


Article

Contactless Vital Sign Monitoring System for In-Vehicle Driver Monitoring Using a Near-Infrared Time-of-Flight Camera

Kaiwen Guo ^{1,*},[†] , Tianqu Zhai ¹, Manoj H. Purushothama ¹, Alexander Dobre ¹, Shawn Meah ¹, Elton Pashollari ¹, Aabhaas Vaish ¹, Carl DeWilde ² and Mohammed N. Islam ^{1,2}

¹ Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109, USA; tianquzh@umich.edu (T.Z.); hpmanoj@umich.edu (M.H.P.); dooreal@umich.edu (A.D.); meah@umich.edu (S.M.); epasholl@umich.edu (E.P.); aabhaas@umich.edu (A.V.); mni@eecs.umich.edu (M.N.I.)

² Omni Science Inc., Dexter, MI 48130, USA; cdewilde@omnisciinc.com

* Correspondence: kaiwguo@umich.edu

[†] Current address: EECS Building, 1301 Beal Avenue, Ann Arbor, MI 48109, USA.

Abstract: We demonstrate a Contactless Vital Sign Monitoring (CVSM) system and road-test the system for in-cabin driver monitoring using a near-infrared indirect Time-of-Flight (ToF) camera. The CVSM measures both heart rate (HR) and respiration rate (RR) by leveraging the simultaneously measured grayscale and depth information from a ToF camera. For a camera-based driver monitoring system (DMS), key challenges from varying background illumination and motion-induced artifacts need to be addressed. In this study, active illumination and depth-based motion compensation are used to mitigate these two challenges. For HR measurements, active illumination allows the system to work under various lighting conditions, while our depth-based motion compensation has the advantage of directly measuring the motion of the driver without making prior assumptions about the motion artifacts. In addition, we can extract RR directly from the chest wall motion, circumventing the challenge of acquiring RR from the near-infrared photoplethysmography (PPG) signal of low signal quality. We investigate the system's performance in various scenarios, including monitoring both drivers and passengers while driving on highways and local roads. Our results show that our CVSM system is ambient light agnostic, and the success rates of HR measurements on the highway are 82% and 71.9% for the passenger and driver, respectively. At the same time, we show that the system can measure RR on users driving on a highway with a mean deviation of -1.4 breaths per minute (BPM). With reliable HR and RR measurement in the vehicle, the CVSM system could one day be a key enabler to sudden sickness or drowsiness detection in DMS.

Keywords: Time-of-Flight; contactless physiological measurement; heart rate monitoring; motion artifacts compensation; driver monitoring system



Citation: Guo, K.; Zhai, T.; Purushothama, M.H.; Dobre, A.; Meah, S.; Pashollari, E.; Vaish, A.; DeWilde, C.; Islam, M.N. Contactless Vital Sign Monitoring System for In-Vehicle Driver Monitoring Using a Near-Infrared Time-of-Flight Camera. *Appl. Sci.* **2022**, *12*, 4416. <https://doi.org/10.3390/app12094416>

Academic Editor:
Rubén Usamentiaga

Received: 9 March 2022

Accepted: 24 April 2022

Published: 27 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In this study, we present a Contactless Vital Sign Monitoring system (CVSM) for in-cabin heart rate (HR) and respiration rate (RR) measurements using a near-infrared indirect Time-of-Flight (ToF) camera. We address the challenges of varying ambient illumination as well as the interference from excessive motion by utilizing the active illumination and the additional depth information from the ToF camera. The synchronized 850 nm active illumination in the ToF camera allows us to operate independently of the ambient light conditions, while the additional depth information can be used to compensate for motion-induced intensity artifacts. As the users move in the vehicle (passively or actively), their relative position against the ToF camera also changes; thus, the additional depth information measured by the ToF camera is correlated with the motion of the driver. We utilize this correlation to compensate for the motion-induced intensity artifact by differentiating whether the grayscale intensity change is due to motion or the blood volume

change from the heartbeat. Furthermore, the depth information is also used to estimate the amount of motion during the measurement period, therefore predicting the quality of the HR measurements. Additionally, the depth change from chest wall motion can be used to directly measure RR, thereby allowing our system to measure both HR and RR with one monolithic sensor. We systematically investigate the performance of our CVSM system by performing HR/RR measurements in various realistic environments with different degrees of motion from the roads and the drivers. We study the benefit of the depth-based motion compensation by comparing the success rate of HR measurements using both grayscale and depth information (“compensated” HR measurements) against HR measurements using only the grayscale information of the same video (“uncompensated” HR measurements). Our study finds that with the depth-based motion compensation, the HR measurement success rate increase from 13.6% (“uncompensated” HR measurements) to 71.9% while the user is driving on the highway. When the user is driving on a local road, the success rate increases to 56% from 12%. As for the RR measurement, our CVSM system measures the RR of the driver on highway scenario with a mean deviation of -1.4 BPM from the reference RR measurements. With the ability to measure two important physiological parameters with one sensor, we believe our ToF based CVSM system can enable a low-cost, compact, and multi-modal driver monitoring system that can be used for applications such as drowsiness or sudden sickness detection in vehicle.

2. Background and Motivation

As Advanced Driver Assistant Systems (ADAS) are increasingly adopted by both car manufacturers and regulators, there is a growing need to monitor the physiological states of drivers/occupants to assist in the engagement of the ADAS system [1–5]. For example, the EuroNCAP regulation requires the driver/occupants monitoring system (DMS/OMS) to detect sudden sickness as well the presence of a younger child in the vehicle [6]. Both HR and RR are shown to be useful indicators to predict the awareness level of the driver, which could, in turn, be used to determine if ADAS functions such as automatic lane keeping or front collision avoidance should be activated to ensure the safety of the driver [3–5,7]. Additionally, with emerging autonomous services such as robotic taxis gaining traction, an unobtrusive CVSM system can be a valuable addition to ensure the well-being of the passengers [3].

Several sensors, such as cameras, radars, electrocardiograms (ECG), and ballistocardiograph sensors, have been studied to extract HR and RR from the driver [4,5,8–16]. Table 1 summarizes the characteristics of various sensors commonly used for HR and RR monitoring of the driver/occupants in the vehicle. Among various sensors, the electrocardiogram measures HR from the electrical potential of the heart, while the ballistocardiograph sensor measures HR through the micro-motion of the body caused by the heartbeat. In comparison, both radars and cameras promise unobtrusive and contactless monitoring of HR and RR. Between these two non-contact sensors, cameras hold an advantage by offering other unique functions such as facial recognition and gaze detection that cannot be achieved with radars. With a conventional 2D camera, HR is usually acquired through imaging photoplethysmography (PPG), where the volumetric change of blood flow from the heartbeat causes different amounts of light to be absorbed by the user’s face [17–22]. Several past studies (such as chrominance (CHROM) or independent component analyses (ICA) based method) have shown the robustness of using a camera to contactlessly measure HR in controlled environments [23–27] by utilizing the RGB color channels in an RGB camera.

Compared to measuring HR and RR in a controlled indoor environment, additional challenges from the varying illumination as well as the motion artifacts need to be addressed when measuring HR and RR while driving. In a moving vehicle, the illumination level can vary significantly based on factors such as weather or time of day. Any imaging-based sensor will need to be unaffected by the ambient light. Secondly, during driving, both the passive motions from the road or active motions from the users can interfere with the measurement result as well. Thus, for a camera-based sensor, a method for rejecting motion-induced artifacts is important to ensure reliable HR/RR measurement performance.

Recently, several studies have discussed the use of a camera to measure HR in a vehicle [11–14]. For example, Huang et al. demonstrated HR monitoring in a driving vehicle with a RGB camera [12]. Nowara et al. demonstrated an in-vehicle HR monitoring system (Autospase PPG) using an NIR camera with LED illuminations, where a sparse frequency estimation method is introduced to improve the HR measurement reliability against external artifacts [11]. However, for the Autospase PPG method to be effective, uniform illumination of the user is required, which may not be satisfied when a compact sensor with a localized/point illumination source (such as Vertical Cavity Surface Emitting Lasers (VCSEL)) is used. In our previous study [28], we demonstrate a system using an indirect NIR ToF camera to monitor HR and RR in an indoor environment. We expand upon the previous study and demonstrate a CVSM system for HR and RR measurements in a moving vehicle. In this study, we improve upon the methods proposed in [28] and focus on the more complex application scenario of in-vehicle HR and RR monitoring. Specifically, we not only use the depth information to distinguish intensity changes caused by body motion versus heartbeat but also to assess the quality during different sections of the HR measurements, which further increase the robustness of the HR measurements against motion-induced artifacts [28]. Since our depth-based method does not make the sparse frequency assumption of the PPG signal, this system would work even if a compact/localized VCSEL illuminator is used.

Furthermore, even though RR has been shown to be correlated with the drowsiness or stress level of the driver [8], the existing camera-based system usually does not have an adequate signal quality to extract RR from the HR signal [29,30]. To circumvent the limitation of low signal quality, in this study, we use the ToF camera to measure RR directly by measuring the chest wall displacement [28,31], avoiding the potential complexity of fusing the data from the radar sensor and the camera, respectively, and reducing the cost for the entire system.

To the best of our knowledge, this is the first study where a ToF camera is used to simultaneously perform HR and RR measurements in a vehicle. Indirect ToF cameras are becoming increasingly popular on various platforms and are already deployed in vehicles for functions such as Smart Restraint Control Systems (Smart-RCS). Given the benefits brought by the active illumination and depth information, our CVSM system would enable compact, reliable, and multi-modal physiological signal monitoring in a vehicle, improving the safety and the driving experience for both the drivers and passengers.

Table 1. Summary of HR/RR measurement techniques used in vehicle.

Type of Sensor	HR?	RR? *	Contactless?	HR Robust against Motion?	Ambient Light Resistance?	Independent of Skin Color?
Electrocardiogram [9,32]	++	Indirect	N	++	++	++
Radar [15,16]	–	Direct	Y	–	++	++
Ballistocardiograph [33]	–	Direct	N	–	++	++
Thermography [8,20]	N/A	Direct	Y	N/A	++	++
PPG with RGB Camera [12,14]	++	Indirect	Y	+	–	–
PPG with IR Camera [11,13]	+	N/A	Y	–	+	+
PPG with indirect ToF Camera [This Study]	+	Direct	Y	+	++	+

* Direct measurement measures chest movement meanwhile an indirect RR measurement uses the modulation of the HR signal to extract RR; +/– indicates the system has advantage/disadvantage in a certain category.

3. Experiment Setup

3.1. Hardware Configuration

The indirect ToF camera used in this study is an amplitude modulated ToF camera with a resolution of 640×480 with a 149° diagonal field of view (FoV). This large FoV can accommodate different seating positions and postures of the driver and can potentially be used to monitor multiple passengers in the vehicle. The active illumination is provided by a pair of 850 nm (we find both 940 and 850 nm illumination to have similar performance in terms of HR measurements) VCSELs with diffusers at eye-safe power. As a prototype, the ToF camera has a form factor of $15 \times 5 \times 5$ cm and consumes 5 W average power, which could be further reduced with future packaging improvements. The ToF camera is inherently robust against ambient illumination change: Firstly, the lens of the ToF camera is coated with a bandpass filter at 850 nm to suppress the out-of-band background illumination. At the same time, the underlying lock-in detection principle of the indirect ToF sensor can suppress the residual in-band background illumination [34].

In our experiment, the ToF camera is mounted in a 2015 Jeep Cherokee near the rearview mirror using an adjustable suction cup mount either facing the driver side of the cabin or the passenger side of the cabin. The ToF camera is positioned at roughly eye level and 50 cm away from the participants. Due to the wide FoV, the ToF can capture the users' face and chest region at the same time (Figure 1). During our experiments, the ToF camera operated at 30 frames per second (FPS) and each measurement records 20 s of video. The frame rate and the measurement time were chosen due to the bandwidth as well as the processing power of the laptop. Longer measurements could be chosen to improve the resolution of the HR/RR measurements at the cost of the recording and processing time. Even though in this study, the HR was measured discretely after each 20 s measurement, the HR measurements can be extracted in a real-time fashion by recording and processing the 3D video in parallel with a rolling time window. The reference HR was measured with an ECG-based Polar H10 chest strap (± 1 BPM resolution) (Polar Electro Oy, Kempele, Finland), while the reference RR was measured through a NUL236 Respiration Monitor Belt (± 3 BPM resolution based on the recording duration) (Neulog, Israel). The brightness level in the cabin was also recorded with a light meter (MT912, URCERI).

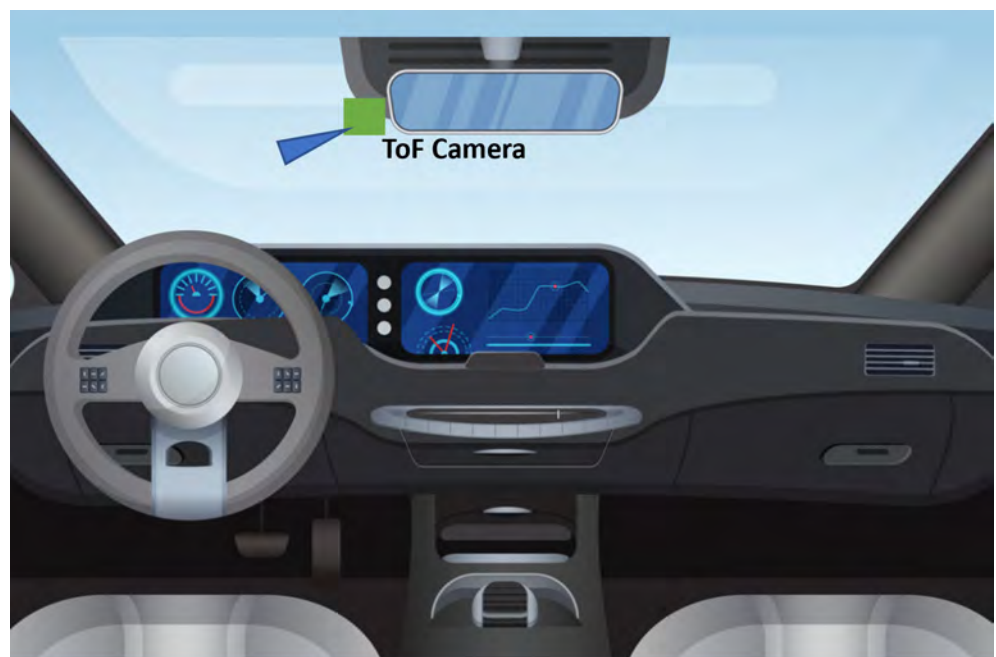


Figure 1. Experimental configuration for HR/RR measurements in a vehicle with a ToF camera.

3.2. HR and RR Extraction

Our system extracts HR using both the average grayscale intensity and depth of ROI (Region of Interest)-1 on the right cheek region while measuring RR with just the average depth information of ROI-2 on the chest. Compared to our previous study [28], the ROI-1 is selected based on the need to adapt to more complex scenarios and the mounting angle of the camera. We only select the smooth region on the right cheek so that the signal is less likely to be affected by the facial hair or the uneven shape of the face. To adapt and track ROI-1 and ROI-2 (Figure 2) from various head poses and seating positions that users could potentially take in a car, a more sophisticated face mesh detection algorithm is used [35] to extract 468 facial landmarks (compared to the 66 landmarks used in [28]) from the user's face. The landmarks that encircle the right cheek/nose area are selected as the corner points of ROI-1. After the facial landmarks are extracted, we apply a pose estimator to find the shoulder of the user and define the ROI-2 on the chest using the relative positive between the jawline and the shoulder. When the user sits 50 cm from the ToF camera, ROI-1 contains roughly 500 pixels while ROI-2 contains roughly 8000 pixels.



Figure 2. Position of both ROI-1 and ROI-2 as seen by the ToF camera. The grayscale intensity and depth information from ROI-1 are used for HR measurement, and the depth information from ROI-2 is used for RR measurements.

Figure 3a shows a typical intensity as well as the depth captured from ROI-1 while the user is driving on the highway. When a localized illuminator (such as VCSEL used in this study) is used, motions from both the driver and the road create artifacts that are several times larger than the underlying HR signal (Figure 3a). If left uncompensated, such artifacts will lead to inaccurate HR measurements (unwanted peaks in the frequency domain) (Figure 3b).

In this study, HR is extracted in three major steps (Figure 4): 1. Extract average grayscale intensity and depth information from ROI-1; 2. compensate for motion-induced intensity artifacts using depth information; 3. use depth information to assess the HR signal quality in different sections during the 20 s video and extract HR from the processed signal. Because the user in a moving vehicle will encounter more random motion, compared to the previous study [28], step 3 is added to further differentiate signal sections with more motion interference from signal sections with less motion interference.

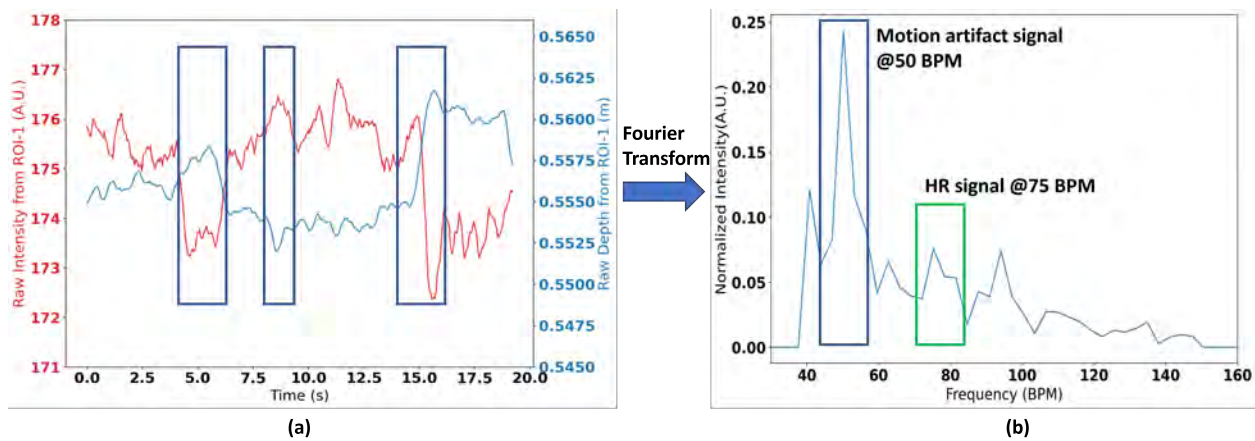


Figure 3. (a) Raw intensity (red) and depth (blue) captured by the ToF camera. Depth change induced by motion (blue rectangle) is inversely proportional to the intensity, which is a key source of erroneous HR reading. (b) When motion artifacts are present, they create artifact frequency components that overwhelm the actual HR signal.

In step 1, before extracting grayscale information and depth information from each frame, we first remove the background in each frame by only selecting pixels that are 40 to 70 cm from the camera. The removal of the background in each frame eliminates the interference from the passengers in the backseats. After the background removal, we apply the face mesh detection algorithm to find ROI-1 and average the grayscale intensity and the depth across every pixel in ROI-1 [35]. While the face mesh is being obtained, we also determine the head orientation (yaw, pitch, and roll) of the user in each frame using a head pose estimation algorithm [36]. If more than 20% of the frames show excessive head orientation (yaw and pitch deviate more than 10 degrees from head orientation at start), we discard this measurement, as the PPG signal can be heavily polluted by head rotation and our depth-based motion compensation method is less effective against intensity changes caused by large head rotation motion [18,28].

If less than 20% of total frames exhibit excessive head rotation in the video stream, the system moves on to step 2 for the removal of motion artifacts. When an active illumination source such as VCSEL lasers are used, the amount of light that is back scattered to the camera is nonlinearly correlated to the distance between the camera and the participants' face. The voluntary motion from the driver or the involuntary motion caused by uneven roads changes the distance between the driver's face and the light source, resulting in intensity artifacts that could corrupt the HR measurements. Because the ToF camera measures both intensity and depth information, the depth information can be used to compensate for the motion-induced artifacts. The depth information obtained from ROI-1, D_{raw} should contain no heart rate component (heartbeats do not cause depth change detectable by ToF camera), while the raw grayscale information from ROI-1, I_{raw} contains both the heartbeat induced intensity change as well as the motion-induced intensity change. After the motion artifacts are compensated, the compensated signal I_{comp} should have minimum correlation with D_{raw} . I_{comp} is calculated using Equation (1), where a and b are the coefficients. We iterated through a range of a and b to find the minimum correlation between D_{raw} and $\frac{I_{raw}}{a * (D_{raw})^{-b}}$. In this study, the range of the nonlinear coefficient is set between 0.2 to 5 while the linear coefficient is set as 1 for all measurements taken in this study. More details on the motion compensation method can be found in [28].

$$I_{comp} = \frac{I_{raw}}{a * (D_{raw})^{-b}}; \text{ where } a, b = \underset{a, b}{\operatorname{argmin}} \operatorname{Correlation}\left(\frac{I_{raw}}{a * (D_{raw})^{-b}}, D_{raw}\right) \quad (1)$$

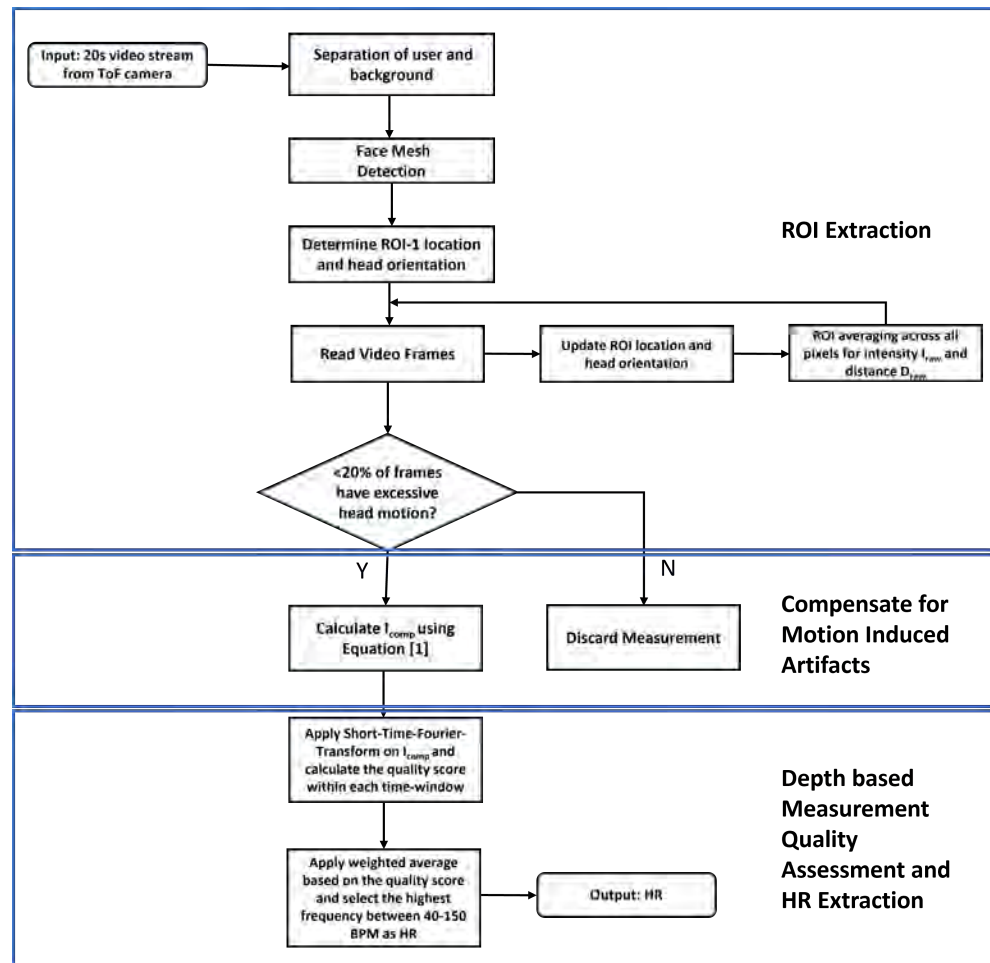


Figure 4. Processes to extract HR while driving using depth and grayscale information from a ToF camera.

After the motion artifacts have been compensated for, we then assess the quality of different sections within the 20 s video using both the depth information and Short Time Fourier Transform (STFT). The goal of the quality assessment is to find the sections of videos that either suffer lower interference from motion or contain stronger HR signal. Ideally, for clean HR measurements without artifacts and noise, the measured signal in the frequency domain will contain one peak in the frequency domain (which we select as HR). However, in real-life measurements, various motion artifacts/noise create additional frequency peaks in the frequency domain. Even though we try to compensate for motion-induced artifacts in step 2, some large motion or head rotation cannot be effectively compensated [28]. These large motions can come from the driver turning their head or from driving across potholes on the road. If those types of motion are large enough, the amplitude of artifact frequency peak can be higher than the HR frequency component, creating erroneous results. Therefore, we use the depth information to determine if such artifacts could exist in a particular section of the recorded signal. The quality assessment is implemented with following steps: 1. The 20 s measurements are split into eleven 10 s windows with a 9 s overlap between two consecutive windows. 2. We perform Fourier transform to the 10 s window and find the amplitudes of highest peak (A_{1st}) and the second highest peak (A_{2nd}) in the frequency domain. 3. We calculate a motion score of this 10 s window using Equation (2), where “var” stands for the variance and “depth” represent the average depth measured by the ToF camera in ROI-1. 4. The final spectrum is calculated as the weighted average of the spectrum of every window with the inverse of the motion score being the weights (Equation (3)). $X_i(f)$ is the

spectral content of each time window and n is the total number of windows (in this study $n = 11$).

$$\text{motion score} = \frac{(\text{var}(\text{depth}(t)) - \text{mean}(\text{depth}(t)))}{A_{1st} / A_{2nd}} \quad (2)$$

$$X_{final}(f) = \sum_0^n X_i(f) \times (\text{motion score}_i)^{-1} \quad (3)$$

Lower the motion score, lower the possibility of motion interference within that section and vice versa. Therefore, by applying such weighted average, we reward sections with low motion scores, during which either the user is moving less, or a clear highest peak can be found in the frequency domain. After the averaged spectrum is calculated, we select the highest frequency component between 40 and 150 BPM as the measured HR.

Compared to HR measurements, RR measurements only rely on the depth information from ROI-2. In this study, we extract RR in three steps (Figure 5): First, with the pose estimator, we define ROI-2 using the relative position between the chin and the shoulder of the participants. The location of ROI-2 is then tracked in every frame in the 20 s video. Then, we calculate the average depth across all pixels within ROI-2. Finally, we apply a bandpass filter (5 to 30 BPM) and Fourier transform to the extracted depth signal from ROI-2, and the highest peak in the frequency domain is then chosen as the RR value.

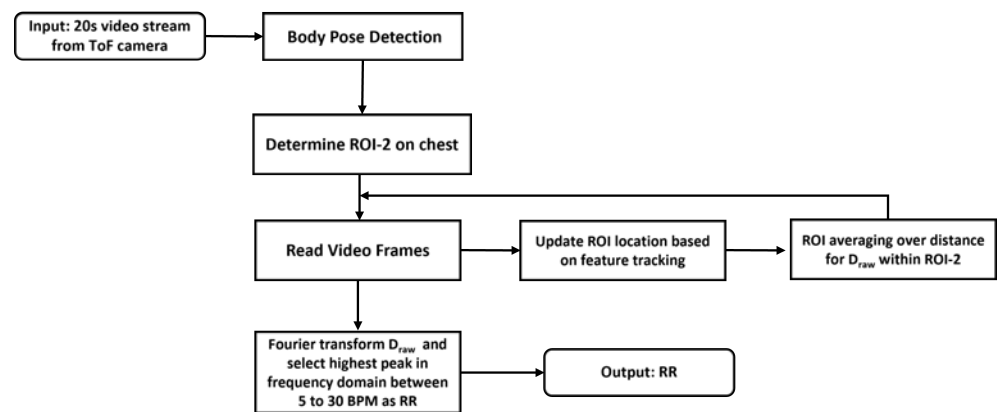


Figure 5. Processes to extract RR while driving using depth and grayscale information from a ToF camera.

3.3. In-Vehicle Testing of the System

To test the system performance and to understand the influences of motion artifacts from various sources, we conducted a series of in-vehicle tests, both on the highway as well as local roads. Our cascaded tests comprise 4 scenarios. Between each scenario, we add one extra source of interference and bring the test scenario closer to a realistic use case. In the first scenario, participants are asked to sit in the lab with their faces toward the camera. In scenario two, we change the viewing angle of the ToF camera and add the vibration from the engine to the test. Participants are asked to sit in a parked car with the engine running. In the third scenario, participants sit on the passenger seat while being driven on a highway/local road (participants asked to stay still). Compared to the previous scenario, we add the “passive motion artifacts” such as driving on an uneven road as well as making turns. Finally, in the last scenario, we add the voluntary motion from the driver such as checking the traffic conditions and adjusting their seating position while driving. Participants are asked to drive the car themselves on a highway/local road, mimicking the real-life operation scenario for which the system intended to be used. In total, 6 of the study team members participated in scenarios 1–3, while only 5 participated in scenario 4. During the measurement of each team member, roughly 25 measurements of HR/RR were taken (actual number of samples depends on time the participants spend to drive the designated route), leading to more than 125 measurements for each scenario. In terms of

RR measurements, we measure the RR of the participants as they drive on the highway while wearing the respiration belt. Table 2 shows the details of the cascaded driving test. To protect the privacy of the study team members, all data recorded in this study are only stored in local hard disk drives, and only study team members in charge of data analysis have access to the hard disk drive.

Table 2. Testing condition for each of the scenario tested.

Parameters Measured	Test Condition	Number of Participants	Number of Measurements	Total Miles Driven
HR	Highway-Passenger	6	141	>70
HR	Highway-Driver	5	151	>60
HR	Local-Passenger	6	141	>15
HR	Local-Driver	5	143	>12
HR	Highway-Driver	5	135	>60

4. Experiment Results

4.1. HR and RR Measurements in Bright and Dark Lighting Conditions

One advantage of using the ToF camera is that the 850 nm active illumination provided by the VCSEL allows the system to operate independent of the ambient light conditions. Such benefits can be attributed to both the bandpass filter as well as the underlying lock-in detection schematics of the indirect ToF camera [34,37]. Figure 6 shows the images seen by the ToF camera in both bright and dark environments. In the bright environment, the brightness level in the cabin exceeds 1800 Lux, while in the dark environment, the brightness level is less than 5 Lux.



Figure 6. The user's face seen by an RGB camera (a) in both bright (top) and dark (bottom) conditions versus the user's face seen by the ToF camera in both bright (top) and dark (bottom) conditions.

In both bright and dark environments, the images seen by the ToF camera are very similar, and HR can be extracted in both cases (Figure 7).

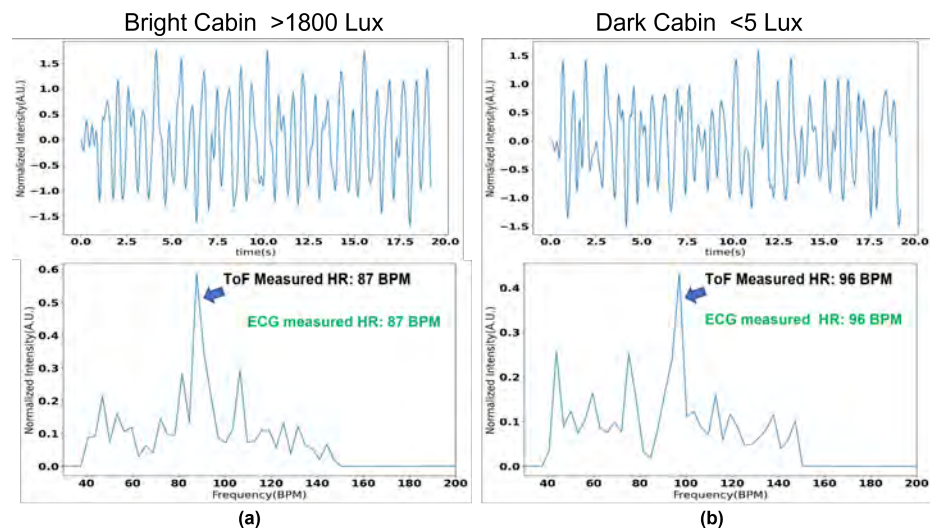


Figure 7. Time (Top) and frequency (Bottom) domain of the motion-compensated HR signal in both bright (a) and dark (b) environments.

4.2. On-the-Road Testing of HR and RR

Figure 8 shows a waterfall chart of the success rate tested under each of the scenarios (Table 2). An HR measurement from our CVSM system is defined as ‘successful’ if the measurement is within 10% of the HR measured by our reference device, a Polar H10 chest strap. The success rates of HR measurements on highway conditions outperform the success rates on the local road conditions in all cases. Such results are expected since the local road represents a more complex driving scenario. Compared to a highway, local roads have more potholes and busier traffic, which leads to more frequent active and passive motions of the driver. With the adding of more potential sources of motion artifacts in each consecutive scenario, we can see a degradation of HR measurement performance by different amounts. In the case of the highway scenario, the success rate drops from 90% when measured in the lab to 71.9% when the driver is driving the vehicle on the highway, with the most significant performance drop occurring when we add the voluntary head motions from the driver (i.e., the case when driver drives the car him/herself). As a comparison, the success rate drops to a lower 56% when the driver is driving on the local road. However, the most significant decrease in performances for the local road scenarios seems to arise from the road surface and traffic conditions rather than the driver’s active motion (the success rate drops most when we started to measure passengers in a moving vehicle). In reality, the highway scenario will be the more important use case for the CVSM system, due to higher speed as well as the longer time of driving, leading to drowsiness and potentially more catastrophic accidents [38]. We will discuss in detail the system performance under each in-vehicle scenario in the following paragraphs.

Figure 9 compares the distribution of error rates of the HR measurements with or without depth-based motion compensation when the user is sitting in a parked car with the engine running. In the following analysis, “motion-compensated” refers to HR measurements utilizing both grayscale intensity and depth information for motion compensation. The “uncompensated” measurements refer to the HR measurement using the same grayscale intensity as the “compensated” HR measurements but not using the depth information for motion compensation (as is the case of a normal 2D camera). With the depth-based motion compensation, the success rate is 86% while the success rate drops to 36.5% when the motion compensation is not used. Even though the participant is sitting still in their seat, there still exists involuntary motion, which can obfuscate the HR signal if left uncompensated.

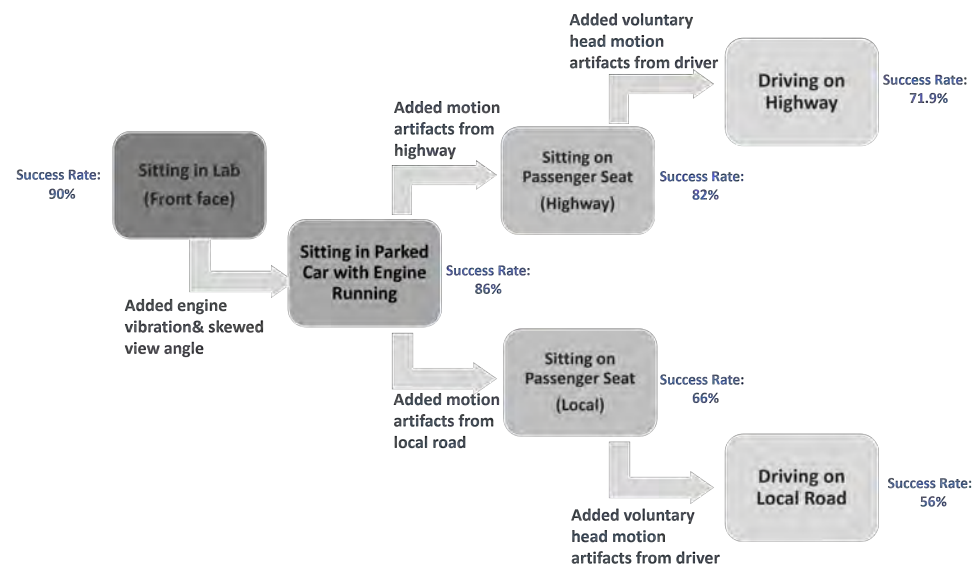


Figure 8. Waterfall diagram of the success rate of the HR measurements under different testing scenarios on highways and local roads.

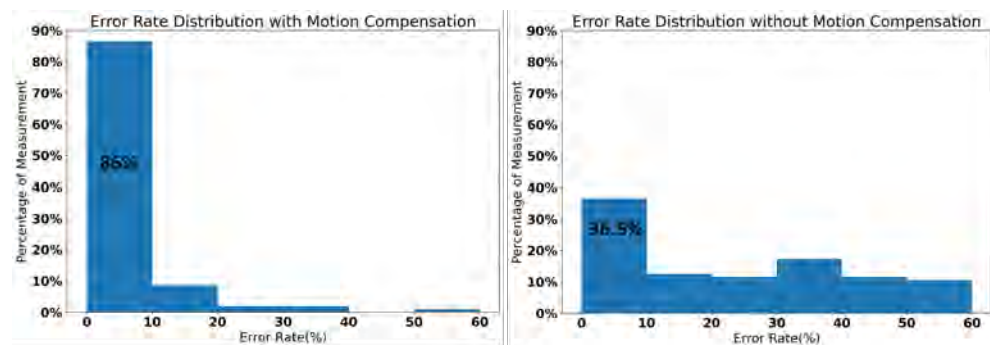


Figure 9. HR measurement error rates when participants sit in a parked vehicle with motion compensation (left) and without motion compensation (right).

Figure 10 shows the Bland–Altman plot of the HR measurements in the parked vehicle with the engine running. With depth-based motion compensation enabled, the mean deviation of the HR measurements from is -2.9 BPM and 95% of the measurement deviations fall between ± 19.6 BPM of the mean deviation. When depth-based compensation is not used, we observe a much higher negative bias of -16.3 BPM with 95% of deviation falling between ± 35 BPM of the mean. Such larger bias is usually caused by the lower frequency motion artifacts being mistakenly recognized as HR.

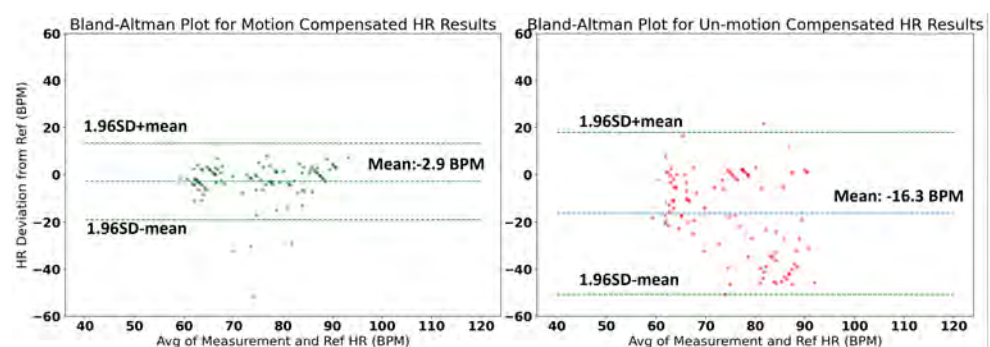


Figure 10. Bland–Altman plot of HR measurements when participants sit in a parked vehicle with motion compensation (left) and without motion compensation (right).

Testing on Highway:

Once on the road, the car motion from the road bumpiness starts to affect the performance of HR measurements. Figure 11 shows the histogram of the error rates when the participants are sitting in the passenger seat while the car is driving on the highway. The participants are asked to not move their heads during the measurements to suppress active head motion. Since there is less road roughness and fewer distractions from traffic signs/lights or pedestrians on the highway, we did not observe significant degradation of the HR measurement success rate. Figure 12 shows the Bland–Altman plot for the HR measurements under the same conditions. The mean deviation for the motion-compensated case is 0.13 BPM (absolute mean deviation at 5.8 BPM) and the 95% of deviation within ± 22 BPM from the average deviation. When motion compensation is not used, the success rate drops to 34% while the mean deviation increases to -19 BPM (absolute mean deviation at 20.5 BPM).

Figures 13 and 14 show the error rate distribution as well as the Bland–Altman plot for the HR measurements when the participants drive the vehicle on the highway. Because the participants need to turn their heads occasionally to check traffic or change lanes, these head motions translate into large motion artifacts that add further complexity to HR measurements. Even though such motions only last for a few seconds, they could still induce motion artifacts that are large enough to disrupt the HR measurements. With the depth-based motion compensation enabled, the HR measurements success rate is 71.9 (absolute mean deviation: 7.7 BPM), while the success rate with no compensation is only 13.6 (absolute mean deviation: 25 BPM). From Figure 14, we can see more of the wrong HR readings clustered at the bottom of the graph, which are the results of the additional motion artifacts from the driver's motion. We will discuss potential ways of mitigating such outliers later in the discussion section.

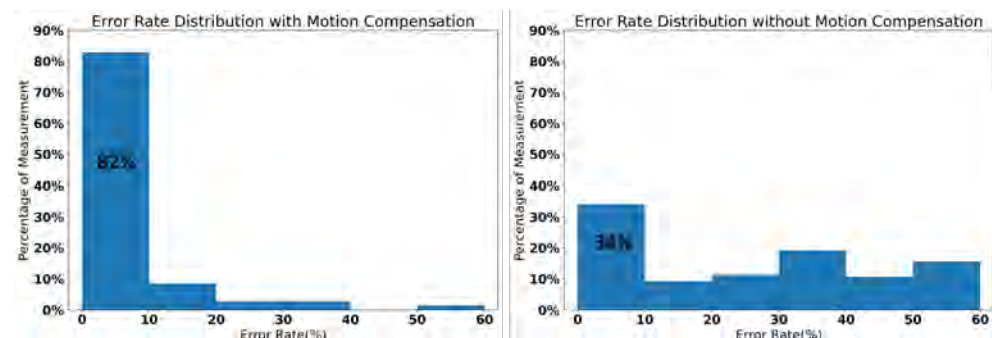


Figure 11. HR measurement error rate when participants sit in the passenger seat on the highway with motion compensation (left) and without motion compensation (right).

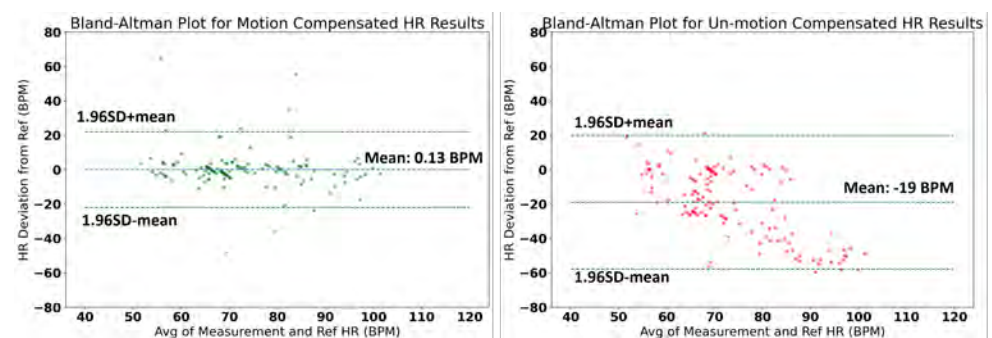


Figure 12. Bland–Altman plot of HR measurements when participants sit in the passenger seat on the highway with motion compensation (left) and without motion compensation (right).

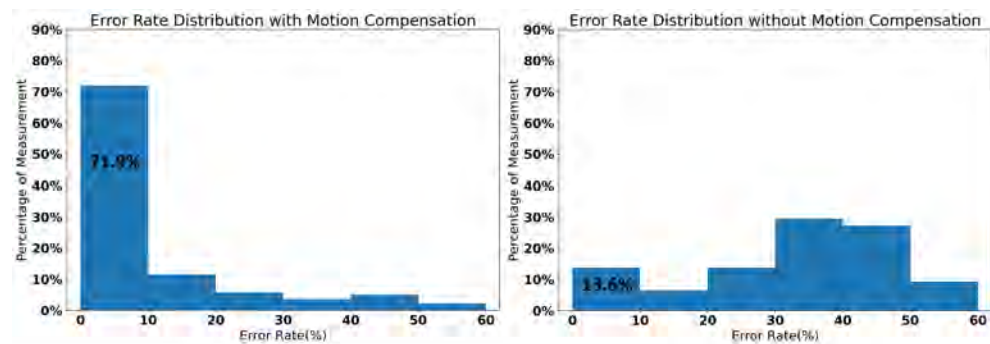


Figure 13. HR measurements error rate when participants drive on the highway with motion compensation (left) and without motion compensation (right).

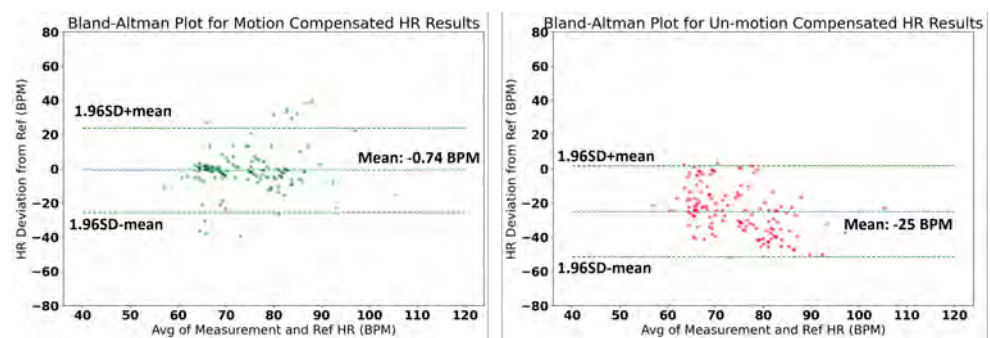


Figure 14. Bland–Altman plot of HR measurements when participants drive on the highway with motion compensation (left) and without motion compensation (right).

Testing on Local Road:

When testing the CVSM system on the local road, we observe the same performance improvement brought by the depth-based motion compensation. Figures 15 and 16 show the error rate as well as the Bland–Altman plot for HR measurements for the local road and participants in the passenger seat scenario. Because both the road surface condition (roughness, potholes) and the traffic (number of stop signs, number of cars) are much worse compared to the highway scenario, the success rate degradation for the HR measurements is also more severe. With the participants sitting on the passenger seat, the success rate is 66% (absolute mean deviation 8.4 BPM) with motion compensation and only 22% (absolute mean deviation 22 BPM) for HR measurements without motion compensation.

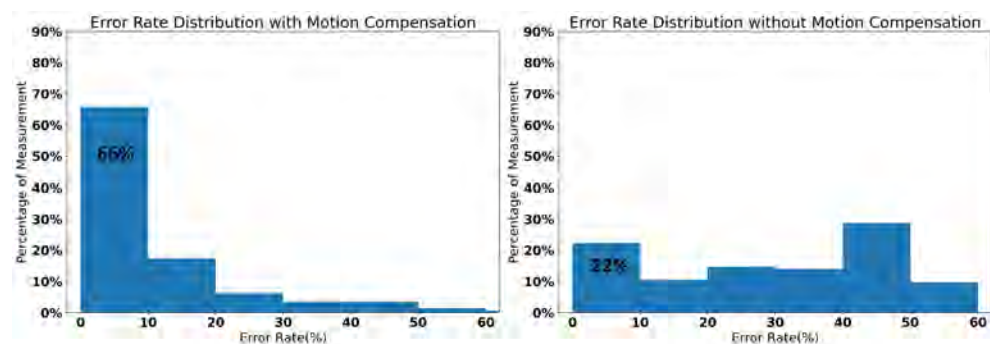


Figure 15. HR measurement error rate when participants sit in the passenger seat on local road conditions with motion compensation (left) and without motion compensation (right).

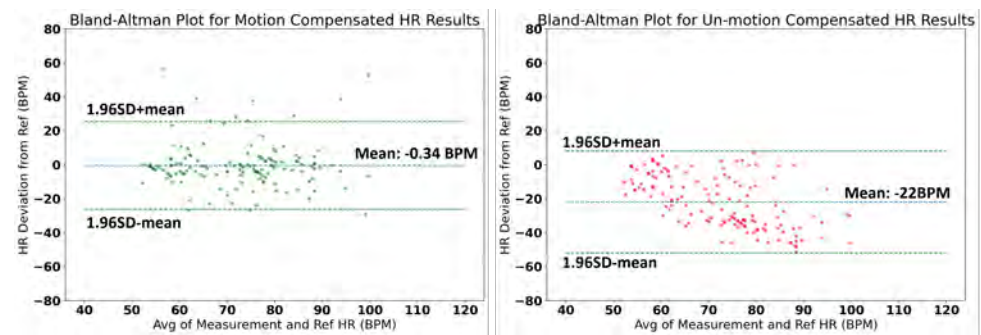


Figure 16. Bland–Altman plot of HR measurements when participants sit in the passenger seat on local road conditions with motion compensation (left) and without motion compensation (right).

When the participants move on to drive on the local road, we see similar HR measurement degradation (10%) as the highway scenario (Figures 17 and 18). With motion compensation, the success rate of the HR measurements is 56% (absolute mean deviation 10.5 BPM) when depth-based motion compensation is used and only 12% (absolute mean deviation 25 BPM) when the depth-based motion compensation is not used.

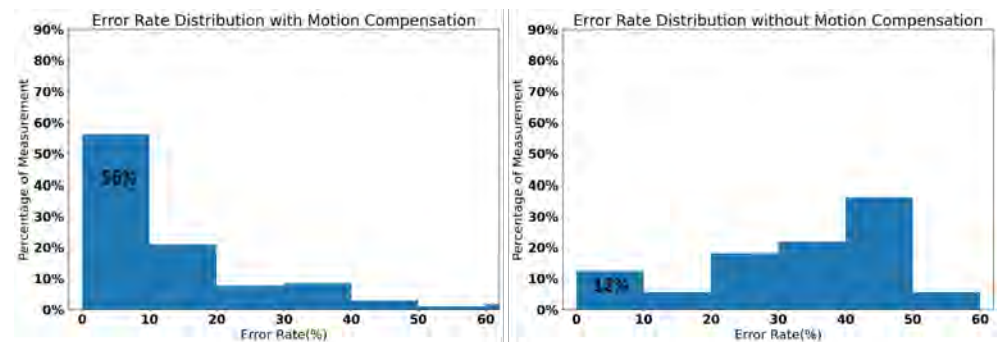


Figure 17. HR measurement error rate when participants drive on a local road with motion compensation (left) and without motion compensation (right).

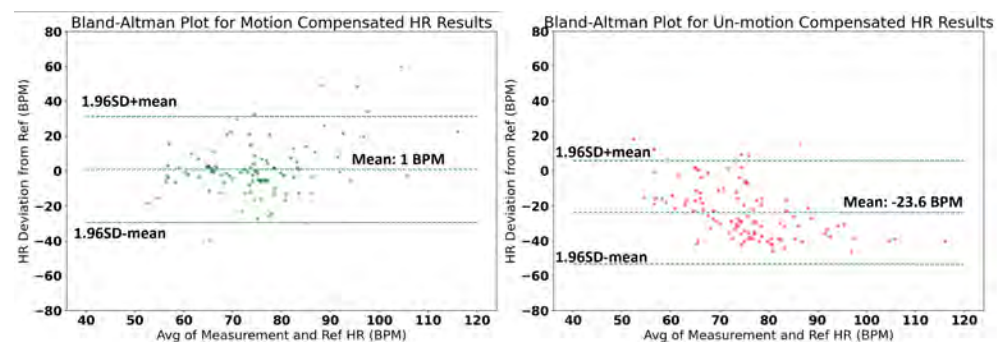


Figure 18. Bland–Altman plot of HR measurements when participants drive on a local road with motion compensation (left) and without motion compensation (right).

RR Measurements on Highway:

In this study, RR is also measured on participants while driving on the highway. Figure 19 shows typical respiration motion acquired on the chest region compared with the reference chest belt on the highway. The ToF camera measures a typical chest wall depth change of ± 2 mm, with the pattern matching the pressure change measured by the reference respiration belt.

Figure 20 shows the distribution of the deviation of RR measurements from the reference respiration monitoring belt. The mean deviation from the reference reading is found to be -1.4 BPM with a total of 71% of measurements falling within the error of

± 3 BPM. It should be noted that since RR is extracted through the Fourier transform of a 20 s measurement, the reference RR measurements are limited to a resolution of 3 BPM. In our measurements, we could sometimes see some extreme outliers with errors as large as -10 BPM. These errors could be attributed to events such as driving over potholes or the participants adjusting themselves for a more comfortable position while driving.

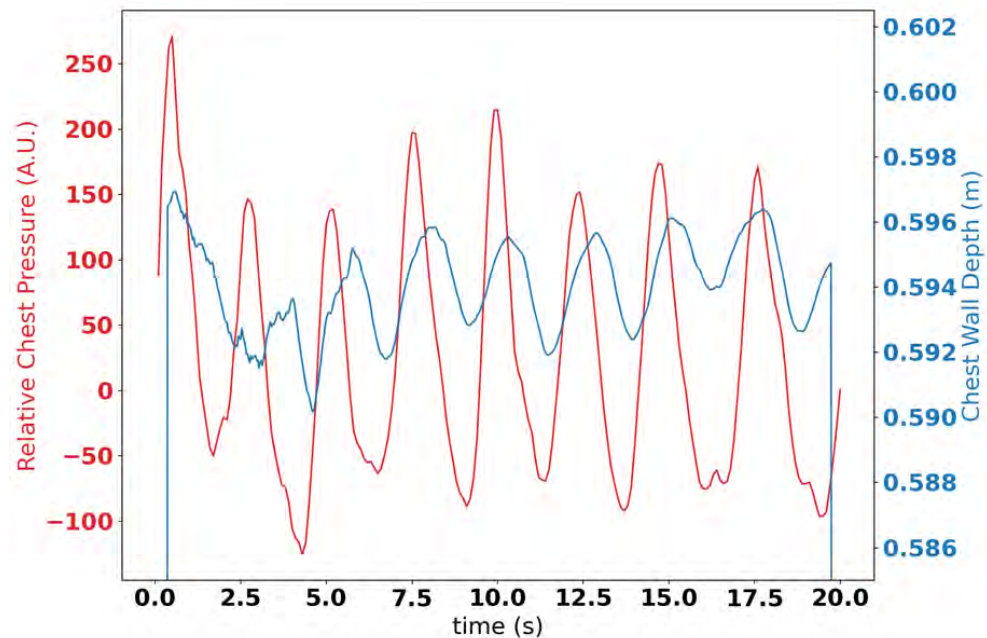


Figure 19. Typical RR motion measured by a ToF camera (blue), and the RR measured by the reference respiration belt (red).

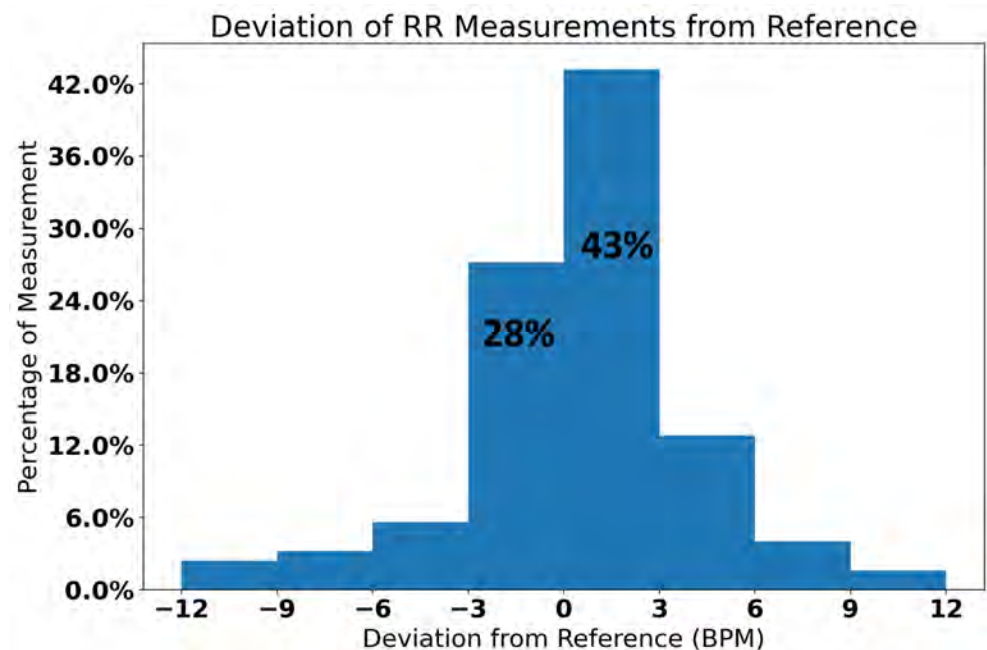


Figure 20. Percentage of ToF-based RR measurements versus deviation from the reference respiration belt.

5. Discussion

In the previous sections, we have shown the benefits of using both the active illumination and depth information from an indirect ToF camera for HR and RR measurement in a vehicle. However, because of the complex environments in a vehicle cabin, it is worth discussing the potential limitations of the current methods and exploring potential approaches that could be used to improve the robustness and accuracy of our CVSM system.

Potential Limitations and Improvements for HR measurements:

To obtain HR reliably, the system needs to keep track of the same ROI accurately across all frames. Facial expression or facial occlusions could all interfere with the acquisition of ROIs on the face. If the vertices of the ROI are jittering from frame to frame, the system could mistake the ROI jittering frequency as the HR. For example, Figure 21 shows an example where the blinking of the participant interferes with the HR measurement. The participant is asked to blink his eye at the same frequency of an external metronome at 115 BPM (selected to be out of the normal range of the participant's resting HR). The periodic bright reflection from the eye causes the ROI-1 locations to jitter at the same frequency, which introduces an artifact frequency that overwhelms the true HR frequency component. Since this type of error rises from the jittering of ROI locations, they could potentially be mitigated by using a face tracking algorithm that is less sensitive to external interference. Furthermore, one possible solution is to check the grayscale intensity change against the position change of the ROIs to make sure the dominating frequency peak in the frequency domain is not from the jittering of ROIs [11].

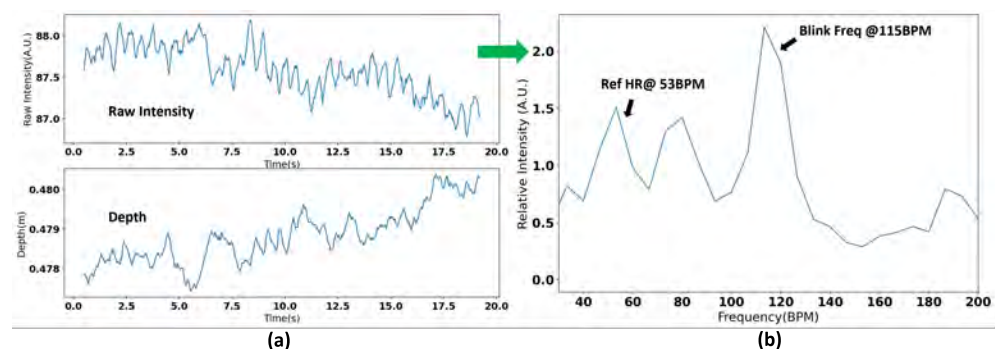


Figure 21. Periodic reflection from participant blinking could affect the location of the ROI-1 as a result of jitter. The jittering can lead to additional intensity artifacts that will not be captured by the (a) depth signal, leading to erroneous HR reading from the (b) frequency domain.

In addition to the interference of the ROI acquisition, subtle motions such as facial expression change and talking could also limit the performance of the CVSM system because these motions could introduce motion artifacts without causing detectable depth changes from the ROI [28]. These subtle artifacts that are confined to localized areas on the face could be alleviated with methods such as the sparse frequency estimation and will be discussed in the later sections.

Moreover, a unique challenge to monitoring HR in a vehicle is the sudden motions from either the driver/passenger or the road conditions. In order to capture these motions accurately, it might be beneficial to use a camera with a higher frame rate. To investigate the potential benefits of using a higher frame rate camera, we measured HR on three participants while driving on the highway at 60 FPS (compared with 30 FPS in the experiment section). The acquired 60 FPS videos are then down-sampled (only using every other frame in the original video) to 30 and 15 FPS to compare the performance of HR measurements at different frame rates. Each frame in the 60/30/15 FPS video shares the same integration time and, therefore, the same background noise. Figure 22 shows the error rate distribution for the HR measurements on the highway. With a higher frame rate (60 FPS), the motion across two consecutive frames is smaller, which makes tracking the

ROI easier when sudden or large motions are present. In our preliminary test, the original 60 FPS videos result in a success rate of 76% while the down-sampled 30 FPS videos leads to a reduced success rate of 68%. When the frame rate is further reduced to 15 FPS, a further degradation of the success rate to 58.7% is observed.

The given benefits of a higher frame rate could be attributed to the smaller motion across two consecutive frames. For example, if we assume the user is sitting at 50 cm away from the camera, for the ToF camera with 149-degree diagonal FOV and half-inch imaging sensor size (diagonal), we could roughly calculate how many pixels the ROI landmarks will shift across two different frames using Equation (4), where D_{user} is the distance between the camera and the driver, θ is the half angle of the horizontal field of view, L is the horizontal size of the imaging sensor, μ is the pitch of the pixel, v is the velocity of the motion and FPS is the frame rate of the camera.

$$N_{pixel} = \text{ceiling}\left(\frac{L}{D_{user} * 2\mu * \tan(\theta)} * \frac{v}{FPS}\right) \quad (4)$$

The above equation makes the rough estimation that the image fills the entire sensor, and the active area of the pixel is small. In this case, we consider the head of the user is moving at 50 cm/s (typical for a fast head turning), and Table 3 shows the number of pixels the ROI landmark would transit at different frame rates.

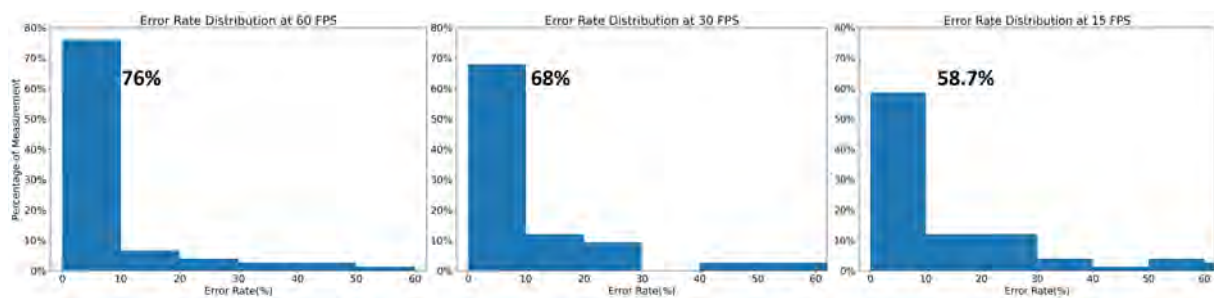


Figure 22. HR measurement error rate when participants drive on the highway recorded at 60 FPS (left), 30 FPS (middle) and 15 FPS (right).

Table 3. Success rate and number of pixels transitioned across two consecutive frames at different frame rates.

Frame Rate (FPS)	15	30	60	105
Npixel	7	4	2	1
Success Rate	58.7%	68%	76%	N/A

The lower the frame rate, the larger the motion (more pixels transitioned) seen by the face tracking algorithm between two frames and, therefore, a higher possibility for the face tracker to output erroneous estimations of facial landmark positions [11]. The erroneous landmarks can cause the detected ROI-1 to change over time, leading to time-varying intensity changes in the measured signal. When the frame rate reaches 105 fps, the motion across two consecutive frames will not move more than 1 pixel across different frames. Of course, a higher frame rate does not solve all the challenges of HR measurements in NIR, but it should alleviate the burden of the motion compensation as well as the face tracking algorithm. In the future, with ToF cameras of higher resolution and the frame rate being developed [39], the robustness of tracking fast motion could be increased and reduce the effects of the driver's active motion on the HR measurements.

Another potential improvement for the CVSM system proposed in this study is to combine the depth-based motion compensation with other post-processing methods to improve the reliability of the HR measurements. For example, in the 2020 study from Nowara et al. [11], a sparse matrix estimation method is used to further reject noise from

the remote PPG signal. The sparse matrix estimation method makes the assumption that the HR signal is common across different ROIs on the face while the noises/artifacts will differ within different ROIs. Therefore, by extracting the common signal across multiple ROIs, one can differentiate the HR signal from the external interferences. However, such method may not be directly applicable to the CVSM demonstrated in this study because the localized VCSEL source causes intensity from all regions on the face to change in a similar pattern against the motion from the user. If different ROIs on the face share the same artifacts/noise, the sparse frequency matrix method from [11] could still pick the artifacts/noises as the HR signal. For example, Figure 23 shows an example of an erroneous HR measurement using the sparse matrix estimation method. We extract intensity signals from three ROIs from cheek, nose and forehead, respectively, and apply the Autosparse PPG method from [9] to the extracted signals to find the common frequency components across these three ROIs. Figure 23a shows that the signals from the cheek, nose and forehead show similar patterns of artifacts. Therefore, even though the sparse frequency estimation method correctly identified the HR at 62 BPM, it also included the artifacts that are common to the three ROIs in the extracted frequencies (Figure 23b).

One potential solution to this challenge is to also utilize the fact that HR is usually slowly varying in time. One could potentially find the sparse frequency that is common across both multi regions on the face as well as within several consecutive HR measurements in time. By combining both time and multiple ROIs on the face, such method could hopefully further reject noises from the compensated HR signal and improve the signal reliability.

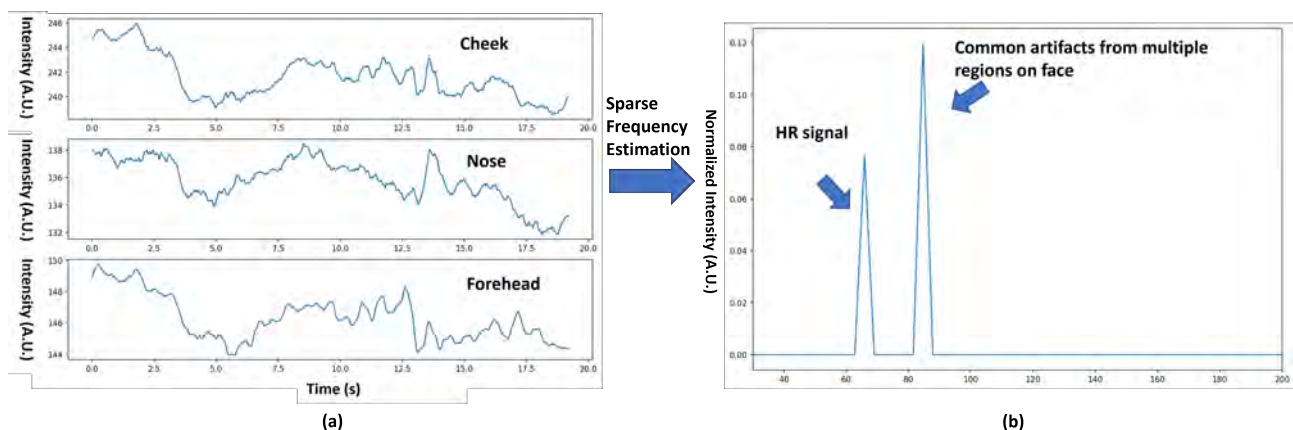


Figure 23. (a) Grayscale signal from three different ROIs on the face captured with the VCSEL illuminated ToF camera, showing similar patterns of artifacts among all three ROIs. (b) Extracted sparse frequency component using the method in [11], showing both HR and artifact frequency components.

Finally, in this study, the open-source face landmark identification and tracking algorithms [35] are trained mostly with RGB images in various environments and, therefore, may not be optimized for videos that are taken in the NIR spectrum. In the future, commercially available face tracking algorithms (i.e., Visage, Seeing Machine, Emotion3D, etc.) specifically designed for facial landmark tracking in NIR videos in driving environments might be used to improve the robustness against sudden motion or reflection glares. With the help of a more optimized face tracking algorithm, one could alleviate these problems mentioned in the previous discussion, which in turn improve the reliability and accuracy of the HR measurements in the vehicle.

Potential Limitations and Improvements for RR measurements:

Regarding using a ToF camera for RR measurements, a natural question to ask is how would different types of clothing (that could block or dampen the motion of the chest wall) affect the performance of RR measurement using the ToF camera. Even though the ToF

camera offers the capability of directly measuring RR similar to a millimeter-wave radar, one drawback is that the 850 nm light cannot penetrate clothes compared to the millimeter wave used in radar. To test the effect of clothing, we measure the depth change on ROI-2 while the participants wear clothes of different degrees of thickness (Figure 24). Fortunately, in all cases, even including the case of a thick down jacket, we were able to detect the chest wall motion from respiration in various commonly worn clothes (Figure 25). Only when the participant is purposefully wearing multi-layer “puffy” clothes with air gaps in between did we start to see the errors of RR measurements increasing, as the puffy clothing can block the chest motion from the camera (Table 4). Finally, it should still be noted that the typical amplitude of the chest motion is only ± 2 mm, and the RR is derived solely from the depth data. In the existence of large external motion, it is still possible for the excessive motion to corrupt the RR measurements. Such limitation could be mitigated by adding complementary RR measurements methods such as RR extraction from HR variability or through extra sensors such as thermal cameras that are co-equipped in the vehicle [20,21,29].



Figure 24. Types of clothes worn by participants in this study. Participants were asked to sit in a parked car with the engine running and breathe naturally during the measurements.

Table 4. Absolute mean error of RR measurement when the participant is wearing different types of clothing.

Clothing Type	T-Shirt	Sweater	Rain Jacket	Wool Coat	Down Jacket	Multi-Layer *
Absolute Mean Error (BPM)	1.77	1.56	1.51	1.02	1.75	3.57

* “Multi-layer” clothing refers to wearing several different clothes with air gaps in between.

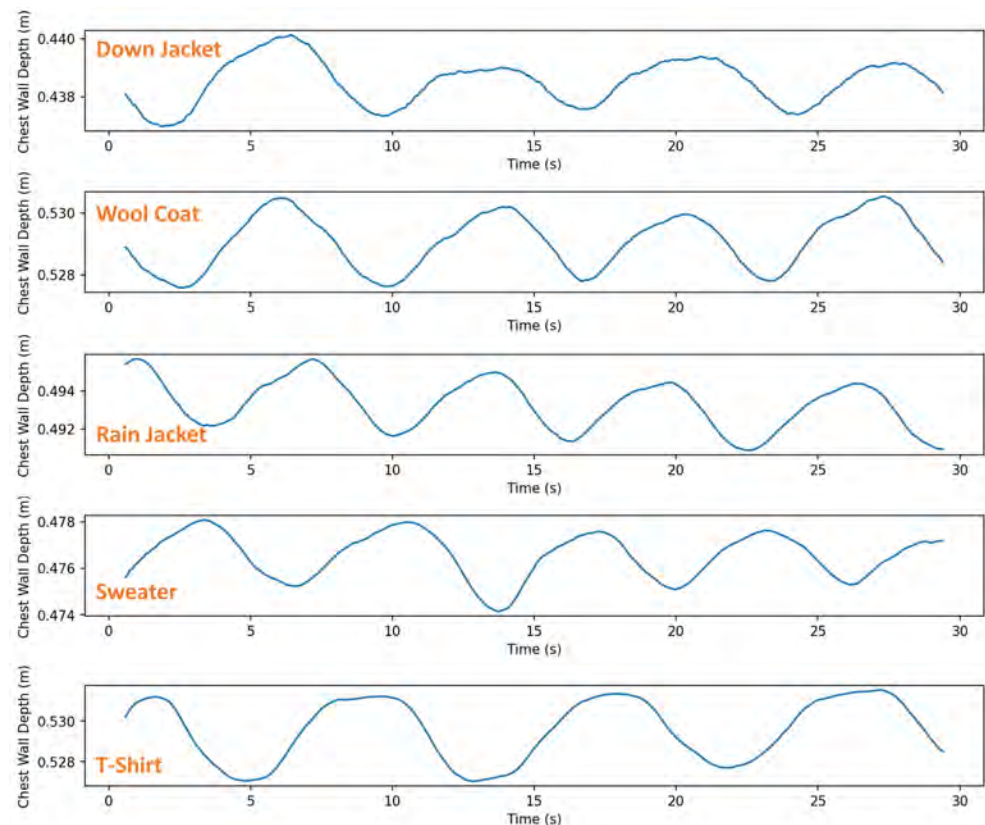


Figure 25. Respiration motion on chest measured by the ToF camera while the participants were wearing different types of clothing. Clear respiration patterns can be seen in all cases.

6. Summary

In this study, we first present a non-distracting CVSM system that can monitor the HR and RR for both passengers and drivers in a vehicle using an indirect ToF camera. The ToF camera allows us to mitigate two major challenges for measuring HR in a moving vehicle, namely, varying illumination level and motion artifacts [11]. Compared to a conventional 2D camera, the ToF camera is inherently robust against the ambient light fluctuation that is commonly encountered in a moving vehicle. Moreover, in this study, the depth information is first used to compensate for the motion-induced artifact and then to assess the quality of the HR signal within different sections of the HR measurement (i.e., within the 20 s video). Additionally, with the depth information from the ToF camera, the CVSM system can measure RR directly from chest wall motion.

We conduct a series of on-the-road testing to evaluate the performance of the CVSM system in realistic operation environments. We first show that, because of the 850 nm active illumination, the system can measure HR independent of ambient light conditions. We then show that the depth-based motion assessment/compensation greatly improves the HR measurement success rate. When the user is sitting in the parked vehicle with the engine running, the CVSM system can measure HR with 86% success rate with motion compensation and only 36.5% success rate when only the 2D grayscale information is used. Such improvement becomes more evident in more complex driving scenarios. When measuring the passenger while driving on the highway and local roads, the success rate is 82% (34% without compensation) and 66% (22% without compensation), respectively. Ultimately, in the case of measuring users driving the vehicle, the compensated HR measurements achieve a success rate of 71.9% and 56% on highways and local roads, respectively, as compared to just 13.6% and 12% when only the grayscale information is used. Even with motion compensation, the HR measurement on the local road remains the most challenging task since both the road unevenness and the driver's active motion are

more significant in this case. Moreover, we show RR measurement on highway driving scenarios with an average deviation of -1.4 BPM.

In the end, we explore the limitations as well as potential improvements to the CVSM system. The limitations of the HR measurements using the CVSM system could come from effects such as facial expression, sudden motion or even reflection glares from the driver's eye. These limitations could potentially be alleviated using a higher frame rate ToF camera or using other post-processing algorithms to further reject these intensity artifacts. As for the RR measurement, even though the 850nm illumination does not penetrate as deep as a millimeter-wave radar, our CVSM system is still able to measure the respiration pattern of the chest wall when the user is wearing various types of clothing. With its robustness against both ambient light and motion artifacts, the ToF-based CVSM system could be a key enabler in low-cost and compact driver/occupant monitoring systems in the vehicle. The physiological information measured by the system could one day empower a safer and more functional ADAS system.

Author Contributions: Conceptualization, K.G., M.N.I., C.D. and T.Z.; methodology, K.G., T.Z. and E.P.; software, K.G., M.H.P., A.V. and S.M.; validation, K.G. and E.P.; data analysis, K.G. and E.P.; data curation, K.G., E.P., M.H.P. and A.D.; writing—original draft preparation, K.G.; writing—review and editing, K.G., E.P. and M.N.I.; visualization, K.G.; supervision, M.N.I.; project administration, M.N.I.; funding acquisition, M.N.I. All authors have read and agreed to the published version of the manuscript.

Funding: This study received partial funding from Omni Sciences. Inc. and from e-HAIL (e-Health and Artificial Intelligence), University of Michigan.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy concerns.

Acknowledgments: The authors of this study thank Ioulia Kolveman, Fred Terry, and Jenna Wiens for their support in data collection, study design, and data analysis.

Conflicts of Interest: Mohammed N. Islam is the founder and chief technology officer of Omni Science Inc.

References

1. Ziebinski, A.; Cupek, R.; Grzechca, D.; Chruszczyk, L. Review of advanced driver assistance systems (ADAS). In *AIP Conference Proceedings*; AIP Publishing LLC.: Melville, NY, USA, 2017; Volume 1906, p. 120002.
2. Ziebinski, A.; Cupek, R.; Erdogan, H.; Waechter, S. A survey of ADAS technologies for the future perspective of sensor fusion. In *International Conference on Computational Collective Intelligence*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 135–146.
3. Béquet, A.J.; Hidalgo-Muñoz, A.R.; Jallais, C. Towards mindless stress regulation in advanced driver assistance systems: A systematic review. *Front. Psychol.* **2020**, *11*, 609124. [CrossRef] [PubMed]
4. Wang, J.; Warnecke, J.M.; Haghi, M.; Deserno, T.M. Unobtrusive health monitoring in private spaces: The smart vehicle. *Sensors* **2020**, *20*, 2442. [CrossRef] [PubMed]
5. Patel, M.; Lal, S.K.; Kavanagh, D.; Rossiter, P. Applying neural network analysis on heart rate variability data to assess driver fatigue. *Expert Syst. Appl.* **2011**, *38*, 7235–7242. [CrossRef]
6. EuroNCAP. Test and Assessment Protocol-Child Presence Detection. 2021. Available online: <https://cdn.euroncap.com/media/64101/euro-ncap-cpd-test-and-assessment-protocol-v10.pdf> (accessed on 27 February 2022).
7. Vicente, J.; Laguna, P.; Bartra, A.; Bailón, R. Drowsiness detection using heart rate variability. *Med Biol. Eng. Comput.* **2016**, *54*, 927–937. [CrossRef] [PubMed]
8. Kiashari, S.E.H.; Nahvi, A.; Homayounfard, A.; Bakhoda, H. Monitoring the variation in driver respiration rate from wakefulness to drowsiness: A non-intrusive method for drowsiness detection using thermal imaging. *J. Sleep Sci.* **2018**, *3*, 1–9.
9. Fujiwara, K.; Abe, E.; Kamata, K.; Nakayama, C.; Suzuki, Y.; Yamakawa, T.; Hiraoka, T.; Kano, M.; Sumi, Y.; Masuda, F.; et al. Heart rate variability-based driver drowsiness detection and its validation with EEG. *IEEE Trans. Biomed. Eng.* **2018**, *66*, 1769–1778. [CrossRef]
10. Leonhardt, S.; Leicht, L.; Teichmann, D. Unobtrusive vital sign monitoring in automotive environments—A review. *Sensors* **2018**, *18*, 3080. [CrossRef]
11. Nowara, E.M.; Marks, T.K.; Mansour, H.; Veeraraghavan, A. Near-infrared imaging photoplethysmography during driving. *IEEE Trans. Intell. Transp. Syst.* **2000**, *23*, 3589–3600. [CrossRef]

12. Huang, P.W.; Wu, B.J.; Wu, B.F. A heart rate monitoring framework for real-world drivers using remote photoplethysmography. *IEEE J. Biomed. Health Inform.* **2020**, *25*, 1397–1408. [\[CrossRef\]](#)
13. Magdalena Nowara, E.; Marks, T.K.; Mansour, H.; Veeraraghavan, A. SparsePPG: Towards driver monitoring using camera-based vital signs estimation in near-infrared. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 1272–1281.
14. Zhang, Q.; Zhou, Y.; Song, S.; Liang, G.; Ni, H. Heart rate extraction based on near-infrared camera: Towards driver state monitoring. *IEEE Access* **2018**, *6*, 33076–33087. [\[CrossRef\]](#)
15. Lee, Y.S.; Pathirana, P.N.; Steinfort, C.L.; Caelli, T. Monitoring and analysis of respiratory patterns using microwave doppler radar. *IEEE J. Transl. Eng. Health Med.* **2014**, *2*, 1–12. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Zhao, P.; Lu, C.X.; Wang, B.; Chen, C.; Xie, L.; Wang, M.; Trigoni, N.; Markham, A. Heart rate sensing with a robot mounted mmwave radar. In Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA), Virtual Confernece, 31 May–31 August 2020; pp. 2812–2818.
17. Verkruysse, W.; Svaasand, L.O.; Nelson, J.S. Remote plethysmographic imaging using ambient light. *Opt. Express* **2008**, *16*, 21434–21445. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Kumar, M.; Veeraraghavan, A.; Sabharwal, A. DistancePPG: Robust non-contact vital signs monitoring using a camera. *Biomed. Opt. Express* **2015**, *6*, 1565–1588. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Wang, W.; Den Brinker, A.C.; Stuijk, S.; De Haan, G. Algorithmic principles of remote PPG. *IEEE Trans. Biomed. Eng.* **2016**, *64*, 1479–1491. [\[CrossRef\]](#)
20. Sun, G.; Nakayama, Y.; Dagdanpurev, S.; Abe, S.; Nishimura, H.; Kirimoto, T.; Matsui, T. Remote sensing of multiple vital signs using a CMOS camera-equipped infrared thermography system and its clinical application in rapidly screening patients with suspected infectious diseases. *Int. J. Infect. Dis.* **2017**, *55*, 113–117. [\[CrossRef\]](#)
21. Negishi, T.; Abe, S.; Matsui, T.; Liu, H.; Kurosawa, M.; Kirimoto, T.; Sun, G. Contactless vital signs measurement system using RGB-thermal image sensors and its clinical screening test on patients with seasonal influenza. *Sensors* **2020**, *20*, 2171. [\[CrossRef\]](#)
22. Nahler, C.; Feldhofer, B.; Ruether, M.; Holweg, G.; Druml, N. Exploring the usage of Time-of-Flight Cameras for contact and remote Photoplethysmography. In Proceedings of the 2018 21st Euromicro Conference on Digital System Design (DSD), Prague, Czech Republic, 29–31 August 2018; pp. 433–441.
23. De Haan, G.; Jeanne, V. Robust pulse rate from chrominance-based rPPG. *IEEE Trans. Biomed. Eng.* **2013**, *60*, 2878–2886. [\[CrossRef\]](#)
24. Cheng, J.; Chen, X.; Xu, L.; Wang, Z.J. Illumination variation-resistant video-based heart rate measurement using joint blind source separation and ensemble empirical mode decomposition. *IEEE J. Biomed. Health Inform.* **2016**, *21*, 1422–1433. [\[CrossRef\]](#)
25. Feng, L.; Po, L.M.; Xu, X.; Li, Y.; Ma, R. Motion-resistant remote imaging photoplethysmography based on the optical properties of skin. *IEEE Trans. Circuits Syst. Video Technol.* **2014**, *25*, 879–891. [\[CrossRef\]](#)
26. van Gastel, M.; Stuijk, S.; de Haan, G. Motion robust remote-PPG in infrared. *IEEE Trans. Biomed. Eng.* **2015**, *62*, 1425–1433. [\[CrossRef\]](#)
27. Yu, S.; Hu, S.; Azorin-Peris, V.; Chambers, J.A.; Zhu, Y.; Greenwald, S.E. Motion-compensated noncontact imaging photoplethysmography to monitor cardiorespiratory status during exercise. *J. Biomed. Opt.* **2011**, *16*, 077010.
28. Guo, K.; Zhai, T.; Pashollari, E.; Varlamos, C.J.; Ahmed, A.; Islam, M.N. Contactless Vital Sign Monitoring System for Heart and Respiratory Rate Measurements with Motion Compensation Using a Near-Infrared Time-of-Flight Camera. *Appl. Sci.* **2021**, *11*, 10913. [\[CrossRef\]](#)
29. Charlton, P.H.; Bonnici, T.; Tarassenko, L.; Clifton, D.A.; Beale, R.; Watkinson, P.J. An assessment of algorithms to estimate respiratory rate from the electrocardiogram and photoplethysmogram. *Physiol. Meas.* **2016**, *37*, 610. [\[CrossRef\]](#) [\[PubMed\]](#)
30. van Gastel, M.; Stuijk, S.; de Haan, G. Robust respiration detection from remote photoplethysmography. *Biomed. Opt. Express* **2016**, *7*, 4941–4957. [\[CrossRef\]](#) [\[PubMed\]](#)
31. Gleichauf, J.; Herrmann, S.; Hennemann, L.; Krauss, H.; Nitschke, J.; Renner, P.; Niebler, C.; Koelpin, A. Automated non-contact respiratory rate monitoring of neonates based on synchronous evaluation of a 3D Time-of-Flight camera and a microwave interferometric radar sensor. *Sensors* **2021**, *21*, 2959. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Matsuda, T.; Makikawa, M. ECG monitoring of a car driver using capacitively-coupled electrodes. In Proceedings of the 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Vancouver, BC, Canada, 20–24 August 2008; pp. 1315–1318.
33. Wusk, G.; Gabler, H. Non-invasive detection of respiration and heart rate with a vehicle seat sensor. *Sensors* **2018**, *18*, 1463. [\[CrossRef\]](#)
34. Foix, S.; Alenya, G.; Torras, C. Lock-in Time-of-Flight (ToF) cameras: A survey. *IEEE Sens. J.* **2011**, *11*, 1917–1926. [\[CrossRef\]](#)
35. Lugaresi, C.; Tang, J.; Nash, H.; McClanahan, C.; Uboweja, E.; Hays, M.; Zhang, F.; Chang, C.L.; Yong, M.G.; Lee, J.; et al. Mediapipe: A framework for building perception pipelines. *arXiv* **2019**, arXiv:1906.08172.
36. Ruiz, N.; Chong, E.; Rehg, J.M. Fine-grained head pose estimation without keypoints. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 2074–2083.
37. Zanuttigh, P.; Marin, G.; Dal Mutto, C.; Dominio, F.; Minto, L.; Cortelazzo, G.M. Time-of-flight and structured light depth cameras. *Technol. Appl.* **2016**, 978–983. [\[CrossRef\]](#)

-
38. NHTSA. Traffic Safety Facts (2018 Data): Rural/Urban Comparison of Traffic Fatalities. 2020. Available online: <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812957> (accessed on 27 February 2022).
 39. Keel, M.S.; Kim, D.; Kim, Y.; Bae, M.; Ki, M.; Chung, B.; Son, S.; Lee, H.; Jo, H.; Shin, S.C.; et al. A 4-tap 3.5 μm 1.2 Mpixel Indirect Time-of-Flight CMOS Image Sensor with Peak Current Mitigation and Multi-User Interference Cancellation. In Proceedings of the 2021 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 13–22 February 2021; Volume 64, pp. 106–108.