Exploring Proprioceptive Take-Over Requests for Highly Automated Vehicles

Sarah Faltaous
University of Duisburg-Essen
Essen, Germany
sarah.faltaous@uni-due.de

Chris Schönherr
University of Duisburg-Essen
Essen, Germany

Henrik Detjen
University of Applied Science Ruhr West
Mülheim, Germany
henrik.detjen@hs-ruhrwest.de

Stefan Schneegass
University of Duisburg-Essen
Essen, Germany
stefan.schneegass@uni-due.de

ABSTRACT
The upsurging levels of autonomous vehicles allow the drivers to shift their attention to non-driving tasks while driving (i.e., texting, reading, or watching movies). However, these systems are prone to failure and, thus, depending on human intervention becomes crucial in critical situations. In this work, we propose using human actuation as a new mean of communicating take-over requests (TOR) through proprioception. We conducted a user study via a driving simulation in the presence of a complex working memory span task. We communicated TORs through four different modalities, namely, vibrotactile, audio, visual, and proprioception. Our results show that the vibrotactile condition yielded the fastest reaction time followed by proprioception. Additionally, proprioceptive cues resulted in the second best performance of the non-driving task following auditory cues.

CCS CONCEPTS
• Human-centered computing → Interaction techniques;

KEYWORDS
Take-Over Requests, Autonomous Vehicles, Electrical Muscle Stimulation, Proprioception

1 INTRODUCTION
The near future-offered autonomous vehicles would mostly belong to level 3 automation, which allows the driver to have "eyes-off" driving granting the chance to be engaged in a non-driving tasks (e.g., reading a book or watch a movie [18]). This level, nonetheless, still requires the driver to intervene whenever the system fail to operate under specific conditions. In other words, if the system accuracy is below certain threshold the driver would be notified to take-over. Research explored various aspects of such take-over requests (TOR). Particularly the modality used to communicate the TOR has been explored in detail. Currently, research focuses on classical modalities (e.g., visual, auditory [2, 15, 24]) that are also engaged by the secondary task that can be done while the car drives autonomously. Thus, the driver might not immediately react to the TOR since that the used communication channel might be overloaded.

In this paper, we explore actuating the human body as a novel way of communicating take-over request in autonomous vehicles. For this purpose, we use electrical muscle stimulation (EMS) to actuate the drivers’ arms to be directed to the driving wheel. This is communicated through proprioception to the user. We conduct a user study comparing the proprioceptive approach to currently used modalities. We found that the proprioceptive approach results in the second fastest reaction time and the best performance in the non-driving task.
1.1 Contribution

The contribution of this work is two-fold: (1) We present the concept of proprioception as a novel modality to communicate take-over requests in highly automated vehicles. (2) We report on the results obtained in a user study, comparing proprioception to conventional modalities (i.e., Audio, Visual, and Vibro-tactile).

2 BACKGROUND

With the rise of Level 3 of automation, the role of human intervention in case of automation failures became crucial. Hence, researchers have been focusing on designing new techniques to communicate take over request (TOR). The main challenge facing the implementation of such techniques is how to get a disengaged driver back into the "the driving-loop". They, therefore, have been trying to improve the quality of TORs either by implementing new techniques or improving the existing ones. One major quantifiable attribute that determines the quality of the TOR has been identified as time to react (TTR). Note that this is also known as time buffer [5] or time budget [10]. To be able to measure it several aspects are being investigated, namely driving scenario, non-driving task impact, TOR issuing time and modality. Not only that they are correlated but also each of them has a direct impact on the TTR.

2.1 Non-Driving Task Impact

In one study by Gold et al. [8] they showed that drivers are likely to intervene faster in urgent situations when they have their focus completely shifted to the road and don’t have the automated feature on. This arose the issue of workload induced by non-driving tasks, as in another study [7] they inspected the drivers performance under four different tasks (i.e., visual-motoric SURT task and cognitive 2-Back task, cognitive-motoric task and a laptop-based fill-in the blank text) compared to a base line condition with no automation. Results showed that the different tasks had a small role to play with respect to the TTR and the conclusion was similar to the previously mentioned study. However, they highlighted the importance of the situation complexity.

2.2 Driving Scenario

Situation complexity, or in other words the driving scenario, was further examined in several studies under similar conditions. For example, Radlmayr et al. [19] investigated the drivers performance in critical take over scenario (i.e., high density traffic). They deployed cognitive 2 back task and visual SURT task, reaching the same conclusion of Gold et al. [7], furthermore they indicated the impact of the driving situation on the overall performance. Later, Gold et al. [10] confirmed similar findings using a different secondary task (i.e., verbal 20-Questions Task) and under various traffic densities. In a recent study from Borjeni et al. [20], using a motion simulator, they showed that TOR in curves affects the take over quality more than that of straight roads. This goes in accordance with what was previously reported by the participants in Walch et al. [24] examining similar situations. Borjeni et al. [20] further recommended that the TOR should be adapted to the road and the driver responses. That would mean that both the right timing and a suitable modality should be conveyed.

2.3 Take-Over Request Modality and Timing

To inspect the issuing time, Gold et al. [9] compared the drivers’ performance given two different TTRs (i.e., 5s and 7s). Their results showed that the shorter the TTR, the worse the driver is likely to perform (i.e., swerve). While the timing would remain always relevant to the driving situation and system’s limitation, the communicated modality would totally depend on the drivers state. In a survey by Bazilinskyy et al. [1] with 3000 replies across 102 countries, participants were presented with 5 different driving situations and were asked to indicate their preferred modality to receive a TOR. Their preferences indicated mostly multi-modal systems, combining either two of these: visual, audio or vibro. These qualitative preferences have been further confirmed in recent studies. As auditory-vibro multi modal yielded to better performance and ratings compared to either [15].

3 PROPRIOCEPTIVE TAKE-OVER REQUESTS

To date most of the research done in the field of TOR design deploys either uni-modal or multi-modal cues to notify the driver. Targeting one or multiple human senses (i.e., hearing, sight, and touch), they focus mainly on using audio, visual, or vibro-tactile feedback signals.

Given the potential workload generated in future highly automated cars (e.g., through reading or texting and hearing music [18]), the driver’s visual and auditory channels might already be occupied [25]. Similarly, the vibration in a driving vehicle might already put load on the tactile perception. Thus, we propose targeting the human proprioception as a new way of communicating TOR. According to research work in neuroscience [6], proprioception consists of the sense of position (limb position sense) and movement of the limbs (kinaesthesia) in the absence of vision. This means that humans are capable of perceiving the position and movement of their limbs without visually focusing on them.

In this work, we propose using electrical muscle stimulation (EMS) [22] to actuate the drivers arm. Researchers use EMS in multiple different scenarios from communicating navigation cues to the leg [17], drawing computer calculated graphs [12], or communicating affordance of physical objects [11]. In this TOR scenario, we strive to pull the arm of the drivers upwards to the height of the driving wheel so that they can easily grab the driving wheel and steer it in the communicated direction. This is achieved by actuating the biceps muscle (see Figure 1). The direction is communicated by the actuated arm. We actuate the left arm to communicate a movement to the left and the right arm to communicate a movement to the right.

4 DRIVING SIMULATOR AND TAKE-OVER REQUESTS

To evaluate the idea of proprioceptive TORs, we first developed a driving simulation environment. This simulation environment consists of a hardware setup and a driving scenario. Second, we designed four different ways to communicate a TOR using visual, auditory, vibro-tactile, and proprioceptive cues respectively.


4.1 Hardware Setup

We present the simulation environment on three displays mimicking the view to the front (24 inches) and sides (22 inches) as depicted in Figure 2. We used a Logitech G29 steering wheel fixed at the table in front of the driver.

4.2 Driving Scenario

We chose a collision avoidance scenario as a driving scenario which is commonly used to evaluate take-over requests [4, 10]. We designed a 3-lanes highway road, where we placed the autonomous vehicle in the middle one (cf., Figure 3). The vehicle drives on the middle lane and accelerates up to a velocity of 100 km/h. Every 1.2-2.4 kilometers (i.e., every 50-100 seconds), an obstacle appears, either in the left and the middle, right and the middle, or middle lane only. These obstacles are accidents which block certain lanes. Once the obstacle appears a TOR is communicated to the drivers using a single modality (i.e., audio, video, vibro-tactile, or proprioception). This gives the drivers five seconds to react, which according to a previous study [9] is enough time to react and take-over the vehicle. Besides reacting in time, they also need to identify the appropriate lane that is safe to use (e.g., if the right and middle lanes are blocked, the driver should take the left lane). The drivers react by starting to steer in a certain direction they want to steer to by turning the steering wheel. The simulation interprets this interaction with the steering wheel in a binary way similar to the work of Borojeni et al. [2].

4.3 Take-Over Request Design

For the TOR design we used four different modalities. The goal of the TORs is to communicate the potential upcoming obstacle and the potential direction that can be used to avoid the obstacle. If the obstacle is in the left and middle lanes, the drivers receive a TOR that directs their attention to the right lane, which is a free way to continue driving in. The warning signal would not disappear unless either the drivers react or the car does an emergency break to avoid a collision.

We integrated the visual and the audio feedback in the driving simulator using Unity. In the visual TOR we show either an arrow pointing to the left, the right, or both sides. As for the auditory a beep sound would be played, through external speakers, either from the left, right, or both speakers.

The vibro-tactile feedback alarms the user via vibro-motors mounted on a bracelet that touches directly the drivers’ wrists. The simulation controls the vibro-motors via a NodeMCU micro-controller connected through WiFi.

For the proprioceptive feedback, we used the Let Your Body Move (LYBM) toolkit [16]. LYBM is composed of an Arduino, an Android application, and an off-the-shelf EMS signal generator. Again, the simulation controls the LYBM toolkit through commands send via Bluetooth to the Android application. Whenever the TOR to be communicated is issued either the For communicating left, the left, and for communicating right, the right, and, for both, both arms are pulled upwards.

5 USER STUDY

We conducted a user study to compare the proprioceptive approach to regularly used modalities to communicate take-over requests. For this we used visual, auditory, and vibro-tactile notifications as this are modalities commonly used to communicate take-over requests [2, 15, 24].

5.1 Secondary Task

To create a realistic scenario and to ensure that the drivers’ eyes are off the road, we chose a span task implemented by BrainTurk 1, displayed on a an android tablet. The task addresses the working memory, where the application displays a task separated by a span option. Overall, the application used 2 tasks along with 2 different kinds of spans. In these tasks, the drivers had to verify the correctness of words (i.e., appears for 5 seconds) or the symmetry of shapes which don’t have an appearance limit. The words or shapes were then separated by letters or a grid with highlighted

1https://www.brainturk.com/games
We invited 12 participants (4 females and 8 males) age range from 21-70 years ($\mu$ (33), $\sigma$ (15)). After welcoming the participants, we explained the purpose of the study and the participants provided written informed consent. In a first step, participants familiarized themselves with the driving simulator and the non-driving task. The study consisted of four conditions displayed in four different blocks, separated by at least 5 minutes breaks, and lasted about one hour. In each block, we used one modality. To avoid any pattern, we ordered the modalities using Latin square. Before the beginning of each block, the modalities were prepared and again briefly explained to the participants. In the vibrotactile condition, the vibrating wristbands were connected and attached to the wrist. In the EMS condition, the experimenter calibrates the EMS-electrodes by continuously increasing the intensity of the signal until the participant performed the desired movement. With the consent of participants, we video-recorded the whole study for post-hoc analyses.

At the beginning of each condition, a welcome screen would appear, where the experimenter sets the participant ID and chooses the present modality (i.e., visual, audio, vibrotactile, or EMS). The experimenter then placed the tablet with the secondary task on the participant’s lap (cf., Figure 2). The participant presses first the start button of the simulation and then starts with the non-driving task. We measured the reaction time from presenting the feedback until the participant started steering, as well as the steering direction.

5.2 Participants and Procedure

We invited 12 participants (4 females and 8 males) age range from 21-70 years ($\mu$ (33), $\sigma$ (15)). After welcoming the participants, we explained the purpose of the study and the participants provided written informed consent. In a first step, participants familiarized themselves with the driving simulator and the non-driving task. The study consisted of four conditions displayed in four different blocks, separated by at least 5 minutes breaks, and lasted about one hour. In each block, we used one modality. To avoid any pattern, we ordered the modalities using Latin square. Before the beginning of each block, the modalities were prepared and again briefly explained to the participants. In the vibrotactile condition, the vibrating wristbands were connected and attached to the wrist. In the EMS condition, the experimenter calibrates the EMS-electrodes by continuously increasing the intensity of the signal until the participant performed the desired movement. With the consent of participants, we video-recorded the whole study for post-hoc analyses.

At the beginning of each condition, a welcome screen would appear, where the experimenter sets the participant ID and chooses the present modality (i.e., visual, audio, vibrotactile, or EMS). The experimenter then placed the tablet with the secondary task on the participant’s lap (cf., Figure 2). The participant presses first the start button of the simulation and then starts with the non-driving task. We measured the reaction time from presenting the feedback until the participant started steering, as well as the steering direction.

6 RESULTS

Next, we report the results of the user study. Prior to the presented analysis, we corrected the recorded data to the delay of the wireless communication. For this, we measured the delay induced and subtract it from the measured time. Specifically, we subtract 200ms from both the proprioceptive cues (WiFi to the mobile phone and Bluetooth between mobile phone and LYBM-toolkit) and the vibrotactile cues (WiFi to the NodeMCU). We had to exclude two TORs from 2 participants (i.e., P11 and P12) in which the proprioceptive cue was not delivered due to technical issues. When the participants did not perceive the feedback, we removed these measurements as well from the calculation of the time to take-over.

6.1 Time to Take-Over

The results show that the participants needed least time in the vibrotactile condition ($\mu$ (1.76), $\sigma$ (0.36)), followed by proprioception ($\mu$ (1.86), $\sigma$ (0.28)), audio ($\mu$ (1.93), $\sigma$ (0.28)), and visual ($\mu$ (2.62), $\sigma$ (0.76)). A repeated measures analysis of variance shows statistically significant differences between the four conditions, $F ((3, 33)) = 11.89, p < .001$. Follow-up Holm-Bonferroni-corrected t tests show that the differences between audio - visual, $t ((11)) = -3.446, p = .022$, and EMS - visual, $t ((11)) = -3.579, p = .022$ and vibro-visual, $t ((11)) = -3.691, p = .021$ are statistically significant (cf., Figure 4).

6.2 Error Rate and Directional Guidance

Overall, the participants reacted to all TORs except in the visual condition in which a total of 16 TORs were not perceived by the user. The direction, however, was always chosen as communicated through each cue. We further inspected the bi-directional trials ($N = 36$) in which we placed the obstacle only in the middle leaving the participant to either steer to the left or to the right. The video analysis showed that the drivers used both hands equally in all the trials under EMS condition ($N_{EMS_L} = N_{EMS_R} = 18$), unlike the other conditions where the right hand was the dominant ($N_{Vibro_R} = 20, N_{Audio_R} = 21, N_{Visual_R} = 26$) (cf., Figure 5).
7 DISCUSSION

7.1 Take Over Times

The obtained score from the non-driving task reveals a trade-off between the reaction time and performance. A previous study by Zhou et al. [26] showed that vibrotactile feedback opposes the cognitive load. However, that wasn’t the case with the proprioception condition. As mentioned by Schmidt and Young [21], the challenge sometime lies in how to do an action more that what to do for an action. Hence, by deploying the EMS actuation, the amount of cognitive load induced is less than that of the other conditions.

7.2 User Study and Task

Comparing the lab-based setup we used in the user study with a real car, it becomes apparent that the background noise of the car as well as the vibration from driving is not simulated. Given that these would mainly, negatively, influence the audio and vibrotactile condition, we assume that the proprioception condition would perform better in a realistic environment. Similarly, the result is influenced by the non-driving related task. We used a span task that is commonly used for such studies [3, 7] and reflects potential future activities [18]. Using a task that focuses more on the haptics and less on the visuals would perhaps result in different results. Thus, the type of feedback might in the future be chosen based on the drivers current activity.

7.3 Proprioceptive Cues

In the user study we calibrated the proprioceptive cues and made sure that, in a relaxed condition, the drivers’ forearms move upwards to reach the driving wheel. However, observing the participants reactions in the recorded videos, it did not always show the same movement. As also indicated by multiple participants that while doing the study they didn’t always feel the same actuation as in the calibration phase. Since that the proprioception, as previously mentioned, is related to the sense of position as well as the awareness of the motion [6], having the drivers solving the non-driving task during the study, resulted in 2 different types of motions. The first one is the one directed by their brain to touch the correct answers buttons and move their hands. The other one is the artificial one that we suddenly induce. That might have resulted in a conflict leading to a delay in the reaction, yet not the signal perception.

8 LIMITATION AND FUTURE WORK

We acknowledge the following limitation to our work. We used the EMS to actuate the drivers both arms, which might have had an impact on the reaction time. As the drivers might have countered the incoming signals with their own. Hence, we recommend in the future work to actuate a different muscles other than the ones needed for the non-driving task. Another option would be to change the nature of the non-driving task (e.g., audio span task). Furthermore, using the current EMS devices, we made sure visually in the calibration process to have the arms of the participants actuated upwards. However, for future work to result in more accurate movements further ems actuation devices need to be developed to achieve more controlled movements (i.e., EMS actuation based on EMG). While the results of our study suggest that feedback based on proprioception is feasible, an evaluation with a larger sample size should be conducted to strengthen our findings.
9 CONCLUSION

In this work, we investigated the TOR performance in a L2/L3 autonomous vehicle. We deployed a collision avoidance system and tested a new TOR modality that addresses the human proprioception via EMS. The concept of EMS is to induce an electrical signal on a targeted muscle to actively pull it to reach the same height of the driving wheel. While there are several modalities to communicate TORs to drivers (e.g., visual, auditory, and vibrotactile), none of them influence their actions without requiring additional attentional shift and, thus, inducing additional cognitive load. We found that the vibrotactile resulted in the fastest RT, followed by the proprioception condition, audio, and at last visual. The participants, however, performed the best in the non-driving task in the audio condition, then proprioception, vibrotactile, and visual conditions. We reflected on the obtained results and demonstrated the advantages of addressing the human proprioception.

REFERENCES


cal organisation and assessment. Journal of Neurology, Neurosurgery 

& Psychiatry 73, 5 (2002), 473–477. https://doi.org/10.1136/jnnp.73.5.473


