Wireless CSI-Based Head Tracking in the Driver Seat

Xiufeng Xie, Kang G. Shin, Hamed Yousefi, Suining He
The University of Michigan
{xiufeng,kgshin,hyousefi,suiningh}@umich.edu

ABSTRACT

In recent years, augmented reality (AR) has drawn significant attention in the automotive industry and shown great potential for a variety of driver-assistance applications. Tracking the driver's head is vital to seamlessly merge the AR content in the driver's view but still remains an open problem. Specifically, most existing in-vehicle AR solutions rely on cameras for head tracking, which suffer from low sampling rate, weak performance at night, and even high computation costs. Wearing a dedicated headset, on the other hand, is intrusive and inconvenient for daily driving.

To overcome these limitations, we propose ViHOT, a novel wireless CSI-based predictive & device-free head tracking system for in-vehicle use. Given that drivers usually mount phones on the dashboard for navigation, ViHOT leverages the CSI of the phone's WiFi signal to track the driver's head, with a light-weight design suited for real-time driving assistance. Thanks to the high WiFi frame rate, ViHOT achieves more than 10× sampling rate over conventional camera-based approaches and thus eliminates motion blur. Moreover, ViHOT's novel tracking algorithm accurately translates CSI phase readings to head orientations (merely 4°–10° median error) without relying on any head-mounted device.

CCS CONCEPTS

• **Networks** → **Wireless access points, base stations and infrastructure**; **Wireless local area networks**; • **Human-centered computing** → **Mixed / augmented reality**; **Interaction devices**;

KEYWORDS

In-vehicle Augmented Reality; Off-the-shelf Wi-Fi; Wireless Sensing; Head Tracking

1 INTRODUCTION

Recent years have witnessed an increasing interest in head tracking technologies, which are the foundations of augmented reality (AR) systems, and an explosive growth of AR development in both automotive and consumer electronics industries. For instance, the automotive AR-related market is projected to grow steadily at a compound annual rate of 30% to reach a market size of $8 billion by 2025 [31]. AR-enabled car windshields are emerging, and head-up displays (HUD) from Continental become an important part of the Human-Machine-Interface of the Audi, BMW, and Mercedes-Benz [8]. The AR technology augments the physical world, e.g., critical road lanes, traffic signs, other cars and pedestrians in the driver's head direction with virtual tags, notifications or alarms in real time, providing a safe and pleasant driving experience. Despite the success in each individual area of head tracking and in-vehicle AR, the problem of in-vehicle head tracking remains largely unexplored.

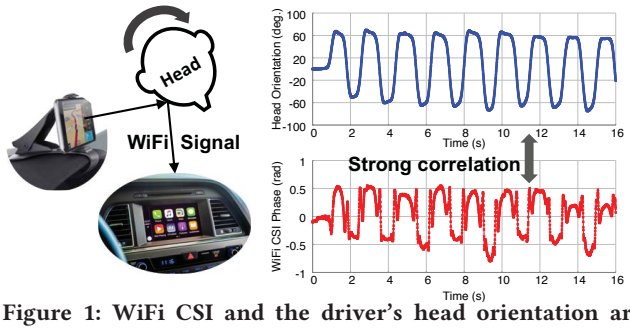


Figure 1: WiFi CSI and the driver's head orientation are strongly correlated at a certain head position.

In-vehicle head tracking has also shown great potential in a variety of advanced driver-assistance systems (ADAS). At a corner-side of night time, the car's headlight can follow driver's head orientation before making a sharp turn to avoid blind spots. It also enables inspection of the side mirror and no fatigue/distracted driving. Furthermore, at an intersection, head turning crowdsourced from drivers can help teach self-driving cars which direction to pay more attention to.

Existing head-tracking systems still suffer from several limitations that hinder their practical in-vehicle use. Generally, they are either camera-based or wearable solutions. For camera-based solutions, their low frame-rate and the resulting motion blur prevents smooth tracking. Besides, the brightness in a car cabin varies greatly with driving scenarios. Unfortunately, the frame quality of a typical camera drops significantly in the dark, while the infra-red cameras like those on the Kinect cannot function well in bright sunlight [41]. High-end cameras and their image processing may mitigate the above problems, but they incur higher financial cost, computational latency, and energy consumption. On the other hand, the wearable solutions [14, 15, 21] require the driver to wear an intrusive and inconvenient headset, which is also unsafe in a vehicle with airbags. Meanwhile, the IMU sensors in the headset are interfered by the vehicle steering [7] and unable to isolate the driver's head motions. To overcome the limitation of existing solutions, *can we accurately track the driver's head using a light-weight & device-free system with a high sampling rate?*

CoNEXT '18, December 4–7, 2018, Heraklion, Greece

ACM ISBN 978-1-4503-6080-7/18/12…$15.00
https://doi.org/10.1145/3281411.3281414

In this paper, we present ViHOT, a device-free Head Orientation Tracking system customized for in-vehicle scenarios. Given the facts that many people drive with their phones on the dashboard for navigation and that modern cars start to incorporate built-in WiFi while adding aftermarket WiFi to an older car is cheap [5], our solution leverages the in-vehicle WiFi links, and tracks the driver's head based on the WiFi signals reflected from her/his head. Specifically, ViHOT leverages the phase of wireless CSI (Channel State Information) collected from the in-vehicle WiFi links to accurately track the driver's head (Fig. 1 shows an example of the CSI-orientation relation). It first collects a *CSI profile* that contains the relation between the head pose (position & orientation) and CSI phase, then matches the run-time CSI reading with the profile to obtain current head pose. This way, *ViHOT's performance is not limited by the light condition, the driver has no need to wear a headset, and the high WiFi frame rate ensures a high sampling rate that makes ViHOT ideal for vehicular applications requiring smooth tracking without motion blur.*

The realization of ViHOT faces several unique challenges:

(i) CSI phase is determined by not only the head orientation, but also the head position: After a driver slightly shifts her/his head position, a different CSI phase will be observed under the same head orientation, thus we cannot simply map current CSI phase to an estimated head orientation following a fixed CSI-orientation relation like in Fig. 1.

To meet this challenge, ViHOT first employs a *position-orientation joint profiling* mechanism (Sec. 3.3) which can quickly build a CSI profile of the CSI-orientation relation at a wide range of head positions. To achieve fast and flexible profiling, we let the driver turn her/his head continuously at each of the different head positions, while the ground truths of head position & orientation can be collected in real time using either headsets or camera on the phone.

ViHOT then employs a *position-orientation joint tracker* (Sec. 3.4) in the run-time, which first estimates the rough head position based on the stable CSI phase before the head turning, and then translates the subsequent phase readings to the orientations according to the CSI-orientation relation at that particular position. The complex in-vehicle signal reflection may lead to the same CSI phase value at different head orientations, making direct CSI-orientation mapping difficult. To uniquely find the present orientation, ViHOT uses both the current and historical CSIs to form a short time-series and then searches for the unique best match of this series in the CSI profile, while applying Dynamic Time Warping (DTW) to cope with different head-turning speeds.

(ii) The strong phase noise in the CSI measurement from the commodity WiFi hardware: The measured CSI phase contains an unknown phase offset caused by the Carrier Frequency Offset (CFO) and Sampling Frequency Offset (SFO) of WiFi hardware [47]. ViHOT handles this challenge by leveraging the multiple antennas of the commodity WiFi device. Since the CSI measurements from different receiving antennas suffer from the same CFO and SFO [38], we compute the CSI phase difference between two receiving antennas to cancel the CFO and SFO noises (Sec. 3.2).

(iii) Various sources of CSI distortions in the cabin besides the driver's head motions: Interference like the passengers' motions and the driver's hand motion when steering the car further complicates the relation between the driver's head pose and CSI phase.

ViHOT handles the CSI polluted by various motions from multiple aspects: First, it leverages the multiple antennas on the commodity WiFi device to suppress the motion interference from the passenger side while retaining the CSI fluctuation caused by the driver's head motions (Sec. 3.5). Second, it leverages the phone's IMU sensors to distinguish the phase variations caused by the large-scale hand steering motion and can fall back to backup solutions like camera-based tracking when CSI is severely polluted (Sec. 3.6).

ViHOT is prototyped on a Dell laptop with Intel 5300 WiFi NIC which uses the 802.11n CSI tool [16] to obtain the raw CSI, while our phone-side implementation runs on commodity Android smartphones (Sec. 4). We evaluate the performance of our prototype under various scenarios by varying drivers, passengers, road conditions, and WiFi interference conditions. Our experimental results (Sec. 5) demonstrate a low error of 4°–10° in terms of median orientation angle estimation, and 400Hz of the sampling rate for head orientation tracking, which is more than 10× conventional camera-based approaches. Furthermore, we observe that ViHOT works reliably under the interfering motions from passengers and hand steering. Our experiment also reveals ViHOT's robustness over a long time between profiling and run-time, which ensures infrequent re-profiling and low maintenance costs. The related works are discussed in Sec. 6, and the paper concludes in Sec. 8.

To our knowledge, ViHOT is the first CSI-based driver head tracking system, with the following contributions:

- Design of ViHOT, a light-weight device-free head tracking system for the in-vehicle scenario, laying a foundation for various in-vehicle AR and ADAS applications.
- Identifying and resolving the unique challenges of realizing CSI-based driver head tracking, including the CSI's sensitivity to the head position and the interference of various motions in the cabin.
- ViHOT is compatible with the commodity WiFi hardware. Our concept of CSI-based head tracking can also be used for dedicated hardware to support more challenging scenarios like tracking the pilot's head in an airplane cockpit.

2 BACKGROUND AND MOTIVATION

2.1 Limitation of Existing Solutions

Head tracking has been a hot research topic for years (Sec. 6). In-vehicle head tracking, however, remains largely unexplored, and existing solutions fail to handle this important use-case.

On one hand, the headset-based solutions require the driver to wear a headset throughout an entire trip, which is intrusive and inconvenient. Furthermore, despite their popularity for indoor VR/AR applications [14, 15, 21], headsets are unsafe during driving as they can be hazardous if the airbag pops out.

On the other hand, the device-free solutions typically use cameras for head tracking. These solutions suffer from low sampling rate and motion blur due to the limitation of rolling-shutter cameras [58], thus failing to track head turning smoothly. For instance, if the driver quickly turns head, camera-based solutions (*e.g.*, the popular FaceRig App [9]) may temporarily lose track of the head. Although the slow-motion cameras incur less motion blur, their wide deployment is still hindered by the high device cost and the heavy processing overhead due to high video frame rates. Moreover, typical cameras perform poorly in the nighttime while the infrared

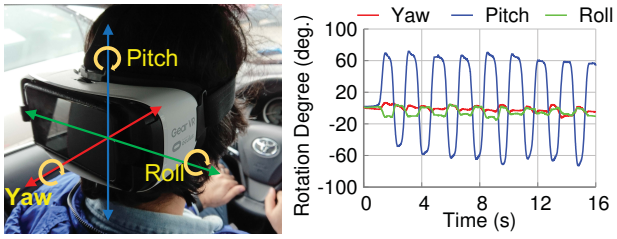


Figure 2: The driver's head typically turns within 2D horizontal plane (headset, merely for ground-truth collection, is not necessary when ViHOT is running).

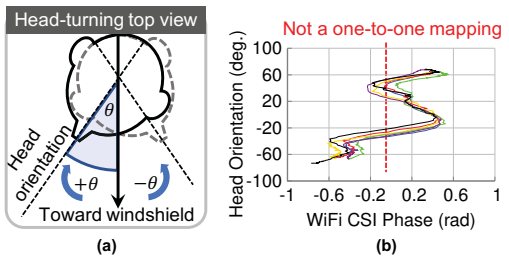


Figure 3: CSI phase vs. head orientation (colors indicate different rounds).

cameras cost much more. Finally, the camera-based solutions require complicated image processing algorithms that are unsuitable for real-time ADAS applications.

2.2 Why WiFi for Head Tracking?

To overcome the limitation of existing head tracking solutions in ADAS, we propose WiFi CSI-based head tracking based on the following two critical observations. First, it is a common practice to mount the smartphones on the dashboard for navigation and hands-free phone control, and almost every phone nowadays has a WiFi interface. Second, most new cars come with built-in WiFi, and it is also cheap to add portable WiFi devices on an old car.

Inspired by these two practical observations, we propose to track the driver's head orientation based on the CSI of the WiFi link between the smartphone on the dashboard and the car's WiFi receiver. Such a CSI-based solution is preferable to existing camera- or headset-based approaches for the following reasons. First, it is device(headset)-free, readily deployable, and also minimizes safety risk and inconvenience. Second, it can support a more than 10× sampling rate over a typical camera thanks to the high WiFi frame rate, and the 2.4GHz WiFi carrier frequency ensures a very small Doppler frequency shift under the human head rotation speed. Therefore, our CSI-based solution is free from the motion blur, unlike the camera-based solutions. Finally, CSI-based head tracking runs a simple matching algorithm, and thus renders an ideal solution for real-time ADAS.

Note that our algorithm is general enough to be extended to other emerging wireless techniques like 60GHz sensing [28, 54, 57], despite its unavailability on existing commodity smartphones for our prototyping at this time.

2.3 Understanding Head Turning & CSI

First, we focus on the head turning rather than eye movement in this paper, as existing measurement studies [24, 44] have shown that the vehicle steering is strongly correlated with the driver's head turning, while the driver's eyes mostly stay close to the head axis ($< 5^{\circ}$) and the eye movements have very weak correlation with the vehicle steering. Although the human head can turn to various orientations in the 3D space, *a driver usually turns her/his head horizontally since most roads are flat*. To validate this, we asked a volunteer driver to wear a VR headset reversely on the back of his head (a detailed setup in Sec. 5) and turn his head repeatedly to check the roadside objects on both sides. The IMU sensor in the headset measure the driver's head-turning angle against the 3 different axes as shown in Fig. 2. The results in the same figure confirm that the head turns mostly in the horizontal plane with only small projections on the other two planes. Therefore, we focus on the 2D head rotation. Our solution can also extend to 3D cases like in the aircraft cockpit, which is beyond the scope of this paper.

In the same experiment, we also simultaneously recorded the resultant WiFi CSI phase during the periodic head turning (more detailed experiment setup in Sec. 4). Before introducing the results, we first define the *head position* i and *head orientation* θ: the head position i is the relative position of the driver's head center *w.r.t.* the WiFi transmitter (the phone), while the head orientation θ is the angle between the direction where the driver faces and the direction from the car's back to the front, as shown in Fig. 3(a). Our results in Fig. 3(b) illustrate that the relation between the CSI phase and head orientation depends not only on the head orientation but also on the head position. Although the driver's head position typically does not vary much during a trip, the CSI is sensitive to slight head position and turning trajectory variations, leading to a set of parallel curves in Fig. 3(b).

To perform head tracking, we first build a *CSI profile* by recording the set of CSI-orientation curves as shown in Fig. 3(b) and label each curve with its corresponding head position. Then in the run-time, our head tracking algorithm can estimate the head position first thus locate the curve corresponding to current head position from the parallel curves in Fig. 3 (Sec. 3.4.1), then translate current CSI reading to the head orientation based on that CSI-orientation curve (Sec. 3.4.2).

Although the CSI phase varies continuously with the horizontal head rotation in Fig. 3(b) at a certain head position, the same phase value can be observed at different head orientations, even within a single head-turning round. In other words, the mapping from head orientation to phase is a non-injective function, and thus the inverse mapping from phase to head orientation does not return a unique output, which is unacceptable for head tracking applications. Therefore, our design must be able to uniquely determine the head orientation even if a current CSI phase reading can be observed at different head orientations in the CSI profile (Sec. 3.4.3).

3 VIHOT DESIGN

In this section, we introduce the design details of ViHOT system. Fig. 4 outlines ViHOT's workflow that involves two devices: a *smartphone* mounted in front of the driver with an HUD-style phone holder and another *WiFi receiver* on the car (a laptop in our prototype). ViHOT operates in two stages: *profiling* and *run-time*. In the profiling stage, it employs a *position-orientation joint profiling* mechanism to build the *CSI profile* of a driver, which collects *(i)* the

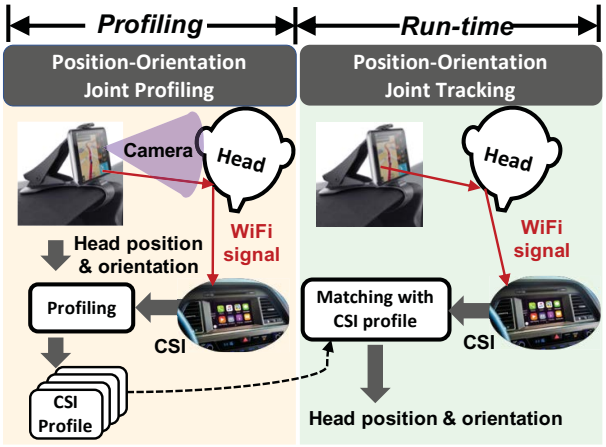


Figure 4: ViHOT system architecture.

ground-truth head position and orientation using the phone's front camera[1] and *(ii)* the coinciding WiFi CSI reading. The profiling can be done quickly (within 100 seconds in our prototype) as our design allows real-time data collection during head rotation. Such profiling only needs to be done infrequently as ViHOT can tolerate a long time gap between profiling and run-time tests (Sec. 5.2.4). Then in the run-time, ViHOT uses a *position-orientation joint tracker* to map current CSI phase reading to the head position and orientation.

3.1 WiFi CSI Model in the Cabin

WiFi sends data over multiple OFDM subcarriers. Let $X_f(t)$ denote the symbol transmitted over subcarrier f at time t, and $Y_f(t)$ denote the corresponding received symbol. We have $Y_f(t) = H_f(t) \cdot X_f(t)$, where the complex-valued factor $H_f(t) = A_f(t)e^{j\phi_f(t)}$ is the CSI of subcarrier f at time t with its amplitude $A_f(t)$ characterizing the channel's attenuation, and phase $\phi_f(t)$ characterizing the phase distortion.

In a car cabin, the WiFi signal reflects on the driver's head and other interior objects, so there are multiple signal paths between the WiFi sender and receiver, and then we have:

$$H_f(t) = \sum_{k=1}^{K} A_f^k(t)e^{j\phi_f^k(t)} = \sum_{k=1}^{K} A_f^k(t)e^{j2\pi\frac{d_k(t)}{\lambda_f}}, \quad (1)$$

where K denotes the total number of paths, $A_f^k(t)$ is the signal attenuation, $d_k(t)$ is the length of propagation path k, and λ_f is the WiFi signal wavelength on channel f.

As in Fig. 4, the driver's head motion changes the length d_{k^*} of a reflection path k^* from the driver's head, thus the phase $\phi_f(t)$ changes along with the head orientation $\theta(t)$ and head position i. Similarly, the motion of the driver body and passengers also cause CSI phase change, which is essentially the interference for the driver head tracking. Sec. 3.6 details how ViHOT handles such interference. Since other car interior objects are stationary[2] with no relative movement *w.r.t.* the WiFi sender and receiver, the signals reflected on them do not contribute to the CSI phase variations.

In summary, the head position i and orientation $\theta(t)$ jointly determine the reflection path length $d_{k^*}(t)$ and finally the CSI phase $\phi_f(t)$, while the target of ViHOT is to estimate i and $\theta(t)$ based on the observed $\phi_f(t)$.

3.2 Removing the CFO & SFO Noise in CSI

Due to the imperfection of commercial hardware, there is a carrier frequency offset (CFO) between the WiFi sender and receiver, causing an unknown phase offset $\beta(t)$ [47]. The sender and receiver also have different A/D sampling clocks [39], and this sampling frequency offset (SFO) causes a time lag Δt and a phase error $2\pi\frac{f}{N}\Delta t$ that increases linearly with the subcarrier index f. The CFO and SFO jointly cause a noisy CSI phase measurement $\hat{\phi}_f(t)$:

$$\hat{\phi}_f(t) = \phi_f(t) + 2\pi\frac{f}{N}\Delta t + \beta(t) + Z_f, \quad (2)$$

where Z_f is the measurement (e.g., thermal) noise.

To remove the phase noise caused by CFO and SFO, we leverage the multiple antennas on the WiFi receiver. Most WiFi devices nowadays support 802.11n/ac, and come with at least two antennas. When receiving packets, these antennas share the same oscillator and sampling clock, thus suffering from exactly the same CFO and SFO *w.r.t.* the single-antenna sender (the phone). In other words, the phase measured on RX antennas 1 and 2 (denoted as $\hat{\phi}_f^1(t)$ and $\hat{\phi}_f^2(t)$) share the same $\beta(t)$ and Δ_t, thus the noises caused by both CFO and SFO can be removed by computing the difference between the two antennas' noisy phase measurements, *i.e.*,

$$\hat{\phi}_f^1(t) - \hat{\phi}_f^2(t) = \phi_f^1(t) - \phi_f^2(t) + (Z_f^1 - Z_f^2). \quad (3)$$

To further reduce the measurement noise $Z_f^1 - Z_f^2$, we compute the average phase difference $\phi(t)$ across all K WiFi subcarriers as $\phi(t) = \frac{1}{K}\sum_{f=1}^{K}\left(\phi_f^1(t) - \phi_f^2(t)\right)$.

As a result, ViHOT utilizes the *phase difference* $\phi(t)$ between the two RX antennas for head tracking. For simplicity, we still use "phase" to denote $\phi(t)$ for the rest of this paper.

3.3 Position-Orientation Joint Profiling

Before running head tracking, ViHOT needs a *CSI profile* of the driver that associates the WiFi CSI with the driver's head position & orientation. Although the objects in the cabin are mostly stationary, if some objects move (like the driver seat adjustment), the CSI profile needs to be updated. As a result, our design calls for a quick CSI profiling with low time overhead. To that end, ViHOT employs a position-orientation joint profiling mechanism that enables a quick and easy profiling process within 100 seconds. Instead of collecting the CSI fingerprint at a discrete set of head positions & orientations as in existing studies, ViHOT obtains the ground truth head orientation & position in real time using wearable headsets, and then labels them with the CSI measured at the same moment. Therefore, the driver can perform the profiling by simply turning her/his head in the horizontal plane to scan through all possible orientations at a certain head position, and then repeating that at different possible positions by leaning head forward/backward.

During the profiling process, the phone keeps streaming tiny packets to the WiFi receiver on the car that extracts CSI phase

[1] Note that the camera is used only for the profiling, and is turned off during run-time. Our prototype uses a headset with IMU sensors to obtain the ground truth of head position and orientation for the evaluation purpose.

[2] In practice, there are many micro-motions and vibrations in the car cabin, but our measurements in Sec. 5.3.1 demonstrate that they only cause tiny CSI phase variations and do not affect ViHOT's performance. They can be even metal objects and lead to strong reflection.

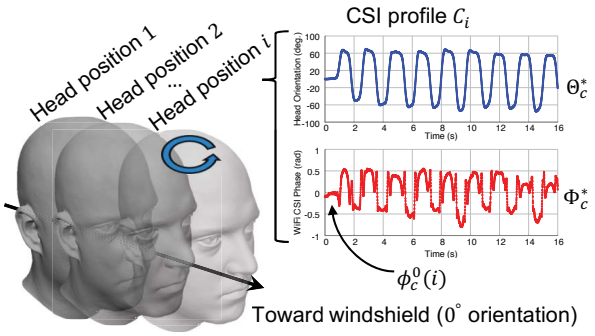


Figure 5: Position-orientation joint profiling.

$\phi_c^*(\tau)$ from each packet received at time τ (For clarity, we use τ for the time index of profiling data and t for the time index at run-time). Meanwhile, the driver turns head from the anatomic leftmost head orientation to the rightmost at a certain head position i and the instant head orientation $\theta_c^*(\tau)$ is measured by the phone based on its front camera[3] in real time. It is worth noting that the head rotation in the CSI profiling stage can be made slow intentionally to guarantee the accuracy of the camera-based head tracking. In the run-time, we use the WiFi CSI phase to estimate the head orientation & position based on the collected CSI profile $\mathcal{P}$, and the head rotation can be much faster than the profiling stage as WiFi sensing has a high sampling rate.

This way, we obtain a time-series of CSI measurements $\Phi_c^* = \{\phi_c^*(\tau)\}$ and another time-series of head orientations $\Theta_c^* = \{\theta_c^*(\tau)\}$. By labeling each CSI measurement with the corresponding head orientation sharing the same timestamp τ, we get a profile set $C_i = \{\Phi_c^*, \Theta_c^*\}$ that contains a seamless CSI-to-orientation mapping at a head position i.

In typical driving, head turns at different time have slightly different head positions and turning trajectories, and the relation between CSI phase and head orientation also varies slightly as shown in Fig. 3. ViHOT addresses this problem by letting the driver repeat the above profiling process for different head positions[4] as illustrated in Fig. 5. This way, we finally obtain a driver's CSI profile $\mathcal{P} = \{C_1, C_2, \ldots, C_i, \ldots\}$ that covers a wide range of possible head positions and orientations. Meanwhile, for head position i, we also record the CSI phase before the head rotation (when the head is at center with 0° head orientation) which is denoted as $\phi_c^0(i)$ and later used as the fingerprint of head position i in Sec. 3.4. Since the CSI profile $\mathcal{P}$ is from a discrete set of head positions & orientations, the run-time head position/orientation to estimate can be drawn from $\mathcal{P}$, i.e., picking the one in $\mathcal{P}$ with the closest CSI features (Sec. 3.4.3).

Intuitively, if the CSI profile $\mathcal{P}$ contains more head positions, the system tends to be more robust. In other words, there is a trade-off between the profiling overhead and system robustness. From our evaluation (Sec. 5), we found that 10 different head positions measured within 100 seconds can already provide a high accuracy in typical driving scenarios. ViHOT also allows to keep updating a driver's CSI profile by adding new traces after each trip so that the system performance can be timely improved after each use.

[3]In order to show the full potential of the CSI-based head tracking, our evaluation still uses a headset to provide the ground truth of the head position & orientation.

[4]Our CSI profiling does not include the abnormal poses like head tilting. Such poses are rare during driving and last only for a short time, ViHOT handles these temporary poses with the algorithm in Sec. 3.4.3.

3.4 Position-Orientation Joint Tracking

Given the driver's CSI profile $\mathcal{P}$, we can run ViHOT head tracking based on the instantaneous CSI phase reading $\phi_r(t)$. The CSI phase is estimated based on the WiFi packets from the smartphone to the vehicle's built-in WiFi. To guarantee a fine-grained phase measurement, when the application data is insufficient to sustain a continuous WiFi packet stream, dummy packets will be inserted/transmitted to maintain a small packet interval. ViHOT adopts a two-level design for head tracking: it first estimates the driver's current head position i based on $\phi_r(t)$ and CSI profile $\mathcal{P}$, and then estimates the driver's head orientation based on $\phi_r(t)$ and the profiled mapping $C_i \in \mathcal{P}$ corresponding to current head position i.

3.4.1 Estimating driver head position based on CSI. Since the CSI phase depends on both the driver head position and orientation, at a first glance, it is difficult to uniquely identify the driver head position based on CSI phase when the head orientation is unknown. However, ViHOT can still realize that thanks to a unique and reasonable feature of the real-world driving scenario — drivers have to always focus on the road in the front for safety, and they will never keep the neck twisted for a long time for comfort. In other words, *if current CSI phase measurement is stable, we know the driver is facing the front with* 0° *head orientation at this time.* We denote the CSI measurement at this time as ϕ_r^0.

Based on this unique feature, ViHOT estimates the driver head position by comparing the run-time phase ϕ_r^0 measured under the 0° head orientation with the phase $\phi_c^0(i)$ corresponding to head orientation i and the same 0° orientation in the CSI profile. Finally, the head position estimation i^* will be the one with the closest profiled phase to ϕ_r^0:

$$i^* = \arg\min_i \left|\phi_c^0(i) - \phi_r^0\right|. \tag{4}$$

3.4.2 Estimating driver head orientation based on CSI & head position. Given current head position estimation i^*, we can then estimate the head orientation based on the CSI profile $C_{i^*} \in \mathcal{P}$, which is essentially the CSI-orientation relation at head position i^*. C_{i^*} contains a pair of time series: the WiFi CSI samples $\Phi_c^* = \{\phi_c^*(\tau) : \tau \in [\tau_S, \tau_E]\}$ and the corresponding head orientations $\Theta_c^* = \{\theta_c^*(t) : t \in [\tau_S, \tau_E]\}$, where τ_S is the starting time of the profiling stage and τ_E is its ending time. Given this CSI profile $C_i = \{\Phi_c^*, \Theta_c^*\}$ and current CSI phase reading $\phi_r(t)$, one may assume that we immediately get a mapping function $\mathcal{R}(\cdot)$ for

$$\theta_c^* = \mathcal{R}(\phi_c^*), \tag{5}$$

Then the current head orientation $\theta_r(t)$ can be directly estimated by mapping the current CSI phase $\phi_r(t)$ to $\hat{\theta_r}(t) = \mathcal{R}(\phi_r(t))$. Unfortunately, our measurements reveal that $\mathcal{R}(\cdot)$ is not a one-to-one mapping (Fig. 3). Although a particular head orientation θ_r^1 leads to a unique CSI phase measurement ϕ_r, another $\theta_r^2 \neq \theta_r^1$ may result in the same ϕ_r, rendering the aforementioned idea inapplicable.

3.4.3 Overview: Series Matching Algorithm. To address the above problem, we design a light-weight series-matching algorithm that estimates the instantaneous head orientation $\theta_r(t)$, taking into account not only the current CSI phase reading $\phi_r(t)$, but also the

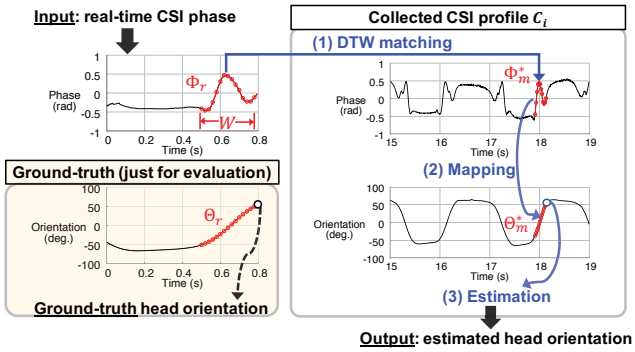


Figure 6: DTW-based series matching algorithm.

historical CSI samples in a time window $\mathcal{T} = [t-W, t]$, where W is the time length of the window. If current $\phi_r(t)$ is unable to uniquely identify its corresponding head orientation $\theta_r(t)$, the historical CSI phase values in $\mathcal{T}$ help us narrow down the scope and obtain a precise estimation of $\theta_r(t)$ without ambiguity. In other words, we transform the problem of head-orientation tracking into a time-series matching problem that matches a time-series containing both current and historical reading ($\Phi_r = \{\phi_r(t) : t \in \mathcal{T}\}$) with the CSI profile series $\Phi_c^* = \{\phi_c^*(\tau) : \tau \in [\tau_S, \tau_E]\}$, rather than the single-point mapping like Eq. (5). Such a series matching algorithm also makes the ViHOT system more robust under bursty motions that cause unintended variations in the WiFi CSI, as a few noisy samples in the input time-series will not change the matching result.

Fig. 6 outlines our matching algorithm:

(Step.1) CSI Series Matching: We first find the best match of the input CSI series Φ_r in the CSI profile series Φ_c^*. Since the CSI sampling interval is random due to WiFi CSMA, we resample Φ_r and Φ_c^* to the same sampling rate before matching them. We denote this best match as $\Phi_m^* = \{\phi_c^*(\tau) : \tau \in [\tau_s, \tau_e]\}$, where the starting and ending time of Φ_m^* are denoted by τ_s and τ_e, respectively. Thus, Φ_m^* is essentially a fragment of the CSI profile series Φ_c^*.

(Step.2) Mapping to Head Orientation: Given Φ_m^* and its time span $[\tau_s, \tau_e]$, we segment the head orientation series Θ_c^* in the CSI profile based on this time span to get the corresponding head orientation series of Φ_m^*, which is denoted as $\Theta_m^* = \{\theta_c^*(\tau) : \tau \in [\tau_s, \tau_e]\}$.

(Step.3) Head Orientation Estimation: Since the current CSI measurement $\phi_r(t)$ is the last sample of the input CSI series $\Phi_r = \{\phi_r(t) : t \in [t-W, t]\}$, the estimation $\hat{\theta}_r(t)$ of current head orientation $\theta_r(t)$ should also be the last sample of Θ_m^*, *i.e.*, $\hat{\theta}_r(t) = \Theta_m^*(\tau_e)$.

3.4.4 Challenge: Mismatched Head-Turning Speed. The different head-rotation speed between profiling and run-time poses a new challenge — given the current input Φ_r and its length W, the length of its best match Φ_m^* remains obscure due to the unknown head-rotation speeds, and we cannot find Φ_m^* without knowing its length.

To handle this speed mismatch, ViHOT employs Dynamic Time Warping (DTW) [35]. Specifically, in order to accommodate the different head-rotation speed between the profiling and run-time, we assume the potential length L_m of Φ_m^* is within a wide range from $0.5W$ to $2W$. We then enumerate a set of candidate lengths $L_n \in [0.5W, 2W]$. For each L_n, ViHOT performs DTW matching to find the potential best match (Φ_n^*) of Φ_r, and then singles out the length $L_m \in \{L_n\}$ leading to the best matching with Φ_r.

Algorithm 1 DTW-based head orientation estimation.

Input: W – Time length of the CSI input window
Φ_r – CSI phase series in current input window
Φ_c^* – Entire CSI phase series in the CSI profile
Θ_c^* – Head orientation series in the CSI profile
Output: $\hat{\theta}_r(t)$ – Current head orientation

1: A setup time of W to fully fill the input window
2: **//Find the best match (Φ_m^*) of Φ_r in Φ_c^***
3: List all possible lengths of Φ_m^* as $L_n \in [0.5W, 2W]$ with a step ΔL
4: **for all** $L_n \in [0.5W, 2W]$ **do**
5: **for all** $\tau_j \in [\tau_S, \tau_E]$ **do**
6: $\Phi_n^*(L_n, \tau_j) = \{\phi_c^*(\tau) : \tau \in [\tau_j, \tau_j + L_n]\}$
7: $d(\Phi_n^*) = \mathbf{DTW}(\Phi_r, \Phi_n^*)$
8: $\Phi_m^*(\tau_s, L_m) = \arg\min d(\Phi_n^*)$
9: The ending time of Φ_m^* is $\tau_e = \tau_s + L_m$
10: $\Theta_m^* = \{\theta_c^*(\tau) : \tau \in [\tau_s, \tau_e]\}$
11: $\hat{\theta}_r(t) = \Theta_m^*(\tau_e)$

3.4.5 Algorithm Design. As summarized in Algorithm.1, the system first needs an initialization time of W to fill the input CSI window (Line 1). Note that this does not add any delay to head tracking after the initialization is finished ($t > W$). Then, given the input CSI series of length W (Line 2), its best match has many potential lengths $L_n \in [0.5W, 2W]$ (Line 4). For each L_n, we search through all possible segments with length L_n in the profile CSI series to find the one that best matches current input $\Phi_r = \{\phi_r(t) : t \in [t-W, t]\}$, *i.e.*, the segment with the minimum distance to Φ_r under DTW (Lines 5–8).

In this way, we obtain a list of best-matching segments $\{\Phi_1^*, \Phi_2^*, \ldots\}$ with different lengths from $0.5W$ to $2W$. Among all elements in this list, the one with the minimum DTW distance to Φ_r is our target Φ_m^* — the overall best match of Φ_r.

After obtaining Φ_m^* along with its starting time τ_s and length L_m, we immediately get its corresponding head-orientation series $\Theta_m^* = \{\theta_c^*(\tau) : \tau \in [\tau_s, \tau_e]\}$, and the last sample of Θ_m^* is our estimation of current head orientation θ_t as discussed in Sec. 3.4.3 (Lines 9–11).

3.4.6 Head Orientation Forecasting. Besides estimating the current head orientation $\theta(t)$, ViHOT is also capable of predicting the head orientation $\theta(t + t_h)$ in the near future, where t_h is the prediction horizon. This way, ViHOT can enable advanced functionalities like smoothing the AR experience by masking the computational delay for AR contents with speculative computation.

Intuitively, if we have found the best match segment $\Phi_m^* = \{\phi_s^*(\tau) : \tau \in [\tau_s, \tau_e]\}$ of current input CSI series $\Phi_r = \{\phi_r(t) : t \in [t-W, t]\}$ in the CSI profile series Φ_c^*, the trend that Φ_r evolves in near future should be similar to how Φ_m^* evolves in the CSI profile series Φ_c^* as both corresponds to the driver head rotations.

Based on this idea, we design the following head-orientation prediction algorithm: the CSI input has length W while its best match has length L_m, then the ratio $\frac{L_m}{W}$ is essentially the turning-speed ratio between the profiling and run-time. With this ratio, predicting t_h at run-time is equivalent to moving forward $t_h \cdot \frac{L_m}{W}$ in the CSI profile. As a result, the CSI phase at $t + t_h$ can be predicted by $\Phi_c^*(\tau_e + t_h \cdot \frac{L_m}{W})$, and its corresponding head orientation is:

$$\hat{\theta}(t + t_h) = \Theta_c^*(\tau_e + t_h \cdot \frac{L_m}{W}). \quad (6)$$

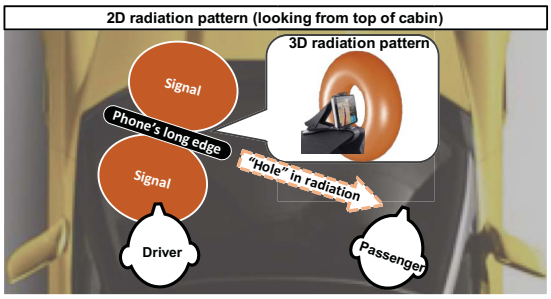


Figure 7: Leveraging the 3D antenna radiation pattern to suppress the signal reflection from the passenger

3.5 Handling the Passenger Interference

In many situations, people do not drive alone, and the passenger sitting by the side of driver can cause strong interference in the CSI variation (the reflected signals from passengers sitting on the back seats are too weak to interfere with head tracking as they are too far away from the smartphone). In theory, if the smartphone can roughly beamform the WiFi signal to the driver, the reflection from the passengers can be significantly reduced. However, commodity smartphones cannot perform transmit beamforming yet. As a readily-deployable solution, ViHOT must handle the passenger interference under omnidirectional WiFi transmission.

To address this challenge, we leverage the 3D radiation pattern of the WiFi antenna inside the phone. Specifically, the phone's WiFi antenna is typically a wire hidden in the long edge of the phone, According to the antenna theory [19], the radiation pattern of such an antenna is like a "donut" as illustrated in Fig. 7. The 2D radiation pattern is omnidirectional in the plane orthogonal to the phone's screen, which is designed to fit how most people hold the phone. However, this radiation pattern is not omnidirectional in the 3rd dimension which is parallel to the antenna. In other words, there is a "hole" in the 3D radiation pattern, and the radiation is the weakest in the direction to which the phone's short edge points. Therefore, the driver can place the phone on the dashboard so that the phone screen points to her/his head to maximize the intended signal strength, while the short edge points to the passenger to suppress the reflection signal from the passenger.

3.6 Handling the Steering Interference

Besides head turning, the driver also frequently turns the steering wheel during a trip, which also alters the WiFi signal reflection path and changes the CSI phase. Since the head orientation can only change continuously, the jumpy estimation caused by a small & bursty steering motion to keep the car straight can be easily filtered out. However, the large-scale steering event like turning at intersection can significantly change the CSI phase in a large time window and ViHOT must handle this challenge.

To demonstrate the impact of the hand movement during the vehicle steering input, we let a test driver keep turning his head back and forth without hand movement, and then start to turn the steering wheel back and forth without any head motion. In the meantime, we record the time-series of WiFi CSI phase and the driver's head orientation following the setup described in Sec 4. From the results in Fig. 8, we can observe that when the driver turns the steering wheel rather than his head, the head orientation

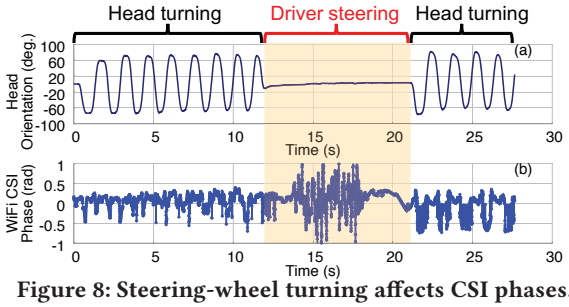


Figure 8: Steering-wheel turning affects CSI phases.

remains unchanged but the CSI phase varies significantly. As a result, if we simply apply the design discussed in Sec. 3.4.2 for head tracking, such CSI phase variation caused by the driver's turning of steering wheel may trigger the false alarm of head turning and cause inaccurate tracking result.

3.6.1 Head Rotation vs. Steering Input. In order to distinguish the driver's turning of the steering wheel from head turning in the CSI measurements, we first investigate the differences between them.

First, existing measurement studies [24, 44] revealed a strong correlation between the driver's head rotation and the car's rotation around one second later during a turn, because the driver must turn her/his head to visually check the road before turning the steering wheel, so head turning and driver steering occur at different times. In other words, a large CSI variation comes from either head turning or driver steering, and ViHOT needs to separate these two possibilities. Besides the different timing, the head turning and driver steering also have different impacts on the car's movement. Specifically, the car body will turn only if the driver's hand turns the steering wheel. In contrast, turning the head cannot directly steer the car. This feature enables ViHOT to distinguish the head turning from driver steering based on the car's movement.

3.6.2 Identifying Steering Motion with Phone . ViHOT uses a smartphone to generate the WiFi signal that probes head motions, here we leverage the IMU sensors in the same phone to detect the turning of the car body and identify the driver's steering events [7]. In particular, the smartphone is mounted rigidly on the dashboard, thus the car body's motion will be captured by the phone's IMU sensors. As discussed in Sec. 3.6.1, steering the wheel will immediately redirect the vehicle's movement while turning the head will not, so we can leverage the phone's IMU sensors to estimate whether the car is turning and also whether the driver is steering the wheel, thus finally identify whether a CSI variation at this moment is caused by hand or head movement.

The workflow of the driver steering identifier can be summarized as follows: *(i)* On detecting a CSI variation, ViHOT reads the phone's IMU sensor and estimates whether the car is turning. If so, the CSI variation is from the steering wheel motion. Otherwise, it is from the driver's head rotation. *(ii)* If the CSI variation is from the head rotation, ViHOT updates its estimation of head orientation based on Sec. 3.3; Otherwise, the system falls back to the video-based head tracking using the phone's front camera, as the driver directly faces the phone (WiFi sender) in ViHOT design. More details of this fallback mode are in Sec. 4. It is worth noting that the fallback mode is rarely triggered as the sharp turns are infrequent thus ViHOT system still enjoys significant energy saving

Figure 9: ViHOT prototype in a Toyota Camry car.

and lighter computation than conventional camera-based solutions. Besides, this fallback mode in our design is to compensate for the limitation of the commodity WiFi hardware, which is unable to separate the signal reflections from the steering wheel and those from the driver's head. The next-generation WiFi, 802.11ax [4], will support uplink MU-MIMO, and the receiver can perform beamforming to suppress the signal reflection from the steering wheel. Therefore, the fallback mode is no longer necessary when 802.11ax is available on future cars.

4 VIHOT IMPLEMENTATION

Our ViHOT prototype includes two devices: a smartphone mounted on the dashboard in front of the driver as the WiFi sender, following the placement discussed in Sec. 3.5, and a laptop as the WiFi receiver. We attach two external WiFi antennas to the laptop to test various antenna placements in the vehicle (Sec. 5.2.2), which are connected to the laptop's NIC via the SMA to U.FL converters [17]. Fig. 9 shows a running ViHOT prototype in a Toyota Camry car and the phone is a Sony Xperia XZs.

Profiling: To collect the CSI profile for a driver, we configure the smartphone as a WiFi hotspot without password protection and run an Iperf client on it, while the laptop connects to this hotspot and operates as an Iperf server. We use NTP to roughly synchronize the phone and the laptop. This way, the phone keeps sending a continuous stream of UDP packets to the laptop which uses an Intel 5300 NIC[16] to measure the CSI from its received packets in real time. To obtain the ground-truths of head position & orientation, our implementation uses a Samsung GearVR headset with a smartphone inside, which uses the builtin IMU sensor to get the head position and orientation (Sec. 5.1). Note that, as a general framework, ViHOT can easily employ other head-orientation tracking schemes (e.g., Apple Face ID [3]) to achieve more accurate ground-truth labeling, which is beyond the scope of this paper.

Run-time: We keep using the aforementioned Iperf client-server setting during the run-time and measure the CSI on the laptop in real time. Moreover, current CSI reading and the collected CSI profile are already available on the laptop. Thus, it can easily run the fine-grained mapping algorithm in Sec. 3.6 to track the head based on current CSI phase reading in real time. Meanwhile, to combat the interference caused by turning the steering wheel, the real-time IMU sensor measurements from the WiFi sender phone are required, which are UDP-streamed to the laptop along with the dummy Iperf packets. The video-based head tracking used in the fallback mode is implemented over the cross-platform `dlib` [22] library. For the performance benchmark, we use a VR headset to obtain the ground-truth head orientation and compare that with the estimation of our ViHOT system.

5 EVALUATION

In this section, we evaluate the performance and efficacy of ViHOT. We first introduce our experiment setup (Sec. 5.1), then conduct extensive experiments to validate our design under various system configurations (Secs. 5.2.1–5.2.5), and report our evaluation results under a wide range of practical factors (Secs. 5.3.1–5.3.5) such as passenger sitting beside the driver and antenna vibration caused by car motion.

5.1 Evaluation Setup

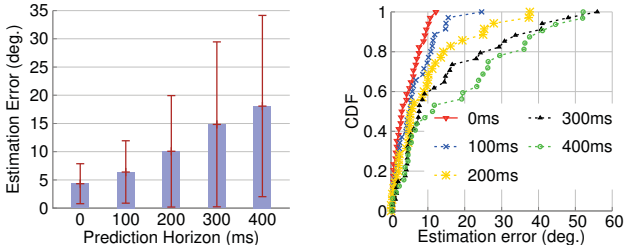
Experiment settings. Unless stated otherwise, we use the following default settings: Our evaluation involves multiple test drivers. For each driver, we first build his unique CSI profile by asking him to move head from front to back through 10 discrete positions as in Fig. 5, and keep turning head back and forth for 10 seconds at each head position. In the run-time, we run each test for 60 seconds, and repeat the test session 10 times. The driver is alone in the vehicle for our experiments unless stated otherwise. Since head tracking focuses on the relative movement between the driver's head and car body, the car speed is unrelated. Meanwhile, the WiFi signals reflected from other vehicles are much weaker than those reflected from the driver's head, as they are far away from the in-vehicle WiFi antennas. Therefore, in our experiments, we drive the car on a campus road with light traffic at a safe speed below 15mph only to create enough vibrations of a moving vehicle.

System configuration. By default, we use a 100ms CSI input window, a 0ms prediction horizon (deactivating head orientation forecast), 1-minute interval between profiling and test, normal head-turning speed around $100^\circ/s$–$120^\circ/s$, and antenna placement layout illustrated in Fig. 9. Furthermore, since our ViHOT prototype operates on the 2.4GHz ISM band, we turn off the car's Bluetooth by default and evaluate the impact of the ISM band interference in Sec. 5.3.5.

Performance metric & benchmark. We assess the angular deviation (absolute estimate error) between ViHOT's head orientation estimation and the *ground truth obtained by a Samsung GearVR headset* worn on the back of the driver's head as in Fig. 2 (to avoid blocking the eyesight and avoid the unintended WiFi signal reflection from the headset rather than the head), and then compute the CDF of the deviations across multiple head-turning events in different experiments.

5.2 Testing Various Configurations

5.2.1 Prediction Horizon. VR/AR applications incur heavy computation overhead and suffer from high rendering latency. Therefore, in existing VR/AR studies, speculative rendering facilitated by the head pose prediction is often used to mask such delays [6, 25]. To support in-vehicle AR, ViHOT is capable of predicting the head orientation in the near future (Sec. 3.4.6.) We evaluate such predictive tracking performance under different prediction horizons. Following the experiment setup in Sec. 5.1, we benchmark ViHOT's head orientation estimations, which are measured under a list of



(a) Mean prediction error. (b) Prediction error CDF.

Figure 10: Orientation prediction accuracy (angular).

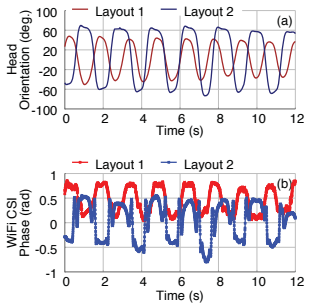


Figure 11: Antenna placement affects the CSI-orientation mapping.

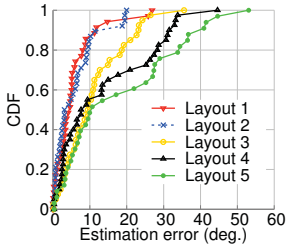


Figure 12: Tracking accuracy under various antenna placements.

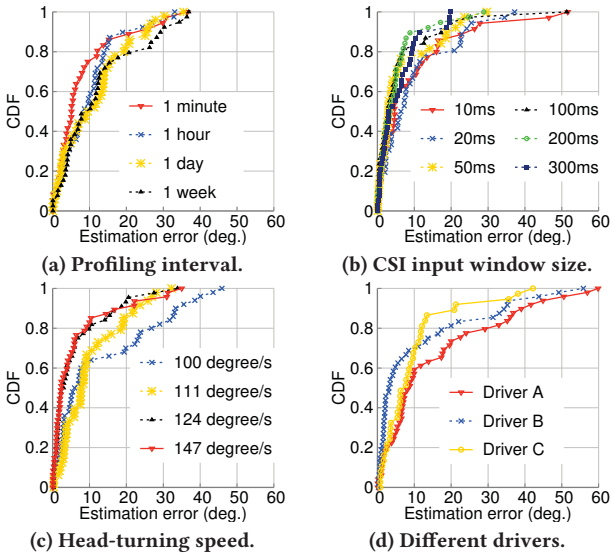
prediction horizons from 0*ms* to 400*ms*, against the ground truth, then plot the average estimation error in Fig. 10a with error bars showing the standard deviation, and CDF of the estimation error across multiple experiment sessions in Fig. 10b.

From the results in Fig. 10a, we first observe that the fine-grained head orientation estimation is very accurate under a small prediction horizon, e.g., only 4° mean angular error for the 0*ms* horizon and 6° for the 100*ms* horizon. This figure also shows that the average estimation accuracy gradually decreases as the prediction horizon increases, but even the aggressive prediction of 400*ms* still has a small 18° estimation error. Such performance degradation is illustrated with more details in Fig. 10b. We see that the maximum error for the 0*ms* horizon is only around 10° and occurs rarely. Similarly, even for the aggressive prediction horizons, notable errors happen very rarely and never exceed 60°.

These results validate ViHOT's predictive tracking design and demonstrate its potential in applications requiring predictive & accurate head orientation tracking.

5.2.2 Antenna Placement. ViHOT needs a WiFi TX (the phone) in front of the driver, which is easy to achieve given the popular phone mounts, but a vehicle manufacturer can install the RX antennas at various places. How does the RX antenna placement affect the ViHOT's performance, and are there any guideline for placing these antennas? We answer these questions by conducting various experiments.

We first compare the relation between CSI phase & head orientation under two different placements: the Layout 1 follows Fig. 9 while Layout 2 has two RX antennas on the center console. From the results in Fig. 11, we observe very different shapes of the CSI phase curves for these two placement layouts even the head-turning patterns are similar. We then conduct more experiments under 5



(a) Profiling interval. (b) CSI input window size.
(c) Head-turning speed. (d) Different drivers.

Figure 13: Tracking accuracy vs. configurations.

different placements and plot the CDFs of the head orientation estimation error in Fig. 11. One can see the performance variations between different layouts: the best one has less than 5° median error while this error can reach 20° for the worst layout. Among the tested layouts, Layout 1 in Fig. 9 offers the best performance, so we use it in all other experiments.

We then investigate why Layout 1 leads to good accuracy: since the head rotation merely changes the signal reflection paths, the head orientation is "modulated" only in the phase of such reflection paths from the head, rather than the line-of-sight (LOS) path between the TX and RX antennas. Meanwhile, for noise cancellation, the CSI phase used in ViHOT is essentially the phase difference between the two RX antennas (Sec. 3.1). As a result, the ideal layout should have one RX antenna receiving only the reflected signals (non-LOS path), and the other RX antenna seeing a LOS path, so the phase difference between these two antennas can retain most of the phase variation caused by the head rotation. The layout in Fig. 9 follows this principle—the direct path from the phone to one RX antenna is blocked by the driver's head, and the other antenna has a LOS path.

5.2.3 CSI Input Window Size. We further evaluate the effect of input window size of the CSI series to find the best trade-off between the performance and the system initialization overhead. Under the other setups in Sec. 5.1, we vary the window size from 10*ms* to 300*ms*. The results in Fig. 13b reveal that a longer CSI input window leads to higher accuracy as it is more robust to the bursty motion noise and hardware thermal noise. Even if a bursty motion or hardware thermal noise pollute a few phase samples in the window, the best matching of the CSI data in the window is less likely to go wrong with more samples in a longer window.

Despite the slightly lower performance under a smaller input window, ViHOT still achieves high accuracy under all these configurations. We also note that even the smallest 10*ms* CSI input window can achieve only 7° of estimation error while eliminating all the ambiguity. In other words, the ViHOT's performance is insensitive to the CSI measurement window size, and we can thus use a small

window to reduce the initialization time and DTW computation overhead in deployment while retaining good accuracy.

5.2.4 Profiling-Runtime Time Interval. Despite the fast profiling, we still wonder how the preceding temporal accuracy degradation would behave after a certain period from the profiling. Besides the default setups, we test 4 different intervals from 1 minute to 1 week. The results in Fig. 13a show ViHOT to be robust under a long time gap as it can still achieve a low median estimation error around 10° even for the very long 1-week case. We also observe that although the accuracy drops with a longer interval, the performance of those long intervals from 1 hour to 1 week shares a similar median estimation error of around 10°.

The reason for this is the driver's sitting position—for the 1-minute test interval, the driver did not leave the seat so that the head positions during the profiling and test remained almost the same and accuracy is the highest, while for other tests with 1 hour to 1 week of profiling-runtime interval, the driver left the seat and when he returned, the head position changed slightly from that in the profiling, leading to a higher median estimation error. Irrespective of how long the interval is, as long as the driver leaves the seat before the test, the level of head position change is similar, while the other objects in the cabin usually do not change much, explaining why the long profiling-runtime intervals from 1 hour up to 1 week share a similar median estimation error and why ViHOT's performance is insensitive to the profiling-runtime interval.

5.2.5 Different Drivers & Head-Turning Speeds. ViHOT allows each driver to quickly build her/his unique CSI profile before running the system. However, a driver's unique features such as the head size, sitting pose and head turning pattern still affect the tracking accuracy. Therefore, we need to compare ViHOT performance under different drivers. In this experiment, following the default setup, we measure the tracking accuracy for 3 different drivers (heights: 170*cm* – 182*cm*) under the same condition, and the results in Fig. 13d demonstrate that ViHOT performs well for all 3 different drivers, with its median tracking error always below 10°.

One major reason of the different tracking accuracy in Fig. 13d is the head turning speed difference due to the drivers' habits. Furthermore, the head-turning speed of the same driver may vary with different driving conditions. Hence, we then focus on the impact of head-turning speed on the system performance. With the default setup and one unchanged driver, we analyze the head orientation tracking accuracy under various head-turning speeds. Fig. 14 showcases the CSI phase and head orientation under two different head-turning speeds for the same driver, and the CDFs across multiple running instances are plotted in Fig. 13c, showing the accuracy improves with the head-turning speed. Although the median estimation errors are always under 10°, we see a higher maximum estimation error for lower head-turning speed. In fact, since our algorithm focuses on finding the best match of the windowed CSI input in the CSI profile, a small head-turning speed yields fewer features in the scope of the fixed 300*ms* sliding window used in this experiment. The shortage leads to more candidates for the matching series with very close DTW distances, and finally degrades the accuracy.

Besides, the results in Fig. 13c also concur with our conjecture that the Doppler frequency shift caused by head rotation is not

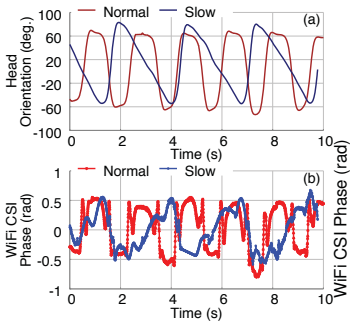


Figure 14: Rotation speed affects CSI curve shape.

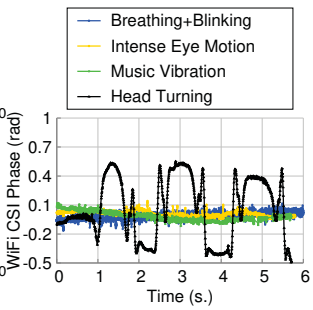


Figure 15: Phase variations vs. micro-motions.

strong enough to devastate the system performance (Sec. 2.2), the accuracy under a higher rotation speed is actually better, *i.e.*, no motion blur problem. We note that the head-turning speed is always above 120°/*s* in typical driving as the driver should always focus on the road in front and turn her/his head very quickly to minimize the time looking away from the road. Therefore, ViHOT always achieves high accuracy[5] in typical driving.

5.3 ViHOT under Various Practical Factors

5.3.1 Micro-Motions in Car Cabin. There are various micro-motions inside the car cabin, such as the natural human breathing and eye blinking, the controlled eye movements when scanning the road condition, and the vibration of surfaces when playing music. How do these micro-motions alter the CSI phase and affect the ViHOT operation? In this experiment, we seek the answer to this question.

We measure the time-series of CSI phase under the aforementioned micro-motions and head turning, and then compare the results in Fig. 15. The driver's head turning is found to cause much stronger phase variations than these micro-motions, so ViHOT remains robust under such micro-motions.

5.3.2 Antenna Vibration. The micro-motions discussed in Sec. 5.3.1 change the WiFi CSI by displacing the reflecting surface of the WiFi signal. The bumpy road condition may also incur a slight vibration of the WiFi TX and RX antennas, thus the WiFi signal path may also change as one or both ends of the path moves. Will that affect ViHOT's performance?

To answer this question, we test both the cases with and without antenna vibration. To evaluate the upper-bound of the antenna vibration's effect, we use the long and soft coil antennas shown in Fig. 9, which vibrate a lot on bumpy roads and in practice the car's built-in antennas have less severe vibration. From the results in Fig. 16, we first see two almost parallel curves for the CSI phase measured under the antenna vibration, meaning that the antenna's vibration has a regular pattern. We can also observe that these two curves have almost the same shape, and the gap between them is very small, implying that the performance should not be affected much. Our performance evaluation results in Fig. 17a concur with

[5]Our evaluation uses a headset to obtain the ground truth of head orientation, but the headset may temporarily slip away during rotation, causing a high but rare error. In future, we plan to improve the evaluation method by using dedicated head tracking devices for the ground truth.

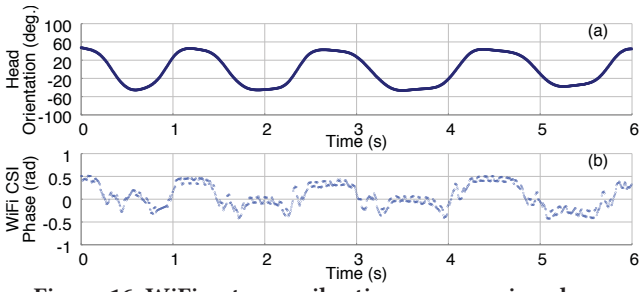
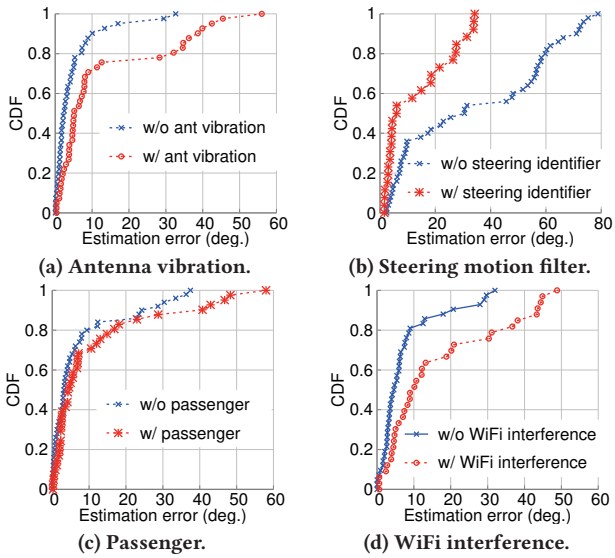


Figure 16: WiFi antenna vibration causes noisy phase.



(a) Antenna vibration. **(b) Steering motion filter.**

(c) Passenger. **(d) WiFi interference.**

Figure 17: Tracking accuracy w/&w/o various factors.

this conjecture. Although the antenna vibration reduces the accuracy, ViHOT's mediate estimation error is still a very small 6°, demonstrating that ViHOT remains robust even under the most extreme antenna vibration.

5.3.3 Driver's Steering Movement. We now validate the driver steering identifier design (Sec. 3.6) that is used to identify and filter out the CSI variations caused by the driver's steering movement. To show the efficacy of this module, we compare the head orientation accuracy measured under the same setup as in Sec. 5.1, while enabling/disabling the driver steering identifier in each experiment.

The results in Fig. 17b first show that the driver's steering movement can affect the head-tracking accuracy, and the estimation error can rise up to 80°. We can also see the significant accuracy improvement when the driver steering identifier is turned on to detect the steering-related CSI phase variations, demonstrating the efficacy of this module.

5.3.4 Presence of Passengers. In many situations, people do not drive alone, and ViHOT has considered the interference of the passenger's movements by attaching the phone (WiFi sender) in front of the driver, so the WiFi signal reflected on the driver's head is much stronger than the signal reflected on the passenger due to a shorter propagation path. The results in Fig. 17c validate this design choice by comparing ViHOT's accuracy measured with and without a passenger sitting by the driver. The participating volunteer acts as a normal passenger who turns his head infrequently to look at roadside scenes. Fig. 17c shows very similar performance for these two cases with close median estimation error. Only a few of head-turning instances yield high estimation error which corresponds to the moments when the passenger turns his head, and the error never exceeds 60°.

5.3.5 Nearby WiFi Traffic. ViHOT is built on the commodity WiFi interface, thus other devices sharing the ISM band may cause interference, e.g., WiFi in the roadside building may interfere with ViHOT. To create the interfering WiFi traffic, we parked the car by the side of a building, then put another laptop on the passenger seat playing a Youtube video through the WiFi connection to the AP in the building, and finally performed experiments according to the setup in Sec. 5.1 during the video streaming.

The results in Fig. 17d show that ViHOT's accuracy degrades under the WiFi interference. Since WiFi has CSMA to avoid packet collisions, the collected CSI remains clean even under the interfering WiFi traffic; then, what causes the performance degradation? Its reason lies in the WiFi frame rate or the CSI sampling frequency. When there is no WiFi interference, ViHOT sends around 500 frames per second at a 34*ms* maximum interval between frames, so the CSI sampling frequency is 500Hz, while this sampling frequency drops to 400Hz in the presence of WiFi interference, with a maximum frame interval of 49*ms*. Since ViHOT resamples current CSI time-series to do the matching with the CSI profile (Sec. 3.4.2), resampling the data that contains a large time gap may lead to inaccurate results and a wrong matching, which, in turn, decreases the accuracy as in Fig. 17d. Despite this performance degradation, ViHOT still achieves an acceptable median angular error of around 10° in the presence of interfering WiFi traffic.

6 RELATED WORK

Device-based Head Tracking: Head tracking shows great potential in the automotive industry, including in-vehicle AR/VR and many subsequent applications [13, 27, 36, 42, 43, 51]. Using embedded inertial measurement units (IMU) like accelerometer and gyroscope, commercial head-mounted displays, represented by Hololens[33], Oculus and HTC VIVE[18], achieve high accuracy of head-orientation tracking to facilitate a smooth VR/AR experience indoors. Since some of the AR/VR headsets are equipped with an internal WiFi transceiver, the authors of [23] proposed to accurately estimate the VR user's head position (rather than the head orientation) by tracking the WiFi transceiver in the headset. Another direction is to leverage the microphone on the headset and the car's speakers to perform acoustic ranging [53]. Unfortunately, wearing a headset during driving is not only inconvenient but also dangerous (Sec. 2.1), and in-vehicle head tracking calls for a device-free solution.

Device-free Head Tracking: Many device-free solutions rely on cameras and the associated image/video processing algorithms to measure the head orientation, and their algorithms mainly fall into two categories: *geometric/shape feature-based* and *appearance/texture feature-based*. The shape-feature-based approaches analyze the face model along with the geometric configuration of facial features (eyes, lip corners, and the bounding box of the face [40]; pupils, nose bottom, and pupil glints [20]; eye corners, nose corners, and the nose tip [32]), while the texture-based approaches directly use raw input data to train a wide range of learning models [2, 10–12, 34].

Wi-Fi-based Human Sensing: Besides the camera-based solutions, human activity detection based on RF signal, especially the prevalent WiFi signal, has drawn significant attention in recent years [59]. Despite simplicity in fingerprinting human activities, Received Signal Strength Indicator (RSSI) hardly handles complex environments with channel dynamics and multi-path [55]. Compared to RSSI, the Channel State Information (CSI) provides more comprehensive information of the wireless channel including the signal attenuation, phase distortion and frequency selectivity, thus enables a broad range of ubiquitous sensing applications for fine-grained human activity detection. It has been shown to capture the large-scale/"macro" human activities, such as detecting human falling down[1, 49], playing exergames [38], recognizing daily household activities [48], detecting human presence [26], and counting people in a crowd [52]. CSI has also been explored in sensing the small-scale/"micro" movements, such as detecting a speaker's mouth shape [45], eavesdropping a conversation [50], identifying the movements of different hands [56], monitoring a person's sleep [29], gauging the breathing rate [30, 46], and recognizing the human gestures [37]. Unfortunately, such existing solutions are not designed for in-vehicle use and are unable to handle the head tracking challenges in a vehicle cabin.

7 DISCUSSION AND FUTURE WORK

Choice of radio frequency: Our current prototype only operates in the 2.4GHz band due to the limitation of the CSI extraction tool [16]. However, the concept of CSI-based head tracking is not limited to the choice of RF signal frequency. At the deployment stage, ViHOT may also work with 5GHz or even 60GHz frequency as long as the phone supports these bands. In fact, we expect a higher ViHOT's performance on the 5GHz or 60GHz band. First, the shorter wavelength of these bands leads to less diffraction, thus improving the tracking accuracy. Besides, the higher propagation loss of the 5GHz and 60GHz signals also reduces the interference from unintended motions like passengers and cars nearby.

More accurate CSI profiling: ViHOT needs the ground-truth head position & orientation in the CSI profiling, which can be obtained by the phone's front camera. However, the camera-based head tracking results still contain errors. In our evaluation, we obtained more accurate ground truth using a headset to avoid underestimating our CSI-based solution due to the inaccurate ground truth. In future, we would like to improve the CSI profiling design with more advanced but costly ground-truth collection. Specifically, we can design a hybrid system that uses an accurate but heavy tracking solution for profiling while using a lightweight tracking solution like CSI-based tracking at run-time. For instance, the new phone models (like Sony Xperia XZs) are already equipped with a slow-motion camera operating at around 1, 000FPS but cannot keep running due to overheating. ViHOT may leverage such slow-motion cameras to achieve more accurate offline profiling and improve the run-time accuracy without extra energy cost or overheating.

3D head tracking: Current ViHOT prototype only tracks the head orientation in the 2D horizontal plane. Although this is reasonable as a driver typically turns head in that way (Sec. 2.3), we still consider extension of ViHOT to the 3D head tracking scenarios like tracking the pilot's head in the plane cockpit. ViHOT is prototyped for 2D tracking due to the limited number of antennas on the 802.11n devices (1 TX and 2 RX antennas are used). Since 802.11ac is gaining popularity, up to 8 antennas may soon become available on the cars and can then be leveraged to beam the WiFi signal to periodically scan the 3D spaces for more accurate head tracking.

Combining with cameras: ViHOT achieves good performance on its own. But it can also be combined with the cameras. Compared to the CSI-based solutions, cameras still have their unique advantages, e.g., they are more robust under various motions in the cabin. In future, we can also build a hybrid system that uses sensor fusion and energy-aware scheduling to make the most of both the CSI-based and camera-based solutions.

Filtering passenger movements: By leveraging the 3D antenna radiation pattern, current ViHOT prototype already mitigates the impact of passenger movements by placing the smartphone (WiFi sender) in front of the driver as described in Sec. 3.5. Since the next-generation WiFi with MU-MIMO capability may soon be available in future vehicles, we also consider use of multiple receiving antennas to perform RX beamforming—applying different beamforming weights to different antennas to strengthen the WiFi signal from the driver side and cancel that from the passenger side, so that we can filter out the WiFi signal reflected from the passenger and neutralize the interference of the passenger's movements.

Computational & energy cost: ViHOT requires only 1D time-series matching for head tracking, while camera-based solutions require 2D image processing. Hence, ViHOT naturally incurs much less computation than camera-based solutions, thus being more suitable for real-time ADAS. Furthermore, in ViHOT design, most computation is done on the WiFi receiver (expected to be embedded soon in the vehicle) rather than the phone, and thus will not drain the phone's battery. The phone only has simple tasks including periodic WiFi packet transmissions to trigger CSI measurement and sharing its IMU readings with the WiFi receiver, consuming much less energy than video processing.

8 CONCLUSION

We have presented ViHOT, a novel WiFi-CSI-based driver head tracking system customized for in-vehicle use. Without requiring any head-mounted device, ViHOT turns the driver's smartphone and the in-vehicle WiFi into a small radar system. It leverages the WiFi signals reflected on the driver's head to accurately track & predict her/his head orientation regardless of the head position. With accurate head orientation estimation with 4°–10° of median angular error and a more than 10× sampling rate than the conventional camera-based solutions, ViHOT enables a wide range of in-vehicle AR/ADAS applications with accurate and smooth head-orientation tracking. Besides the in-vehicle use cases, its novel concept of CSI-based head tracking can be applied to many practical and important scenarios, including indoor/in-cockpit head tracking.

ACKNOWLEDGMENTS

We appreciate the anonymous reviewers for their insightful comments. The work reported in this paper was supported in part by NSF under Grants CNS-1505785 and CNS-1646130.

REFERENCES

[1] Fadel Adib, Zachary Kabelac, Dina Katabi, and Robert C. Miller. 2014. 3D Tracking via Body Radio Reflections. In *11th USENIX Conference on Networked Systems Design and Implementation*. 317–329.

[2] Byungtae Ahn, Jaesik Park, and In-So Kweon. 2014. Real-Time Head Orientation from a Monocular Camera Using Deep Neural Network. In *12th Asian Conference on Computer Vision*. 82–96.

[3] Apple. 2018. About Face ID Advanced Technology. https://support.apple.com/en-us/HT208108 [Online].

[4] Aruba. 2018. 802.11ax White Paper. https://www.arubanetworks.com/assets/wp/WP_802.11AX.pdf [Online].

[5] AT&T. 2018. Connected Car. https://www.att.com/plans/connected-car.html [Online].

[6] Kevin Boos, David Chu, and Eduardo Cuervo. 2016. Flashback: Immersive Virtual Reality on Mobile Devices via Rendering Memoization. In *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 291–304.

[7] Dongyao Chen, Kyong-Tak Cho, Sihui Han, Zhizhuo Jin, and Kang G Shin. 2015. Invisible Sensing of Vehicle Steering with Smartphones. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*. ACM.

[8] Continental. 2018. Continental Head-up-Displays. http://continental-head-up-display.com/. [Online].

[9] FaceRig team. 2018. FaceRig. https://facerig.com/ [Online].

[10] Gabriele Fanelli, Matthias Dantone, Juergen Gall, Andrea Fossati, and Luc Gool. 2013. Random Forests for Real Time 3D Face Analysis. *Int. J. Comput. Vision* 101, 3 (2013), 437–458.

[11] G. Fanelli, J. Gall, and L. Van Gool. 2011. Real Time Head Pose Estimation With Random Regression Forests. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'11)*. 617–624.

[12] Gabriele Fanelli, Thibaut Weise, Juergen Gall, and Luc Van Gool. 2011. Real Time Head Pose Estimation from Consumer Depth Cameras. In *33rd International Conference on Pattern Recognition*. 101–110.

[13] Luke Fletcher and Alexander Zelinsky. 2009. Driver Inattention Detection Based on Eye Gaze-Road Event Correlation. *Int. J. Rob. Res.* 28, 6 (2009), 774–801.

[14] J. L. Gabbard, G. M. Fitch, and H. Kim. 2014. Behind the Glass: Driver Challenges and Opportunities for AR Automotive Applications. *Proc. IEEE* 102, 2 (2014), 124–136.

[15] J. Grubert, T. Langlotz, S. Zollmann, and H. Regenbrecht. 2017. Towards Pervasive Augmented Reality: Context-Awareness in Augmented Reality. *IEEE Transactions on Visualization and Computer Graphics* 23, 6 (2017), 1706–1724.

[16] Daniel Halperin, Wenjun Hu, Anmol Sheth, and David Wetherall. 2011. Tool Release: Gathering 802.11n Traces with Channel State Information. *ACM SIGCOMM Comput. Commun. Rev.* 41, 1 (Jan. 2011), 53.

[17] HIROSE ELECTRIC CO., LTD. 2018. SMT Ultra-Miniature Coaxial Connectors - Product Specifications. http://www.farnell.com/datasheets/307202.pdf [Online].

[18] HTC. 2018. HTC VIVE. https://www.vive.com/us/ [Online].

[19] Richard C Johnson and Henry Jasik. 1984. Antenna Engineering Handbook. *New York, McGraw-Hill Book Company* (1984).

[20] Jeremy Yrmeyahu Kaminski, Dotan Knaan, and Adi Shavit. 2009. Single Image Face Orientation and Gaze Detection. *Mach. Vision Appl.* 21, 1 (2009), 85–98.

[21] H. Kim, J. L. Gabbard, A. M. Anon, and T. Misu. 2018. Driver Behavior and Performance with Augmented Reality Pedestrian Collision Warning: An Outdoor User Study. *IEEE Transactions on Visualization and Computer Graphics* (2018).

[22] Davis E. King. 2009. Dlib-ml: A Machine Learning Toolkit. *Journal of Machine Learning Research* 10 (2009), 1755–1758.

[23] Manikanta Kotaru and Sachin Katti. 2017. Position Tracking for Virtual Reality Using Commodity WiFi. In *Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on*. IEEE, 2671–2681.

[24] 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.

[25] Kyungmin Lee, David Chu, Eduardo Cuervo, Johannes Kopf, Yury Degtyarev, Sergey Grizan, Alec Wolman, and Jason Flinn. 2015. Outatime: Using Speculation to Enable Low-Latency Continuous Interaction for Mobile Cloud Gaming. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 151–165.

[26] C. C. Li and S. H. Fang. 2016. Device-free Human Detection Using WiFi Signals. In *5th Global Conference on Consumer Electronics*. 1–3.

[27] L. Li, Y. Chen, and Z. Li. 2009. Yawning Detection for Monitoring Driver Fatigue Based on Two Cameras. In *12th International IEEE Conference on Intelligent Transportation Systems*. 1–6.

[28] Jaime Lien, Nicholas Gillian, M. Emre Karagozler, Patrick Amihood, Carsten Schwesig, Erik Olson, Hakim Raja, and Ivan Poupyrev. 2016. Soli: Ubiquitous Gesture Sensing with Millimeter Wave Radar. *ACM Trans. Graph.* 35, 4 (2016), 1–19.

[29] X. Liu, J. Cao, S. Tang, and J. Wen. 2014. Wi-Sleep: Contactless Sleep Monitoring via WiFi Signals. In *IEEE Real-Time Systems Symposium*. 346–355.

[30] X. Liu, J. Cao, S. Tang, J. Wen, and P. Guo. 2016. Contactless Respiration Monitoring Via Off-the-Shelf WiFi Devices. *IEEE Transactions on Mobile Computing* 15, 10 (2016), 2466–2479.

[31] Markets and Markets. 2018. Augmented Reality Automotive Market worth 7.98 Billion USD by 2025. https://www.marketsandmarkets.com/PressReleases/automotive-augmented-reality.asp. [Online].

[32] Sujitha Martin, Ashish Tawari, Erik Murphy-Chutorian, Shinko Y. Cheng, and Mohan Trivedi. 2012. On the Design and Evaluation of Robust Head Pose for Visual User Interfaces: Algorithms, Databases, and Comparisons. In *4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. 149–154.

[33] Microsoft. 2018. Microsoft HoloLens. https://www.microsoft.com/en-us/hololens [Online].

[34] S. S. Mukherjee and N. M. Robertson. 2015. Deep Head Pose: Gaze-Direction Estimation in Multimodal Video. *IEEE Transactions on Multimedia* 17, 11 (2015), 2094–2107.

[35] Meinard Müller. 2007. Dynamic Time Warping. *Information Retrieval for Music and Motion* (2007), 69–84.

[36] E. Murphy-Chutorian and M. M. Trivedi. 2010. Head Pose Estimation and Augmented Reality Tracking: An Integrated System and Evaluation for Monitoring Driver Awareness. *IEEE Transactions on Intelligent Transportation Systems* 11, 2 (2010), 300–311.

[37] Qifan Pu, Sidhant Gupta, Shyamnath Gollakota, and Shwetak Patel. 2013. Whole-home Gesture Recognition Using Wireless Signals. In *19th Annual International Conference on Mobile Computing and Networking*. 27–38.

[38] Kun Qian, Chenshu Wu, Zimu Zhou, Yue Zheng, Zheng Yang, and Yunhao Liu. 2017. Inferring Motion Direction Using Commodity Wi-Fi for Interactive Exergames. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 1961–1972.

[39] Souvik Sen, Božidar Radunovic, Romit Roy Choudhury, and Tom Minka. 2012. You Are Facing the Mona Lisa: Spot Localization Using PHY Layer Information. In *10th International Conference on Mobile Systems, Applications, and Services*. ACM, 183–196.

[40] P. Smith, M. Shah, and N. da Vitoria Lobo. 2003. Determining Driver Visual Attention With One Camera. *IEEE Transactions on Intelligent Transportation Systems* 4, 4 (2003), 205–218.

[41] Jesus Suarez and Robin R Murphy. 2012. Using the Kinect for Search and Rescue Robotics. In *Safety, Security, and Rescue Robotics (SSRR), 2012 IEEE International Symposium on*. IEEE, 1–2.

[42] A. Tawari, S. Sivaraman, M. M. Trivedi, T. Shannon, and M. Tippelhofer. 2014. Looking-in and Looking-out Vision for Urban Intelligent Assistance: Estimation of Driver Attentive State and Dynamic Surround for Safe Merging and Braking. In *IEEE Intelligent Vehicles Symposium*. 115–120.

[43] R. Valenti, N. Sebe, and T. Gevers. 2012. Combining Head Pose and Eye Location Information for Gaze Estimation. *IEEE Transactions on Image Processing* 21, 2 (2012), 802–815.

[44] 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.

[45] G. Wang, Y. Zou, Z. Zhou, K. Wu, and L. M. Ni. 2016. We Can Hear You with WiFi! *IEEE Transactions on Mobile Computing* 15, 11 (2016), 2907–2920.

[46] Hao Wang, Daqing Zhang, Junyi Ma, Yasha Wang, Yuxiang Wang, Dan Wu, Tao Gu, and Bing Xie. 2016. Human Respiration Detection with Commodity Wifi Devices: Do User Location and Body Orientation Matter?. In *ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 25–36.

[47] Wei Wang, Alex X Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. 2015. Understanding and Modeling of WiFi Signal Based Human Activity Recognition. In *21st Annual International Conference on Mobile Computing and Networking*. ACM, 65–76.

[48] Yan Wang, Jian Liu, Yingying Chen, Marco Gruteser, Jie Yang, and Hongbo Liu. 2014. E-eyes: Device-free Location-oriented Activity Identification Using Fine-grained WiFi Signatures. In *20th Annual International Conference on Mobile Computing and Networking*. 617–628.

[49] Y. Wang, K. Wu, and L. M. Ni. 2017. WiFall: Device-Free Fall Detection by Wireless Networks. *IEEE Transactions on Mobile Computing* 16, 2 (2017), 581–594.

[50] Teng Wei, Shu Wang, Anfu Zhou, and Xinyu Zhang. 2015. Acoustic Eavesdropping through Wireless Vibrometry. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. ACM, 130–141.

[51] M. Wollmer, C. Blaschke, T. Schindl, B. Schuller, B. Farber, S. Mayer, and B. Trefflich. 2011. Online Driver Distraction Detection Using Long Short-Term Memory. *IEEE Transactions on Intelligent Transportation Systems* 12, 2 (2011), 574–582.

[52] W. Xi, J. Zhao, X. Y. Li, K. Zhao, S. Tang, X. Liu, and Z. Jiang. 2014. Electronic Frog Eye: Counting Crowd Using WiFi. In *IEEE Conference on Computer Communications*. 361–369.

[53] Jie Yang, Simon Sidhom, Gayathri Chandrasekaran, Tam Vu, Hongbo Liu, Nicolae Cecan, Yingying Chen, Marco Gruteser, and Richard P Martin. 2011. Detecting Driver Phone Use Leveraging Car Speakers. In *Proceedings of the 17th annual*

international conference on Mobile computing and networking. ACM, 97–108.
[54] Zhicheng Yang, Parth H. Pathak, Yunze Zeng, Xixi Liran, and Prasant Mohapatra. 2016. Monitoring Vital Signs Using Millimeter Wave. In *ACM International Symposium on Mobile Ad Hoc Networking and Computing*. 211–220.
[55] Zheng Yang, Zimu Zhou, and Yunhao Liu. 2013. From RSSI to CSI: Indoor Localization via Channel Response. *ACM Comput. Surv.* 46, 2 (2013), 25:1–25:32.
[56] Yunze Zeng, Parth H. Pathak, Chao Xu, and Prasant Mohapatra. 2014. Your AP Knows How You Move: Fine-grained Device Motion Recognition Through WiFi. In *1st ACM Workshop on Hot Topics in Wireless*. 49–54.
[57] Yunze Zeng, Parth H. Pathak, Zhicheng Yang, and Prasant Mohapatra. 2016. Human Tracking and Activity Monitoring Using 60 GHz mmWave: Poster Abstract. In *15th International Conference on Information Processing in Sensor Networks*. 1–2.
[58] Lan Zhang, Cheng Bo, Jiahui Hou, Xiang-Yang Li, Yu Wang, Kebin Liu, and Yunhao Liu. 2015. Kaleido: You Can Watch It but Cannot Record It. In *21st Annual International Conference on Mobile Computing and Networking*. ACM, 372–385.
[59] Y. Zou, W. Liu, K. Wu, and L. M. Ni. 2017. Wi-Fi Radar: Recognizing Human Behavior with Commodity Wi-Fi. *IEEE Communications Magazine* 55, 10 (OCTOBER 2017), 105–111.