



A workload adaptive haptic shared control scheme for semi-autonomous driving

Ruikun Luo^{a,1}, Yifan Weng^{b,1}, Yifan Wang^c, Paramsothy Jayakumar^d, Mark J. Brudnak^d, Victor Paul^d, Vishnu R. Desaraju^e, Jeffrey L. Stein^b, Tulga Ersal^{b,*²}, X. Jessie Yang^{f,**²}

^a Robotics Institute, University of Michigan, Ann Arbor, MI, United States

^b Mechanical Engineering, University of Michigan, Ann Arbor, MI, United States

^c Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI, United States

^d U.S. Army Ground Vehicles System Center, Warren, MI, United States

^e Toyota Research Institute, Ann Arbor, MI, United States

^f Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI, United States



ABSTRACT

Haptic shared control is used to manage the control authority allocation between a human and an autonomous agent in semi-autonomous driving. Existing haptic shared control schemes, however, do not take full consideration of the human agent. To fill this research gap, this study presents a haptic shared control scheme that adapts to a human operator's workload, eyes on road and input torque in real time. We conducted human-in-the-loop experiments with 24 participants. In the experiment, a human operator and an autonomy module for navigation shared the control of a simulated notional High Mobility Multipurpose Wheeled Vehicle (HMMWV) at a fixed speed. At the same time, the human operator performed a target detection task. The autonomy could be either adaptive or non-adaptive to the above-mentioned human factors. Results indicate that the adaptive haptic control scheme resulted in significantly lower workload, higher trust in autonomy, better driving task performance and smaller control effort.

1. Introduction

Autonomous driving technology is currently limited in its scope and reliability, giving rise to the semi-autonomous mode of driving. In this mode, the driving task is shared between the human and the autonomy. Thus, properly allocating the control authority between these two agents becomes critical for safety and performance. Managing this allocation is a challenging problem referred to as the shared control problem.

As the literature review in "Background on shared control" section shows, researchers have recognized this challenge and proposed many strategies for managing control authority in shared control. However, one very important consideration has been overlooked in the strategies developed to date: namely, the human operator's workload. The human operator's workload may change over time as the driving conditions change, or as a result of the human getting involved in secondary tasks. These variations in workload have direct implications on the management of control authority, as the human operator may or may not be ready to seize or relinquish control depending on the current workload.

However, the relationship between human workload and control authority management has not yet been explored. This work aims to fill this gap by developing a shared control scheme that adapts to the human operator's workload in real time.

1.1. Background on shared control

Based on the hierarchy of authority, prior literature on shared control can be broadly classified into two categories: supervisory and cooperative.

In supervisory shared control, one agent supervises the behavior of the other agent and determines the final control input to the vehicle. One example is a vehicle with Level 2 automation (SAE International, 2016), where the human acts as the supervisor. Human operator monitors the status of the vehicle and decides when to engage/disengage the automated driving function. On the other hand, in some schemes, the autonomy serves as the supervisor to monitor human operator's control commands and modify them as needed (Erlien et al., 2016; Schwarting

* Corresponding author at: 1231 Beal Ave, Ann Arbor, MI 48109, United States.

** Corresponding author at: 1205 Beal Avenue, Ann Arbor, MI 48109, United States

E-mail addresses: tersal@umich.edu (T. Ersal), xjyang@umich.edu (X.J. Yang).

¹ These authors have contributed equally.

² These authors have contributed equally.

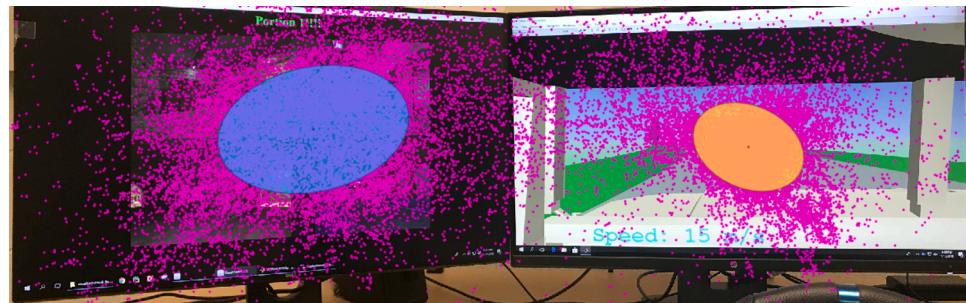


Fig. 1. Example of using the hidden Markov model to model gaze trajectory to estimate workload. Magenta dots: gaze points. Ellipsoids: Multivariate normal distributions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

et al., 2017; Storms et al., 2017). In supervisory shared control, the control authority transfers entirely from one agent to the other in a discrete manner. Therefore, only one of the agents has the final control of the vehicle.

In co-operative shared control, both agents can affect the final control input. One type of co-operative shared control directly blends the steering angle inputs from both the human and autonomy through a designed arbitrator (Anderson et al., 2011). This scheme has the property that the loop between the human and autonomy is closed after the steering wheel; i.e., the human will be able to feel the impact of autonomy input only after the resultant steering command takes effect and through the response of the vehicle. The other type of co-operative shared control is haptic shared control, in which the human and autonomy can negotiate the steering angle through the torques they apply to the steering wheel (Griffiths and Gillespie, 2005; Petermeijer et al., 2015; Nguyen et al., 2018). In this scheme, the human operator can directly feel the torque from the autonomy and can choose to yield to or fight with it by adjusting his/her torque on the steering wheel. Researchers developed and tested a haptic shared control framework, and showed that haptic control improved the driving performance while reducing visual demand or shortening the reaction time of the secondary task (Griffiths and Gillespie, 2005). Others used the haptic control framework with the bandwidth guidance version and the continuous guidance version, and showed that both helped reduce driver errors (Petermeijer et al., 2015). Haptic shared control has also been compared to the two supervisory control modes mentioned above and found to yield better overall performance (Bhardwaj et al., 2020).

The impedance of autonomy in a haptic shared control scheme can be considered as a natural tuning parameter through which adaptability can be introduced. Indeed, even though earlier haptic shared control schemes used a fixed impedance (Griffiths and Gillespie, 2005; Mulder et al., 2008), later works started investigating adaptive impedance schemes. Some schemes adopt vehicle-performance-based switching rules as adaptation mechanisms, such as turning shared control on when the lateral error of the vehicle exceeds a designed threshold (Petermeijer et al., 2015). Others consider human-performance-based metrics to adapt impedance continuously, such as human's input torque and attention as the guideline for designing control authority allocation (Nguyen et al., 2018). However, workload, an important human factor, has not yet been considered for adaptation purposes.

Therefore, in the present work, we developed an adaptive haptic shared control scheme, which considers the human operator's workload, eyes on road and torque input in the control authority allocation.

1.2. Background on workload estimation

Workload can be measured offline or online. Offline retrospective measures are commonly used after a human operator completes a task via questionnaires such as the NASA TLX (Task Load Index) (Hart and Staveland, 1988a). Their offline nature prohibits their utilization for real-time adaptation. Online real-time measures of workload are assessed while the human operator is performing the task and thus could be used for designing adaptive systems. Online real-time measures of workload are usually based on task performance or human operator's physiological signals. The underlying rationale for performance-based measures is that under high workload, human operator's task performance would be harmed. Performance-based measures, however, are not applicable if the task performance is ambiguous or is not available immediately. Physiological measures rely on changes in the human physiological signals. Various types of physiological signals have been used to estimate workload, including heart rate, electroencephalogram (EEG), eye-related measures, galvanic skin response (GSR) and near infrared spectroscopy (NIRS). Please refer to (Heard et al., 2018) for a review.

Among all the physiological measures, some could be intrusive (e.g., EEG (Liu et al., 2017a; Diaz-Piedra et al., 2021) or could be easily affected by body movements (e.g., heart rate (Chen et al., 2015)). Eye-tracking emerges as a less intrusive and robust technique. Research efforts have been spent on using eye-related measurements to assess operators' workload, including pupil diameter (Recarte and Nunes, 2003; Lu et al., 2019), gaze distribution (Reimer, 2009), gaze trajectory (Wang et al., 2014; Fridman et al., 2018), and blink rate (Halverson et al., 2012).

To assess workload online using physiological measurements, previous studies largely adopted statistical methods to show the relationships between certain physiological signals and workload. Recently, researchers started to use machine learning techniques to classify mental workload into different levels. For example, some researchers used a decision tree to classify the drivers' workload into two levels using 30 s driving data and the pupil diameter data (Zhang et al., 2004). Others proposed a deep neural network to analyze a 6 s video of the eye and classified operators' workload into three categories in real time (Fridman et al., 2018). However, such online measures of workload have not yet been incorporated into the shared control schemes.

2. Present study

In the present study, we developed a dual-task shared control platform. Using the experimental platform, the human operator and the

autonomy shared the control of a simulated notional High Mobility Multipurpose Wheeled Vehicle (HMMWV) at a fixed speed. At the same time, the human operator performed a target detection task for surveillance. We used the hidden Markov model (HMM) to estimate the human's workload by analyzing 4 s gaze trajectory data. We then designed and tested the adaptive shared control scheme by regulating the assistance level of the autonomy based on the estimated workload, and the operator's eyes on road and input torque.

2.1. Workload estimation with HMM

An HMM is a probabilistic model of the joint probability of a collection of random variables $\{O_1, O_2, \dots, O_T, S_1, S_2, \dots, S_T\}$. S_t is a discrete variable that represents the hidden state at time step t . S_t can take values from $\{1, 2, \dots, N\}$, where N is the number of hidden states. O_t represents the observations at time step t . T represents the termination time step. An HMM also contains a tuple of parameters as $\Theta = (\pi, A, B)$. $\pi \in \mathbb{R}^N$ is the prior distribution of $P(S_1)$. $A \in \mathbb{R}^{N \times N}$ is the stochastic transition matrix, where $A = \{a_{ij}\} = P(S_t=j|S_{t-1}=i)$. $B = \{b_j(\cdot)\}$ is a set of observation model for every hidden state $j \in \{1, 2, \dots, N\}$, where $b_j(\mathbf{o}_t) = P(O_t = \mathbf{o}_t | S_t = j)$ and \mathbf{o}_t is a given observation at time step t .

In the present study, the observations \mathbf{o}_t are the gaze points, i.e., locations of where the human is looking at relative to the external world coordinate shown as the magenta dots in Fig. 1. The observation models are a set of multivariate distributions over the gaze points, i.e., $b_j(\mathbf{o}_t) = P(O_t = \mathbf{o}_t | S_t = j) \sim \mathcal{N}(\mu_j, \Sigma_j)$, shown as the ellipsoids in Fig. 1. Thus $B = \{\mu_j, \Sigma_j\}$.

We trained two HMMs, one for the high workload and one for the moderate workload. For each workload level w , we collected a set of L gaze trajectories $D_w = \{\mathcal{O}_l | \mathcal{O}_l = \{\mathbf{o}_1^l, \mathbf{o}_2^l, \dots, \mathbf{o}_T^l\}\}$, where $l = \{1, 2, \dots, L\}$. Thus, the learning process learns two sets of HMM parameters $\Theta_w = (\pi, A, B)$, one for each workload level using data D_w . The parameters of the HMMs were learned by the expectation maximization (EM) algorithm using the open source implementations from [Rozo et al. \(2016\)](#) and [Calinon \(2016\)](#). To learn the parameters, we define four important probabilities:

$$\begin{aligned} P(O_1 = \mathbf{o}_1^l, \dots, O_l = \mathbf{o}_l^l, S_l = i | \Theta^k) &= \alpha_i^l(t)^k \\ P(O_{l+1} = \mathbf{o}_{l+1}^l, \dots, O_T = \mathbf{o}_T^l | S_l = i, \Theta^k) &= \beta_i^l(t)^k \\ P(S_l = i | \mathcal{O}_l, \Theta^k) &= \gamma_i^l(t)^k \\ P(S_l = i, S_{l+1} = j | \mathcal{O}_l, \Theta^k) &= \xi_{ij}^l(t)^k \end{aligned} \quad (1)$$

where k represents the k th iteration in the EM algorithm. The EM algorithm is then:

E-step:

Recursively update α :

$$\begin{aligned} \alpha_i^l(1)^{k+1} &= \pi_i^k \mathcal{N}(\mathbf{o}_1^l; \mu_i^k, \Sigma_i^k) \\ \alpha_j^l(t+1)^{k+1} &= [\sum_{i=1}^N \alpha_i^l(t)^{k+1} a_{ij}^k] \mathcal{N}(\mathbf{o}_{t+1}^l; \mu_j^k, \Sigma_j^k) \end{aligned}$$

Recursively update β :

$$\begin{aligned} \beta_i^l(T)^{k+1} &= 1 \\ \beta_i^l(t)^{k+1} &= \sum_{j=1}^N a_{ij}^k \beta_j^l(t+1)^{k+1} \mathcal{N}(\mathbf{o}_{t+1}^l; \mu_j^k, \Sigma_j^k) \end{aligned}$$

Update γ :

$$\gamma_i^l(t)^{k+1} = \frac{\alpha_i^l(t)^{k+1} \beta_i^l(t)^{k+1}}{\sum_{j=1}^N \alpha_j^l(t)^{k+1} \beta_j^l(t)^{k+1}}$$

Update ξ :

$$\xi_{ij}^l(t)^{k+1} = \frac{\gamma_i^l(t)^{k+1} a_{ij}^k \beta_j^l(t+1)^{k+1} \mathcal{N}(\mathbf{o}_{t+1}^l; \mu_j^k, \Sigma_j^k)}{\beta_i^l(t)^{k+1}}$$

M-step:

$$\begin{aligned} \mu_i^{k+1} &= \frac{\sum_{l=1}^L \sum_{t=1}^T \gamma_i^l(t)^{k+1} \mathbf{o}_t^l}{\sum_{l=1}^L \sum_{t=1}^T \gamma_i^l(t)^{k+1}} \\ \Sigma_i^{k+1} &= \frac{\sum_{l=1}^L \sum_{t=1}^T \gamma_i^l(t)^{k+1} (\mathbf{o}_t^l - \mu_i^{k+1})(\mathbf{o}_t^l - \mu_i^{k+1})^T}{\sum_{l=1}^L \sum_{t=1}^T \gamma_i^l(t)^{k+1}} \\ \pi_i^{k+1} &= \frac{\sum_{l=1}^L \gamma_i^l(1)^{k+1}}{L} \\ a_{ij}^{k+1} &= \frac{\sum_{l=1}^L \sum_{t=1}^T \xi_{ij}^l(t)^{k+1}}{\sum_{l=1}^L \sum_{t=1}^T \gamma_i^l(t)^{k+1}} \end{aligned}$$

The two steps iterate until convergence. The number of hidden states was determined by the Bayesian Information Criterion (BIC) ([Calinon and Billard, 2005](#); [Schwarz et al., 1978](#)).

Given a gaze trajectory $\mathcal{O} = \{\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_T\}$, we computed the likelihood of $P(\mathcal{O} | \tilde{\Theta}_w)$ via the forward algorithm, where $\tilde{\Theta}_w$ represents parameters for different learned HMMs for the high workload and moderate workload. The forward algorithm is similar to the Recursive update of α in the E-step of the EM algorithm. We have $P(\mathcal{O} | \tilde{\Theta}_w) = \sum_{i=1}^N \tilde{\alpha}_i(T)$. To estimate the workload of \mathcal{O} , we found the HMM with the maximum likelihood, i.e., $\text{argmax}_w P(\mathcal{O} | \tilde{\Theta}_w)$.

Our adaptive shared control scheme is based on the human operator's real-time workload, eyes on road, and input torque. We used the gaze point data from a 4 s time window captured by the Tobii eye tracker (30 Hz sampling rate) to estimate participants' workload and eyes on road. Thus, $T = 120$. Let w_t represent a human operator's workload at time t , $w_t = c_1 \text{argmax}_w p(\mathcal{O}_t | \tilde{\Theta}_w) + c_2$, where c_1, c_2 are scaling and offset factors such that $w_t = 50$ represents moderate workload, and $w_t = 100$ represents high workload. A human operator's eye on road is defined as the percentage of time that s/he is looking at the driving task. Let e_t denote the human operator's eyes on road. e_t is calculated as the average number of times that a participant's gaze points fall on the driving screen within the time window T .

Due to the large mass and high center of gravity of the simulated military vehicle (see "Experiment 1 – Method" section), a rapid change of control commands resulting from a rapid change of w_t and e_t could trigger a rollover. Therefore, we applied a moving average filter with a 1 s time window and downsampled w_t and e_t to 10 Hz.

2.2. Adaptive shared control scheme

2.2.1. Autonomy design

In the present work, the chosen scenario is a single-vehicle mission where there are no other vehicles in the environment. We designed a fixed speed scenario, with the vehicle traveling at 15 mph. Hence, autonomy only needed to provide the steering torque for haptic feedback. We used the nonlinear model predictive control (NMPC) method to generate the steering wheel angle commands, which can track the given centerline for the vehicle. A proportional-integral-derivative (PID) controller is then used to convert the steering angle commands to steering torque. The PID controller acts on the difference between the steering angle commands resulting from the NMPC framework as the reference trajectory and the current steering angle measurement.

[Liu et al. \(2017\)](#) and [Febbo et al. \(2017\)](#) describe the formulation of the NMPC in detail. We used the same bicycle model representation of the vehicle within the NMPC framework with the same states and control constraints as [Liu et al. \(2017\)](#) and [Febbo et al. \(2017\)](#). We tailored the cost function to fit to our problem, because unlike the scenarios for which the original NMPC formulation was developed, our context was path following, the vehicle speed was fixed, and there were no obstacles on the path. Specifically, the cost function was defined as

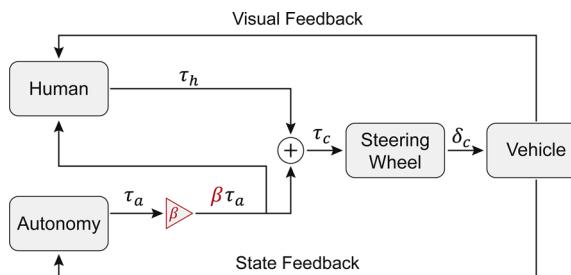


Fig. 2. Block diagram for haptic shared control. τ_h and τ_a represent the torque from human and autonomy, respectively. τ_c and δ_c are the actual control torque and actual control steering angle. β is the assistance level, which is always 1 in the baseline non-adaptive scheme, whereas it varies in the proposed adaptive scheme.

$$\begin{aligned} J = & w_1 \int_{t_0}^{t_p} (y_{\text{ref}}(x(t)) - y(t))^2 dt + w_2 \int_{t_0}^{t_p} \gamma^2 dt \\ & + w_3 \int_{t_0}^{t_p} \left(\tanh\left[\frac{a - F_{z,\text{rl}}}{b}\right] + \tanh\left[\frac{a - F_{z,\text{rr}}}{b}\right] \right) dt \end{aligned} \quad (2)$$

The cost function comprises three terms. The first term is designed to penalize the distance from the position of the vehicle $y(t)$ to the given position on the centerline $y_{\text{ref}}(x(t))$. The second term is designed to regularize the steering rate γ , which ensures the steering angle command changes smoothly. The third term is a soft constraint that increases the cost when the vertical load in either rear tire, $F_{z,\text{rl}}$ or $F_{z,\text{rr}}$, is close to the lowest allowable threshold. This soft constraint is used to prevent the vehicle from operating at its dynamic limits unnecessarily (Liu et al., 2017b; Liu et al., 2018). The allowable threshold is chosen such that even with small violations of this soft constraint, autonomy does not cause tire lift-off. Only the rear tires are considered, since they experience the lift-off first due to the longitudinal weight distribution of the vehicle. The weights w_1 , w_2 and w_3 are set to achieve a trade-off between the three terms in the cost function. In this formulation, t_0 is the time when the prediction horizon starts, while t_p marks the end time of the prediction horizon. $t_p = t_0 + T_p$, where T_p is the fixed prediction horizon and it equaled 6.5 s in this work. We used the open-source nonlinear optimal control package NLOptControl (Febbo, 2017), which uses the Legendre–Gauss–Radau collocation method to transfer the continuous optimal control problem into a nonlinear program. We then solved the nonlinear program by using the solver package IPOPT (Wächter and Biegler, 2006). This optimization process generates a series of steering angle commands through the whole control horizon T_p , and we use only the first 3 s worth of commands. While executing the previous control command series, the system formulates and solves a new optimal control

problem with a receded horizon, and the resulting new command series are applied as soon as they are available.

2.2.2. Non-adaptive haptic shared control

Haptic shared control combines the torques applied by the autonomy and human operator on the steering wheel. It creates a smooth control authority transfer between the human operator and autonomy. The implementation is visualized in Fig. 2, where $\beta = 1$ for the baseline non-adaptive case.

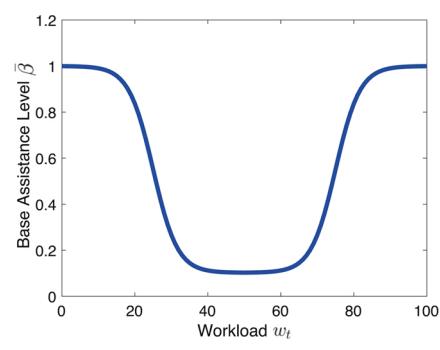
When there is no input from the human operator, the autonomy follows the reference centerline it perceives. The perceived reference centerline may be different from the actual centerline. When there is an input from the human operator that deviates the vehicle from the centerline autonomy perceives, the autonomy applies extra torque to bring the vehicle back to the perceived centerline. The human operator can hence feel the intention of the autonomy and decide whether s/he would agree with it and let autonomy have more control authority (yield), or claim more control authority (fight). The resultant torque applied on the steering wheel, which is the summation of the torques from the human operator and autonomy, determines the final steering angle applied to the vehicle.

2.2.3. Adaptive haptic shared control

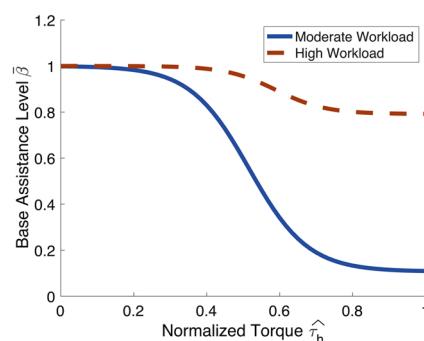
We designed our adaptive shared control scheme based on three different features: workload, torque from the human operator, and eyes on road. The resultant torque τ_c in the adaptive scheme is $\tau_c = \tau_h + \beta(w_t, e_t, \hat{\tau}_h)\tau_a$, where the term β is referred to as assistance level and it determines the strength of assistance torque from autonomy. $\hat{\tau}_h$ is the normalized human torque that is calculated by dividing the input torque from the human operator by the maximum torque a human operator can apply. The implementation of the adaptive scheme is shown in Fig. 2. This scheme is in contrast to the direct blending of the input torques from both the human operator and autonomy as in the non-adaptive haptic shared control scheme. Specifically, β is always 1 in the baseline non-adaptive haptic shared control scheme, whereas it varies in the proposed adaptive scheme.

In our heuristic design for the assistance level, β was separated into two parts: base assistance level $\bar{\beta}$ and assistance level increment $\Delta\beta$; i.e., $\beta = \bar{\beta}(w_t, \hat{\tau}_h) + \Delta\beta(w_t, e_t)$. The base assistance level $\bar{\beta}$ considers the impact from workload and input torque from the human operator, while the assistance level increment $\Delta\beta$ considers the combined effect of eyes on road and workload.

The base assistance level $\bar{\beta}$ was designed according to the principles illustrated in Fig. 3 and explained next. On the one hand, when the torque from the human operator is held constant, the relationship between the assistance level and workload is shown in Fig. 3a. The



(a) Relationship between base assistance level $\bar{\beta}$ and workload w_t



(b) Relationship between base assistance level $\bar{\beta}$ and normalized human torque $\hat{\tau}_h$

Fig. 3. Illustration of base assistance level design principles.

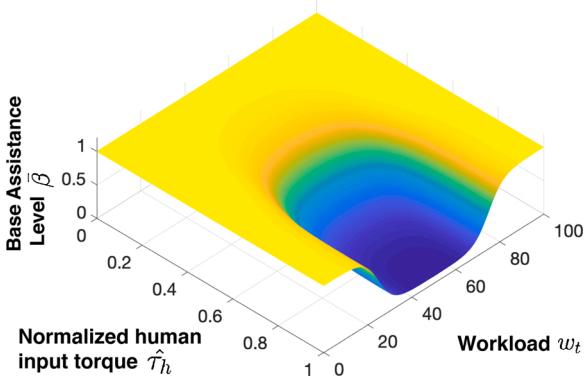


Fig. 4. Relationship between base assistance level $\bar{\beta}$, workload w_t , and normalized human input torque $\hat{\tau}_h$.

designed curve for assistance level matches the study of Flemisch et al. (2010), which shows the assistance from the autonomy should be high when the workload is either very high (overloaded) or very low (underloaded). When the subject has a moderate workload, the assistance from the autonomy should be lower. We set the workload value w_t as 0 when the subject is underloaded, 50 when the subject experiences moderate workload and 100 when the subject is overloaded. We heuristically set the assistance level as 0.1 for moderate workload ($w_t = 50$) and as 1 for very high workload ($w_t = 100$). We fit a sigmoid function to create the smooth transition from $w_t = 50$ to $w_t = 100$. We then mirror the function when workload w_t ranges from 0 to 50 and obtain the curve for the whole workload spectrum.

On the other hand, when the workload of the human operator is held constant, the relationship between the assistance level β and normalized human torque $\hat{\tau}_h$ is shown in Fig. 3b. There are two critical properties of the designed curve. When the human torque is small, the assistance level is kept at a high level ($\bar{\beta} = 1$) to filter out some unintended input torque from the human operator. The assistance level starts to drop after the normalized human torque passes a threshold, which increases as workload increases from moderate workload ($w_t = 50$) to high workload ($w_t = 100$). The threshold value from underloaded to moderate workload mirrors the threshold value when the workload ranges from moderate to high. The threshold is smaller when workload approaches the moderate level, since we assume human would make less mistakes at this workload level based on the results in the literature that show a moderate workload level to be optimal (Flemisch et al., 2010). We heuristically set the threshold value as 0.01 when the human operator experiences a moderate workload ($w_t = 50$), while this value is 0.3

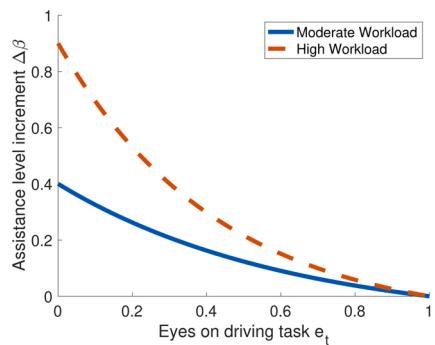
when the human operator is fully overloaded and underloaded ($w_t = 100$ and $w_t = 0$). We fit a quadratic function that is symmetric about the moderate workload case ($w_t = 50$). When the input torque from the human operator becomes greater, the assistance level starts to drop to a lower level. The system is designed to let the human operator have more control authority when there is a strong intention for intervention from the human operator. The value also changes according to the workload. When the human is overloaded, the assistance level for a large torque input remains at a relatively large value. We heuristically set the assistance level for maximum torque as 1 when the human operator is overloaded ($w_t = 100$) and 0.1 when the human operator experiences moderate workload ($w_t = 50$). We then use a modified sigmoid function, connecting the threshold point with the maximum torque point.

Combining those two principles, the formulation of the base assistance level $\bar{\beta}$ is obtained as

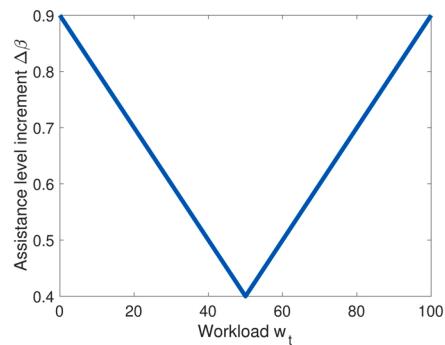
$$\begin{aligned} \bar{\beta}(w_t, \hat{\tau}_h) &= 1 - \left[1 - \left(\frac{0.9e^{0.3(|w_t-50|-25)}}{e^{0.3(|w_t-50|-25)} + 1} + 0.1 \right) \right] \\ &\times \left[\frac{\frac{72\hat{\tau}_h - 36.6 - 15(\frac{w_t}{50} - 1)^2}{5.9 - 2.5(\frac{w_t}{50} - 1)^2}}{\frac{72\hat{\tau}_h - 36.6 - 15(\frac{w_t}{50} - 1)^2}{5.9 - 2.5(\frac{w_t}{50} - 1)^2} + 1} \right] \end{aligned} \quad (3)$$

The corresponding 3D plot showing the relationship between the baseline assistance level, the workload and the normalized human torque is shown in Fig. 4.

The assistance level increment $\Delta\beta$ was designed according to the principles illustrated in Fig. 5 and explained next. On the one hand, keeping the workload constant, when the subject focuses on the driving task, i.e., e_t is very close to 1, $\Delta\beta$ is very close to 0, which indicates no additional assistance level is provided based on the eyes on road metric. When the subject directs their attention to the secondary tasks, i.e., e_t is very close to 0, $\Delta\beta$ increases to a high level, which is illustrated in Fig. 5a. An exponential function is used to connect these two points. On the other hand, keeping the eyes on road constant, when workload is high, the increment $\Delta\beta$ is large, while when the workload is moderate, the increment $\Delta\beta$ is small, which is shown in Fig. 5b. We heuristically set the value of $\Delta\beta$ as 0.4 when the subject experiences moderate workload ($w_t = 50$), and as 0.9 when the subject is overloaded or underloaded ($w_t = 100$ or $w_t = 0$). This value is calculated through linear interpolation when the workload is between these critical values.



(a) Relationship between assistance level increment $\Delta\beta$ and eyes on road for different workloads



(b) Relationship between assistance level increment $\Delta\beta$ and workload w_t when $e_t = 0$

Fig. 5. Illustration of assistance level increment design principles.

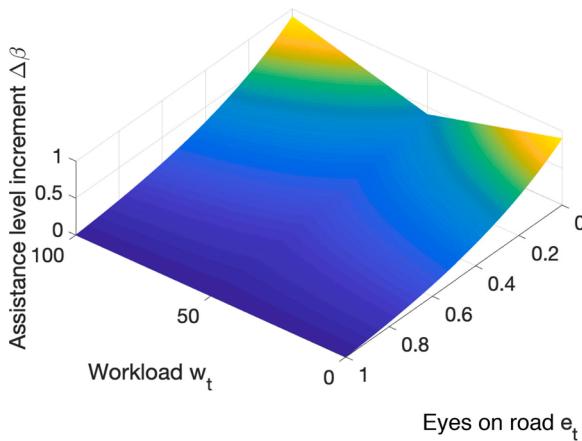


Fig. 6. Relationship between assistance level increment $\Delta\beta$, workload w_t , and eyes on road.

Combining these considerations, the formulation of assistance level increment $\Delta\beta$ is obtained as

$$\Delta\beta(w_t, e_t) = 0.1(0.1|w_t - 50| + 5)^{1-e_t} - 0.1 \quad (4)$$

The corresponding 3D plot showing the relationship between the assistance level increment, the workload and the eyes on road is shown in Fig. 6.

3. Experiment 1

3.1. Introduction

In Experiment 1, we aimed to estimate a human operator's workload in real time by analyzing his/her gaze trajectories. We conducted a human-in-the-loop experiment with 12 participants using a dual-task shared control platform. The participant and the autonomy shared the control of a simulated notional HMMWV. At the same time, the participant performed a surveillance task. The participant wore a pair of Tobii Pro Glasses 2 during the experiment and his/her gazes were captured at 30 Hz. Based on the gaze trajectories, we estimated the participant's workload using HMM.

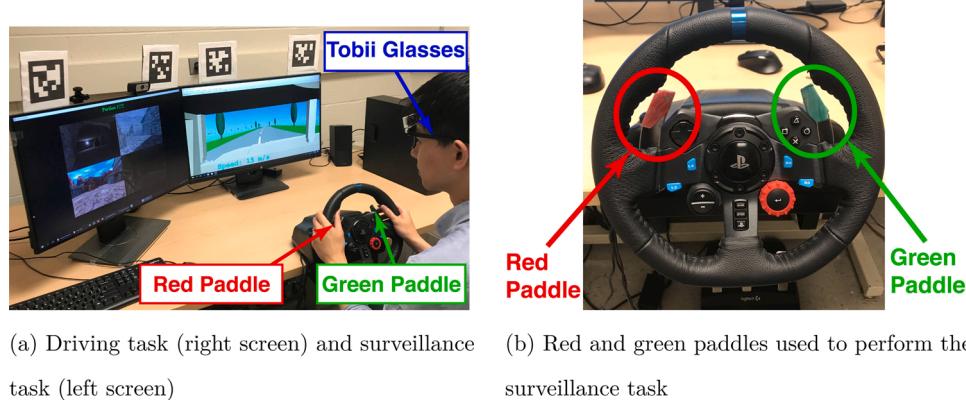


Fig. 7. Dual-task shared control simulation platform.



Fig. 8. Illustration of the surveillance task. Lower left: threat.

3.2. Method

3.2.1. Participants

A total of 13 university students participated in the experiment. Data from one participant were discarded due to equipment malfunction. The remaining 12 participants were on average 26.7 years old ($SD = 3.0$ years) and had an average of 8.3 years of driving experience ($SD = 4.4$ years). All participants had normal or corrected-to-normal vision.

3.2.2. Apparatus and stimuli

A dual-task shared control simulation platform was used in Experiment 1. Participants performed two tasks simultaneously, a driving task and a surveillance task as shown in Fig. 7.

In the driving task, the participant and the autonomy shared the control of the HMMWV at a fixed speed of 15 mph, with the goal to complete a track with minimal deviation from the centerline. To emulate degraded localization due to sensor uncertainty, an offset was introduced such that the autonomy tracked a line which deviated from the centerline by 1 m. The non-adaptive shared control scheme was used in Experiment 1.

In the surveillance task, the participant received image feeds and was asked to identify potential threats (Fig. 8). If the participant identified a threat, s/he pressed the red paddle at the steering wheel to report "danger." Otherwise, the participant pressed the green paddle to report "clear" (Fig. 7b).

Participants received a new set of four images at a fixed time interval, with a 1 s white screen in between, and were responsible for detecting potential threats as accurately as possible. The fixed time interval was varied to manipulate the workload level (see Appendix B for more details).

3.2.3. Experimental design

We manipulated the workload of the experimental tasks (the driving and the surveillance task) by varying the time interval of the surveillance task. During the experiment, the participants drove on 6 different tracks, each lasting for approximately 3 min. Every track was equally segmented into 3 portions and each portion had a different time interval for the surveillance task, 1.5, 2.5, or 6.5 s. The order of presentation for the 3 time intervals on each track was balanced by two 3×3 Latin squares.

3.2.4. Measures

Participants wore a pair of the Tobii Pro Glasses 2 and their gaze points were recorded at 30 Hz.

3.2.5. Experimental procedure

Participants provided a signed informed consent and filled in a demographic survey. After that, they received a training. Participants were first trained on the driving task alone, followed by the surveillance task alone. After that, they performed both the driving and the surveillance task.

After the training session, participants were assisted to wear the eye tracker and underwent the calibration. With the normal room light and without any specific tasks, the experimenter measured each participant's baseline pupil diameter twice, each about 30 s. During the experiment, participants performed the driving task and the surveillance task on 6 different tracks, each lasting approximately 3 min.

3.3. Results

3.3.1. Data processing

Participants drove on 6 different tracks in this experiment. As mentioned above, each track was segmented into 3 portions and each portion had a different time interval for the surveillance task. The portion with the 1.5 s time interval was considered as the high workload portion, and the portion with the 6.5 s time interval the moderate

Table 1

Performance of the HMM.

	F_1	Precision	Recall
HMM	0.664 ± 0.005	0.668 ± 0.005	0.660 ± 0.005

workload portion. The ground truth labels were determined in two pilot studies (see Appendices A and B for details). For each track, we randomly selected 5 sequences of data and each sequence lasted 4 s.

3.3.2. Evaluation of the workload estimation performance

Due to the small dataset of 12 participants, we used the holdout method (Kim, 2009) for cross-validation and tested the performance of our proposed method. In each run of the holdout, we randomly selected data of 3 participants as the testing dataset and data of the remaining 9 participants as the training dataset. To find the best number of hidden states, we varied the number of hidden states from 2 to 10 for the HMM and ran 100 holdouts for each number of hidden states. The results indicate that 2 was the best number of hidden states.

We then ran another 100 holdouts to evaluate the performance of the HMM for workload estimation. Precision, recall and F_1 score were used as performance metrics, where precision = $\frac{\# \text{true positives}}{\# \text{true positives} + \# \text{false positives}}$ and recall = $\frac{\# \text{true positives}}{\# \text{true positives} + \# \text{false negatives}}$. For our multi-classification problem, the precision is the mean precision of all classes and the recall is the mean recall of all classes. $F_1 = 2 \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$. Table 1 shows the mean and standard error of each performance metric. The results show that the HMM model achieved a 0.66 F_1 score, 0.67 precision, and 0.66 recall.

4. Experiment 2

4.1. Introduction

In Experiment 2, we tested two haptic shared control schemes: the adaptive haptic shared control and the non-adaptive haptic shared control schemes. The adaptive haptic shared control scheme adapted to the estimated real-time workload, and the participant's eyes on road and torque input. We used the HMM learned with the data from all the 12 participants to estimate the participant's workload in real time.

4.2. Method

4.2.1. Participants

A total of 13 students participated in the experiment. Data of 1 participant were discarded due to the wrong experiment setup. The remaining 12 participants were on average 22.3 years old ($SD = 3.7$ years) and had an average of 5.7 years of driving experience ($SD = 3.9$ years). All participants had normal or corrected-to-normal vision.

4.2.2. Apparatus and stimuli

The same dual-task shared control simulation platform was used in this experiment as in Experiment 1. Both the adaptive haptic shared control and the non-adaptive haptic shared control were used in this experiment.

4.2.3. Experimental design

The experiment used a within-subjects design with two independent variables. The first independent variable was the haptic shared control scheme (adaptive haptic shared control vs. non-adaptive haptic shared control). The second independent variable was the surveillance task urgency (1.5 vs. 6.5 s). Each participant experienced four tracks in the experiment. On each track, one type of haptic shared control scheme was used. In addition, each track was segmented into two portions, one portion with high urgency surveillance task (1.5 s) and the other with low urgency surveillance task (6.5 s). The resulting four test conditions

Table 2
Four test conditions in Experiment 2.

Condition	Surveillance task urgency		Haptic shared control scheme
	First half of the track	Second half of the track	
1	1.5 s	6.5 s	Non-adaptive
2	1.5 s	6.5 s	Adaptive
3	6.5 s	1.5 s	Non-adaptive
4	6.5 s	1.5 s	Adaptive

Table 3
Cost function weights in NMPC in Experiment 2.

Weight	w_1	w_2	w_3
Value	10.0	1.5	0.001

are shown in Table 2. The presentation of test conditions followed a 4×4 Latin square design to eliminate potential order effects.

The autonomy implemented in this experiment was governed by the weights shown in Table 3.

4.2.4. Measures

Five dependent variables were collected in the experiment: participants' self-reported workload and trust in the shared control autonomy, participants' control effort, driving task performance, and surveillance task performance. After each track, participants reported their workload and trust for the first and the second half of the track using two uni-dimensional scales. The NASA TLX survey (Hart and Staveland, 1988b) and Jian's trust survey (Jian et al., 2000) were presented to the participants such that they understood the meaning of workload and trust. Participants' control effort was calculated as the average torque that a participant applied on the steering wheel. Driving task performance was evaluated by lane keeping error. The lane keeping error is calculated as the mean of the absolute deviation of the vehicle's position from the centerline. The torque from human was estimated from a nonlinear steering wheel model, which captures the relationship between the torque and steering angle based on experimental data. The surveillance task performance was measured using the detection accuracy.

4.2.5. Experimental procedure

Participants provided a signed informed consent and filled in a demographic survey. After that they were assisted to wear the eye tracker with calibration. With the normal room light and without any specific tasks, the experimenter measured each participants' baseline pupil diameter twice, each about 30 s before the training.

During the training session, the participants first performed two trials of driving task only, one with the non-adaptive haptic shared control and one with the adaptive haptic shared control. Each trial took approximately 1.5 min. Then the participants performed three trials of the surveillance task only. Each trial took approximately 60 s. After that, the participants performed four trials of the combined driving and

surveillance task.

During the official experiment, participants performed the driving task and the surveillance task on 4 different tracks with different test cases as described in Table 2. Each trial took approximately 3 min. After each trial, the participants were asked to fill a post survey about the workload and trust during each portion of the track.

4.3. Experiment 2 results

Two-way repeated measures analysis of variance (ANOVA) was conducted with the shared control scheme and the surveillance task urgency as the within-subjects variables. Results are reported as significant for $\alpha < .05$. Table 4 summarizes the mean and standard error (SE) values of the participants' self-reported workload and trust as well as driving task performance, surveillance task performance and their exerted torque.

4.3.1. Participants' workload

Both control scheme and surveillance task urgency influence participants' self-reported workload. With the adaptive shared control, participants reported lower workload ($F(1, 11) = 5.18, p = .044$). When the surveillance task was less urgent, participants reported lower workload ($F(1, 11) = 20.26, p < .001$) (Fig. 9a).

4.3.2. Trust in automation

Participants trusted the shared control autonomy more when the autonomy was adaptive ($F(1, 11) = 12.76, p = .004$). The effect of surveillance task urgency on trust was not significant (Fig. 9b).

4.3.3. Driving task performance

Results revealed that the haptic shared control scheme and the surveillance task urgency significantly affected the driving task performance. Participants had smaller lane keeping errors when using the adaptive shared control autonomy ($F(1, 11) = 7.593, p = .019$), and when the surveillance task was less urgent ($F(1, 11) = 96.33, p < 0.001$) (Fig. 10a). There was also an interactive effect between the control scheme and surveillance task urgency ($F(1, 11) = 6.141, p = .031$). Using adaptive shared control led to a large reduction in lane keeping error when the surveillance task was more urgent. Fig. 11 shows the time history of lane keeping error of a representative subject, which illustrates the general tendency.

4.3.4. Surveillance task performance

For the surveillance task, task urgency significantly influenced the detection accuracy ($F(1, 11) = 6.73, p = .025$). Detection accuracy was higher when the task was less urgent. The effect of the shared control scheme was non-significant (Fig. 10b).

4.3.5. Participants' control effort

There was a significant effect of shared control scheme on participants' control effort ($F(1, 11) = 217.66, p < .001$) (Fig. 12). With adaptive shared control, participants exerted significantly less control effort. The effect of surveillance task urgency on participants' control effort was non-significant. In addition, results revealed a significant

Table 4
Mean and standard error (SE) of workload, trust, lane keeping error, detection accuracy and torque.

Metrics	N	Surveillance task urgency			
		1.5 s		6.5 s	
		Adaptive	Non-adaptive	Adaptive	Non-adaptive
Workload	12	13.96 ± 0.82	14.08 ± 0.87	7.83 ± 0.81	8.71 ± 0.97
Trust	12	4.04 ± 0.37	3.63 ± 0.30	3.92 ± 0.32	3.29 ± 0.38
Lane keeping error (m)	12	0.28 ± 0.033	0.36 ± 0.045	0.21 ± 0.03	0.26 ± 0.04
Detection accuracy (%)	12	93.43 ± 1.38	91.86 ± 1.13	94.30 ± 1.77	96.54 ± 1.18
Torque (Nm)	12	0.36 ± 0.03	0.73 ± 0.03	0.30 ± 0.02	0.79 ± 0.01

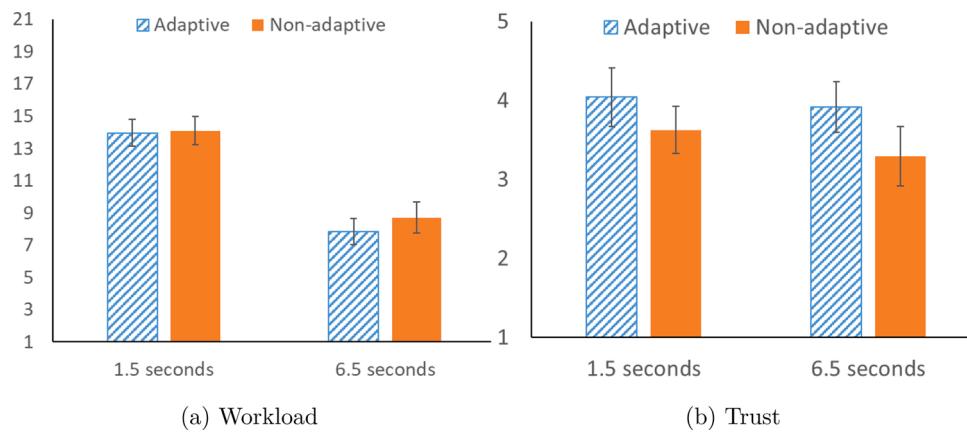


Fig. 9. Mean and standard error (SE) values of self-reported workload and trust.

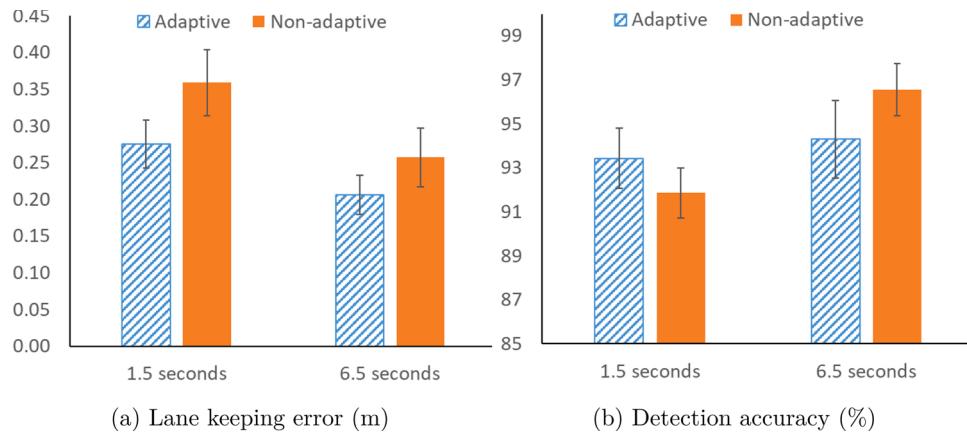


Fig. 10. Mean and standard error (SE) values of lane keeping error (m) and surveillance task detection accuracy (%).

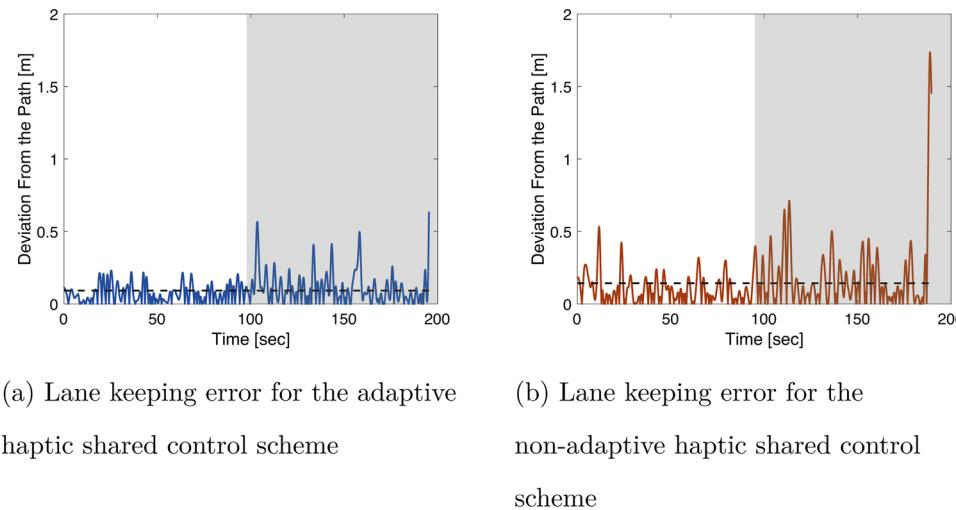


Fig. 11. Lane keeping performance in a representative average case. The track is segmented into two portions. The surveillance task urgency is low during the first portion (white background) while it is high in the second portion (shaded background). (a) and (b) show the time history of lane keeping error for the adaptive and non-adaptive haptic shared control scheme, respectively. The blue line shows the adaptive case while the orange line shows the non-adaptive case. The black dashed line represents the average deviation throughout the entire track. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

interaction effect between control scheme and surveillance task urgency ($F(1, 11) = 11.42, p = .006$). When the surveillance task was less urgent (6.5 s), the adaptive shared control scheme led to a larger drop in torque. Fig. 13 shows the representative torque from autonomy and human. It supports the conclusion that participants exerted less control effort in adaptive cases.

5. Discussion

5.1. Participants' workload

Participants' self-reported workload decreased when using the adaptive shared control scheme and when the surveillance task became less urgent. The results could have resulted from the following reasons. First, the 6.5 s surveillance task urgency imposed a smaller temporal

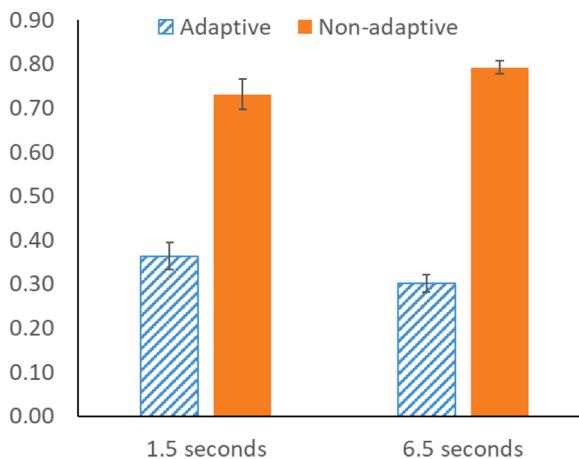


Fig. 12. Mean and standard error (SE) values of participants' torque (N m).

demand on participants. Second, the participants' control effort was smaller with the adaptive control scheme. Third, participants' driving task performance was higher with the adaptive control scheme and when the surveillance task was less urgent.

5.2. Trust in automation

Our result is consistent with prior research that human operators' trust in automation is determined by the autonomy's performance (Yang et al., 2017; Du et al., 2020; Guo and Yang, 2020). Human operators perceived both the driving and the surveillance task performance continuously, based on which they adjusted their trust in automation. As the driving task performance increased with the adaptive control scheme, trust increased accordingly.

5.3. Driving task performance

The results showed that the adaptive shared control scheme benefited the driving task performance, especially when participants were under a high workload. Based on the design of the adaptive haptic shared control scheme, with the same input torque, when the human operator has a high workload and focuses on the surveillance task, the assistance level is increased. The increment in the assistance level is

expected to aid the driving task and reduce the lane keeping error. This design principle was supported by the experimental results.

5.4. Surveillance task performance

As the surveillance task became more urgent and more demanding, the surveillance task performance decreased significantly. This result is consistent with prior research that when workload increased from a moderate to a high level, task performance would decrease (Lu et al., 2019).

5.5. Participants' control effort

Our results indicate that with adaptive shared control participants exerted significantly less amount of control effort in both low and high workload conditions. The results can be explained as follows: First, as the participants' trust toward the adaptive shared control scheme is significantly higher than the non-adaptive control scheme, participants had a higher tendency to yield to the autonomy, resulting in smaller input torque. Second, according to the design of the adaptive shared control scheme, with the same input torque, when the human operator experiences moderate workload and focuses on the driving task, the assistance level is reduced. With a reduced assistance level, regardless of whether the human yields to or fights with the autonomy, the human operator's torque is expected to be smaller.

6. Conclusion

In this study, we developed an adaptive haptic shared control scheme by designing a heuristic function for assistance level considering human's workload, torque, and eyes on road. The results indicate that our adaptive haptic shared control scheme leads to lower self-reported workload, higher trust in automation, lower lane keeping error, and lower human control effort. To our best knowledge, this is the first study in which the human operator's workload was estimated in real time and used as an input to an adaptive haptic shared control scheme.

The findings should be viewed in light of the following limitations. First, a group-level workload estimation model was developed in the study, ignoring potential individual differences. In this study, we assumed that participants experienced high workload when the surveillance task was more urgent and moderate workload when surveillance task was less urgent. This assumption may not hold for different

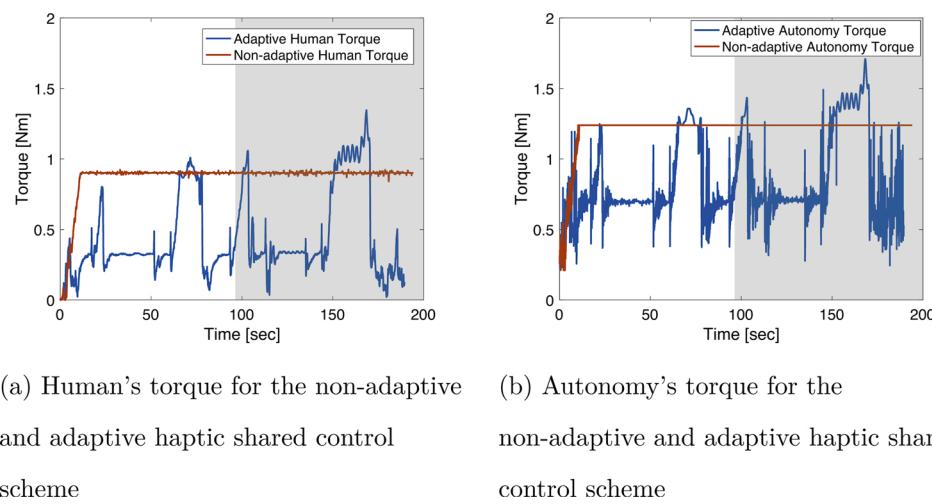


Fig. 13. Human and autonomy's torque in a representative average case. The track is segmented into two portions. The surveillance task urgency is low during the first portion (white background) while it is high in the second portion (shaded background). (a) and (b) show the torque for human and autonomy, respectively, both with the adaptive and non-adaptive haptic shared control scheme. The blue line shows the adaptive case while the orange line shows the non-adaptive case. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

(a) Human's torque for the non-adaptive and adaptive haptic shared control scheme

(b) Autonomy's torque for the non-adaptive and adaptive haptic shared control scheme

individuals. In the future work, we will develop a personalized workload estimation model to account for individual differences, as such personalized models have shown improved performance over group averaged data in terms of other metrics like distraction (Ersal et al., 2010).

Second, the present study estimated workload by using HMM to analyze participants' gaze trajectory. Other types of eye-related features such as pupil diameter (Recarte and Nunes, 2003) and blink rate (Halverson et al., 2012) have also been used to estimate workload. In our future work, we will exploit other features and use our previously proposed Bayesian Inference model (Luo et al., 2019) to combine different machine learning models that work best for different features.

Third, the assistance level adaptation function is manually designed based on heuristics. Future work can investigate other methods to find the optimal assistance level adaptation function.

Fourth, this study focuses on only a single-vehicle scenario with fixed speed. Further research should extend to a mixed-traffic scenario with varying speeds. When other vehicles are present in the environment, the impact from interactions with them will be important and can be captured with various control schemes (Kerner, 2021, 2018b,a). In addition, the impact of various surveillance tasks can also affect the performance when mixed-traffic is considered. Studying these combined effects is subject to future research.

Authors' contribution

Ruikun Luo and Yifan Weng: conceptualization, methodology, software, formal analysis, writing – original draft. Yifan Wang: methodology, software, formal analysis. Paramsothy Jayakumar, Mark J. Brudnak, Victor Paul and Vishnu R. Desaraju: conceptualization, methodology. Jeffrey L. Stein, Tulga Ersal and X. Jessie Yang: conceptualization, methodology, writing – reviewing and editing, resources, supervision, funding acquisition.

Conflict of interest

The authors have no conflict of interest.

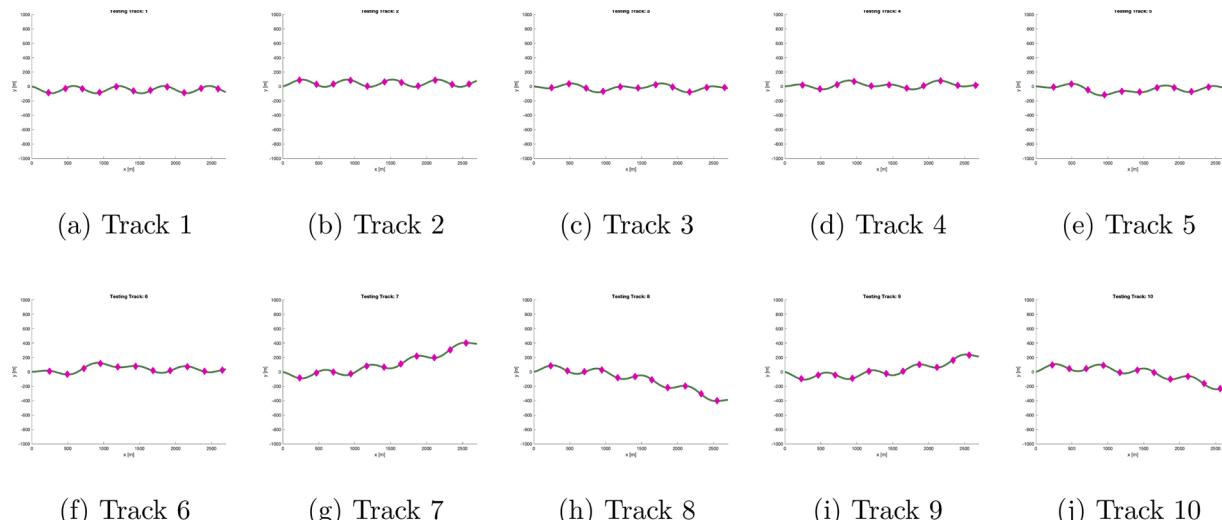


Fig. A1. Candidate tracks. Magenta dots indicate the locations where the participants reported the difficulty of driving. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Acknowledgment

We acknowledge the technical and financial support of the Automotive Research Center (ARC) in accordance with Cooperative Agreement W56HZV-19-2-0001 U.S. Army CCDC Ground Vehicle Systems Center (GVSC), Warren, MI.

Appendix A. Pilot study 1 – track selection

In pilot study 1, we developed and selected 6 driving tracks with two considerations. First, the driving tracks should have the same difficulty. Second, along each track, the difficulty at every point should be roughly the same. The two considerations ensure that the difficulty of the dual-task mission can be easily manipulated by varying the surveillance task urgency, because the difficulty of the driving task is fairly constant.

A.1 Participants

10 participants (Age: Mean = 21.8 years, $SD = 2.7$ years) took part in pilot study 1. All participants had normal or corrected-to-normal vision and hearing, with an average of 4.1 years of driving experience ($SD = 1.7$ years).

A.2 Apparatus and stimuli

Pilot study 1 used the same driving simulator as in Experiment 1 with driving task only and non-adaptive haptic shared control scheme.

A.3 Experimental design

The pilot study used a within-subjects design with 10 different candidate tracks (Fig. A1). The presentation of tracks followed a 10×10 Latin square design to eliminate potential order effects.

A.4 Measures

Along each track, participants reported the difficulty of driving at 11 locations using a 7-point Likert scale (1: easiest; 7: most difficult). After completing each track, participants also evaluated to what extent the track had the same difficulty anywhere along it (i.e., uniformity score), using another 7-point Likert scale (1: the same; 7: significantly different). For each track, the average of the 11 reported difficulty scores was calculated to represent the overall difficulty of the track (i.e., overall difficulty score).

A.5 Result

One-way repeated measures Analysis of Variance (ANOVA) was conducted with the driving track as the within-subjects variable. The results showed a non-significant difference between the ten tracks in their overall difficulty scores ($F(9, 81) = 1.161, p = 0.331$) and in their uniformity score ($F(9, 81) = 0.557, p = 0.828$). Based on the results, we selected track 2, 3, 5, 6, 8, 9 to be used in Pilot study 2 and Experiment 1, and track 2, 3, 6, 9 to be used in Experiment 2.

Appendix B. Pilot study 2 – design of surveillance task

We aimed to manipulate the difficulty of the dual-task mission and hence the human operators' workload by varying the surveillance task urgency. In Pilot study 2, we selected the paces of the surveillance task, so that the difficulty and workload of the dual-task mission can be manipulated.

B.1 Participants

A total of 7 students participated in Pilot study 2. Data of one participant was discarded due to equipment malfunction. The remaining 6 participants were on average 25.3 years old ($SD = 1.6$ years) and had an average of 2.7 years of driving experience ($SD = 1.6$ years). All participants had normal or corrected-to-normal vision.

B.2 Apparatus and stimuli

Pilot study 2 used the same dual-task shared control simulation platform as in Experiment 1. The non-adaptive haptic shared control scheme was applied.

B.3 Experimental design

The pilot study used a within-subject design with six different time intervals of the surveillance task: 1.5, 2.5, 3.5, 4.5, 5.5, and 6.5 s, i.e., participants had to complete the detection task within any given time interval. The six time intervals were selected based on the results from our previous study (Luo et al., 2019). Participants performed both the driving task and the surveillance task on 6 different tracks, each with a constant different time interval. The presentation of surveillance task conditions followed a 6×6 Latin square design to eliminate potential order effects.

B.4 Measures

Participants reported their workload of the dual-task mission using the NASA TLX survey (Hart and Staveland, 1988a) and their perceived difficulty of the dual-task mission.

B.5 Result

One-way repeated measures Analysis of Variance (ANOVA) was conducted with the surveillance time interval as the within-subjects variable. The results showed a significant difference of time interval

on workload ($F(5, 25) = 10.458, p < 0.001$) and difficulty ($F(5, 25) = 13.423, p < 0.001$). We then performed a series of *t* tests between different pairs of time intervals. The results revealed significant differences in workload and difficulty between 1.5 and 2.5 s (workload: $p < .001$, difficulty: $p = .006$), between 1.5 and 3.5 s (workload: $p = .005$, difficulty: $p = .012$), between 1.5 and 4.5 s (workload: $p = .004$, difficulty: $p = .006$), between 1.5 and 5.5 s (workload: $p = .001$, difficulty: $p < .001$), and between 1.5 and 6.5 s (workload: $p = .004$, difficulty: $p < .001$). The differences between any other pairs of time intervals were non-significant.

Based on the results, we selected 1.5 and 6.5 s time intervals to be used in the Experiment 1 and Experiment 2 to induce varying levels of workload. Note in Experiment 1, we also included the 2.5 s time interval, as we were interested to explore participants' performance with a slightly larger time interval compared to the 1.5 s time interval.

References

- Anderson, S.J., Peters, S.C., Pilutti, T.E., Iagnemma, K., 2011. Design and development of an optimal-control-based framework for trajectory planning, threat assessment, and semi-autonomous control of passenger vehicles in hazard avoidance scenarios. In: Pradlir, C., Siegwart, R., Hirzinger, G. (Eds.), Robotics Research. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 39–54.
- Bhardwaj, A., Ghasemi, A., Zheng, Y., Febbo, H., Jayakumar, P., Ersal, T., Gillespie, B., 2020. Who's the boss? Arbitrating control authority between a human driver and automation system. *Transp. Res. Part F: Psychol. Behav.* 68, 144–160.
- Calinon, S., 2016. A tutorial on task-parameterized movement learning and retrieval. *Intell. Serv. Robot.* 9 (1), 1–29.
- Calinon, S., Billard, A., 2005. Recognition and reproduction of gestures using a probabilistic framework combining PCA, ICA and HMM. *Proceedings of the 22nd International Conference on Machine Learning* 105–112.
- Chen, W., Jaques, N., Taylor, S., Sano, A., Fedor, S., Picard, R.W., 2015. Wavelet-based motion artifact removal for electrodermal activity. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* 6223–6226.
- Diaz-Piedra, C., Rieiro, H., Di Stasi, L.L., 2021. Monitoring army drivers' workload during off-road missions: an experimental controlled field study. *Saf. Sci.* 134, 105092.
- Du, N., Huang, K.Y., Yang, X.J., 2020. Not all information is equal: effects of disclosing different types of likelihood information on trust, compliance and reliance, and task performance in human-automation teaming. *Human Factors* 62 (6), 987–1001.
- Erlien, S.M., Fujita, S., Gerdes, J.C., 2016. Shared steering control using safe envelopes for obstacle avoidance and vehicle stability. *IEEE Trans. Intell. Transp. Syst.* 17 (2), 441–451.
- Ersal, T., Fuller, H.J.A., Tsimhoni, O., Stein, J.L., Fathy, H.K., 2010. Model-based analysis and classification of driver distraction under secondary tasks. *IEEE Trans. Intell. Transp. Syst.* 11 (3), 692–701.
- Febbo, H., 2017. Nloptcontrol.jl. <https://github.com/JuliaMPC/NLOptControl.jl>.
- Febbo, H., Liu, J., Jayakumar, P., Stein, J.L., Ersal, T., 2017. Moving obstacle avoidance for large, high-speed autonomous ground vehicles. *2017 American Control Conference (ACC)* 5568–5573.
- Flemisch, F., Nashashibi, F., Rauch, N., Schieber, A., Glaser, S., Temme, G., Kaußner, A., 2010. Towards highly automated driving: intermediate report on the haveit-joint system. *Proc. 3rd Eur. Road Transp. Res. Arena*.
- Fridman, L., Reimer, B., Mehler, B., Freeman, W.T., 2018. Cognitive load estimation in the wild. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* 652.
- Griffiths, P.G., Gillespie, R.B., 2005. Sharing control between humans and automation using haptic interface: primary and secondary task performance benefits. *Human Factors* 47 (3), 574–590.
- Guo, Y., Yang, X.J., 2020. Modeling and predicting trust dynamics in human–robot teaming: a Bayesian inference approach. *Int. J. Soc. Robot.* <https://doi.org/10.1007/s12369-020-00703-3>.
- Halverson, T., Estep, J., Christensen, J., Monnin, J., 2012. Classifying workload with eye movements in a complex task. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 56 (1) 168–172. <https://doi.org/10.1177/1071181312561012>.
- Hart, S.G., Staveland, L.E., 1988a. Development of nasa-tlx (task load index): results of empirical and theoretical research. *Advances in Psychology*, Vol. 52 139–183.
- Hart, S.G., Staveland, L.E., 1988b. Development of nasa-tlx (task load index): results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (Eds.), *Human Mental Workload*, Vol. 52. North-Holland, pp. 139–183.
- Heard, J., Harriott, C.E., Adams, J.A., 2018. A survey of workload assessment algorithms. *IEEE Trans. Human-Mach. Syst.* 48 (5), 434–451. <https://doi.org/10.1109/THMS.2017.2782483>.
- Jian, J.-Y., Bisantz, A.M., Drury, C.G., 2000. Foundations for an empirically determined scale of trust in automated systems. *Int. J. Cogn. Ergonom.* 4 (1), 53–71.
- Kerner, B.S., 2018a. Autonomous driving in framework of three-phase traffic theory. *Proc. Comput. Sci.* 130, 785–790. The 9th International Conference on Ambient Systems, Networks and Technologies (ANT 2018)/The 8th International Conference on Sustainable Energy Information Technology (SEIT-2018)/Affiliated Workshops.

- Kerner, B.S., 2018b. Physics of automated driving in framework of three-phase traffic theory. *Phys. Rev. E* 97, 042303.
- Kerner, B.S., 2021. Effect of autonomous driving on traffic breakdown in mixed traffic flow: a comparison of classical ACC with three-traffic-phase-acc (TPACC). *Phys. A: Stat. Mech. Appl.* 562, 125315.
- Kim, J.-H., 2009. Estimating classification error rate: repeated cross-validation, repeated hold-out and bootstrap. *Comput. Stat. Data Anal.* 53 (11), 3735–3745.
- Liu, J., Jayakumar, P., Stein, J.L., Ersal, T., 2017a. Combined speed and steering control in high speed autonomous ground vehicles for obstacle avoidance using model predictive control. *IEEE Trans. Vehic. Technol.* 66 (10), 8746–8763.
- Liu, J., Jayakumar, P., Stein, J.L., Ersal, T., 2018. A nonlinear model predictive control formulation for obstacle avoidance in high-speed autonomous ground vehicles in unstructured environments. *Vehic. Syst. Dyn.* 56 (6), 853–882.
- Liu, Y., Ayaz, H., Shewokis, P.A., 2017b. Multisubject “learning” for mental workload classification using concurrent EEG, fNIRS, and physiological measures. *Front. Human Neurosci.* 11, 389.
- Lu, S., Zhang, M.Y., Ersal, T., Yang, X.J., 2019. Workload management in teleoperation of unmanned ground vehicles: effects of a delay compensation aid on human operators’ workload and teleoperation performance. *Int. J. Human-Comput. Interact.* 1–11.
- Luo, R., Wang, Y., Weng, Y., Paul, V., Brudnak, M.J., Jayakumar, P., Yang, X.J., 2019. Toward real-time assessment of workload: a Bayesian inference approach. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 63 (1). SAGE Publications, pp. 196–200.
- Mulder, M., Abbink, D.A., Boer, E.R., 2008. The effect of haptic guidance on curve negotiation behavior of young, experienced drivers. *2008 IEEE International Conference on Systems, Man and Cybernetics* 804–809.
- Nguyen, A.-T., Sentouh, C., Popieul, J.-C., 2018. Sensor reduction for driver-automation shared steering control via an adaptive authority allocation strategy. *IEEE/ASME Trans. Mechatron.* 23 (1), 5–16. <https://doi.org/10.1109/TMECH.2017.2698216>.
- Petermeijer, S.M., Abbink, D.A., de Winter, J.C.F., 2015. Should drivers be operating within an automation-free bandwidth? Evaluating haptic steering support systems with different levels of authority. *Human Factors* 57 (1), 5–20.
- Recarte, M.A., Nunes, L.M., 2003. Mental workload while driving: effects on visual search, discrimination, and decision making. *J. Exp. Psychol.: Appl.* 9 (2), 119.
- Reimer, B., 2009. Impact of cognitive task complexity on drivers’ visual tunneling. *Transp. Res. Rec.* 2138 (1), 13–19.
- Rozo, L., Silverio, J., Calinon, S., Caldwell, D.G., 2016. Learning controllers for reactive and proactive behaviors in human-robot collaboration. *Front. Robot. AI* 3, 30.
- SAE International, 2016. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (Standard J3016_201609). SAE International.
- Schwarting, W., Alonso-Mora, J., Pauli, L., Karaman, S., Rus, D., 2017. Parallel autonomy in automated vehicles: safe motion generation with minimal intervention. *2017 IEEE International Conference on Robotics and Automation (ICRA)* 1928–1935.
- Schwarz, G., et al., 1978. Estimating the dimension of a model. *Ann. Stat.*
- Storms, J., Chen, K., Tilbury, D., 2017. A shared control method for obstacle avoidance with mobile robots and its interaction with communication delay. *Int. J. Robot. Res.* 36 (5–7), 820–839.
- Wächter, A., Biegler, L.T., 2006. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.* 106 (1), 25–57.
- Wang, Y., Reimer, B., Dobres, J., Mehler, B., 2014. The sensitivity of different methodologies for characterizing drivers’ gaze concentration under increased cognitive demand. *Transp. Res. Part F: Traff. Psychol. Behav.* 26, 227–237.
- Yang, X.J., Unhelkar, V.V., Li, K., Shah, J.A., 2017. Evaluating effects of user experience and system transparency on trust in automation. *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* 408–416.
- Zhang, Y., Owechko, Y., Zhang, J., 2004. Driver cognitive workload estimation: a data-driven perspective. *Proceedings. The 7th International IEEE Conference on Intelligent Transportation Systems (IEEE Cat. No. 04th8749)* 642–647.