Human Factors in Design, Engineering, and Computing, Vol. 238, 2026, 1552–1559
https://doi.org/10.54941/ahfe1008383

Detection-Aware Generation of Fatigue-Conditioned Ocular and Head Behaviors for Virtual Drivers

Woojin Kim[1], Myeongkyu Lee[2], Seung-Jun Lee[1], Yonghyun Kim[1], and Brandon J. Pitts[2]

[1]Electronics and Telecommunications Research Institute, Daejeon, 34129, Republic of Korea
[2]Purdue University, West Lafayette, IN 47907, USA

ABSTRACT

Driver fatigue induces systematic changes in observable behaviors, which should be represented in virtual drivers beyond predefined animations. This study presents a data-driven framework for generating stochastic ocular and head behaviors conditioned on subjective sleepiness. Empirical behaviors were obtained from a controlled driving experiment and characterized across Alert, Intermediate, and Drowsy states based on the Karolinska Sleepiness Scale (KSS). Significant fatigue-dependent changes were observed in eyelid-closure characteristics, PERCLOS, and head-pitch behavior, and were incorporated into probabilistic behavioral models. A detection-aware generation method was developed to synthesize eyelid trajectories while accounting for discrepancies between intended behavior and behavior re-measured after detection. Monte Carlo validation showed that generated PERCLOS distributions reproduced empirical distributions across all KSS states, with normalized Wasserstein distances of 0.14–0.24 and no significant differences detected by Kolmogorov–Smirnov tests. These results provide quantitative evidence that the proposed generator reproduces fatigue-dependent characteristics and population-level variability observed in real drivers.

Keywords: Virtual driver, Driver fatigue, Behavioral generation, PERCLOS, KSS

INTRODUCTION

Driver fatigue induces measurable changes in ocular and head behaviors, such as prolonged eyelid closure, increased PERCLOS, and changes in head posture. These behavioral cues have been extensively studied for driver monitoring and fatigue estimation (Hassan et al., 2026; Jia et al., 2021; El-Nabi et al., 2024). As driving simulation increasingly incorporates virtual humans to represent realistic driver responses, their behaviors should reflect not only physical actions but also changes associated with internal states such as fatigue (Marekar et al., 2026; Oh et al., 2025; Kim et al., 2024). However, reproducing such fatigue-dependent behaviors in virtual drivers remains challenging, particularly when realistic behavioral variability needs to be represented.

Existing virtual human behaviors are often generated using predefined motions or independently specified parameters. Such approaches may

Received September 18, 2026; Revised October 18, 2026; Accepted October 27, 2026; Available online December 1, 2026

© 2026 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

reproduce individual actions but do not necessarily reflect the behavioral distributions observed in real drivers. A fatigue-controllable virtual driver should instead reproduce how multiple behaviors vary with fatigue while preserving natural variability across drivers.

This study develops a fatigue-conditioned behavioral generator grounded in empirical driver data. Ocular and head behaviors were characterized across Alert, Intermediate, and Drowsy states based on the Karolinska Sleepiness Scale (KSS), and significant behavioral characteristics were represented using probabilistic models. For eyelid behavior, a detection-aware generation method was further introduced because behavioral characteristics specified during generation can change after trajectory smoothing and event detection. Monte Carlo simulations were then used to evaluate whether the generated behaviors reproduce the PERCLOS distributions observed in real drivers. The resulting framework provides a data-grounded approach for generating virtual drivers whose observable fatigue behaviors reflect empirical human behavior.

EMPIRICAL DRIVER DATA

Empirical data were collected in a controlled driving experiment designed to capture driver behaviors across different levels of sleepiness. Facial behavior was recorded using a driver-facing camera, while subjective sleepiness was assessed using the Karolinska Sleepiness Scale (KSS) (Gillberg et al., 1994), ranging from 1 (extremely alert) to 9 (very sleepy), before driving and immediately after driving when fatigue was detected. Figure 1 shows the experimental setup and the sensors used for data collection.

For the present analysis, the initial and terminal 30-s segments of each driving session were used and paired with the corresponding pre- and post-drive KSS ratings, respectively. The KSS ratings were grouped into three states: Alert (KSS 1–3), Intermediate (KSS 4–6), and Drowsy (KSS 7–9). Ocular behavior was obtained from frame-level eyelid measurements, from which closure events and their temporal characteristics were extracted. Head pose measurements were obtained using a stereo depth camera with body tracking and used to characterize changes in head pitch and yaw. These measurements formed the empirical basis for identifying fatigue-dependent behavioral characteristics and constructing the subsequent behavior generator.

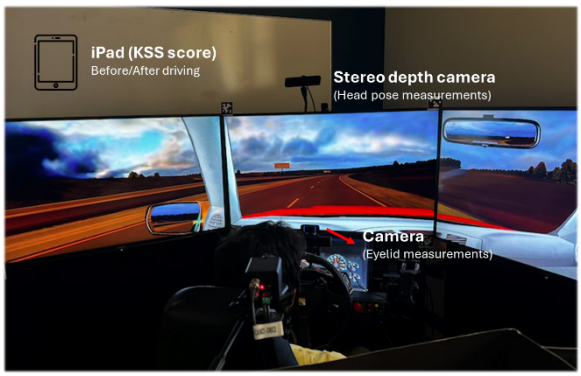


Figure 1: Experimental setup for collecting empirical driver behavior. Driver-facing facial recordings were synchronized with subjective sleepiness ratings using the Karolinska Sleepiness Scale (KSS).

FATIGUE-DEPENDENT BEHAVIOR CHARACTERIZATION

Ocular and head behaviors were compared across the three KSS states to identify fatigue-dependent characteristics for subsequent behavior generation. Group differences were evaluated using the Kruskal–Wallis tests, with Holm correction applied for multiple comparisons (Kruskal and Wallis, 1952; Holm, 1979), and effect sizes were quantified using epsilon-squared (ε^2). Table 1 summarizes the behavioral features showing significant differences or relevance to the subsequent generator.

Ocular behavior showed pronounced changes with increasing sleepiness. Mean and median closure durations increased from 0.176 and 0.161 s in the Alert state to 0.693 and 0.256 s in the Drowsy state, respectively. The increase was particularly evident in the upper tail of the duration distribution, with the 90th-percentile closure duration increasing from 0.247 to 1.769 s. Consistently, the probability of prolonged closures increased from 0.143 to 0.556. Closure depth also decreased with increasing sleepiness, indicating deeper eyelid closures. Among the head-behavior features, mean head pitch showed a significant state-dependent shift, whereas pitch variability did not differ significantly across states. This suggests that drowsiness was associated primarily with a sustained head-pitch bias rather than increased head-motion variability.

Derived PERCLOS increased markedly from 2.84% in the Alert state to 23.71% in the Drowsy state. PERCLOS was retained as a validation outcome rather than a direct generation parameter, allowing the generated trajectories to be evaluated against an emergent fatigue-dependent characteristic observed in the empirical data.

Table 1: KSS-dependent behavioral characteristics derived from the empirical driver data.

Feature	Alert	Intermediate	Drowsy	Holm p	ε^2
Mean closure duration (s)	0.176	0.212	0.693	<0.001	0.372
Median closure duration (s)	0.161	0.175	0.256	<0.001	0.326
P90 closure duration (s)	0.247	0.257	1.769	<0.001	0.305
Prolonged closure probability	0.143	0.188	0.556	<0.001	0.269
Median closure depth	0.529	0.508	0.418	0.001	0.176
Head pitch mean	0.003	0.017	0.149	<0.001	0.212
Head pitch SD	0.023	0.053	0.056	0.179	0.030
Head yaw mean	0.000	−0.014	0.065	0.025	0.086
Derived PERCLOS (%)	2.84	4.67	23.71	<0.001	0.267

FATIGUE-CONDITIONED BEHAVIOR GENERATOR

Probabilistic Behavior Model

The fatigue-dependent characteristics identified from the empirical data were represented as KSS-conditioned probabilistic models. Rather than assigning deterministic behavior to each fatigue state, the generator samples a set of behavioral parameters according to

$$\theta_{ocular} \sim p\left(\theta_{ocular} | KSS\right),$$

where θ_{ocular} includes closure duration, prolonged-closure probability, and closure depth. Figure 2 summarizes the empirical characteristics and their corresponding probabilistic representations.

Eyelid-closure duration was modeled separately for typical and prolonged closure events. For typical closures, the empirical median duration of each KSS state was used as the distribution anchor. For prolonged closures, defined as closures lasting at least 0.25 s, the empirical 90th-percentile (*P90*) duration was used to represent upper-tail behavior. Event durations were generated using lognormal distributions around these KSS-dependent anchors (Fig. 2a–b), allowing the model to represent both ordinary closures and the substantially longer closures observed with increasing sleepiness.

The probability of prolonged closure, P_{pro}, was modeled for each KSS state using a boundary-corrected reflected kernel density estimate (KDE) on [0,1] (Fig. 2c). Because P_{pro} and *P90* closure duration were correlated in the empirical data, they were not sampled independently. The generator first samples

$$P_{pro} \sim p\left(P_{pro} | KSS\right), \qquad T_{90} \sim p\left(T_{90} | P_{pro}, KSS\right),$$

where the conditional T_{90} distribution was constructed from neighboring empirical observations with similar prolonged-closure probabilities. This preserves the observed tendency for drivers with more frequent prolonged closures to also exhibit longer upper-tail closure durations.

Closure depth was modeled from the empirical distributions of median closure depth (Fig. 2d). Candidate parametric distributions were fitted separately for each KSS state, resulting in Normal distributions for Alert and Drowsy and a Weibull distribution for Intermediate. Head behavior was generated separately using the KSS-conditioned empirical characteristics identified in Section 3. In particular, head pitch incorporated the state-dependent mean postural shift, while yaw was treated as an auxiliary stochastic head-motion component.

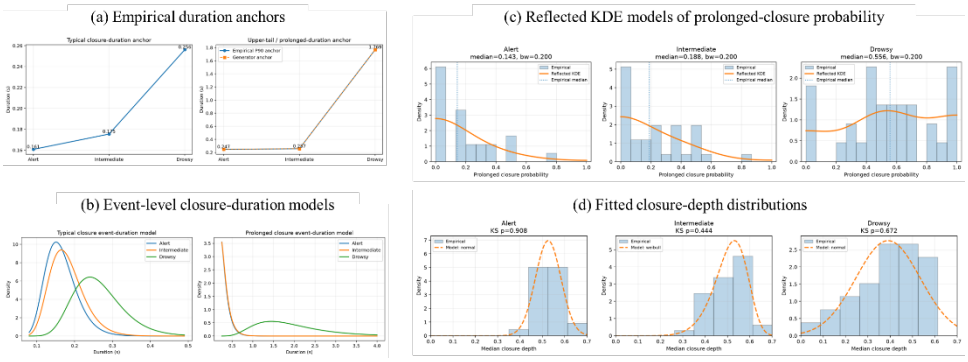


Figure 2: KSS-conditioned probabilistic models of ocular behavior derived from empirical driver data. (a) Empirical duration anchors, (b) event-level closure-duration models, (c) reflected KDE models of prolonged-closure probability, and (d) fitted closure-depth distributions.

Detection-Aware Generation

The sampled behavioral parameters were converted into continuous eyelid-opening trajectories. However, directly embedding a target closure duration or depth into a trajectory does not guarantee that the same characteristic will be recovered after behavioral detection. To maintain consistency with the empirical analysis, the same eyelid-event detector used for extracting the real-driver behaviors was applied to the generated trajectories. Eye opening was normalized using a baseline estimate and temporally smoothed over 0.12 s. A closure event was detected using hysteresis thresholds of 0.65 for closure and 0.75 for reopening, with a minimum event duration of 0.05 s. Closures lasting at least 0.25 s were classified as prolonged closures.

Consequently, smoothing and event-detection rules can cause the characteristics measured from a generated trajectory to differ from the originally sampled targets:

$$\theta_{detected} = D\left[x\left(t; \theta_{target}\right)\right] \neq \theta_{target},$$

where $x(t)$ is the generated eyelid-opening trajectory and $D(\cdot)$ denotes the eyelid-event detector described above. To reduce this discrepancy, the generated trajectory was repeatedly evaluated by the detector, and its generation parameters were adjusted based on the difference between the target and re-detected characteristics. This process continued until the detected behavior sufficiently approximated the sampled target:

$$\theta_{detected} \approx \theta_{target}.$$

Head pitch and yaw trajectories were generated separately from their KSS-conditioned empirical characteristics and were not included in the detection-aware adjustment.

VALIDATION OF GENERATED BEHAVIOR

To evaluate whether the proposed generator reproduces the population-level characteristics observed in real drivers, Monte Carlo simulations were conducted for each KSS state. A total of 200 independent 30-s trajectories were generated for each of the Alert, Intermediate, and Drowsy states. Behavioral parameters were independently sampled for each run from the corresponding KSS-conditioned models, and PERCLOS was subsequently calculated from the generated eyelid trajectory using the same detector applied to the empirical data. Importantly, PERCLOS was not directly specified as a generation target; it emerged from the generated temporal patterns of eyelid closure.

The generated PERCLOS distributions were compared with the empirical distributions using the Wasserstein distance (Khan and Zhang, 2022) and the two-sample Kolmogorov-Smirnov (KS) test (Lin et al., 2010) (Table 2). Median generated PERCLOS values were 3.00%, 4.42%, and 18.50% for Alert, Intermediate, and Drowsy states, respectively, compared with empirical medians of 2.84%, 4.67%, and 23.71%. Normalized Wasserstein distances

ranged from 0.142 to 0.244. The KS tests did not detect significant differences between the empirical and generated PERCLOS distributions for any KSS state ($p > 0.05$).

Table 2: Comparison of empirical and generated PERCLOS distributions across KSS states.

KSS State	Empirical Median (%)	Generated Median (%)	Normalized Wasserstein Distance	KS p
Alert	2.84	3.00	0.194	0.963
Intermediate	4.67	4.42	0.244	0.449
Drowsy	23.71	18.50	0.142	0.103

Figure 3 further illustrates the distributional agreement between empirical and generated PERCLOS. The generator reproduced the progressive shift toward higher PERCLOS with increasing sleepiness as well as substantial population-level variability, particularly in the Drowsy state. Although the generated Drowsy median was lower than the empirical median, its upper distributional range remained broadly represented.

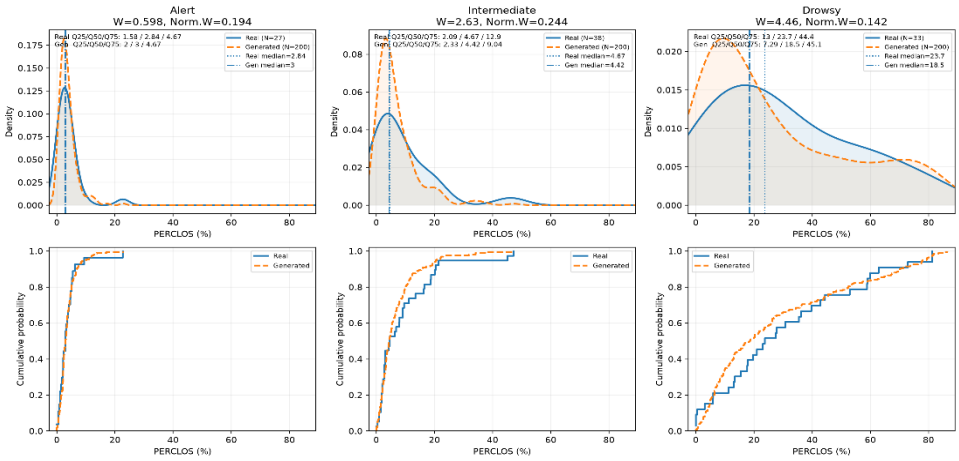


Figure 3: Empirical and generated PERCLOS distributions across KSS states. Empirical distributions are compared with 200 Monte-Carlo-generated 30-s trajectories for each fatigue state.

Finally, representative 30-s trajectories are shown in Fig. 4 to illustrate how the probabilistic models are expressed as continuous virtual-driver behavior. Compared with the Alert state, the Drowsy trajectory exhibits longer eyelid closures and a sustained downward head-pitch shift, consistent with the fatigue-dependent characteristics identified from the empirical data. These trajectories are illustrative examples rather than additional quantitative validation.

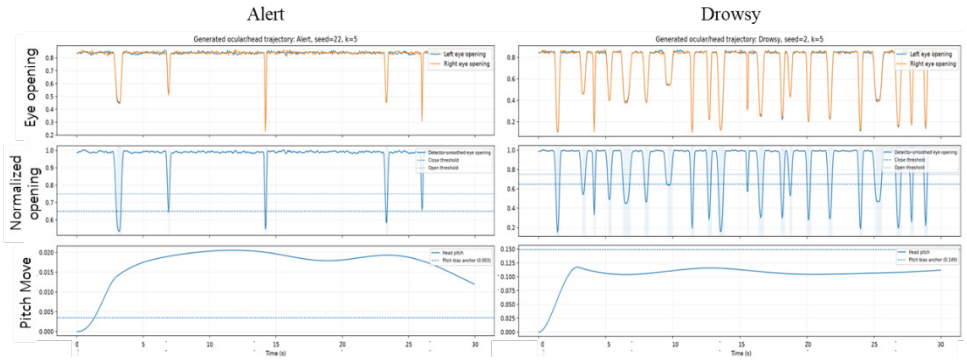


Figure 4: Representative 30-s ocular and head-behavior trajectories generated for Alert and Drowsy states. The examples illustrate longer eyelid closures and a sustained head-pitch shift in the Drowsy state.

DISCUSSION AND CONCLUSION

This study developed a fatigue-conditioned behavior generator grounded in empirical driver data. The results showed systematic changes in ocular behavior with increasing KSS, particularly in closure duration, prolonged-closure probability, and closure depth. Head behavior showed a significant shift in mean pitch, suggesting that increased sleepiness was associated with a sustained postural change rather than simply greater head-motion variability. These empirical characteristics were incorporated into probabilistic models to generate heterogeneous behaviors within each fatigue state.

A key aspect of the proposed approach is that the generated behavior was evaluated after being translated into continuous trajectories and re-detected using the same behavioral detector applied to the empirical data. PERCLOS was not directly imposed as a generation target, but emerged from the generated eyelid behavior. Monte Carlo simulations showed that the resulting PERCLOS distributions followed the fatigue-dependent shift and population-level variability observed in the empirical data, with no significant distributional differences detected by the KS tests across the three KSS states. Nevertheless, some discrepancies remained, including a lower median PERCLOS for generated Drowsy behavior, indicating that the model does not fully reproduce all aspects of the empirical distribution.

The present study focuses on behavioral trajectories rather than photorealistic driver representation or sensor-level simulation. Future work will therefore integrate the generator with embodied 3D virtual drivers and investigate whether the generated fatigue characteristics remain observable under variations in appearance, head pose, camera configuration, and environmental conditions. Overall, the proposed approach provides an empirically grounded method for generating fatigue-controllable virtual drivers and establishes a basis for incorporating realistic human-state variations into driving simulation.

ACKNOWLEDGMENT

This work was supported by Korea Evaluation Institute of Industrial Technology (KEIT) grant funded by the Ministry of Trade, Industry and Energy (MOTIE) (RS-2024-00403108, Development of DMS evaluation system based on virtual driver model).

REFERENCES

El-Nabi, S.A., El-Shafai, W., El-Rabaie, E.S.M., Ramadan, K.F., Abd El-Samie, F.E. and Mohsen, S. (2024), "Machine learning and deep learning techniques for driver fatigue and drowsiness detection: A review", Multimedia Tools and Applications, Vol. 83 No. 3, pp. 9441–9477.

Gillberg, M., Kecklund, G. and Åkerstedt, T. (1994), "Relations between performance and subjective ratings of sleepiness during a night awake", Sleep, Vol. 17 No. 3, pp. 236–241.

Hassan, O.F., Ibrahim, A.F., Makhlouf, M.A., Hafiz, B. and Gomaa, A. (2026), "Intelligent driver monitoring systems: A survey of drowsiness detection technologies for road safety", Artificial Intelligence Review, Vol. 59 No. 6, p. 137.

Holm, S. (1979), "A simple sequentially rejective multiple test procedure", Scandinavian Journal of Statistics, Vol. 6 No. 2, pp. 65–70.

Jia, H., Xiao, Z. and Ji, P. (2021), "Fatigue driving detection based on deep learning and multi-index fusion", IEEE Access, Vol. 9, pp. 147054–147062.

Khan, G. and Zhang, J. (2022), "When optimal transport meets information geometry", Information Geometry, Vol. 5 No. 1, pp. 47–78.

Kim, W., Oh, C., Kang, D.W., Lee, S.J., Kim, H.S., Jang, E.H. and Yoon, D. (2024), "System architecture for evaluating driver monitoring based on virtual driver model", in International Conference on Computational Science and Computational Intelligence, Springer Nature Switzerland, Cham, pp. 197–202.

Kruskal, W.H. and Wallis, W.A. (1952), "Use of ranks in one-criterion variance analysis", Journal of the American Statistical Association, Vol. 47, pp. 583–621.

Lin, P.C., Wu, B. and Watada, J. (2010), "Kolmogorov-Smirnov two sample test with continuous fuzzy data", in Integrated Uncertainty Management and Applications, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 175–186.

Marekar, A., Fehr, J., Haringer, M. and Blaume, B. (2026), "Advancing digital human modeling: real-time simulation and physics performance in Unreal Engine", in International Conference on Digital Human Modeling, Springer Nature Switzerland, Cham, pp. 170–188.

Oh, C., Kim, W. and Yoon, D. (2025), "Practical considerations in building a simulation-based driver monitoring evaluation system: Focus on virtual driver generation", in 2025 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE, pp. 6051–6054.