Technical Disclosure Commons

Defensive Publications Series

July 15, 2019

PREDICTING FUTURE GLANCE BEHAVIOR FROM GAZE PATTERN HISTORY

Michael Keller

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Recommended Citation
Keller, Michael, "PREDICTING FUTURE GLANCE BEHAVIOR FROM GAZE PATTERN HISTORY", Technical Disclosure Commons, (July 15, 2019)
https://www.tdcommons.org/dpubs_series/2354

This work is licensed under a Creative Commons Attribution 4.0 License.
This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

PREDICTING FUTURE GLANCE BEHAVIOR FROM GAZE PATTERN HISTORY

ABSTRACT

A computing system predicts where a gaze of a vehicle operator ("driver") will be in the near future to infer driver attention rather than inferring driver attention based on currently detected gaze data, which may be stale by the time the computing system processes the data and infers attention. The gaze may be predicted based on data from a camera that captures an image of the driver's eyes, head, posture, or other aspects. In some examples, the computing system uses a machine-learned model (ML model), such as a recurrent neural network (RNN) model to predict future glance behavior. The machine-learned model is trained using machine learning. The computing system may provide the predicted future driver gaze or glance information to application modules requiring driver gaze information, such as applications that monitor driver gaze and infer driver attention. In some examples, the computing system may use the gaze information based on the gaze prediction to take an action, such as to prompt the driver to return attention to the driving task.

DESCRIPTION

Operating a vehicle, such as an automobile, motorcycle, aircraft, marine craft, and the like, requires attention from the vehicle operator. Safe driving requires high levels of situational awareness of the driving task and of influences on the driving task. To maintain a high level of situational awareness for driving, the operator should pay attention to the road and be actively thinking about the driving task. When an operator engages in a secondary task, the operator may divert one or more of their eyes, hands, and mind away from the primary driving task. Diverting attention from the primary driving task too long may result in dangerous driving behavior, and potentially cause a vehicle accident. In addition, some vehicles offer driver assist features (e.g., lane keeping, adaptive speed control). These features may improve safety but carry a risk that the driver will lose vigilance over time and over-trust the automation provided by the driver assist features. Some regulatory bodies have begun mandating driver attention monitoring for vehicles to be manufactured in the future.

Some approaches to monitoring driver gaze may use camera systems that compute instantaneous gaze location and infer attention from the instantaneous value. Such approaches suffer significant problems. Each instantaneous frame may be inherently quite noisy. Automotive grade cameras have limited resolution and the angle from which they are mounted is restricted by the overall design of the vehicle. Second, predicting a driver's gaze with perfect accuracy from any single frame of data may be difficult or impossible. Some approaches attempt to address the noise issue by filtering (sometimes quite heavily) the instantaneous frames of video. However, this leads to a second problem of lag. Specifically, there is a processing delay in computing the images received from the cameras. If the output is filtered as is necessary to address the noise issue, the prediction is delayed further. Finally, there is inherent delay from when the vehicle system detects a problem to when the vehicle system can attempt countermeasures and when the vehicle operator responds to these measures. The total end-to-end delay may be on the order of hundreds of milliseconds.

As described herein, a vehicle computing system may predict where a gaze of a vehicle operator ("driver") will be in the near future rather than assessing and monitoring driver attention based on where the driver's gaze was in the past. In some examples, the computing system predicts gaze or glance behavior using a machine-learned model that predicts where a driver's gaze will be in the near future rather than assessing where it was in the past. This approach may significantly improve real-time attention monitoring assessment. In some examples, the computing system uses a baseline gaze model, and a driver-specific gaze model. The computing system may perform one or more actions in response to detecting that the operator is likely giving insufficient attention the driving task, and may take no action in response to detecting that the operator is likely giving sufficient attention to the driving task.

The example system shown in Figure 1 provides driver gaze information based on predicted driver gaze, such as to one or more applications requiring current driver gaze information. The system may use the driver gaze information to improve driving safety.

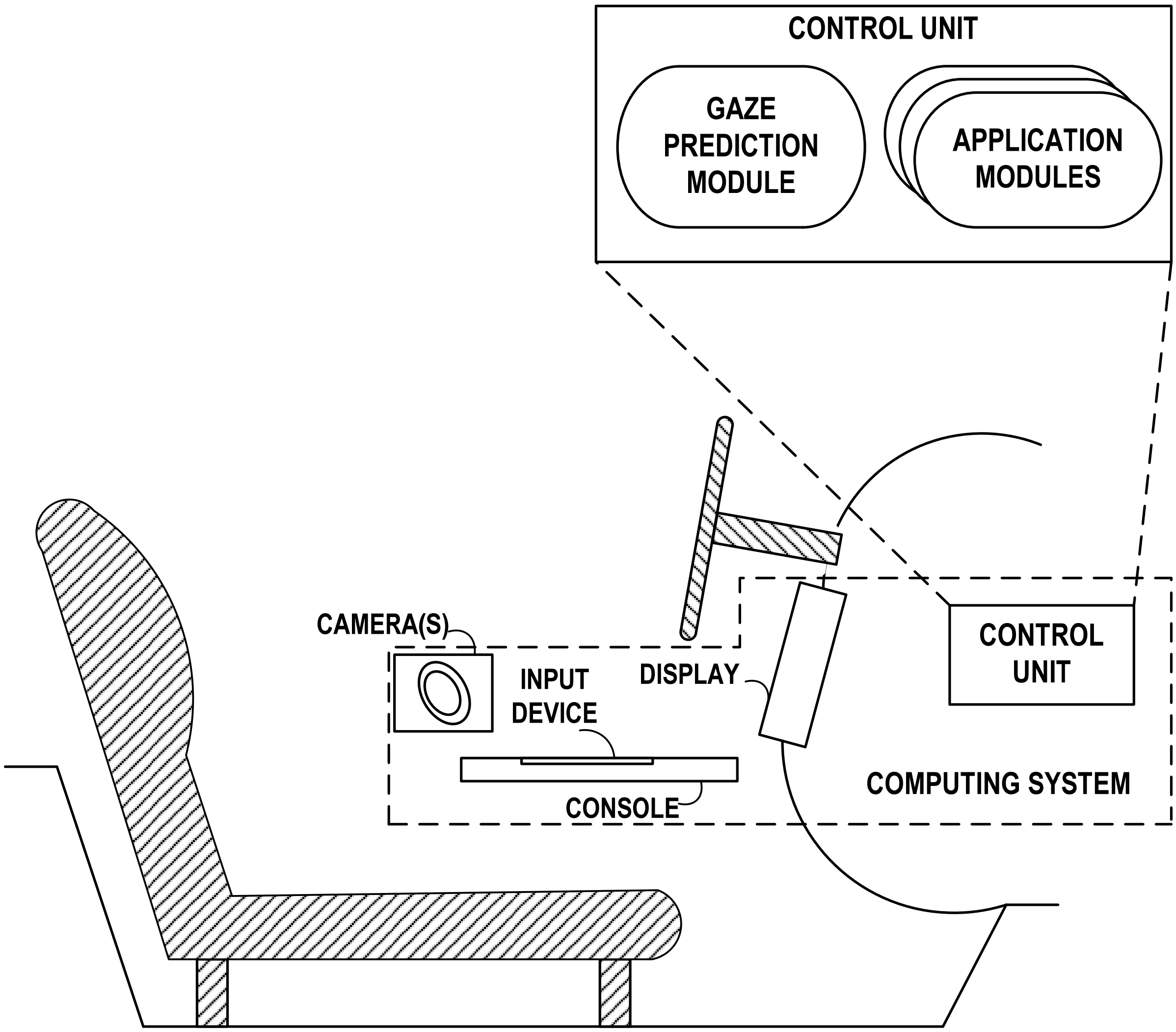

Figure 1

Figure 1 is a conceptual diagram illustrating an interior of a vehicle that includes a computing system configured to predict gaze behavior of an operator of the vehicle, in accordance with one or more techniques of this disclosure. As illustrated in Figure 1, the vehicle includes a computing system, seat, steering wheel, and dashboard. As described in further detail, the computing system is configured to use one or more machine-learned models to determine predicted future driver gaze or glance behavior based on raw gaze data, and provides information to application modules requiring driver gaze information, such as applications that monitor driver gaze and infer driver attention. A trained machine-learning model is referred to a machine-learned model. The machine-learned model is trained using machine learning techniques.

As illustrated in Figure 1, the vehicle may be an automobile, but aspects of the present disclosure may also be applicable to other types of vehicles, including trucks, motorcycles, aircraft, watercraft, trains, bicycles, or other vehicles. Further, the techniques described herein may be used in other contexts besides vehicle operation that require monitoring gaze behavior. For example, the techniques described herein may be used for control station monitoring (e.g., air travel controllers).

A collection of devices, components, and modules that may each be included in the computing system is also shown in Figure 1. The computing system includes, but is not limited to, one or more cameras (e.g., video cameras), an input device such as a presence-sensitive panel, a display and a control unit. One or more components of the computing system, such as the input device may be directly and physically accessible to occupants seated in the front driver and front passenger seats of the vehicle, and may be located within, near, or on a console (e.g., a center console). The input device may include a presence-sensitive panel that receives inputs for the computing system. In some examples, the input device may be integrated into the display such that the display may be a presence-sensitive display.

Although described for purposes of example as a component or system installed in a vehicle, in some examples, some or all aspects described with respect to the computing system may be housed within a computing device separate from the vehicle, such as a mobile computing device. For instance, the mobile computing device may be a mobile computing device (e.g., a smart phone) of an operator of the vehicle. In some examples, the computing device may be held by the operator or mounted to the dashboard or other portion of the vehicle (e.g., removably mounted). The computing device may allow the operator to control the computing device and/or applications executing on the computing device, such as controlling via one or more of a presence-sensitive display of the computing system and voice commands.

The control unit may provide an operating environment or platform for one or one more modules, such as a combination of hardware, firmware, and software. For instance, the control unit may include one or more processors and storage devices that may execute instructions and store data of one or more modules. The control unit may also be operably coupled to one or more other software and/or hardware components, including the cameras, the input device, and the display to control, configure, and/or communicate information with the components, to name only a few example operations.

In some examples, the computing system may include applications that operate to assist, inform, entertain, or perform other tasks that require user interactions with occupants of a vehicle. In some examples, the computing system may execute an in-vehicle infotainment (IVI) system, or may be a subcomponent thereof. In some examples, the computing system may include applications providing driver assist functions, including autonomous or semi-autonomous driving functions, that control or affect operation of the vehicle. For example, the computing system may include one or more application modules that perform functions or process information on behalf of one or more occupants of the vehicle, including the vehicle operator. In some examples, the computing system may be controlled through input detected by the input device, and/or through input detected by one or more additional input devices. The display may function as an output device, such as a display device. Based on user input, the display may present output to a user.

As described above, the computing system may include a gaze prediction module and one or more application modules. The gaze prediction module and application modules may perform operations described herein using software, hardware, firmware, or a mixture of both hardware, software, and firmware residing in and executing by the computing system or at one or more other remote computing devices. As such, the gaze prediction module and application modules may be implemented as hardware, software, and/or a combination of hardware and software. The computing system may execute the gaze prediction module and application modules, or one or more other modules as or within a virtual machine executing on underlying hardware. The gaze prediction module and application modules may be implemented in various ways. For example, the gaze prediction module and application modules may be implemented as a downloadable or pre-installed application or "app." In another example, the gaze prediction module and application modules may be implemented as part of an operating system of the computing system.

The vehicle may include a wide variety of sensors, which may be configured to provide output to the computing system. For instance, the vehicle may include a speed sensor, an acceleration sensor, a position sensor, and the like. In some examples, the computing system may be configured to communicate with the sensors via a network or bus of the vehicle, such as a component area network (CAN) bus.

In accordance with one or more techniques of this disclosure, the gaze prediction module of the computing system may predict where a gaze of a vehicle operator ("driver") will be in the near future rather than assessing where the driver's gaze was in the past. In some examples, the gaze prediction module uses a machine-learned model, such as a recurrent neural network (RNN) model. The recurrent neural network model receives one or more inputs, such as from sensors of the computing system, sensors of the vehicle, and applications executed by the computing system, for example.

Example inputs to the recurrent neural network model may include initial data indicative of the driver's gaze, such as data obtained by one or more cameras and/or sensors within the vehicle. The initial data indicative of the driver's gaze may be obtained by the computing system in any of a variety of ways, and may reflect, for example, positioning of the driver's body, head, eyes, or other data. Any suitable underlying gaze tracker technology may be used. In some examples, the recurrent neural network model may accept as sequential inputs the last N seconds (e.g., 10 or 30 seconds) of raw gaze data. Other example inputs to the recurrent neural network model may include environmental data (e.g., traffic, current mode) for each sequence of time. The recurrent neural network model then outputs to the gaze prediction module a gaze prediction indicating where the driver's gaze is expected to be X milliseconds (e.g., 100 ms) in the future. The recurrent neural network model implicitly handles instantaneous noise from the gaze detector inputs and dramatically improve the real-time predictive performance.

In some examples, the gaze prediction module may employ a baseline gaze model and a driver-specific gaze model. The baseline gaze model and the driver-specific gaze model may be implemented as one or more recurrent neural network models as described above, for example. Using a driver-specific gaze model allows for individualizing the recurrent neural network model to individual drivers. Research conducted in naturalistic driving suggest that different people will exhibit subtly different gaze sequential behavior while driving, but that for any specific person such behavior remains relatively constant. For example, different people may move their heads in different ways, may move their eyes in different ways, may have different durations for which they look at different areas, and other variations.

In some examples, the gaze prediction module starts with a baseline gaze model for all drivers, which may be trained based on average or typical gaze behavior, and then the gaze prediction module modifies the baseline gaze model for each driver based on the driver's own history of glance/gaze behavior from their real driving to develop the driver-specific gaze model for that driver. The longer a driver uses a vehicle, the better the driver-specific gaze model will be tailored, yielding an improved accuracy in gaze prediction for the driver. The gaze prediction module may save the driver-specific gaze model and may begin using a saved model in response to determining that the associated driver is again driving the vehicle. In some examples, the driver-specific model customization is implemented completely on the computing system, without sending information to a cloud system and without saving any raw video. This ensures the security of any personally identifiable data.

The gaze prediction module may provide the determined predicted driver gaze as current gaze information to application modules requiring such information, such as application modules that monitor driver gaze and infer driver attention. In some examples, the computing system may use the gaze information based on the gaze prediction to prompt the driver to return attention to the driving task. For example, an application module may cause the computing system to output one or more of an audible notification via a speaker associated with the computing system, a visual notification via the display, a tactile notification via the steering wheel or other component of the vehicle, and/or a notification via a computing device separate from the computing system, or by other means. In some examples, the computing system may use the gaze information based on the gaze prediction to change a mode of the device, such as by changing or activating a driver-assist feature of the vehicle. In some examples, the gaze prediction module may infer a level of driver attention based on the predicted gaze information, and may provide driver attention information to the application modules. This may help more accurately satisfy requirements for driver attention monitoring. The prediction-based gaze determination can provide more accurate and more near real-time gaze information to the computing system, which may reduce delay from when the vehicle system detects a problem to when the computing system can attempt measures to address the problem.