BlinkListener: "Listen" to Your Eye Blink Using Your Smartphone

JIALIN LIU, School of Software, Dalian University of Technology, Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province, China and University of Massachusetts Amherst, USA

DONG LI, University of Massachusetts Amherst, USA

LEI WANG*, School of Software, Dalian University of Technology, Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province, China

JIE XIONG, University of Massachusetts Amherst, USA

Eye blink detection plays a key role in many real-life applications such as Human-Computer Interaction (HCI), drowsy driving prevention and eye disease detection. Although traditional camera-based techniques are promising, multiple issues hinder their wide adoption including the privacy concern, strict lighting condition and line-of-sight (LoS) requirements. On the other hand, wireless sensing without a need for dedicated sensors gains a tremendous amount of attention in recent years. Among the wireless signals utilized for sensing, acoustic signals show a unique potential for fine-grained sensing owing to their low propagation speed in the air. Another trend favoring acoustic sensing is the wide availability of speakers and microphones in commodity devices. Promising progress has been achieved in fine-grained human motion sensing such as breathing using acoustic signals. However, it is still very challenging to employ acoustic signals for eye blink detection due to the unique characteristics of eye blink (i.e., subtle, sparse and aperiodic) and severe interference (i.e., from the human target himself and surrounding objects). We find that even the very subtle involuntary head movement induced by breathing can severely interfere with eye blink detection. In this work, for the first time, we propose a system called BlinkListener to sense the subtle eye blink motion using acoustic signals in a contact-free manner. We first quantitatively model the relationship between signal variation and the subtle movements caused by eye blink and interference. Then, we propose a novel method that exploits the "harmful" interference to maximize the subtle signal variation induced by eye blinks. We implement BlinkListener on both a research-purpose platform (Bela) and a commodity smartphone (iPhone 5c). Experiment results show that BlinkListener can achieve robust performance with a median detection accuracy of 95%. Our system can achieve high accuracies when the smartphone is held in hand, the target wears glasses/sunglasses and in the presence of strong interference with people moving around.

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

Additional Key Words and Phrases: eye blink detection, acoustic signals, contact-free sensing, HCI

---

*Lei Wang is the corresponding author (email: lei.wang@ieee.org).

Authors' addresses: Jialin Liu, School of Software, Dalian University of Technology, Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province, Dalian, China, University of Massachusetts Amherst, Massachusetts, USA, jialinliu@umass.edu; Dong Li, University of Massachusetts Amherst, Massachusetts, USA, dli@cs.umass.edu; Lei Wang, School of Software, Dalian University of Technology, Key Laboratory for Ubiquitous Network and Service Software of Liaoning Province, Dalian, China, lei.wang@ieee.org; Jie Xiong, University of Massachusetts Amherst, Massachusetts, USA, jxiong@cs.umass.edu.

---

1 INTRODUCTION

Eye blink detection has long been a hot research topic, attracting interest from both academia and industry. Eye blink detection plays a critical role in many real-life applications such as Human-Computer-Interaction for people with Motor Neuron Disease (MND) [34, 49, 52, 57], drowsy driving prevention [6, 20, 33, 60] and eye disease detection [19, 21, 48]. MND is a progressive disease that causes muscle weakness and stiffness throughout the body. People with MND first experience difficulties moving their limbs and eventually lose their ability to control the mouth muscle to speak. Each year, more than 500,000 people are diagnosed with MND, and eye-based communication through blinking is one of the very few ways these patients can use to communicate with others and interact with devices. Another key application of eye blink detection is drowsy driving prevention. From just 2011 to 2015, over 396,000 traffic accidents are related to drowsy driving in the United States [5]. Since the eye blink pattern is distinctive when drivers become drowsy, eye blink detection is extensively leveraged by in-car anti-drowsiness systems (e.g., Volvo Driver Alert Control [25]) to prevent traffic accidents and save people's lives. Moreover, by continuously monitoring blink intervals, eye blink detection can help people detect eye diseases such as dry eyes and Computer Vision Syndrome (CVS) which are caused by increased visual concentration.

Existing eye blink detection systems mainly rely on different types of sensors, which can be divided into three categories: EOG sensor [53, 64], proximity sensor [21, 31] and camera [27, 61]. EOG and proximity sensors are usually embedded in wearable devices such as virtual reality headsets and eyewears [21, 53, 64]. Although these systems can achieve high accuracy, wearable-based solutions are intrusive and may cause discomfort in the long term. On the other hand, cameras are usually placed remotely to capture images or record videos to detect eye blinks in a contact-free manner [7, 27]. Though promising, the performance of camera-based systems degrades in low lighting conditions and may raise privacy concerns.

In this work, for the first time, we propose a system called BlinkListener to detect eye blinks using acoustic signals. Acoustic sensing is one key component of wireless sensing which gains a tremendous amount of attention in recent years. Different from traditional sensing which relies on sensors, wireless sensing does not require any sensors and the target-reflection wireless signal is utilized to sense the target's information. The key rationale of wireless sensing is that the target movements affect the signal propagation in the air. Thus, by carefully analyzing the signal variations, target information such as movement speed and displacement can be obtained. A large range of wireless sensing applications has been enabled ranging from coarse-grained activity tracking [3, 55] to fine-grained breathing monitoring [69, 80]. Also different wireless signals are exploited including WiFi [72, 75], RFID [18, 74], ultrasound [26, 58], 60 GHz [41, 70] and even visible light [39, 40].

Among these wireless technologies, acoustic signal exhibits unique advantages in terms of sensing granularity due to the inherent low propagation speed (340 m/s) in the air. Researchers have successfully pushed the sensing granularity to millimeter level, enabling applications that require fine-granularity such as finger tracking [50, 73, 79] and breathing monitoring [69, 71, 76]. One interesting trend in favor of acoustic sensing is that more and more commodity devices in our everyday lives have speakers and microphones built in. Besides laptops and smartphones, most TVs nowadays are capable of voice control, and it is reported that around 20% of consumers interact with their TVs or streaming devices via voice commands [46]. Smart speakers (e.g., Amazon Echo) are also becoming increasingly popular at home.

Motivated by this trend, we propose to employ inaudible acoustic signals for eye blink detection in this work. We emit inaudible acoustic signals through a commodity smartphone and analyze the reflected signals to detect eye blink. Although promising progress has been achieved in acoustic sensing, eye blink detection is still a very challenging task due to the following reasons.

- *Extremely subtle motion.* Eye blink is a very subtle motion. Compared with breathing sensing which is already a non-trivial task with a chest displacement of about 5 mm, the eyelid has a thickness of around 0.5 mm [12]. Therefore, during the eye blinking process, the displacement is just 10% of that caused by

breathing. This subtle displacement-caused signal variation is very small and can be easily buried in noise without being detected. Furthermore, the sensing performance is also closely related to the size of the reflection area. Compared with the chest, the reflection area of the eyes is much smaller (i.e., 1%), leading to much weaker reflection signals and accordingly a poorer sensing performance.

- *Severe interference.* The second challenge is the strong interference from surrounding environments and human target himself. When we apply acoustic signals for eye blink detection, besides the reflection signals bounced off the eyes, surrounding objects also reflect back signals which can be stronger than the eye reflection. These signals get mixed together at the microphone. Separating the eye-reflection signals from other interfering reflections is nontrivial. Note that when a smartphone is used for sensing, the direct path signal from the speaker to the microphone is also very strong, which can interfere with the reflection-based sensing. We further observe another interesting challenge that even after we address the interference and obtain the clean reflection signals from eyes, it is still difficult to detect eye blink. This is because the signals reflected from the eyes contain information not only related to eye blink but also related to breathing and heartbeat-caused movements. Even if the target keeps completely stationary, the small involuntary head movements caused by breathing and heartbeat always exist [10, 62] and are strong enough to interfere with eye blink detection. We term this embedded interference and this interference makes eye blink detection even more challenging.

- *Sparsity and aperiodicity.* The sparsity and aperiodicity of eye blink motion require an effective real-time algorithm to detect individual eye blink. Different from breathing and heartbeat which induce body movement all the time, eye blink only happens for 1.5 to 6% of the time.[1] Therefore, eye blink is a sparse motion in the time domain. Further, the blink interval varies significantly from several seconds to tens of seconds, which is determined by the cognitive processes of our brain [11]. While heartbeat and breathing are periodic in a stable state (e.g., sleep), even the time domain information is slightly corrupted, we can still convert signals into the frequency domain to obtain the breathing/heartbeat rate. The unique nature of sparsity and aperiodicity makes frequency domain analysis infeasible for eye blink detection.

To address the first challenge to sense the very subtle movement, we propose a novel viewing position scheme based on an important observation: the sensing performance varies dramatically when we view the movement-induced signal variation from different positions in the I-Q vector space.[2] We show in this paper for the first time that viewing position matters in wireless sensing. By choosing the optimal viewing position, the sensing performance can be greatly boosted and a subtle movement that previously cannot be detected can now be accurately sensed.

To address the interference issue, we propose a two-stage scheme. We first employ the chirp signal design to differentiate the signals reflected by the eyes from the signals reflected by other surrounding objects. Owing to the low propagation speed in the air, the resolution of acoustic signal with a bandwidth of 4 kHz is able to distinguish two signals with a distance difference larger than 4.25 cm. Therefore, signals reflected by the eyes can be clearly separated from signals reflected by surrounding objects. After we obtain the signals reflected from the eyes, we still need to deal with the embedded interference caused by breathing and heartbeat. The signals reflected by the eyes contain not only the eye blink information but also breathing and heartbeat information. Prior studies consider this interference harmful and always try to remove them [4, 56]. In this paper, we show that there is no need to remove them owing to the unique characteristic of eye blink. With the proposed novel scheme, we can successfully turn the embedded interference widely considered harmful to be beneficial. The key rationale is that breathing and heartbeat-induced signal variations are purely caused by path length change. However, the

---

[1]One eye blink takes 100 to 400 ms [65]. Ten eye blinks in one minute corresponds to 1.5 to 6% of the time.

[2]I-Q vector is a 2D representation of the signal after down-conversion. The signal is represented as a vector with both amplitude and phase in the complex plane.

eye blink-induced signal variation is caused not only by the path length change (with vs. without eyelid) but also by a material change of the reflection surface (eyelid vs. eyeball). The reflection coefficients are very different for these two surfaces, leading to a dramatic signal strength variation. Therefore, the eye blink-caused signal variation exhibits a unique property: the signal phase change is small but the signal amplitude change is large. This is very different from breathing and heartbeat which cause a large signal variation in phase but a small signal variation in amplitude. By deeply understanding the underlying principles and modeling the eye blink process in the I-Q vector space, we can detect eye blink without removing the embedded interference. Instead, we employ these interfering variations to help identify the optimal viewing position for eye blink detection.

To address the sparsity and aperiodicity issues, we do not convert the signal to the frequency domain but process the signal variation in the time domain. Since the optimal viewing position shifts during the long-term detection process caused by slight body movements of the human target, we propose a dynamic update strategy to adaptively update the optimal viewing position for eye blink detection, which achieves a high detection accuracy in real-time. To summarize, our contributions are as follows.

(1) We achieve highly accurate eye blink detection in a contact-free manner using a pair of speaker and microphone widely available in commodity devices. Through both theoretical and experimental analysis, we quantitatively model the relationship between signal variations and the subtle movements caused by eye blink and interference. To our best knowledge, BlinkListener is the first system applying acoustic signals for eye blink detection.

(2) For the first time, we demonstrate that the viewing position in the I-Q space affects the sensing performance. We utilize the "harmful" interference to help determine the optimal viewing position, achieving accurate eye blink detection which is previously infeasible due to the extremely subtle signal variation and strong interference. We believe the proposed viewing position scheme can be applied to improve the performance of other sensing applications which involve extremely small signal variations.

(3) We implement BlinkListener on both a research-purpose hardware platform (Bela [66]) and a commodity smartphone (iPhone 5c). Experiment results show that BlinkListener can achieve robust performance with a median detection accuracy of 95%. Also, the proposed system can still achieve high accuracy when the smartphone is held in hand, the target wears glasses/sunglasses, and in low lighting conditions. Even in the presence of strong interference such as people walking around, the proposed system still works well as long as the interfering person is more than 1 m away from the sensing device.

The rest of this paper is organized as follows. Sec. 2 introduces the state-of-the-art literature on eye blink detection and acoustic-based sensing. Sec. 3 presents the principle of applying acoustic signals for eye blink sensing. Sec. 4 introduces different types of interference. Sec. 5 and Sec. 6 present how to handle the interference. Sec. 7 presents the implementation of BlinkListener. Sec. 8 presents the system evaluation. Sec. 9 discusses the limitation of our work followed by a conclusion in Sec. 10.

2 RELATED WORK

In this section, we discuss the literature related to eye blink detection and acoustic sensing.

2.1 Eye Blink Detection

There is a large body of literature related to eye blink detection. We classify them into three categories based on the types of sensors used.

2.1.1 EOG-based. Electrooculography (EOG) sensors have been used in eye movement tracking and eye blink detection [14, 32, 36, 53, 64]. To obtain the electrooculogram signal for eye blink detection, EOG-based systems require users to attach electrodes surrounding their eyes with cables connected to a computing unit, which are intrusive and hinder user mobility [14, 24, 59]. The latest EOG-based solutions integrate EOG sensors into

glasses [13, 36, 63]. There are multiple commercial EOG glasses available on the market, such as JINS MEME glasses [1] and imec eye-tracking glasses [2]. Although these glasses-based solutions are promising, they are still expensive (e.g., the JINS MEME glasses cost $420). In contrast, our solution is hosted on a smartphone without extra cost.

2.1.2 Proximity Sensor-based. Proximity sensors, often referred to as infrared cameras or infrared-reflective sensors, are embedded in head-mounted devices or eyewears for eye blink detection and eye tracking [16, 21, 31]. Proximity sensors have a lot of merits such as low-power, portable, and immune to strong noise. However, it is reported that proximity sensors might cause damage to the eyes due to the emitted heat if not properly configured [23]. Also, eyewears are still intrusive and may cause discomfort in long-term use. Compared with proximity sensor-based solutions, our method detects eye blink in a contact-free and non-intrusive manner, which does not require users to wear a device. Also to make sure it is safe to be used in the long term, our system employs a sound volume far below the allowed exposure limit [15, 29].

2.1.3 Camera-based. There is a vast body of literature on using cameras for eye blink detection [7, 20, 22, 27, 37, 45, 47, 61, 78]. Even though camera-based methods can achieve good performance, the image processing part usually requires complex algorithms for feature extraction, segmentation, and classification, which is computation-intensive and power-hungry [35, 43]. The main drawback of camera-based methods is strict lighting condition requirements [17, 51, 82], which means they cannot work well in the dark. Also, using cameras may raise privacy concerns. Our system provides a complementary solution to camera-based approaches since acoustic sensing is not restricted to lighting conditions, and can still detect eye blink when users wear sunglasses.

2.2 Acoustic Sensing

Acoustic signals have been widely employed for fine-grained motion tracking [26, 38, 69] and localization [44, 68] owing to their low propagation speed in the air. Recent research efforts have pushed the sensing granularity of tracking and sensing to millimeter level. For example, fingerIO [50] successfully tracks fingers using acoustic signals, while LLAP [73] and Strata [79] exploit phase measurement of acoustic signals to achieve a millimeter-level tracking accuracy. Wang *et al.* [71] propose a correlation-based method to enable millimeter-level breathing detection. FM-Track [38] leverages information from multiple dimensions to further improve the sensing granularity for multiple close-by targets. However, millimeter-level sensing granularity is still not enough for eye blink detection since the displacement caused by eye blink is even smaller. The ACG system [56] proposed by Qian *et al.* applies acoustic signals to detect breathing as well as heartbeat rate. However, the frequency domain analysis utilized by ACG does not work for eye blink detection since eye blink is sparse and aperiodic.

3 UNDERSTANDING ACOUSTIC SIGNALS FOR EYE BLINK DETECTION

In this section, we introduce the chirp signal design for BlinkListener followed by the eye blink sensing model.

3.1 Chirp Signal Design

In BlinkListener, we employ chirp signals for eye blink detection. Due to the low propagation speed of acoustic signals in the air, a frequency band of 4 *kHz* can achieve a spatial resolution of 4.25 *cm*. This means if the distance difference of two targets to the device is more than 4.25 *cm*, the reflected signals from these two targets will fall in different frequency bins and can be distinguished.

We analyze the signal propagated from a speaker (transmitter) to the target and then reflected back to a microphone (receiver). The speaker and microphone are co-located. Here, we use the sensing device to refer to the co-located speaker and microphone. The transmitted chirp signals generated by a speaker are shown in Fig. 1a. The frequency f of the signal changes linearly with time as $f = f_0 + \frac{B}{T}t$, where f_0, B, and T are the initial
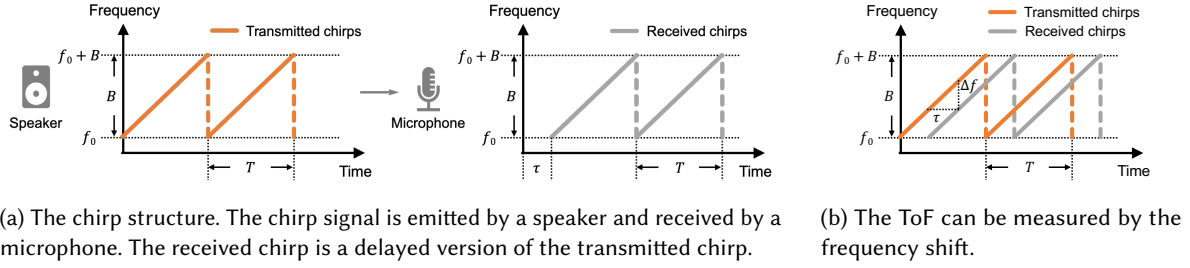
(a) The chirp structure. The chirp signal is emitted by a speaker and received by a microphone. The received chirp is a delayed version of the transmitted chirp.

(b) The ToF can be measured by the frequency shift.

Fig. 1. The chirp signal design and principle.

(a) Signals travel through multiple paths.

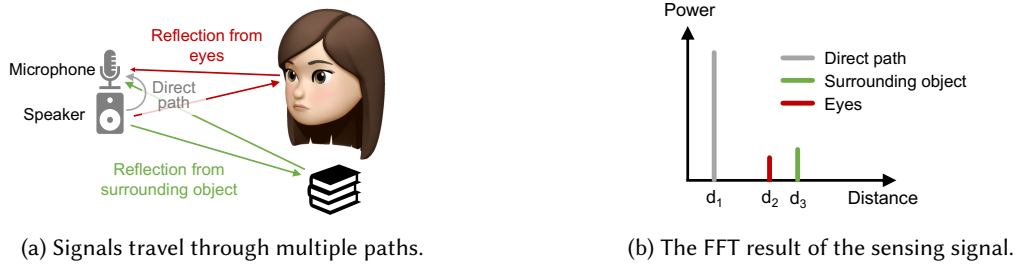(b) The FFT result of the sensing signal.

Fig. 2. The illustration of eye blink detection in a multipath environment, and the corresponding ToFs (frequency shifts) of the multiple paths, respectively.

frequency, bandwidth, and duration of the chirp, respectively. The transmitted chirp can be represented as

$$x(t) = e^{-j2\pi\left(f_0 t + \frac{B}{2T} t^2\right)}. \tag{1}$$

The transmitted chirps arrive at the receiver through multiple paths: direct path, reflection path from eyes, and reflection path from a surrounding object as shown in Fig. 2a. Let us denote the number of paths as N. The received signal is a superposition of all the N paths, and each path is a delayed version of the transmitted signal. The overall received signal can be represented as

$$y(t) = \sum_{i=1}^{N} \alpha_i e^{-j2\pi\left(f_0(t-\tau_i) + \frac{B}{2T}(t-\tau_i)^2\right)}, \tag{2}$$

where α_i and τ_i are signal attenuation and time-of-flight (ToF) of the i-th path signal in the air, respectively. By multiplying the conjugation of transmitted signal, we can obtain the mixed signal below:

$$m(t) = \sum_{i=1}^{N} \alpha_i e^{-j2\pi\left(\frac{B}{T}\tau_i t + f_0\tau_i - \frac{B}{2T}\tau_i^2\right)}. \tag{3}$$

This mixed signal is employed for sensing in BlinkListener. We can see from Fig. 1b that the received chirp is a delayed (shifted) version of the transmitted chirp. The frequency shift Δf_i is proportional to the ToF τ_i of the signal. A larger frequency shift corresponds to a longer ToF. By measuring the frequency shift Δf_i at the receiver, we can obtain ToF as $\tau_i = \frac{\Delta f_i T}{B}$. The distance (range) between the object and sensing device can thus be calculated as $d_i = \frac{c\tau_i}{2} = \frac{c\Delta f_i T}{2B}$, where c is the speed of sound in the air, and the factor 2 accounts for the fact that the reflected signal traverses the path back and forth. The range resolution δd is determined by the frequency bandwidth B of
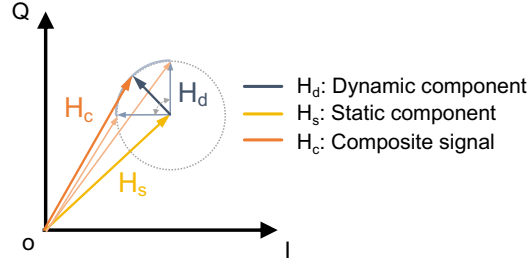
Fig. 3. The composite signal in I-Q vector space.

the chirp signal and can be calculated as $\frac{c}{2B}$. That means if two signals have a very small distance difference ($< \frac{c}{2B}$) with respect to the sensing device, the two signals will fall in the same frequency bin and therefore cannot be differentiated. If the distance difference is larger than $\frac{c}{2B}$, two signals can be distinguished. Without loss of generality, we consider a bandwidth of $B = 4\,kHz$, which yields a range resolution of $\delta d = \frac{c}{2B} = \frac{340m/s}{2 \times 4000} = 4.25\,cm$.

After performing Fast Fourier transform (FFT) over the mixed signal, we can depict different objects located at different distances with respect to the sensing device as shown in Fig. 2b. We can see three different colors corresponding to three different signal paths. The red one denotes the signal reflected from the eyes, and the gray and green colors denote the direct path signal and the reflection from another surrounding object, respectively. The strength of eye-reflection signals is not only weaker than that of the direct path but also weaker than those of the reflection from other objects due to a very small area of the reflective surface. In Sec. 5, we will illustrate how to identify the eye-reflection signals in the presence of other strong reflection signals. We further group signal paths into static paths and dynamic paths. All the static path signals, including the direct path signal and reflections from static surrounding objects, can be represented as a single combined static vector

$$H_s = \sum_{i \in P_s} \alpha_i e^{-j2\pi \left(\frac{B}{T} \tau_i t + f_0 \tau_i - \frac{B}{2T} \tau_i^2 \right)}, \tag{4}$$

where P_s is the set of static paths. The signal reflected from the eyes of the human target can be represented as a dynamic vector

$$H_d = \alpha e^{-j2\pi \left(\frac{B}{T} \tau t + f_0 \tau - \frac{B}{2T} \tau^2 \right)}. \tag{5}$$

We denote the superimposed composite signal as H_c, where H_c is a vector summation of H_s and H_d. We illustrate the static signal vector, dynamic signal vector and composite signal vector in I-Q vector space in Fig. 3. For most small-scale movements, the amplitude of the dynamic vector can be assumed as a constant [81] and only the phase changes. Therefore, the dynamic vector rotates with respect to the static vector, inducing the signal variations in the I-Q vector space.

3.2 Modeling the Eye Blink Process

Now we focus on the dynamic component H_d to deeply understand the underlying principle and quantify the relationship between eye blink motion and the signal variation in the I-Q vector space.

Eye blinking is a process of rapidly closing and then opening of the eyelid within a short period of time (100 to 400 ms). When acoustic signals are applied for eye blink detection in a contact-free manner, the rationale is to capture the subtle variation of the signals reflected from the eyes. According to our analysis, there are two main factors affecting the signal variation during the process of eye blinking: (F1) when the eye is closed, the eyelid causes a subtle path length change; (F2) the reflective surface switches between the eyeball (water texture)

(a) The signal variation between the eye open and closed status.

(b) The resultant phase and amplitude changes.

Fig. 4. The signal model of one eye blink cycle in the I-Q vector space.

and the eyelid (skin texture). The first factor F1 mainly causes a phase shift and the signal amplitude can be considered as a constant.[3] From Eq. 5, we can see the phase of the signal is $\varphi = -2\pi\left(f_0\tau - \frac{B}{2T}\tau^2\right)$. In a typical setting, $f_0 >> \frac{B}{2T}\tau$, so we can omit the quadratic term. Thus, by replacing τ with $\frac{2d}{c}$, the path length change Δd induces a phase change $\Delta\varphi$ which can be represented as

$$\Delta\varphi = -\frac{4\pi f_0 \Delta d}{c}.$$

(6)

The reflective surface change (F2) causes a change on the signal amplitude, that is, the attenuation term α, owing to dramatically different reflection coefficients. We denote φ_{open}, α_{open} as the phase and amplitude of the signal when the eyes are open. Similarly, φ_{closed} and α_{closed} are the phase and amplitude of the signal when the eyes are closed. We then illustrate the eye blink process in Fig. 4. The red and blue dots represent the eye closed status and the eye open status respectively. We divide one complete cycle of eye blink into two stages and discuss the effect on signal variation below.

3.2.1 *Eye Closing Stage.* In the eye closing stage, the status of the eyes changes from open to closed. (i) Phase change: when the eyeball is covered by the eyelid, there is a subtle reflection path length change. The amount of path length change depends on the thickness of the eyelid. According to Eq. 6, a small phase change is induced. In I-Q vector space, the phase change causes the dynamic vector to rotate with respect to the static vector. Without loss of generality, if we consider a zero static vector, the dynamic vector rotates with respect to the origin as shown in Fig. 4a. (ii) Amplitude change: when the eyeball is covered by the eyelid, the reflective surface switches from the water-textured eyeball to the skin-textured eyelid. The amount of sound signal absorption is much more for the eyelid and therefore causing larger signal attenuation. Thus, an amplitude decrease will occur, causing the length of the dynamic vector to be shorter as shown in Fig. 4a.

3.2.2 *Eye Opening Stage.* The second stage is the eye opening stage, in which the status of the eyes changes from closed to open. In this process, the phase and amplitude changes are exactly the opposite of the eye closing stage. The phase change and amplitude change happen at the same time and the resultant signal variation is shown in Fig. 4b. The switch between eye open and eye closed causes the signal samples to move in the I-Q vector space. Note that the signal phase change and amplitude change are in orthogonal directions in the I-Q vector space. The phase change causes a signal variation along the tangential direction of the circle while the

---

[3]For example, if the distance between the eyes and the sensing device is 50 cm, the eyelid-induced distance change is only 0.05 cm, which causes a negligible amplitude change.
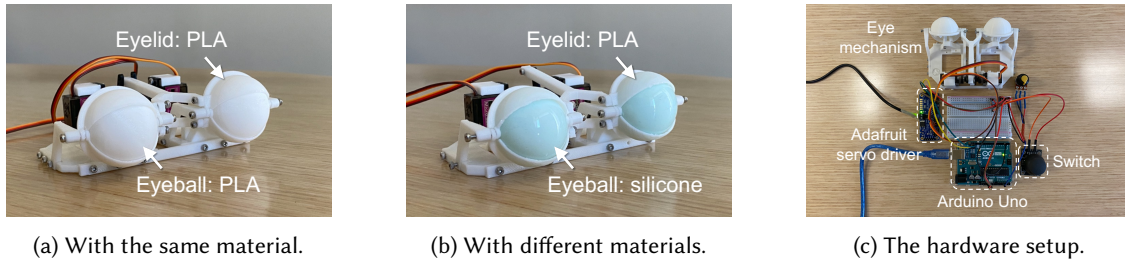
(a) With the same material.

(b) With different materials.

(c) The hardware setup.

Fig. 5. The 3D-printed eye mechanism structure and hardware components.

(a) Two traces of signal samples reflected by two eye mechanisms in the I-Q space, respectively.

(b) Both of the two eye mechanisms will cause phase change in I-Q space, while only the eye mechanism with different materials will also cause amplitude change.

(c) Phase variation results.

(d) Amplitude variation results.
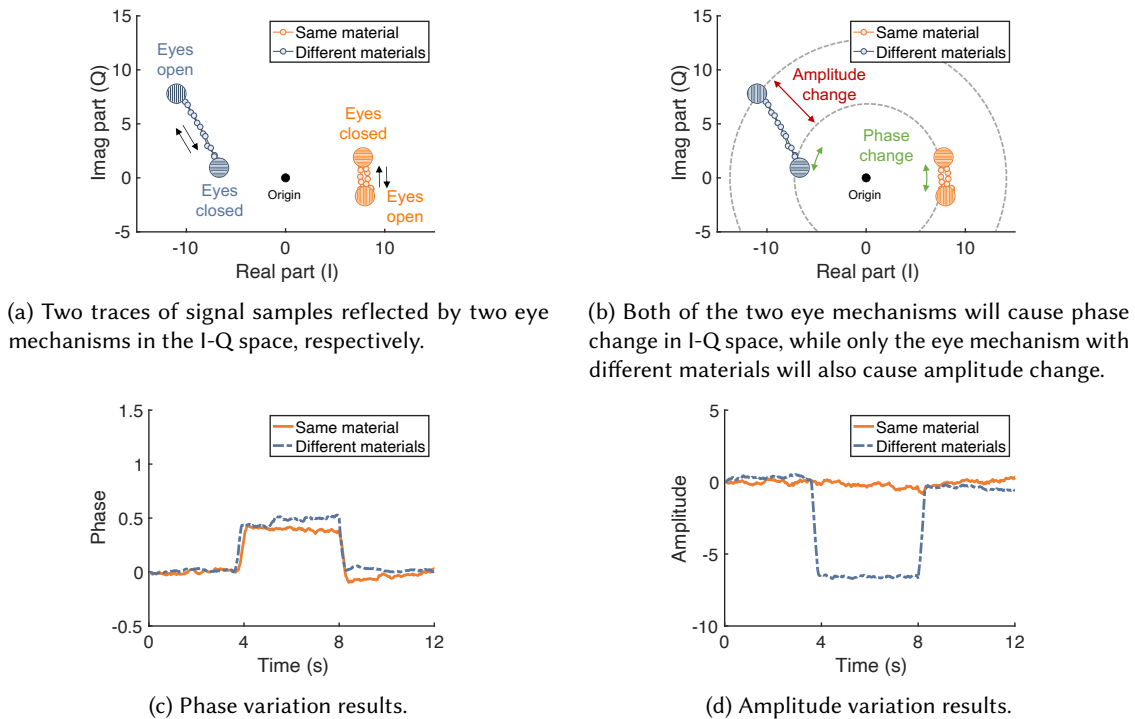
Fig. 6. Experiment results for two eye mechanisms.

amplitude change causes a signal variation along the radial direction as shown in Fig. 4b. This is an important property that we will utilize for eye blink detection in Sec. 5.

3.3 Eye Blink Model Validation

We validate our eye blink model using two 3D-printed eye mechanisms as shown in Fig. 5. Note that we do not conduct the real eye experiment to validate our eye blink model here because the signals reflected by real eyes not only contain the eye blink-induced signal variation but also the heartbeat and breathing-induced variation. Therefore, the real eye experiment is not suitable to clearly demonstrate the correctness of the proposed eye blink
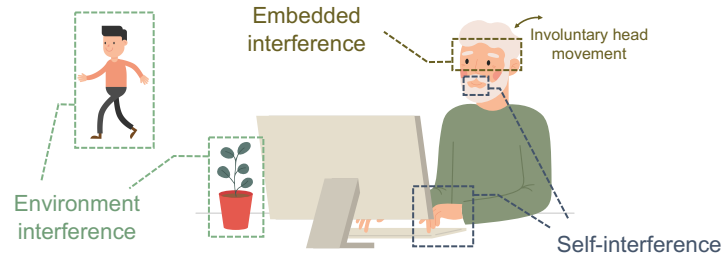
Fig. 7. An Illustrative example of three types of interference.

model. For the first eye mechanism shown in Fig. 5a, the eyelid and the eyeball are both made of PLA material. For the second eye mechanism shown in Fig. 5b, the eyelid is made of PLA and the eyeball is made of silicone. Both of the two eye mechanisms' eyelids have a thickness of 1 *mm*. The eye mechanism is placed at a distance of 0.5 *m* from the sensing device. We operate the eye mechanism by an Arduino Uno and an Adafruit PCA9685 servo driver, and trigger eye blinks by a push-to-make switch as shown in Fig. 5c. We keep the mechanical eyes open for 4 seconds and then closed for 4 seconds. Fig. 6a shows the signal samples during the eye blink cycle. We can see that the distance between eye open and eye close dots for different materials is larger than that for the same material. Fig. 6b further shows that the signal amplitudes are roughly the same for the same material but quite different for different materials. We also plot the corresponding phase and amplitude variations in Fig. 6c and Fig. 6d. We can see that phase changes are significant and are almost the same for both the same material and different materials in Fig. 6c. On the other hand, the amplitude change is only significant for the case of different materials in Fig. 6d. These results match our theoretical analysis very well.

4 UNDERSTANDING THE INTERFERENCE IN EYE BLINK DETECTION

For a subtle movement such as eye blink, even a very small movement for the interference can severely degrade the sensing performance. Therefore, the issue of interference we need to address is much more difficult than that in other wireless sensing applications such as breathing monitoring and gesture recognition.

4.1 Different Types of Interference

As shown in Fig. 7, the interference for eye blink detection can be grouped into three categories, which are described in detail below.

4.1.1 Environment Interference. The sensing signals will travel through different paths from the transmitter to the receiver. The signals reflected from surrounding objects are termed environment interference. The environment interference can come from static objects (e.g., furniture and walls) as well as moving objects (e.g., people moving around).

4.1.2 Self-interference. Besides the signal reflected from the eyes, there also exist some signals reflected from other body parts. The interference caused by signals reflected from other body parts is termed self-interference. Self-interference can come from the hands when the human target is typing a keyboard, or the mouth when the target is speaking.

4.1.3 Embedded Interference. Separating the eye-reflection signal from the environment and self-interference is still not the end. Even if we can obtain the clear eye-reflection signal, it contains multifarious information from not only the eye blink but also the breathing and heartbeat. The reason is that the human head moves

(a) The signal variation for the head movement.

(b) The arc trajectory induced by the head movement.

Fig. 8. The signal model of embedded interference for one head movement cycle in the I-Q vector space.

involuntarily caused by breathing and heartbeat. Specifically, during the breathing process, the head slightly moves with the torso caused by the chest movement. Moreover, there exists an approximate 1 mm head movement synchronous with the heartbeat due to the pumping of blood, which is called ballistocardiography (BCG) [10, 62]. After we completely remove the environment interference and self-interference, the involuntary head movement still exists and is embedded in the signal reflected from the eyes. In this paper, we term it embedded interference, which is the most challenging interference caused by breathing and heartbeat. Note that for a lot of applications such as gesture recognition, this embedded interference can be neglected because it is too small to affect the sensing performance. However, for the detection of subtle eye blinks, this embedded interference cannot be neglected and needs to be addressed before we can detect eye blinks.

4.2 Solutions for Addressing Different Types of Interference

4.2.1 Addressing Environment Interference and Self-interference Using Chirp Design. As mentioned in Sec. 3.1, we employ chirp signal design as a spatial filter to separate reflection signals arriving from different distances into different frequency bins. Owing to the low propagation speed in the air, the resolution of acoustic signal with a bandwidth of 4 kHz is able to distinguish two signals with a distance difference larger than 4.25 cm. Therefore, the signal reflected from the eyes can be clearly separated from signals reflected from surrounding objects which usually have a distance tens of centimeters larger. Note that we do observe a case that cannot be addressed with this chirp design. Specifically, when the lips are moving (e.g., speaking), the reflection from the lips has a propagation distance very similar to the reflection from the eyes and the 4.25 cm resolution is not fine enough to separate them. The good news is that the lip movement causes much larger signal variations which are very different from eye blink and can be easily detected. We can thus remove those signal samples corrupted by the lip movement.

4.2.2 Modeling the Embedded Interference in the I-Q Vector Space. Since breathing and heartbeat are both periodic movements, the induced involuntary head movement is also periodic. As shown in Fig. 8a, to model one cycle of the head movement, we denote φ_{front} and α_{front} as the signal phase and amplitude at the start position of the head, respectively, which are marked in red. Similarly, we denote φ_{back} and α_{back} as the signal phase and amplitude at the end position of the head, respectively, which are marked in blue. The key observation is that, when there is no eye blink, the signal induced by the involuntary head movement forms an arc trajectory in the I-Q vector space as shown in Fig. 8b. Specifically, the signal amplitude change caused by the small head displacement (around 1 mm) is negligible, while the corresponding phase change is significant. Therefore, if we do not consider the static vector, the phase change causes the dynamic vector to rotate with respect to the
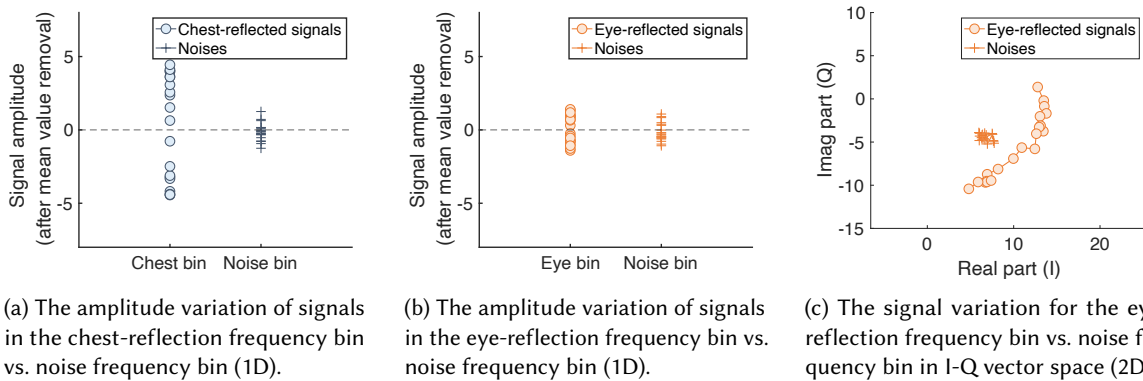
(a) The amplitude variation of signals in the chest-reflection frequency bin vs. noise frequency bin (1D).

(b) The amplitude variation of signals in the eye-reflection frequency bin vs. noise frequency bin (1D).

(c) The signal variation for the eye-reflection frequency bin vs. noise frequency bin in I-Q vector space (2D).

Fig. 9. The comparison between the signal variations reflected from the chest and eyes in amplitude (1D) and I-Q space (2D).

origin, and the radius of the arc is approximately a constant due to the negligible amplitude change. With a deep understanding of the effect of the embedded interference, we further incorporate the embedded interference model into the eye blink signal model and introduce the viewing position scheme in Sec. 6, in which we utilize this embedded interference to help us maximize the eye blink signal variation. Before extracting the subtle eye blink signal variation from the embedded interference, we first introduce our detailed method in addressing the environment interference and self-interference in the next section.

5 IDENTIFYING EYE-REFLECTION FREQUENCY BIN

There are plenty of uninterested interference signals along with the target signal bounced from the eyes. In this section, we introduce how to separate the eye reflection from other surrounding reflections (i.e., environment interference and self-interference) by identifying the frequency bin corresponding to the eyes.

5.1 Problem Description

When we apply the chirp signal design for eye blink detection in a multipath-rich environment, reflectors at different distances induce different amounts of frequency shifts, and the signals fall into different frequency bins. After performing FFT over the sensing signal (mixed signal) in Eq. 3, we can obtain multiple peaks located in different frequency bins corresponding to different reflectors as shown in Fig. 2. Without previous knowledge about the distance between the eyes and the sensing device, we have no idea which peak corresponds to the eye-reflection signals when there exist multiple peaks. The naïve way is to distinguish the eye-reflection signal from other reflections by the power of peaks. However, due to the small reflection area, the power of eye reflection can be weaker than those of other surrounding objects such as walls and furniture even if the eyes are closer to the sensing device.

There are two relevant methods proposed by the prior breathing detection studies to identify the frequency bin corresponding to the chest. One is to find the frequency bin with the highest periodicity [71], and the other is to find the frequency bin with the largest signal amplitude variation [56]. Both methods are effective for breathing detection. On one hand, breathing is a periodic process. On the other hand, the signal variation caused by the chest displacement (e.g., 5 *mm*) is much larger than the noise-caused variation. However, neither the periodicity method nor the signal amplitude variation method works for eye blink detection. Specifically, the nature of sparsity and aperiodicity for eye blink fails the periodicity method, and the small signal variation caused by the eyes fails the signal amplitude variation method. To further illustrate it, we conduct two benchmark experiments.

In the first experiment, we place the sensing device towards the target's chest to collect the chest-reflection signals. In the second experiment, we place the sensing device towards the target's eyes to collect the eye-reflection signals. The signal amplitudes extracted from the two experiments are shown in Fig. 9a and Fig. 9b, respectively. We can observe that the chest-induced signal amplitude variation is large enough to be utilized to identify the chest-reflection frequency bin from the noise bin (i.e., the frequency bin without the target). In contrast, the eye-induced signal variation is much smaller, which is just slightly larger than the noise-caused variation. Therefore, the method proposed for breathing detection does not work well for eye blink detection.

5.2 Eye-reflection Frequency Bin Identification Method in BlinkListener

To quickly identify the frequency bin corresponding to the eyes, we exploit the embedded interference caused by breathing and heartbeat instead of directly depending on the signal variation caused by eye blink. The reason is that, eye blink is a sparse activity and the blink interval varies from several seconds to tens of seconds, which may introduce a large latency if we identify the eye-reflection frequency bin by eye blink. In contrast, the embedded interference exists all the time even if there is no eye blink, which can be leveraged to quickly identify the eye-reflection frequency bin.

Fig. 9c illustrates the signal variation for the eye-reflection frequency bin when there is no eye blink. We have an interesting observation: although the 1D amplitude variation of the eye-reflection signal is small, the signal variation in 2D I-Q vector space is quite large, which forms an arc trajectory due to the embedded interference. As shown in Fig. 9c, the eye-reflection frequency bin can be easily distinguished from the noise bin in the 2D space. Therefore, to identify the frequency bin corresponding to the eyes, we first compute the variance of the 2D signal variations for each frequency bin and then pick out the one with maximum variance. Note that this is the first time we utilize the "harmful" embedded interference to benefit our signal processing for eye blink sensing.

6 EXTRACTING EYE BLINK MOTION

Even after successfully extracting the eye-reflection signal from the identified frequency bin, it is still challenging to obtain the eye blink information. In this section, we first describe the problem in detecting the eye blink from the eye-reflection signal. Furthermore, inspired by our observation that the performance of eye blink detection highly depends on the position from which the signal is viewed in the I-Q space, we propose a novel viewing position scheme that fully utilizes the "harmful" embedded interference to help us maximize the eye blink-induced signal variation. At last, we design the real-time eye blink detection algorithm based on our proposed viewing position scheme.

6.1 Problem Description

According to the eye blink model presented in Sec. 3.2, eye blink can induce both the amplitude and phase change in the 2D I-Q vector space. It is an intuitive way to detect the eye blink using both the phase and amplitude information. Although promising, it still presents some difficulties.

- Embedded interference. The signals reflected from the eyes contain information which is related to not only eye blink but also the embedded interference as introduced in Sec. 4.1. Specifically, even if the target keeps completely stationary, the small involuntary head movements caused by breathing and heartbeat still exist and are strong enough to interfere with eye blink detection.
- Strong residual power from the direct path. Due to the close distance between the speaker and microphone on the commodity smartphones, the power of the direct path from the speaker to the microphone is so high that the extracted eye-reflection signal contains strong residual power from the direct path [56].

To demonstrate the above-mentioned two difficulties, we illustrate the resultant eye-reflection signal variation in Fig. 10a. On one hand, the signal variation induced by eye blink is superimposed on that of embedded
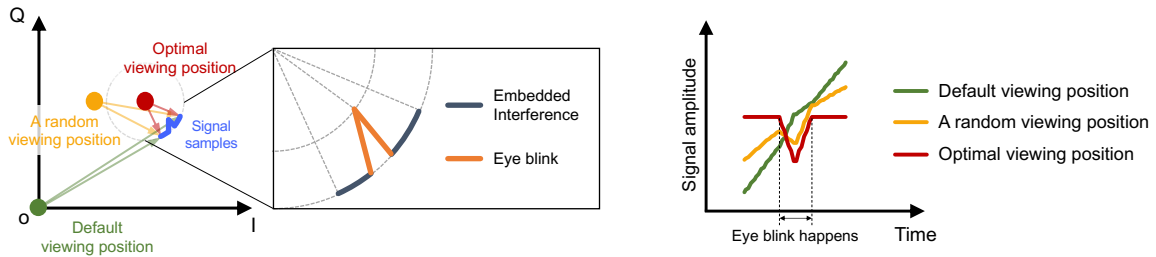
(a) Three different viewing positions and a zoom-in view of the eye-reflected signal containing both eye blink and embedded interference.

(b) The signal amplitude variation from three different viewing positions.

Fig. 10. An illustrative example demonstrating that the performance of eye blink detection varies significantly when we view the signal variations from different positions in the I-Q vector space.

interference, making it difficult to detect eye blink based on the superimposed eye-reflection signal. On the other hand, the strong residual power from the direct path causes a very large static component, making it hard to detect eye blink using the raw signal amplitude computed from the coordinate origin as shown in the green line in Fig. 10b.

6.2 Viewing Position Scheme

To address the above-mentioned problem, we propose a viewing position scheme in the I-Q vector space to maximize the eye blink-induced signal variation. This is inspired by the key observation that the performance of eye blink detection varies significantly when we view the eye blink-induced signal variation from different positions in the I-Q vector space.

Fig. 10a models the signal variation with both embedded interference and eye blink, whose trajectory is an arc with a "bump" on it. The embedded interference induces the arc, and the "bump" on the arc is caused by the reflection switching between the eyeballs and eyelids. To detect eye blink, we need to compute the signal amplitude variation from the trajectory in the I-Q vector space. The traditional approach is to compute the distance between the origin and each signal sample in the trajectory, which can be regarded as viewing the trajectory from the origin. Unfortunately, this approach cannot be applied to detect eye blink since the signal amplitude variation (i.e., the green line in Fig. 10b) cannot clearly show the motion of eye blink. However, if we randomly select the viewing position in the I-Q vector space (i.e., the yellow dot in Fig. 10a), we can observe a slightly clearer eye blink from the signal amplitude variation, which is computed from the new viewing position to each signal sample (i.e., the yellow line in Fig. 10b). By choosing the optimal viewing position (i.e., the red dot in Fig. 10a), we can maximize the signal amplitude variation induced by eye blink (i.e., the red line in Fig. 10b).

One question naturally arises: how to obtain the optimal viewing position for eye blink detection? We observe that the center of the arc formed by the embedded interference is the optimal viewing position. The center of the arc benefits eye blink detection in two aspects. The first one is that it can maximize the signal amplitude variation caused by eye blink. The second one is that it minimizes the impact from embedded interference by ensuring that the signal amplitude variation caused by embedded interference is roughly a constant. Therefore, to identify the optimal viewing position, we apply the arc fitting method to find the center of the arc trajectory induced by the embedded interference. Note that this is the second time we utilize the "harmful" interference to benefit sensing. Recall that we utilize the embedded interference to identify the eye-reflection frequency bin for the first time.
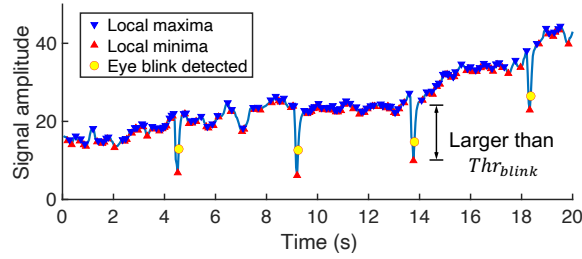
Fig. 11. An illustrative example of our real-time eye blink detection algorithm.

6.3 Real-time Eye Blink Detection Algorithm

Although the above viewing position scheme can achieve good eye blink detection performance, the optimal viewing position changes during the long-term detection process due to the target's slight body movements and surrounding environment changes. Thus, we need to adaptively update the viewing position in real-time to maintain a good sensing performance. Since we obtain the optimal viewing position by the arc fitting method, the more signal samples we have, the more accurate the optimal viewing position is. However, more data samples mean a larger system delay, and the real-time performance can be affected. In the following, we present the design details of our real-time algorithm to maintain a good balance between the detection accuracy and system latency.

Step 1: Initialization. At the beginning, our system collects a number of signal samples for the initialization. Specifically, we accumulate 50 chirps with the default chirp period of 40 ms, which takes 2 s in total. Note that this 2 s is for the cold-start and is a one-time effort. Once the system is initiated, we can output the detection results every 40 ms. Therefore, our proposed system is able to provide real-time detection. We apply the well-known Pratt method [54] for the arc fitting, which is lightweight and robust.

Step 2: Eye blink peak detection. After obtaining the optimal viewing position, BlinkListener continuously traces the relative distance from the viewing position to the newly collected signal sample. Since one eye blink cycle takes $100 - 400$ ms, we apply the Local Extreme Value Detection (LEVD) method to detect the eye blink induced "bumps" in a sliding window. The basic idea of LEVD method is to find the alternative local maxima and minima and compare the difference of two nearby local maxima and minima with a pre-defined threshold Thr_{blink}, which is set as five times of the standard deviation of signal amplitude when there is no eye blink. If the difference of the local maxima and minima is larger than the threshold Thr_{blink}, an eye blink is detected.

Step 3: Viewing position update. Since the optimal view position calculation is very lightweight, we continuously update the viewing position as long as enough samples are accumulated. Note that if too few samples are employed for the arc fitting, the accuracy can be quite low. Therefore, we set Thr_{update} as the minimum number of samples that need to be accumulated before a new viewing position can be calculated. From comprehensive experiments, we set $Thr_{update} = 50$ as the default threshold which corresponds to a period of 2 s.

Step 4: Restart. When a large body movement happens, BlinkListener restarts the whole eye blink detection process and goes back to Step 1. The large body movements are detected by comparing the signal variation with a threshold $Thr_{restart}$, which is set as three times of the eye blink peak detection threshold Thr_{blink}. We illustrate the real-time eye blink detection method using the real data collected in one of our experiments in Fig. 11. We can see that four eye blinks are accurately detected.
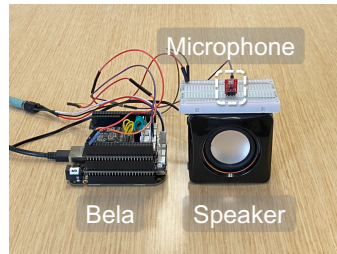
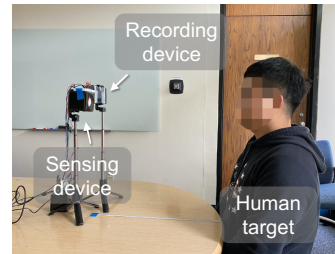Fig. 12. Components of the Bela hardware platform.

Fig. 13. The experiment setup.

7 IMPLEMENTATION

We implement BlinkListener on both a research-purpose hardware platform Bela [66] and a commodity smartphone iPhone 5c [8]. They are used to transmit and receive acoustics signals for eye blink sensing. We employ the Bela platform for benchmark experiments because it is more flexible to vary parameters such as the spacing between the speaker and microphone.

Bela platform: The Bela hardware is connected with a general-purpose speaker to transmit chirp signals and an MEMS microphone to receive chirp signals. The device components are shown in Fig. 12. The speaker and the microphone are bound together. The microphone is placed on the top of the speaker and the distance between them is around 8 cm.

Smartphone: Without loss of generality, we adopt an iPhone 5c to verify the effectiveness of BlinkListener on the smartphone platform. We implement BlinkListener based on the existing acoustic sensing framework LibAS [67]. We choose the built-in speaker and microphone co-located at the earpiece of the phone since they are facing towards the eyes when the human target is using the phone.

Signal parameters: We adopt the frequency band from 18 kHz to 22 kHz which is usually inaudible for human ears with a bandwidth of 4 kHz for sensing. We employ 40 ms as our default chirp duration.

Ground truth: We employ a separate smartphone as the recording device to record the ground truth using the camera. The recording device is held by an adjustable tripod and placed at the same height as the sensing device.

8 EVALUATION

In this section, we comprehensively evaluate the performance of BlinkListener using the Bela platform by varying the parameters under different conditions. We also conduct a series of field studies using a smartphone to demonstrate the feasibility of BlinkListener on the commodity hardware.

8.1 Experiment Setup

We evaluate the performance of BlinkListener with 25 participants (14 male and 11 female). They are diverse in age (from 20 and 56) and race (East-Asian, South-Asian, Caucasian and African). Since the Bela platform is flexible to configure parameters such as the number of microphones and the spacing between the speaker and microphone, we conduct experiments using the Bela platform to evaluate the effectiveness of BlinkListener and vary parameters and conditions to study the impact of different factors. Our experiments are conducted in a 3.3 $m \times 6$ m conference room with furniture and facilities. We collect 3 minutes of acoustic data when the volunteer (target) is sitting in a chair using the Bela platform. Unless specified otherwise, the sensing device (the Bela platform) is placed 0.5 m away from the target at the same height as the target's eyes as shown in Fig. 13.
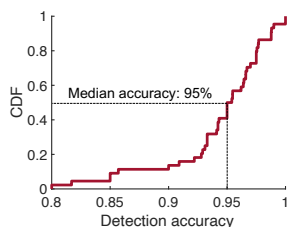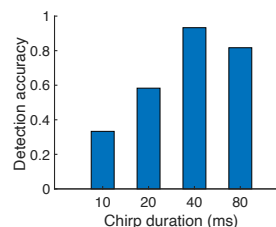
Fig. 14. CDF of the detection accuracy.

Fig. 15. Impact of chirp duration.

8.2 Overall Performance

The eye blink detection accuracy is defined as the number of correctly detected eye blinks over the total number of eye blinks. Fig. 14 plots the cumulative distribution function (CDF) of the overall detection accuracy under different conditions and settings. We see that BlinkListener can achieve a high median detection accuracy of 95%.

8.3 Impact Factors

In this section, we evaluate the performance of our system under different impact factors.

8.3.1 Different Chirp Durations. Chirp duration is an important parameter that affects sensing accuracy. More specifically, a longer chirp duration improves the SNR of the reflection signal and can thus improve the sensing accuracy [38, 44]. However, due to the sparsity of the eye blink, each eye blink cycle only happens within 100 − 400 *ms*. Note that in chirp signal design, we have one distance estimate for each chirp. Therefore, if the chirp duration is too large, the number of estimates becomes too small during one eye blink cycle. For example, if the chirp duration is 100 *ms* and the eye blink duration is 200 *ms*, there are only two estimates during one eye blink cycle. Too few estimates will cause an easy miss of the peak and accordingly degrade the performance of eye blink detection. Therefore, there is an interesting trade-off here. We should choose a chirp duration not too large nor too small. Through comprehensive experiments, we find a chirp duration of 40 *ms* presents us with the optimal performance for eye blink detection. Fig. 15 shows the detailed detection accuracy for different chirp durations.

8.3.2 Different Relative Positions of the Sensing Device with Respect to the Target. To explore the sensing area boundary of BlinkListener, we conduct experiments under different relative positions of the sensing device with respect to the target from four aspects: distance, angle, elevation, and height difference as shown in Fig. 16. Unless otherwise specified, the sensing device is placed at a fixed distance of 0.5 *m*.

Distance: The sensing device is placed at different distances at 0.3, 0.5, 0.8 *m*, respectively. Its orientation is facing towards the human target's eyes. As shown in Fig. 17a, the detection accuracy degrades with the increase of the distance due to the decreasing SNR of the reflected signals. We can achieve a detection accuracy of over 95% within 0.5 *m*. When the distance is increased to 0.8 *m*, the accuracy decreased to 86%. Therefore, we suggest the users keep the device within a distance of 0.5 *m* for high accuracy. Note that beamforming using a microphone array can increase the sensing range. However, most smartphones nowadays are equipped with only 2 microphones.

Angle: The sensing device is placed at different angles from -60 degrees to 60 degrees at a step of 15 degrees. We define the direction of the user's sight as 0 degrees. From Fig. 17b, we can see the detection accuracy decreases with the increase of the angle. When the angle is 0 degree and 15 degrees, the accuracy is higher than 95%. The accuracy degrades severely when the angle is larger than 15 degrees. The reason is that when the sensing device is placed on the side of the human target, the amount of reflection signal power getting back to the receiver
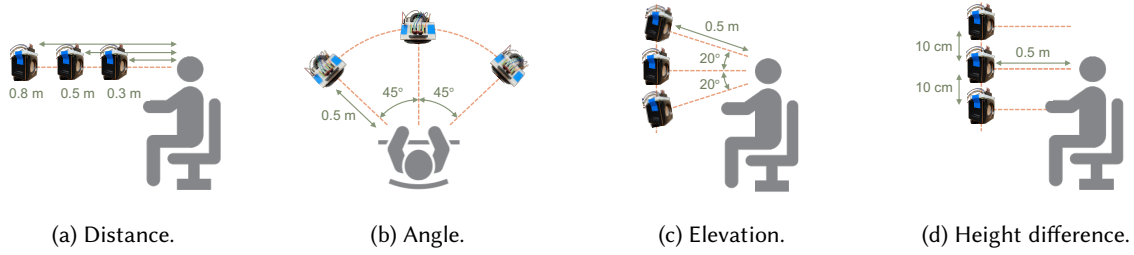
(a) Distance.   (b) Angle.   (c) Elevation.   (d) Height difference.

Fig. 16. Illustration of different relative positions of the sensing device with respect to the human target.

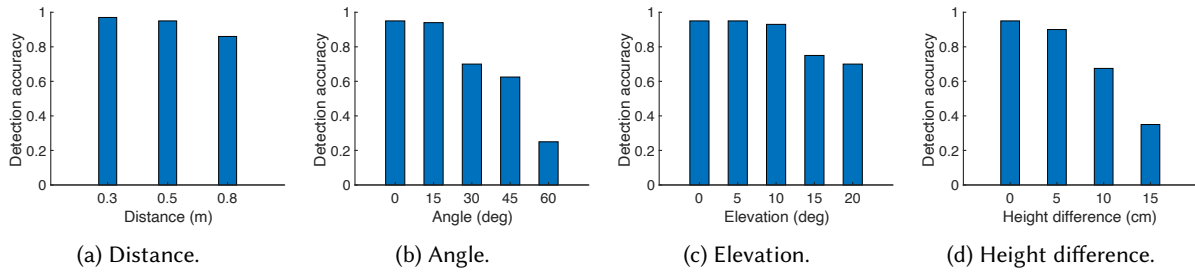(a) Distance.   (b) Angle.   (c) Elevation.   (d) Height difference.

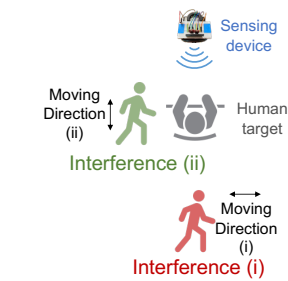Fig. 17. Experiment results under different relative positions of the sensing device with respect to the human target.

becomes less, leading to lower accuracy. This is also related to the signal radiation pattern of the speaker. The signal beam is usually thinner for the inaudible higher frequency band. We suggest the user place the sensing device within the angle range of ±15 degrees to achieve a high detection accuracy.

Elevation: The sensing device is placed at different elevations from -20 degrees to 20 degrees at a step of 5 degrees. The sensing device is placed facing towards the human target's eyes as shown in Fig. 16c. We define the direction of the user's sight as 0 degrees. From Fig. 17c, we can see that the performance decreases as elevation increases. Within ±10 degrees, BlinkListener can achieve a high detection accuracy of 95%.

Height difference: The sensing device is placed at a different height from -15 cm to 15 cm at a step of 5 cm. We define the height of the user's eyes as 0 cm. Different from elevation, the device is not facing towards the target's eyes. From Fig. 17d, we observe that the detection accuracy degrades when the height difference becomes larger. Since the sensing device is not facing towards the eyes and the power of the transmission is confined within a sector range due to the radiation pattern of speakers, the signals can hardly reach the eyes, causing poor sensing performance.

In summary, when the sensing device is exactly facing towards the human target's eyes, BlinkListener can achieve the best performance. As the orientation of the sensing device deviates from the user's sight, the performance decreases. However, as long as the deviation is within a certain range, the performance is still reasonably good.

8.3.3 Different Environment Interference. Since the environment interference coming from static objects can be filtered out by our frequency bin identification method, we only evaluate the impact of environment interference coming from moving objects. We conduct experiments under two different settings as shown in Fig. 18a: (i) A person (interferer) walks behind the human target at a distance of 1 m from the sensing device; (ii) A person walks at the left side of the human target at a distance of 0.5 m from the target. We compare the experiment

(a) Two different interference experiment settings.
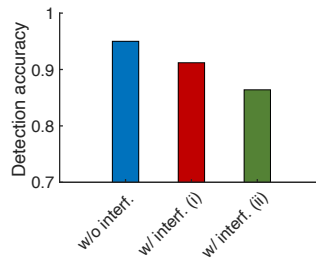
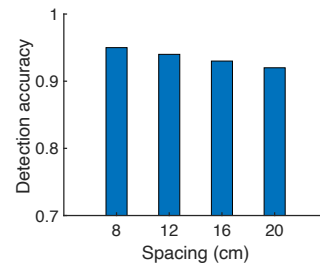(b) Experiment results under different interference cases.

Fig. 18. Environment interference experiment settings and corresponding results.

Fig. 19. Results under different spacing between the speaker and microphone.

results under three different cases: without interference, with interference (i) and with interference (ii). We show the results in Fig. 18b. We can see that under the first case of interference, the detection accuracy is 91%, which is slightly lower than that without interference (95%). This is because the signals from the interferer at a further distance can be filtered out by our frequency bin identification method owing to the high resolution (4.25 *cm*) of the frequency bin, and will not greatly impact sensing of the target. Under the second case of interference, the detection accuracy degrades to 86%. This is because the interferer is at a similar distance as the human target with respect to the sensing device. Therefore, signals reflected by the interferer may fall into the same or adjacent frequency bin as the eye-reflected signals, which will interfere with the eye blink detection. However, we can see that even in the presence of strong interference, the detection accuracy is still reasonably high.
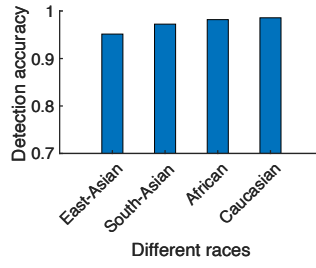
8.3.4 Different Spacing between the Speaker and Microphone. To study the effect of spacing between speaker and microphone, we increase the spacing from 8 to 12, 16 and 20 *cm* on the Bela platform to evaluate the detection accuracy. The speaker and microphone are placed at the same height. The target sits at 0.5 *m* away from the midpoint of the speaker and microphone, and the speaker and microphone are both facing towards the human target's eyes. Fig. 19 shows the results under different spacings between speaker and microphone. Interestingly, we observe that the accuracy slightly decreases with larger spacing. We believe this is because the signal propagation path is actually slightly longer with larger spacing. Also, when the microphone and speaker are further away, the reflected signal may be slightly deviated from the receiver due to the beam-like signal radiation pattern of a commodity speaker.

8.3.5 User Diversity. We recruit 25 participants to evaluate the performance of our system. They are diverse in age (from 20 to 56 years old) and race (East-Asian, South-Asian, Caucasian, and African). Results show that our system achieves an average detection accuracy of 96.2% for the 25 participants. Moreover, we find something interesting: the detection performance does get affected by the eye size and the higher accuracy can be achieved for the larger eye size. We show the detection accuracy with respect to the eye size in Fig. 20a. We can see that for a smaller eye size, the accuracy decreases to around 92%. As the average eye size is related to race, the detection accuracy does slightly vary across races as shown in Fig. 20b. However, we want to point out that even for the smallest eye size (3 *cm* × 0.7 *cm*) in our experiment, the detection accuracy is still above 90%.

8.3.6 Effect of Ambient Sound Noises. We evaluate the performance under different ambient sound noise. We consider two types of sound sources, i.e., music and talk, and evaluate two different volume levels for each sound type. The sound pressure levels (i.e., an indicator of the sound volume) are measured using the Decibel X app [42] on iPhone 11 Pro [9]. Fig. 21 shows the detection accuracy for five different sound levels. We find that
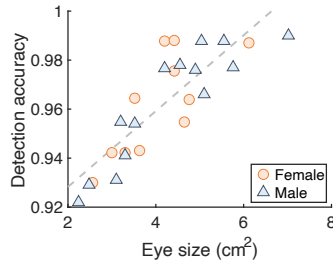
(a) Results with respect to eye sizes.
(b) Results with respect to races.
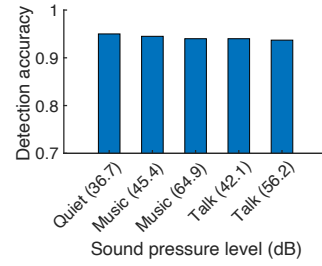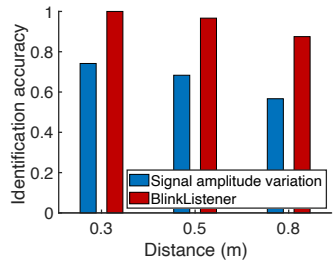
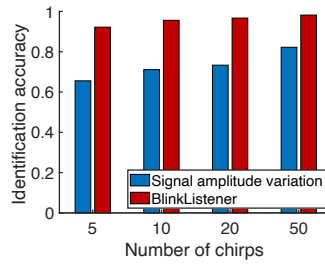Fig. 20. Results of user diversity.

Fig. 21. Effect of different ambient sound noise.

(a) Identification accuracy under different distances.
(b) Identification accuracy under different numbers of chirps.

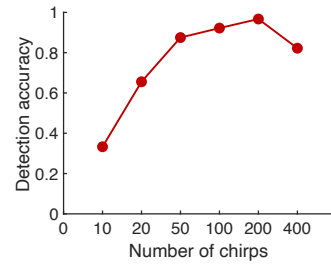Fig. 22. Eye-reflection frequency bin identification accuracy.

Fig. 23. Detection accuracy under different numbers of chirps in our real-time algorithm.

the achieved detection accuracies are very similar, which means the ambient sound noise has little impact on the sensing performance. The reason is that the ambient sound noise is below 14 kHz, while the sensing signal we adopt is in the range of 18-22 kHz, which is much higher than the ambient sound noise.

8.4 Effectiveness of Frequency Bin Identification

Identifying the frequency bin containing the eye-reflection signal is the key enabler for eye blink detection. We evaluate the accuracy of our frequency bin identification method under different distances and different numbers of chirps. We compare the proposed method with the signal amplitude variation method employed in breathing monitoring. From Fig. 22a, we can see that our system can achieve an identification accuracy close to 100% when the distance is less than 0.5 m, which is much higher than the signal amplitude variation method. When the distance is increased to 0.8 m, the identification accuracy slightly decreases to 88% due to weaker signals. This is still much higher than the accuracy (i.e., 57%) achieved with the signal amplitude variation method.

Moreover, the more chirps we employ, the more accurate we can identify the frequency bin with eye-reflection signals. We can see from Fig. 22b that BlinkListener achieves a high identification accuracy of 92.2% with just 5 chirps (i.e., 0.2s), and close to 100% accuracy when more than 20 chirps are employed.

8.5 Effectiveness of Real-time Eye Blink Detection Algorithm

There is a trade-off between detection accuracy and system latency when applying the viewing position method for eye blink detection in our system. Fig. 23 shows the result of the detection accuracy under different numbers

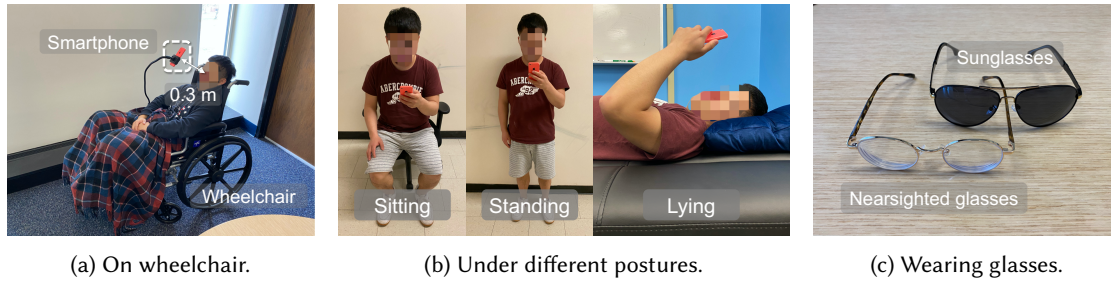(a) On wheelchair. (b) Under different postures. (c) Wearing glasses.
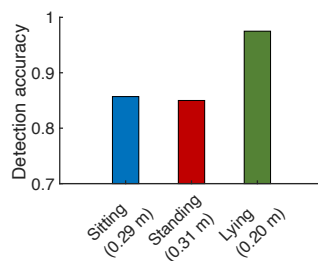
Fig. 24. Field study scenarios.

Fig. 25. Results of detection accuracy under different postures when the sensing device is held by hand.
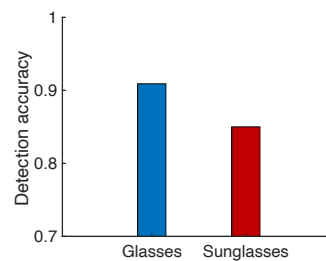
Fig. 26. Results of detection accuracy when user wears nearsighted glasses and sunglasses.

of chirps. We can see that with 10 chirps (i.e., 0.4 s) and 20 chirps (i.e., 0.8 s), the detection accuracy is low. This is because the performance of our arc fitting algorithm highly depends on the number of data samples. If too few samples are available, the obtained optimal viewing position is not accurate, leading to a degraded eye blink detection performance. With an increasing number of chirps, we can obtain a more accurate eye blink detection accuracy. However, when too many chirps are included, a long period of time is needed to accumulate those chirps before a viewing position can be obtained. The bad news is that during this period, the target may have moved and the optimal viewing position is also changed, causing the obtained view position to be outdated. To balance the detection accuracy and system latency, we adopt 50 chirps (i.e., 2 s) for initialization and optimal viewing position update.

8.6 Field Studies with Smartphones

We conduct a series of field studies with a commodity smartphone iPhone 5c to demonstrate the feasibility of BlinkListener.

8.6.1 Eye Blink Detection in a Wheelchair. We mimic eye blink detection for individuals with disabilities in a wheelchair with a leaning back posture as shown in Fig. 24a. We employ a smartphone stand holder to hold the smartphone. The smartphone is placed facing towards the user's eyes at a comfortable viewing distance of 0.3 m. Under this setting, BlinkListener can achieve an eye blink detection accuracy of 96%.

8.6.2 Eye Blink Detection in a Moving Car. We conduct in-the-wild experiments in a moving car. The experiment setup is shown in Fig. 27a. The smartphone is mounted on a stand holder, which is attached to the upper left corner of the windshield. The distance between the smartphone and the target's eyes is around 0.3 m, and the

(a) Experiment settings in the car.
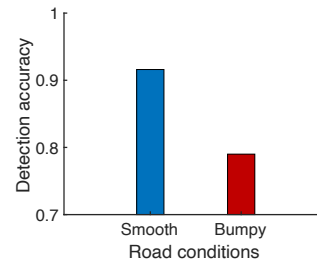
(b) Results under different road conditions.

Fig. 27. In-the-wild experiments in a moving car.

placement of the smartphone does not block the driver's view. The ground truths are recorded by a camera. We drove the car at an average speed of 35 *mph*. The result shows that our system can achieve an average detection accuracy of 87% when the system is inside a moving car. We find that when the road is smooth, our system can achieve 91.6% detection accuracy. When the road is bumpy, the detection accuracy decreases to 79% as shown in Fig. 27b. This is because the large motor vibrations and device displacements vary the distance measurement between the device and the target and accordingly affect the sensing performance.

8.6.3   Eye Blink Detection with a Smartphone Held in Hand. As shown in Fig. 24b, we conduct experiments with a smartphone held in hand under three different postures: sitting, standing and lying. A volunteer is asked to hold the smartphone in hand and keep a natural and comfortable viewing distance. We measure the viewing distance during the experiment process and the average viewing distances of the volunteer under the three postures are 0.29 *m*, 0.31 *m* and 0.20 *m*, respectively, which is consistent with the result reported in research on smartphone viewing distance [77]. The detection accuracy under sitting, standing and lying postures are 86%, 85% and 97% respectively as shown in Fig. 25. We observe some very interesting results here. The eye blink detection accuracy when the smartphone is held in hand is lower than that when the smartphone is mounted at the phone stand holder. We believe this is because when the phone is held in hand, there are muscle tremor and other small body movements which affect the sensing accuracy. However, we do observe a very high accuracy when the user is in a lying posture. We believe this is due to two reasons. First, when a user is in a lying posture, his body is more stable than the case when he is standing/sitting. Second, the viewing distance between the smartphone and eyes in lying posture is smaller than that in sitting and standing postures.

8.6.4   Eye Blink Detection with Glasses on. We conduct experiments to see if the proposed system works when the target wears glasses. We evaluate two different types of glasses: nearsighted glasses and sunglasses as shown in Fig. 24c. The smartphone is fixed in a stand holder at a distance of 0.3 *m* with respect to the human target and facing towards user's eyes. Fig. 26 shows that BlinkListener can achieve 90.9% and 85.0% detection accuracy when the user wears glasses and sunglasses, respectively. Although the accuracy slightly decreases, the proposed system does work with glasses. This result demonstrates the advantage of eye blink detection over camera-based techniques when users wear sunglasses. We believe the slightly lower performance with glasses is mainly due to the signal attenuation through glasses and interference from the metal frame.

9   LIMITATION AND DISCUSSION

In this section, we discuss the limitations of our system and potential future work.

9.1 Limited Angular Sensing Range

Compared with camera-based methods, the proposed acoustic-based method has a smaller angular sensing range. We conclude that this limitation is due to two aspects: (i) limited angular range of speakers, and (ii) the characteristic of reflection-based sensing. First, the sound volume generated from a speaker varies at different angles [30]. We tested two types of general-purpose speakers and the built-in speaker of the iPhone 5c. We found the angular ranges are all within ±30 degrees, which are smaller than that of cameras (i.e., ±45 degrees) [28, 51]. Second, the proposed system relies on the reflection signals for sensing. The size of the human eye is small and therefore the amount of reflection from the eye area is limited. When the eye is located at a relatively large angle with respect to the smartphone (e.g., ±20 degrees), even though there are still signals reaching the eyes, the majority of reflections will be towards other directions due to the symmetrical property of reflection and will not be received at the smartphone. Therefore, the actual sensing range is even smaller than the speaker's angular range. From our experiment results, the effective sensing angular range is ±15 degrees in the horizontal plane and ±10 degrees in the vertical plane.

We position our system as a complementary approach to camera-based methods. Although the camera-based methods can achieve higher accuracy and a larger angular sensing range, the performance degrades in low lighting conditions and may raise privacy concerns. We believe there is no cure-all solution. We should choose the appropriate approach based on the application requirement.

9.2 Impact of Device Vibration/Displacement

As explained in the evaluation part, in a moving car, the performance of system degrades when we drive on a bumpy road. The reason is that vibrations and displacements vary the distance measurement between the smartphone and the human subject and accordingly affect the sensing performance. Device vibration/displacement is a true challenge for wireless sensing because the detected movement information comes from both the target and device. It is non-trivial to separate them to obtain the target information. However, we believe this is an important issue because the sensing device has a good chance to be placed in a car. We leave this interesting yet challenging problem as our important future work.

10 CONCLUSION

In this paper, we enable subtle eye blink detection using acoustic signals on both a research-purpose hardware platform and a commodity smartphone for the first time. Through both theoretical and experimental analysis, we quantitatively model the relationship between signal variations and the subtle movements caused by eye blink and interference. We propose a novel viewing position scheme that fully exploits the "harmful" embedded interference to maximize the subtle signal variation induced by eye blinks. Comprehensive experiment results demonstrate the effectiveness of our system. We believe the proposed viewing position scheme can be applied to improve the performance of other sensing applications which involve extremely small signal variations.

REFERENCES

[1] 2018. *J!NS MEME Eyewear.* https://jins-meme.com/en/

[2] 2020. *Eye-tracking glasses.* https://www.imec-int.com/en/eog

[3] 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.

[4] Fadel Adib, Hongzi Mao, Zachary Kabelac, Dina Katabi, and Robert C Miller. 2015. Smart homes that monitor breathing and heart rate. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 837–846.

[5] National Highway Traffic Safety Administration et al. 2017. Traffic safety facts crash stats: drowsy driving 2015.

[6] Raees Ahmad and JN Borole. 2015. Drowsy driver identification using eye blink detection. *IJISET-International Journal of Computer Science and Information Technologies* 6, 1 (2015), 270–274.

[7] Doğukan Aksu and M Ali Aydın. 2017. Human computer interaction by eye blinking on real time. In *2017 9th International Conference on Computational Intelligence and Communication Networks (CICN)*. IEEE, 135–138.

[8] Apple. 2019. *iPhone 5c.* https://support.apple.com/kb/sp684?locale=en_US

[9] Apple. 2021. *iPhone 11 pro.* https://support.apple.com/kb/SP805?viewlocale=en_US&locale=en_US

[10] Guha Balakrishnan, Fredo Durand, and John Guttag. 2013. Detecting pulse from head motions in video. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3430–3437.

[11] Anna Rita Bentivoglio, Susan B Bressman, Emanuele Cassetta, Donatella Carretta, Pietro Tonali, and Alberto Albanese. 1997. Analysis of blink rate patterns in normal subjects. *Movement disorders* 12, 6 (1997), 1028–1034.

[12] Andrew Bierman, Mariana G Figueiro, and Mark S Rea. 2011. Measuring and predicting eyelid spectral transmittance. *Journal of biomedical optics* 16, 6 (2011), 067011.

[13] Andreas Bulling, Daniel Roggen, and Gerhard Tröster. 2008. It's in your eyes: Towards context-awareness and mobile HCI using wearable EOG goggles. In *Proceedings of the 10th international conference on Ubiquitous computing*. 84–93.

[14] Andreas Bulling, Jamie A Ward, Hans Gellersen, and Gerhard Troster. 2010. Eye movement analysis for activity recognition using electrooculography. *IEEE transactions on pattern analysis and machine intelligence* 33, 4 (2010), 741–753.

[15] Health Canada. 1991. Guidelines for the safe use of ultrasound: Part II—Industrial and commercial applications safety code 24.

[16] Fabio Lo Castro. 2008. Class I infrared eye blinking detector. *Sensors and actuators A: Physical* 148, 2 (2008), 388–394.

[17] Michael Chau and Margrit Betke. 2005. *Real time eye tracking and blink detection with usb cameras.* Technical Report. Boston University Computer Science Department.

[18] Lili Chen, Jie Xiong, Xiaojiang Chen, Sunghoon Ivan Lee, Daqing Zhang, Tao Yan, and Dingyi Fang. 2019. LungTrack: Towards Contactless and Zero Dead-Zone Respiration Monitoring with Commodity RFIDs. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–22.

[19] Tarik Crnovrsanin, Yang Wang, and Kwan-Liu Ma. 2014. Stimulating a blink: reduction of eye fatigue with visual stimulus. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2055–2064.

[20] Taner Danisman, Ian Marius Bilasco, Chabane Djeraba, and Nacim Ihaddadene. 2010. Drowsy driver detection system using eye blink patterns. In *2010 International Conference on Machine and Web Intelligence*. IEEE, 230–233.

[21] Artem Dementyev and Christian Holz. 2017. DualBlink: a wearable device to continuously detect, track, and actuate blinking for alleviating dry eyes and computer vision syndrome. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 1 (2017), 1–19.

[22] Matjaz Divjak and Horst Bischof. 2009. Eye Blink Based Fatigue Detection for Prevention of Computer Vision Syndrome.. In *MVA*. 350–353.

[23] Allen Earman. 2012. Eye Safety for Proximity Sensing Using Infrared Light-Emitting Diodes.

[24] Kyosuke Fukuda, John A Stern, Timothy B Brown, and Michael B Russo. 2005. Cognition, blinks, eye-movements, and pupillary movements during performance of a running memory task. *Aviation, space, and environmental medicine* 76, 7 (2005), C75–C85.

[25] Daniel Golson. 2019. *Every Volvo Will Get In-Car Cameras to Combat Distraction and Drunk Driving.* https://www.caranddriver.com/news/a26893035/volvo-interior-cameras-distraction-drunk-driving/

[26] Sidhant Gupta, Daniel Morris, Shwetak Patel, and Desney Tan. 2012. Soundwave: using the doppler effect to sense gestures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1911–1914.

[27] Seongwon Han, Sungwon Yang, Jihyoung Kim, and Mario Gerla. 2012. EyeGuardian: a framework of eye tracking and blink detection for mobile device users. In *Proceedings of the Twelfth Workshop on Mobile Computing Systems & Applications*. 1–6.

[28] Young-Joo Han, Wooseong Kim, and Joon-Sang Park. 2018. Efficient eye-blinking detection on smartphones: A hybrid approach based on deep learning. *Mobile Information Systems* 2018 (2018).

[29] Mark A Hanson. 2010. *Health effects of exposure to ultrasound and infrasound: Report of the independent advisory group on non-ionising radiation.* Health Protection Agency.

[30] Yaxiong Huang, Simon C Busbridge, and Deshinder S Gill. 2001. Distortion and directivity in a digital transducer array loudspeaker. *Journal of the Audio Engineering Society* 49, 5 (2001), 337–352.

[31] Shoya Ishimaru, Kai Kunze, Koichi Kise, Jens Weppner, Andreas Dengel, Paul Lukowicz, and Andreas Bulling. 2014. In the blink of an eye: combining head motion and eye blink frequency for activity recognition with google glass. In *Proceedings of the 5th augmented*

human international conference. 1–4.

[32] Shoya Ishimaru, Kai Kunze, Yuji Uema, Koichi Kise, Masahiko Inami, and Katsuma Tanaka. 2014. Smarter eyewear: using commercial EOG glasses for activity recognition. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication.* 239–242.

[33] Takehiro Ito, Shinji Mita, Kazuhiro Kozuka, Tomoaki Nakano, and Shin Yamamoto. 2002. Driver blink measurement by the motion picture processing and its application to drowsiness detection. In *Proceedings. The IEEE 5th International Conference on Intelligent Transportation Systems.* IEEE, 168–173.

[34] Sheba Jebakani, J Divya, Megha SM, and Santhosh HV. 2020. Eye Blink to Voice for Paralyzed Patients. *International Journal of Information Technology (IJIT)* 6, 3 (2020).

[35] Moritz Kassner, William Patera, and Andreas Bulling. 2014. Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: Adjunct publication.* 1151–1160.

[36] Nataliya Kosmyna, Caitlin Morris, Thanh Nguyen, Sebastian Zepf, Javier Hernandez, and Pattie Maes. 2019. AttentivU: Designing EEG and EOG compatible glasses for physiological sensing and feedback in the car. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.* 355–368.

[37] Aleksandra Królak and Paweł Strumiłło. 2012. Eye-blink detection system for human–computer interaction. *Universal Access in the Information Society* 11, 4 (2012), 409–419.

[38] Dong Li, Jialin Liu, Sunghoon Ivan Lee, and Jie Xiong. 2020. FM-Track: Pushing the Limits of Contactless Multi-target Tracking using Acoustic Signals. In *Proceedings of the 18th ACM Conference on Embedded Networked Sensor Systems.* 1–14.

[39] Tianxing Li, Chuankai An, Zhao Tian, Andrew T Campbell, and Xia Zhou. 2015. Human sensing using visible light communication. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking.* 331–344.

[40] Tianxing Li, Xi Xiong, Yifei Xie, George Hito, Xing-Dong Yang, and Xia Zhou. 2017. Reconstructing hand poses using visible light. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–20.

[41] 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 Transactions on Graphics (TOG)* 35, 4 (2016), 1–19.

[42] SkyPaw Co. Ltd. 2021. *Decibel X: dB Sound Level Meter.* https://apps.apple.com/us/app/decibel-x-db-sound-level-meter/id448155923

[43] Radosław Mantiuk, Michał Kowalik, Adam Nowosielski, and Bartosz Bazyluk. 2012. Do-it-yourself eye tracker: Low-cost pupil-based eye tracker for computer graphics applications. In *International Conference on Multimedia Modeling.* Springer, 115–125.

[44] Wenguang Mao, Mei Wang, Wei Sun, Lili Qiu, Swadhin Pradhan, and Yi-Chao Chen. 2019. RNN-Based Room Scale Hand Motion Tracking. In *The 25th Annual International Conference on Mobile Computing and Networking.* ACM, 38.

[45] Alvaro Marcos-Ramiro, Daniel Pizarro-Perez, Marta Marron-Romera, Daniel Pizarro-Perez, and Daniel Gatica-Perez. 2014. Automatic blinking detection towards stress discovery. In *Proceedings of the 16th International Conference on Multimodal Interaction.* 307–310.

[46] Denis Metev. 2020. *2020's Voice Search Statistics – Is Voice Search Growing?* https://review42.com/voice-search-stats/

[47] Emiliano Miluzzo, Tianyu Wang, and Andrew T Campbell. 2010. EyePhone: activating mobile phones with your eyes. In *Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds.* 15–20.

[48] Danielle Lumi Miura, Rossen Mihaylov Hazarbassanov, Camila Karim Nakase Yamasato, Francisco Bandeira e Silva, Cléber José Godinho, and José Álvaro Pereira Gomes. 2013. Effect of a light-emitting timer device on the blink rate of non-dry eye individuals and dry eye patients. *British Journal of Ophthalmology* 97, 8 (2013), 965–967.

[49] Assit Prof Aree A Mohammed and Shereen A Anwer. 2014. Efficient eye blink detection method for disabled-helping domain. *Eye* 10, P1 (2014), P2.

[50] Rajalakshmi Nandakumar, Vikram Iyer, Desney Tan, and Shyamnath Gollakota. 2016. Fingerio: Using active sonar for fine-grained finger tracking. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.* 1515–1525.

[51] Md Noman, Talal Bin, Md Ahad, and Atiqur Rahman. 2018. Mobile-based eye-blink detection performance analysis on android platform. *Frontiers in ICT* 5 (2018), 4.

[52] Tanuja Patgar and Ripal Patel. 2020. Emergency Assistive System for Tetraplegia Patient Using Eye Waver Computer Vision Technique. In *Advances in Computational Intelligence Techniques.* Springer, 63–79.

[53] Nhat Pham, Tuan Dinh, Zohreh Raghebi, Taeho Kim, Nam Bui, Phuc Nguyen, Hoang Truong, Farnoush Banaei-Kashani, Ann Halbower, Thang Dinh, et al. 2020. WAKE: a behind-the-ear wearable system for microsleep detection. In *Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services.* 404–418.

[54] Vaughan Pratt. 1987. Direct least-squares fitting of algebraic surfaces. *ACM SIGGRAPH computer graphics* 21, 4 (1987), 145–152.

[55] Qifan Pu, Sidhant Gupta, Shyamnath Gollakota, and Shwetak Patel. 2013. Whole-home gesture recognition using wireless signals. In *Proceedings of the 19th annual international conference on Mobile computing & networking.* 27–38.

[56] Kun Qian, Chenshu Wu, Fu Xiao, Yue Zheng, Yi Zhang, Zheng Yang, and Yunhao Liu. 2018. Acousticcardiogram: Monitoring heartbeats using acoustic signals on smart devices. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications.* IEEE, 1574–1582.

[57] Albara Ah Ramli, Rex Liu, Rahul Krishnamoorthy, Xiaoxiao Wang, Ilias Tagkopoulos, Xin Liu, et al. 2020. BWCNN: Blink to Word, a Real-Time Convolutional Neural Network Approach. *arXiv preprint arXiv:2006.01232* (2020).

[58] Wenjie Ruan, Quan Z Sheng, Lei Yang, Tao Gu, Peipei Xu, and Longfei Shangguan. 2016. AudioGest: enabling fine-grained hand gesture detection by decoding echo signal. In *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing.* 474–485.

[59] Robert Schleicher, Niels Galley, Susanne Briest, and Lars Galley. 2008. Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired? *Ergonomics* 51, 7 (2008), 982–1010.

[60] Jürgen Schmidt, Rihab Laarousi, Wolfgang Stolzmann, and Katja Karrer-Gauß. 2018. Eye blink detection for different driver states in conditionally automated driving and manual driving using EOG and a driver camera. *Behavior research methods* 50, 3 (2018), 1088–1101.

[61] Tereza Soukupova and Jan Cech. 2016. Eye blink detection using facial landmarks. In *21st computer vision winter workshop, Rimske Toplice, Slovenia.*

[62] Isaac Starr, AJ Rawson, HA Schroeder, and NR Joseph. 1939. Studies on the estimation of cardiac ouptut in man, and of abnormalities in cardiac function, from the heart's recoil and the blood's impacts; the ballistocardiogram. *American Journal of Physiology-Legacy Content* 127, 1 (1939), 1–28.

[63] Benjamin Tag, Junichi Shimizu, Chi Zhang, Naohisa Ohta, Kai Kunze, and Kazunori Sugiura. 2016. Eye blink as an input modality for a responsive adaptable video system. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct.* 205–208.

[64] Benjamin Tag, Andrew W Vargo, Aman Gupta, George Chernyshov, Kai Kunze, and Tilman Dingler. 2019. Continuous alertness assessments: Using EOG glasses to unobtrusively monitor fatigue levels In-The-Wild. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* 1–12.

[65] HR Taschenbuch Verlag Schiffman. 2001. Sensation and Perception: An Integrated Approach.

[66] Bela Team. 2020. *Bela Platform.* https://bela.io

[67] Yu-Chih Tung, Duc Bui, and Kang G Shin. 2018. Cross-platform support for rapid development of mobile acoustic sensing applications. In *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services.* 455–467.

[68] Anran Wang and Shyamnath Gollakota. 2019. Millisonic: Pushing the limits of acoustic motion tracking. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* 1–11.

[69] Anran Wang, Jacob E Sunshine, and Shyamnath Gollakota. 2019. Contactless infant monitoring using white noise. In *The 25th Annual International Conference on Mobile Computing and Networking.* 1–16.

[70] Saiwen Wang, Jie Song, Jaime Lien, Ivan Poupyrev, and Otmar Hilliges. 2016. Interacting with soli: Exploring fine-grained dynamic gesture recognition in the radio-frequency spectrum. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology.* 851–860.

[71] Tianben Wang, Daqing Zhang, Yuanqing Zheng, Tao Gu, Xingshe Zhou, and Bernadette Dorizzi. 2018. C-FMCW based contactless respiration detection using acoustic signal. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (2018), 1–20.

[72] Wei Wang, Alex X Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. 2015. Understanding and modeling of wifi signal based human activity recognition. In *Proceedings of the 21st annual international conference on mobile computing and networking.* 65–76.

[73] Wei Wang, Alex X Liu, and Ke Sun. 2016. Device-free gesture tracking using acoustic signals. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking.* 82–94.

[74] Binbin Xie, Jie Xiong, Xiaojiang Chen, Eugene Chai, Liyao Li, Zhanyong Tang, and Dingyi Fang. 2019. Tagtag: material sensing with commodity RFID. In *Proceedings of the 17th Conference on Embedded Networked Sensor Systems.* 338–350.

[75] Yaxiong Xie, Jie Xiong, Mo Li, and Kyle Jamieson. 2019. mD-Track: Leveraging multi-dimensionality for passive indoor Wi-Fi tracking. In *The 25th Annual International Conference on Mobile Computing and Networking.* 1–16.

[76] 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.

[77] Michitaka Yoshimura, Momoko Kitazawa, Yasuhiro Maeda, Masaru Mimura, Kazuo Tsubota, and Taishiro Kishimoto. 2017. Smartphone viewing distance and sleep: an experimental study utilizing motion capture technology. *Nature and science of sleep* 9 (2017), 59.

[78] Chuang-Wen You, Martha Montes-de Oca, Thomas J Bao, Nicholas D Lane, Hong Lu, Giuseppe Cardone, Lorenzo Torresani, and Andrew T Campbell. 2012. CarSafe: a driver safety app that detects dangerous driving behavior using dual-cameras on smartphones. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing.* 671–672.

[79] Sangki Yun, Yi-Chao Chen, Huihuang Zheng, Lili Qiu, and Wenguang Mao. 2017. Strata: Fine-grained acoustic-based device-free tracking. In *Proceedings of the 15th annual international conference on mobile systems, applications, and services.* 15–28.

[80] Youwei Zeng, Dan Wu, Ruiyang Gao, Tao Gu, and Daqing Zhang. 2018. Fullbreathe: Full human respiration detection exploiting complementarity of csi phase and amplitude of wifi signals. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 1–19.

[81] Youwei Zeng, Dan Wu, Jie Xiong, Enze Yi, Ruiyang Gao, and Daqing Zhang. 2019. FarSense: Pushing the range limit of WiFi-based respiration sensing with CSI ratio of two antennas. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–26.

[82] Zhiwei Zhu, Kikuo Fujimura, and Qiang Ji. 2002. Real-time eye detection and tracking under various light conditions. In *Proceedings of the 2002 symposium on Eye tracking research & applications*. 139–144.