity.
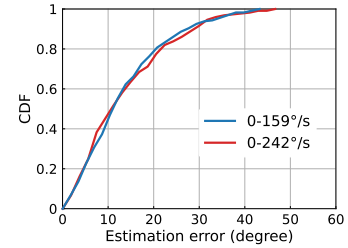
Fig. 29. Evaluation results of different angular velocities.

4.3.6 Impact of Road Bumps. We evaluate the impact of road bumps in a real driving experiment. We focus on the relative motion between the driver's head and the body of the vehicle, and therefore independent of vehicle speed. In our experiment, we drive the car on a heavily trafficked campus road at a safe speed of less than 15 mph as a way to impart enough vibration to the moving vehicle. For safety reasons, a volunteer drive the car and the subject sat in the passenger seat to perform head rotations. The UWB radar is placed on the subject's left side. The experimental results for both parts of the head rotation are shown in Fig. 30(a). The median error is 14.49° for turning over and 13.09° for turning back. As shown in Fig. 30(b), for different rotational directions, the

median errors for up, down, left, and right are 9.29°, 10.34°, 18.92°, and 15.15°, respectively. Since the subject sits in the right seat, the error is only slightly higher than the corresponding error for no road bumps. It proves that UHead can remain stable for road bumps as well.

(a) Two parts.

(b) Different directions.

Fig. 30. CDF of two parts and different directions of driver head rotation of real driving experiments.

5 DISCUSSION

5.1 Limitation of UHead

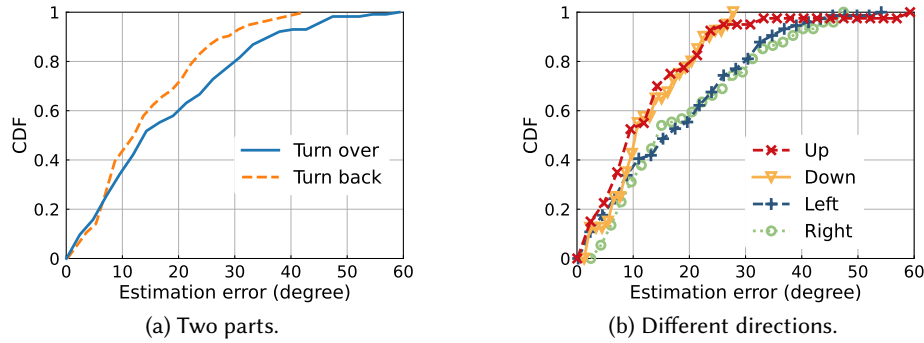We design the UHead to enhance driving comfort and safety. The UHead detects the driver's attention by accurately estimating the direction and angle of 3D rotation of the driver's head on a phone holder, which opens up more possibilities for ADAS. Due to the properties of the Doppler profiles, the UHead does not work well when the mobile phone holder is placed at the angle of 0 degrees to the driver's head. However, the mobile phone may block the driver's view in this case, so this placement of the phone holder is very rare. Moreover, we make the UWB radar face the user to ensure that the user is in the centre of the radar's beam. If the radar is not facing the user, the system effectiveness may be compromised. The impact of system thermal noise is also unclear. In addition, we design UHead as a middleware, so further development is required for customised functionality. Implementing customised functionality on smartphones only requires the development of appropriate software for the smartphone operating system. For car manufacturers, the 3D head rotation information estimated by the mobile phone can be accessed via in-vehicle Bluetooth, allowing further control of the car's driving and interior environment without any changing of the car's production line. It provides more possibilities for enriching the car driving experience.

5.2 Distracted Driving Identification

In recent years, despite a decline in the number of automobile collisions, collisions caused by distracted driving are increasing, resulting in significant morbidity and mortality [47]. Due to the high correlation between driving behaviour and head rotation, 3D head rotation directions and angles offer the potential for more accurate distracted driving identification. When the driver's head rotation angle is greater than a certain threshold, the driver's vision will inevitably lose sight of the road ahead, leading to distraction. It is very dangerous. Accurate 3D head rotation information can be combined with vehicle control to improve driving safety. When driver distraction is detected, the vehicle can make a series of responses to prevent traffic accidents. For example, when the driver's attention is focused in one direction, the vehicle can alert the driver to road conditions in other directions. Moreover, when the driver's head rotates up and down due to fatigue, the vehicle can issue a timely alert to prevent the driver falling asleep, or apply the brakes if necessary.

5.3 Combined with Vehicle System

AR-enabled car windscreens are on the horizon. Headup displays (HUD) are becoming an important component in vehicles such as Audi, BMW and Mercedes-Benz [13]. Combining head rotation information with AR allows vehicles to better adjust to the driver's attention. The HUD can avoid danger by showing real-time road conditions in the blind spot of the driver's field of vision on the windscreen where the driver is concentrating. When driving at night, the vehicle's headlights can follow the direction of the driver's head to give the driver a better view.

5.4 Head Rotation for Health

Previous researches [18, 19, 48] show that head rotation is closely related to driver's physical condition. Specifically, head rotation can lead to changes in blood flow in the carotid, which in turn can lead to health problems such as atherosclerotic disease. Driver health issues are of increasing concern [32, 65, 69] due to the fact that drivers typically spend a lot of time in their vehicles each day, for example, the average American drives 46 minutes a day [3]. It is not easy to establish a link between the head rotation mode and the driver's health. But we think it is an important issue because it can be a good way to improve the usefulness of the system. We take this interesting but challenging issue as our important work for the future.

5.5 More Application Scenarios

With more and more kinds of consumer-grade devices equipped with UWB modules, such as laptops [6], there are more possibilities for UHead applications. When a user's head rotation angle is detected to be too large, indicating that the user's eyes are off the screen, the device can automatically hide the user's private and sensitive data, such as photos and chat logs, to prevent prying eyes. Since RF signals such as UWB cannot access data such as user photos and voice, this solution avoids privacy concerns.

6 RELATED WORK

In this section we discuss existing driver attention monitoring systems.

6.1 Sensor-Based Methods

Many works utilize sensors to monitor driver attention. These jobs require the user to wear additional sensors. Gao et al. [37] detect driver drowsiness based on the spatial-temporal structure of multichannel EEG (electroencephalogram) signals. FDWatch [25] uses PPG (photoplethysmogram) to extract blood oxygen content for driver fatigue detection. SafeDriving [32] collects EMG (electromyogram) sensor signals to detect abnormal driving behavior. However, these sensors are intrusive, which can make drivers feel uncomfortable.

6.2 Vision-Based Methods

There are some works based on visual methods. These works usually deploy a camera at the front of the vehicle to capture the facial information of the driver. Sober-Drive [60] uses a camera to detect driver fatigue by capturing information such as the percentage of closure of the eyelid, blink duration, and blink frequency. Chuang et al. [28] estimate the driver's gaze direction from the video to assess the attention of the driver. CarSafe [66] use the front camera of the cell phone to detect whether the driver is distracted or not. Zhao et al. [68] and Mayer et al. [46] use camera for head pose estimation. However, photographing driver information raises privacy issues, resulting in the driving process no longer being in a private environment. Meanwhile, the system does not always work well in different lighting conditions. Different clothing and glasses also affect the effectiveness of the system.

6.3 Ultrasonic-Based Methods

There are a number of works that utilise ultrasound for driver attention monitoring. FaceOri [56], EHTrack [38] and MOM [36] track the user's head by using the speaker to emit ultrasound waves to measure the movement of the user's earphones. But these methods require the user to wear the earphones in both ears, which can lead to driver discomfort and may cause the driver to ignore the honking of surrounding vehicles. DriverSonar [42] detects risky driving behavior by recognizing head and hand movements using the speaker and microphone of smartphone. VeCare [67] uses ultrasound to detect the breathing and movements of vehicle occupants. However, ultrasound is susceptible to background noise while driving. Besides, weather conditions such as temperature and humidity also have certain influence on the detection accuracy.

6.4 RF-Based Methods

Due to the advantages of RF signals such as high privacy and immune to environmental interference of lighting and weather, a number of driver attention monitoring systems based on RF signals have been proposed. WiDrive [23] recognizes seven in-vehicle activities based on CSI. CARIN [22] further recognizes all possible activity combinations of driver and passenger. V^2iFi [69] estimates driver respiratory rate, heart rate, and heart rate variability using windshield-mounted UWB radar. BlinkRadar [40] performs driver blink monitoring via UWB radar. ER [63] uses audio devices to recognize inattentive driving events. G^3-Guard [59] detects drowsy driving by acoustic sensors. However, these systems can only recognise the possibility of a change in driver attention and are unable to accurately monitor driver attention, which largely limits the widespread use of the system. ViHOT [58] utilizes the driver's prior CSI profiling for head tracking. However, it requires the driver to perform a complex measurement (10 different head positions) in advance, which severely impacts the user experience. Moreover, ViHOT can only track 2D motions of the head. Besides, it requires multiple antennas to be placed in the vehicle, which is difficult to actually deploy in practice. In contrast, UHead provides accurate driver attention monitoring by estimating the direction and angle of 3D head rotation without requiring the driver to take any prior measurements.

7 CONCLUSION

In this paper, we present the design and implementation of UHead, a system for accurately monitoring driver attention placed on a phone holder. We address several challenges, including removing dynamic noise, acquiring high-resolution time-frequency representation and accurately estimating the rotation angle. We design the head structure-based head rotation angle estimation algorithm, which successfully extracts the rotation information from time-varying complex reflections at a certain reflection points. The system is evaluated in real driving scenarios and real roads, and the effectiveness of the components is demonstrated. For our future work, we plan to capture driver health status from head rotation patterns and use UHead for richer user scenarios. This will further extend the usefulness of the system.

REFERENCES

[1] 2006. LS460 achieves a world-first in preventative safety. https://web.archive.org/web/20070927004541/http://www.newcarnet.co.uk/Lexus_news.html?id=5787

[2] 2010. Mobility Transport - Road Safety. https://road-safety.transport.ec.europa.eu/statistics-and-analysis/statistics-and-analysis-archive/young-people/distraction_en

[3] 2015. New Study Reveals When, Where and How Much Motorists Drive. https://newsroom.aaa.com/2015/04/new-study-reveals-much-motorists-drive/

[4] 2017. Single-Chip Radar Sensors with Sub-mm Resolution - XETHRU. https://www.xethru.com/

[5] 2019. Attention Assist: Drowsiness-Detection System Warns Drivers to Prevent Them Falling Asleep Momentarily - Daimler Global Media Site. https://media.daimler.com/marsMediaSite/en/instance/ko/ATTENTION-ASSIST-Drowsiness-detectionsystem-warns-drivers-to-prevent-them-falling-asleep-momentarily.xhtml?oid=9361586

[6] 2020. Lenovo adds Novelda UWB sensor for human presence detection in ThinkPad. https://www.embedded.com/lenovo-adds-novelda-uwb-sensor-for-human-presence-detection-in-thinkpad/

[7] 2020. Xiaomi Introduces groundbreaking UWB Technology. https://www.mi.com/global/discover/article?id=2287

[8] 2021. Google reiterates that Pixel 6 will have UWB as it works to expand support in Android 13. https://9to5google.com/2021/08/25/google-pixel-6-uwb-mention/

[9] 2021. Nearby Interaction. https://developer.apple.com/documentation/nearbyinteraction

[10] 2021. What is Ultra-Wideband (UWB) technology on Samsung Phones? How is it helpful? https://www.smartprix.com/bytes/phones-with-uwb-ultrawideband-connectivity/

[11] 2022. Automotive electronics revolution requires faster, smarter interfaces. https://www.embedded.com/automotive-electronics-revolution-requires-faster-smarter-interfaces/

[12] 2023. https://www.epd.gov.hk/eia/register/report/eiareport/eia_2522017/EIA/html/Appendix/Appendix%2011.1.pdf

[13] 2023. https://www.continental-automotive.com/en-gl/Passenger-Cars/Information-Management/Display-Solutions

[14] 2023. Driver Alert Control. https://www.voxc90.com/volvo_xc90_driver_alert_control-1768.html

[15] 2023. State Cellphone Use While Driving Laws. https://www.ncsl.org/transportation/distracted-driving-cellphone-use

[16] Andrea F Abate, Carmen Bisogni, Aniello Castiglione, and Michele Nappi. 2022. Head pose estimation: An extensive survey on recent techniques and applications. *Pattern Recognition* 127 (2022), 108591.

[17] Fadel Adib, Zach Kabelac, Dina Katabi, and Robert C Miller. 2014. 3D tracking via body radio reflections. In *11th USENIX Symposium on Networked Systems Design and Implementation (NSDI 14)*. 317–329.

[18] Nicolas Aristokleous, Ioannis Seimenis, Georgios C. Georgiou, Yannis Papaharilaou, Brigitta C. Brott, Andrew Nicolaides, and Andreas S. Anayiotos. 2014. Impact of Head Rotation on the Individualized Common Carotid Flow and Carotid Bifurcation Hemodynamics. *IEEE Journal of Biomedical and Health Informatics* 18, 3 (2014), 783–789. https://doi.org/10.1109/JBHI.2014.2305575

[19] Nicolas Aristokleous, Ioannis Seimenis, Yannis Papaharilaou, Georgios C. Georgiou, Brigitta C. Brott, Eleni Eracleous, and Andreas S. Anayiotos. 2011. Effect of Posture Change on the Geometric Features of the Healthy Carotid Bifurcation. *IEEE Transactions on Information Technology in Biomedicine* 15, 1 (2011), 148–154. https://doi.org/10.1109/TITB.2010.2091417

[20] Muhammad Awais, Nasreen Badruddin, and Micheal Drieberg. 2017. A hybrid approach to detect driver drowsiness utilizing physiological signals to improve system performance and wearability. *Sensors* 17, 9 (2017), 1991.

[21] Yunhao Bai, Guoyu Chen, and Xiaorui Wang. 2022. Fusing WiFi Signals and Camera for Driver Activity Recognition based on Deep Learning. In *2022 IEEE 19th International Conference on Mobile Ad Hoc and Smart Systems (MASS)*. 18–26. https://doi.org/10.1109/MASS56207.2022.00011

[22] Yunhao Bai and Xiaorui Wang. 2020. CARIN: Wireless CSI-based driver activity recognition under the interference of passengers. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 1 (2020), 1–28.

[23] Yunhao Bai, Zejiang Wang, Kuangyu Zheng, Xiaorui Wang, and Junmin Wang. 2019. WiDrive: Adaptive WiFi-based recognition of driver activity for real-time and safe takeover. In *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 901–911.

[24] Venkatesh Balasubramanian and K Adalarasu. 2007. EMG-based analysis of change in muscle activity during simulated driving. *Journal of Bodywork and Movement Therapies* 11, 2 (2007), 151–158.

[25] Yetong Cao, Fan Li, Xiaochen Liu, Song Yang, and Yu Wang. 2023. Towards reliable driver drowsiness detection leveraging wearables. *ACM Transactions on Sensor Networks* 19, 2 (2023), 1–23.

[26] Victor C Chen, Fayin Li, S-S Ho, and Harry Wechsler. 2006. Micro-Doppler effect in radar: phenomenon, model, and simulation study. *IEEE Transactions on Aerospace and electronic systems* 42, 1 (2006), 2–21.

[27] Yanjiao Chen, Meng Xue, Jian Zhang, Runmin Ou, Qian Zhang, and Peng Kuang. 2022. DetectDUI: An In-Car Detection System for Drink Driving and BACs. *IEEE/ACM Transactions on Networking* 30, 2 (2022), 896–910. https://doi.org/10.1109/TNET.2021.3125950

[28] Meng-Che Chuang, Raja Bala, Edgar A Bernal, Peter Paul, and Aaron Burry. 2014. Estimating gaze direction of vehicle drivers using a smartphone camera. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*. 165–170.

[29] Sanjay Dey, Sami Ahbab Chowdhury, Subrina Sultana, Md Ali Hossain, Monisha Dey, and Sajal K Das. 2019. Real time driver fatigue detection based on facial behaviour along with machine learning approaches. In *2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON)*. IEEE, 135–140.

[30] Anup Doshi and Mohan Manubhai Trivedi. 2009. On the roles of eye gaze and head dynamics in predicting driver's intent to change lanes. *IEEE Transactions on Intelligent Transportation Systems* 10, 3 (2009), 453–462.

[31] Konstantin Dragomiretskiy and Dominique Zosso. 2013. Variational mode decomposition. *IEEE transactions on signal processing* 62, 3 (2013), 531–544.

[32] Yuanzhao Fan, Fei Gu, Jin Wang, Jianping Wang, Kejie Lu, and Jianwei Niu. 2021. SafeDriving: an effective abnormal driving behavior detection system based on EMG signals. *IEEE Internet of Things Journal* 9, 14 (2021), 12338–12350.

[33] Patrick Flandrin and Bernard Escudié. 1984. An interpretation of the pseudo-Wigner-Ville distribution. *Signal Processing* 6, 1 (1984), 27–36.

[34] Eddie E Galarza, Fabricio D Egas, Franklin M Silva, Paola M Velasco, and Eddie D Galarza. 2018. Real time driver drowsiness detection based on driver's face image behavior using a system of human computer interaction implemented in a smartphone. In *Proceedings of the International Conference on Information Technology & Systems (ICITS 2018)*. Springer, 563–572.

[35] Marco Galvani. 2019. History and future of driver assistance. *IEEE Instrumentation Measurement Magazine* 22, 1 (2019), 11–16. https://doi.org/10.1109/MIM.2019.8633345

[36] Zhihui Gao, Ang Li, Dong Li, Jialin Liu, Jie Xiong, Yu Wang, Bing Li, and Yiran Chen. 2022. Mom: Microphone based 3d orientation measurement. In *2022 21st ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE, 132–144.

[37] Zhongke Gao, Xinmin Wang, Yuxuan Yang, Chaoxu Mu, Qing Cai, Weidong Dang, and Siyang Zuo. 2019. EEG-based spatio–temporal convolutional neural network for driver fatigue evaluation. *IEEE transactions on neural networks and learning systems* 30, 9 (2019), 2755–2763.

[38] Linfei Ge, Qian Zhang, Jin Zhang, and Huangxun Chen. 2023. EHTrack: Earphone-Based Head Tracking via Only Acoustic Signals. *IEEE Internet of Things Journal* (2023).

[39] Magnus R Hestenes. 1969. Multiplier and gradient methods. *Journal of optimization theory and applications* 4, 5 (1969), 303–320.

[40] Jingyang Hu, Hongbo Jiang, Daibo Liu, Zhu Xiao, Schahram Dustdar, Jiangchuan Liu, and Geyong Min. 2022. BlinkRadar: non-intrusive driver eye-blink detection with UWB radar. In *2022 IEEE 42nd International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 1040–1050.

[41] Weijia Jia, Hongjian Peng, Na Ruan, Zhiqing Tang, and Wei Zhao. 2018. WiFind: Driver fatigue detection with fine-grained Wi-Fi signal features. *IEEE Transactions on Big Data* 6, 2 (2018), 269–282.

[42] Hongbo Jiang, Jingyang Hu, Daibo Liu, Jie Xiong, and Mingjie Cai. 2021. Driversonar: Fine-grained dangerous driving detection using active sonar. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–22.

[43] Michael F Land and Benjamin W Tatler. 2001. Steering with the head: The visual strategy of a racing driver. *Current biology* 11, 15 (2001), 1215–1220.

[44] Jialin Liu, Dong Li, Lei Wang, and Jie Xiong. 2021. BlinkListener: " Listen" to Your Eye Blink Using Your Smartphone. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 2 (2021), 1–27.

[45] Nikos Makris, Leonardo Angelone, Seann Tulloch, Scott Sorg, Jonathan Kaiser, David Kennedy, and Giorgio Bonmassar. 2008. MRI-based anatomical model of the human head for specific absorption rate mapping. *Medical & biological engineering & computing* 46 (2008), 1239–1251.

[46] Sven Mayer, Gierad Laput, and Chris Harrison. 2020. Enhancing mobile voice assistants with worldgaze. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–10.

[47] Tiffany L Overton, Terry E Rives, Carrie Hecht, Shahid Shafi, and Rajesh R Gandhi. 2015. Distracted driving: prevalence, problems, and prevention. *International journal of injury control and safety promotion* 22, 3 (2015), 187–192.

[48] Yannis Papaharilaou, Nicolas Aristokleous, Ioannis Seimenis, Mohammad Iman Khozeymeh, Georgios C Georgiou, Brigitta C Brott, Elena Eracleous, and Andreas S Anayiotos. 2013. Effect of head posture on the healthy human carotid bifurcation hemodynamics. *Medical & biological engineering & computing* 51 (2013), 207–218.

[49] John G Proakis. 2007. *Digital signal processing: principles, algorithms, and applications, 4/E.* Pearson Education India.

[50] Muneeba Raja, Viviane Ghaderi, and Stephan Sigg. 2018. WiBot! In-vehicle behaviour and gesture recognition using wireless network edge. In *2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS)*. IEEE, 376–387.

[51] E Donald Sussman, H Bishop, B Madnick, Robert Walter, et al. 1985. Driver inattention and highway safety. (1985).

[52] Johannes Van den Berg. 2006. Sleepiness and head movements. *Industrial health* 44, 4 (2006), 564–576.

[53] Peter M Van Leeuwen, Stefan De Groot, Riender Happee, and Joost CF De Winter. 2017. Differences between racing and non-racing drivers: A simulator study using eye-tracking. *PLoS one* 12, 11 (2017), e0186871.

[54] Fangxin Wang, Jiangchuan Liu, and Wei Gong. 2019. WiCAR: WiFi-based in-car activity recognition with multi-adversarial domain adaptation. In *Proceedings of the International Symposium on Quality of Service*. 1–10.

[55] Fengyu Wang, Xiaolu Zeng, Chenshu Wu, Beibei Wang, and K. J. Ray Liu. 2022. Driver Vital Signs Monitoring Using Millimeter Wave Radio. *IEEE Internet of Things Journal* 9, 13 (2022), 11283–11298. https://doi.org/10.1109/JIOT.2021.3128548

[56] Yuntao Wang, Jiexin Ding, Ishan Chatterjee, Farshid Salemi Parizi, Yuzhou Zhuang, Yukang Yan, Shwetak Patel, and Yuanchun Shi. 2022. Faceori: Tracking head position and orientation using ultrasonic ranging on earphones. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–12.

[57] Yanxue Wang, Fuyun Liu, Zhansi Jiang, Shuilong He, and Qiuyun Mo. 2017. Complex variational mode decomposition for signal processing applications. *Mechanical systems and signal processing* 86 (2017), 75–85.

[58] Xiufeng Xie, Kang G Shin, Hamed Yousefi, and Suining He. 2018. Wireless CSI-based head tracking in the driver seat. In *Proceedings of the 14th International Conference on emerging Networking EXperiments and Technologies*. 112–125.

[59] Yadong Xie, Fan Li, Yue Wu, Song Yang, and Yu Wang. 2020. Real-time detection for drowsy driving via acoustic sensing on smartphones. *IEEE Transactions on Mobile Computing* 20, 8 (2020), 2671–2685.

[60] Lunbo Xu, Shunyang Li, Kaigui Bian, Tong Zhao, and Wei Yan. 2014. Sober-Drive: A smartphone-assisted drowsy driving detection system. In *2014 International conference on computing, networking and communications (ICNC)*. IEEE, 398–402.

[61] Weiye Xu, Wenfan Song, Jianwei Liu, Yajie Liu, Xin Cui, Yuanqing Zheng, Jinsong Han, Xinhuai Wang, and Kui Ren. 2022. Mask does not matter: Anti-spoofing face authentication using mmWave without on-site registration. In *Proceedings of the 28th Annual International Conference on Mobile Computing And Networking*. 310–323.

[62] Xiangyu Xu, Jiadi Yu, Yingying Chen, Yanmin Zhu, Linghe Kong, and Minglu Li. 2019. Breathlistener: Fine-grained breathing monitoring in driving environments utilizing acoustic signals. In *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*. 54–66.

[63] Xiangyu Xu, Jiadi Yu, Yingying Chen, Yanmin Zhu, Shiyou Qian, and Minglu Li. 2017. Leveraging audio signals for early recognition of inattentive driving with smartphones. *IEEE Transactions on Mobile Computing* 17, 7 (2017), 1553–1567.

[64] Xiangyu Xu, Jiadi Yu, Chengguang Ma, Yanzhi Ren, Hongbo Liu, Yanmin Zhu, Yi-Chao Chen, and Feilong Tang. 2022. mmECG: Monitoring Human Cardiac Cycle in Driving Environments Leveraging Millimeter Wave. In *IEEE INFOCOM 2022 - IEEE Conference on Computer Communications*. 90–99. https://doi.org/10.1109/INFOCOM48880.2022.9796912

[65] Mervyn VM Yeo, Xiaoping Li, Kaiquan Shen, and Einar PV Wilder-Smith. 2009. Can SVM be used for automatic EEG detection of drowsiness during car driving? *Safety Science* 47, 1 (2009), 115–124.

[66] Chuang-Wen You, Nicholas D Lane, Fanglin Chen, Rui Wang, Zhenyu Chen, Thomas J Bao, Martha Montes-de-Oca, Yuting Cheng, Mu Lin, Lorenzo Torresani, et al. 2013. Carsafe app: Alerting drowsy and distracted drivers using dual cameras on smartphones. In *Proceeding of the 11th annual international conference on Mobile systems, applications, and services*. 13–26.

[67] Yi Zhang, Weiying Hou, Zheng Yang, and Chenshu Wu. 2023. VeCare: Statistical Acoustic Sensing for Automotive {In-Cabin} Monitoring. In *20th USENIX Symposium on Networked Systems Design and Implementation (NSDI 23)*. 1185–1200.

[68] Zuopeng Zhao, Sili Xia, Xinzheng Xu, Lan Zhang, Hualin Yan, Yi Xu, and Zhongxin Zhang. 2020. Driver distraction detection method based on continuous head pose estimation. *Computational Intelligence and Neuroscience* 2020 (2020), 1–10.

[69] Tianyue Zheng, Zhe Chen, Chao Cai, Jun Luo, and Xu Zhang. 2020. V2iFi: In-vehicle vital sign monitoring via compact RF sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 2 (2020), 1–27.

[70] Tianyue Zheng, Zhe Chen, Shujie Zhang, Chao Cai, and Jun Luo. 2021. More-fi: Motion-robust and fine-grained respiration monitoring via deep-learning uwb radar. In *Proceedings of the 19th ACM conference on embedded networked sensor systems*. 111–124.

A   COMPLEX VARIATIONAL MODE DECOMPOSITION (CVMD)

CVMD extends VMD for processing complex-valued time series, which enables the processing of the positive and negative frequency parts of the signal separately. For a complex signal $s(t)$, it extracts the corresponding positive and negative frequency components by an ideal band-pass filter:

$$B\left(e^{j\omega}\right) = \left\{ \begin{array}{l} 1, 0 \leq \omega < \pi \\ 0, -\pi \leq \omega < 0 \end{array} \right. \tag{16}$$

Two analytic signals are generated by the ideal band-pass filter in Eq. (16), i.e., $B\left(e^{j\omega}\right) S\left(e^{j\omega}\right)$ and $B\left(e^{j\omega}\right) S^*\left(e^{-j\omega}\right)$ in frequency domain, where $S\left(e^{j\omega}\right)$ is the Fourier transform of $s(t)$. It will not lose information that the real part of analytic signals is dealt with only. Thus, we can deduce two real signals using Eq. (17) and (18)

$$s_+(t) = \Re\left\{\mathcal{F}^{-1}\left(B\left(e^{j\omega}\right) S\left(e^{j\omega}\right)\right)\right\} \tag{17}$$

$$s_-(t) = \Re\left\{\mathcal{F}^{-1}\left(B\left(e^{j\omega}\right) S^*\left(e^{-j\omega}\right)\right)\right\} \tag{18}$$

where * denotes complex conjugate operation and $\Re\{\bullet\}$ is the operator that extracts the real part of a complex function, and $\mathcal{F}^{-1}(\bullet)$ is the inverse Fourier operator. According to the analytic characteristics of $B\left(e^{j\omega}\right) S\left(e^{j\omega}\right)$ and $B\left(e^{j\omega}\right) S^*\left(e^{-j\omega}\right)$, the original information can be captured by $s_+(t)$ and $s_-(t)$.

Then, $s_+(t)$ and $s_-(t)$ are decomposed respectively. We non-recursively decompose the real-valued multi-component signal into a discrete number of band-limited sub-signals m_k, which called Intrinsic Mode Functions

(IMFs). Each IMF is compact around a center pulsation ω_k and its bandwidth is estimated using $\mathcal{H}^1$ Gaussian smoothness of the shifted signal.

Firstly, the real valued IMF m_k is turned into an analytic signal m_k^+ with sided frequency spectrum:

$$m_k^+(t) = \left(\delta(t) + \frac{j}{\pi t}\right) * m_k(t) \tag{19}$$

where $\delta(t)$ is the Dirac distribution. Then, after shifting the one-sided spectrum to the 0-frequency baseband, the effective bandwidth of the signal can be determined by calculating the $\mathcal{L}_2$-norm of its time derivative. It can be written as a constrained variational problem:

$$\min_{\{m_k\},\{\omega_k\}} \left\{\sum_{k=1}^{K} \left\|\partial_t \left[m_k^+(t)e^{-j\omega_k t}\right]\right\|_2^2\right\}, \text{s.t.} \sum_{k=1}^{K} m_k(t) = f(t) \tag{20}$$

Introducing Lagrangian multipliers $\lambda(t)$ and a quadratic penalty can resolve the constraint in Eq. (20). The augmented Lagrangian is

$$\mathcal{L}\left(\{m_k\}, \{\omega_k\}, \lambda\right) = \alpha \sum_{k=1}^{K} \left\|\partial_t \left[m_k^+(t)e^{-j\omega_k t}\right]\right\|_2^2 + \left\|f(t) - \sum_{k=1}^{K} m_k(t)\right\|_2^2 + \left\langle\lambda(t), f(t) - \sum_{k=1}^{K} m_k(t)\right\rangle \tag{21}$$

where α is the balancing parameter of the "data-fidelity" constraint. The Alternate Direction Method of Multipliers (ADMM) [39] is used to solve the corresponding unconstrained problem in Eq. (21). All the IMFs with $i < k$ gained from solutions in the frequency domain are updated by Wiener filtering using a filter tuned to the current center frequency on the positive part of the spectrum (i.e., $\omega \geq 0$), that is

$$\hat{m}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i<k} \hat{m}_i^{n+1}(\omega) - \sum_{i>k} \hat{m}_i^n(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha\left(\omega - \omega_k^n\right)^2} \tag{22}$$

where the center frequency ω_k^n is accordingly updated as the center of gravity of the corresponding IMF's power spectrum $\hat{m}_i^{n+1}(\omega)(\omega \geq 0)$. Wiener filtering improves robust for sampling and noise. The ω_k^{n+1} can be computed as

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega \left|\hat{m}_k^{n+1}(\omega)\right|^2 d\omega}{\int_0^\infty \left|\hat{m}_k^{n+1}(\omega)\right|^2 d\omega} \tag{23}$$

If the decomposition level is N, each real signal $s_+(t)$ and $s_-(t)$ is decomposed into N IMFs as

$$s_+(t) = \sum_{i=1}^{N} s_i(t) \tag{24}$$

$$s_-(t) = \sum_{i=-N}^{-1} s_i(t) \tag{25}$$

where $\{s_i(t)\}_{i=1}^{N}$ and $\{s_i(t)\}_{i=-N}^{-1}$ denotes sets of IMFs corresponding to $s_+(t)$ and $s_-(t)$ respectively.