講演番号	241

文献番号	20226241

A Driver's Gaze Prediction Model Based on a Saliency Estimation Model Considering Self-Motion Speed

Yuya HOMMA [1)] **Takeshi KOHAMA** [1)]

1) Graduate School of Biology-Oriented Science and Technology
930 Nishimitani, Kinokawa-shi, Wakayama, 649-6493 Japan
(E-mail: †{2133730006j,kohama}@waka.kindai.ac.jp)

KEY WORDS: **human engineering , driver condition , recognition/judgment , saliency estimation model(C2)**

The purpose of this study is to develop a model for predicting the attention loci with higher accuracy for first-person video images taken inside a car at various driving speeds. We added an algorithm that dynamically changes the frame interval for calculating motion features to the model proposed by Kodama and Kohama (2017) and modified the center-surround antagonistic receptive field response model. Using the proposed model, we estimated the attention loci in these videos and verified the accuracy of the estimation.

The area MT model's processing was modified to detect the frame interval at which the motion vector is maximized. The formula for calculating the frame interval is shown in Equation (1). The tuning property for the selectivity to motion velocity was defined as the frame interval with the most prominent motion vector and the weighted addition of one frame before and after the maximum motion vector, which was used to calculate the area MT responses. Equation (2) shows the weights used in this calculation.

$$dt_{Max} = \underset{dt}{argmax} \sum_{d} \sum_{x,y} MT\ (dt\,;\, d\,), \tag{1}$$

$$W_{dt} = \begin{cases} 0.5 & if \quad dt = dt_{Max,} \\ 0.25 & else\ if\ \ dt = dt_{Max} - 1, dt_{Max} + 1, \\ 0 & others. \end{cases} \tag{2}$$

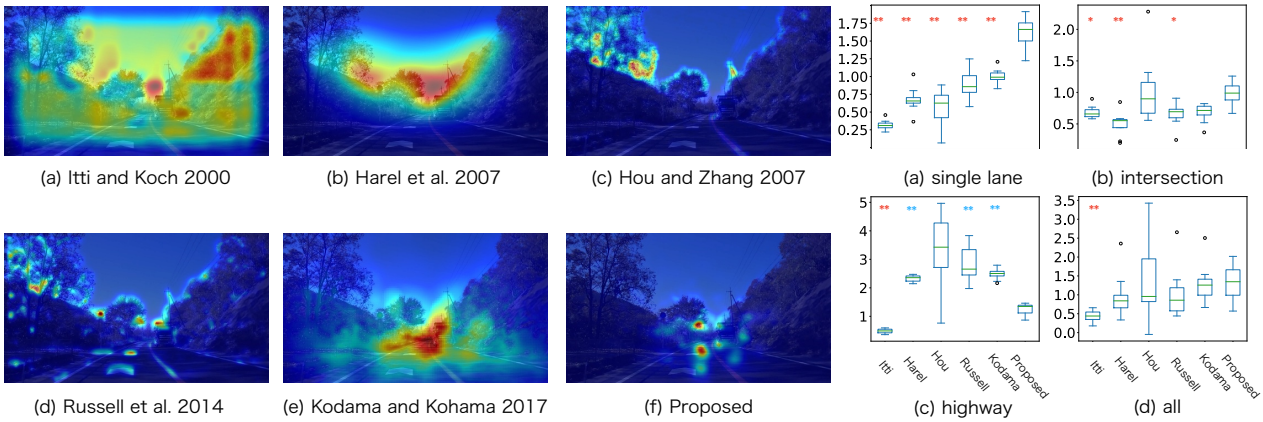

Figure1: The simulation results for each model for a single lane scene and the M-NSS scores.

Figure 1 left shows the simulation results for each model in a single lane scene, and Figure 1 right shows the M-NSS (moving normalized scanpath saliency) scores. The red asterisks indicate the combinations where the proposed model is significantly higher, and the blue asterisks indicate the combinations where the proposed model is significantly lower ($*: p < 0.1,\ **: p < 0.05$).

Simulation results show that the M-NSS scores of the proposed model were significantly higher than those of the other models in the narrow and intersection scenes. On the other hand, the M-NSS scores of the proposed model were significantly lower than those of the other models for highway scenes, indicating room for improvement. In these scenes, Hou's model (Hou and Zhang, 2006) and Russell's model (Russell et al., 2007), which are superior in object detection, scored well. Therefore, introducing an object detection mechanism is expected to improve the prediction accuracy.

The results obtained by totaling all scenes showed no statistically significant differences. However, the median value of the proposed model was higher than that of the other models, indicating the superiority of the proposed model.