講演番号	252

文献番号	20226252

Analysis of Driver's Predictive Characteristics using Eye Tracking and a Deep Learning Model that Mimics Human Vision

Takaaki Seki [1)] Mizuki Amamiya [1)] Masataka Kato [1)] Koichi Emura [1)] Eiji Watanabe [2)]

1) Panasonic Automotive Systems Co., Ltd., Research and Development Division
600 Saedocho, Tsuzuki-ku, Yokohama, Kanagawa pref., 224-8539, Japan (E-mail: seki.takaaki@jp.panasonic.com)
2) National Institute for Basic Biology, National Institutes of Natural Sciences.
5-1 Higashiyama, Myodaiji, Okazaki, Aichi pref., 444-8787, Japan

KEY WORDS: **human engineering, human error, driver model, driver attention, cerebral nerve system, vision system,** machine learning, deep neural networks, predictive coding, visual illusions, eye tracking [C2]

A human predictive characteristic may affect the traffic accident factors such as prediction failure. In our latest research, we observed the area in front view where a human pays attention unconsciously while driving, using deep neural networks which reproduce human vision perception. In this paper, we analyzed the relation between the observed area and the eye movement of a person who is looking at front-view recordings. As a result, the hypothesis that large prediction errors attract stronger bottom-up attention is validated by measuring the gaze of 14 participants. As shown in Fig. 1, the prediction error image at time t_{20} is obtained as a difference image between the real image at time t_{20} and the predicted image at time t_{20} output from the prediction model by inputting a group of video frames from time t_1 to time t_{19}. The prediction error value at the gaze point is extracted for each frame in the video, and the percentage of the value that exceeds the threshold value when the threshold value is changed from the minimum value to the maximum value is calculated. In Fig.2, the vertical axis plots the percentage of the value above the threshold at the gaze point, and the horizontal axis plots the percentage of the value above the threshold at random point in each video frame. A graph of 58 video and 14 participants is shown in Fig. 2. The closer the curve of the graph in Fig. 2 is to the upper left than the line of $y = x$, the higher the probability that the prediction error values of the target pixels gazed at are larger than that of the randomly selected pixels in each video frame, indicating that the gaze tends to place where the prediction error value is larger. In addition, we found that inexperienced drivers focus their gaze on the areas significantly greater prediction error values than experienced drivers.

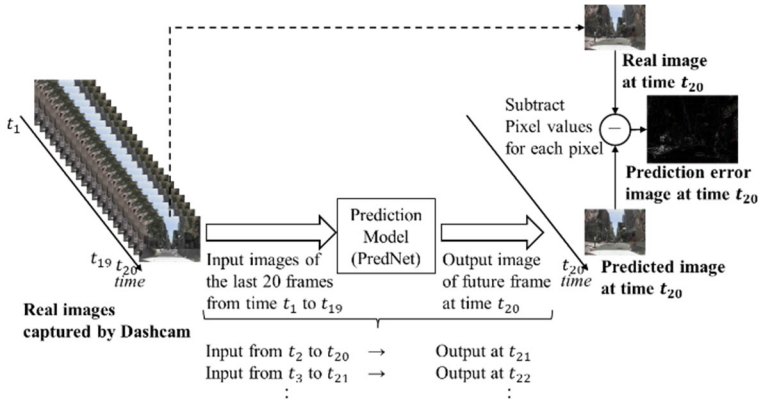

Fig. 1 Calculation process of prediction error images

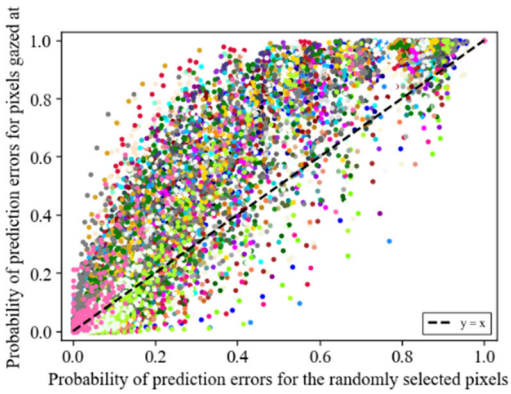

Fig. 2 Probability of prediction errors for pixels gazed at vs. for the randomly selected pixels