Driver Drowsiness Estimation Using Facial Wrinkle Feature

Taro NAKAMURA* Tatsuhide MATSUDA Akinobu MAEJIMA Shigeo MORISHIMA†
Waseda University

1 Introduction

In recent years, the rate of fatal motor vehicle accidents caused by distracted driving resulting from factors such as sleeping at the wheel has been increasing. Therefore, an alert system that detects driver drowsiness and prevents accidents as a result by warning drivers before they fall asleep is urgently required. Non-contact measuring systems using computer vision techniques have been studied, and in vision approach, it is important to decide what kind of feature we should use for estimating drowsiness.

In a related study, by analyzing a video of captured facial expressions, [Hachisuka et al. 2010] focused on the movement of facial feature points on the eyes, eyebrows, and mouth. In this study, the drowsiness degree is evaluated on the basis of changes in these points with reference to an awakened state. However, it not clear whether this feature is enough or not to estimate drowsiness. To realize more precise drowsiness estimation, we need to select other effective feature. We detected a new drowsiness feature by comparing video image and CG model which are applied the existing feature point information. Moreover, we estimated drowsiness degree using new facial wrinkle feature, which has not been studied to date, and we confirmed that it improves the estimation accuracy.

2 Our Approach

2.1 Definition of drowsiness level: We recorded a video (60 fps frame rate and 320 x 240 pixels resolution) of 10 drivers' faces while they were operating a driving simulator. The drowsiness degree was divided into five levels on the basis of the captured video: "level 1: not sleepy," "level 2: slightly sleepy," "level 3: sleepy," "level 4: rather sleepy," "level 5: very sleepy." Two assessors who had studied the psychological measurement technique estimated drowsiness every 5 s. Traffic accidents occur when the level is 3 or higher. Therefore, our objective is to accurately estimate the transition from level 2 to level 3.

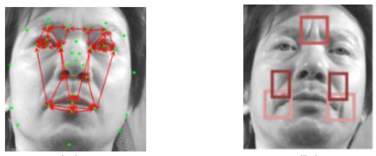
2.2 Existing drowsiness estimation feature: Owing to drowsiness the eyes of the subject are partially closed. To capture changes of such facial expressions, features based on feature point have been used. We also used the variation of the distance between two feature point positions from one awakened state to another as feature parameter. Eighteen feature points were detected by Irie's technique [Irie et al. 2011] in each video frame. We calculate 38 sets of distances between two feature points related to eyes, eyebrows, mouth, etc. The feature parameters are the lengths of arrows in Fig. 1.

2.3 k-NN based learning and estimation: Drowsiness estimation, which classifies the current state into five drowsiness levels, is performed using the k-NN method with the variation of the distance between two feature points. Feature vectors corresponding to each frame of the videos are calculated and averaged every 300 frames (5 s) in both learning and testing stages. The videos were divided into levels on the basis of the assessed ground truth level, and each video was further divided into alternate halves. One half is used for learning data, while the other half is used for testing data. We used the following two criteria for the evaluation index.
- Correct answer rate: The correct answer rate is defined as the percentage of the number of correctly estimated data compared to the ground truth level, which is the integer level from 1 to 5.
- RMSE: The RMSE is the root mean square error between the estimated score and ground truth score.

e-mail: {*taronoaddress@akane., †shigeo@}waseda.jp

SIGGRAPH 2013, July 21 – 25, 2013, Anaheim, California.

ACM 978-1-4503-2261-4/13/07



(a) (b)

Figure 1: (a) Distance parameters, **(b)** Filtering area

(a) (b) (c)

Figure 2: ***Detecting new feature*** **(a)** Awake, **(b)** Drowsy, **(c)** CG face model

2.4 Detecting new feature: To realize more precise drowsiness estimation, we need to select other effective feature. To detect new feature, we made drowsy CG face model which was applied feature point information. We deformed the shape of 661 points 3D mesh according to feature points of drowsy face, and mapped an awakened texture to this deformed model frame by frame. We compared original images and CG models, and we could clearly consider a feature which we should capture. Wrinkles appeared on brows, mouth, and nasolabial fold on original image's face when a subject resists drowsiness. Therefore, we introduced edge-intensity features which are calculated by applying Laplacian filtering at 5 areas on the face, as shown in Fig. 1: between the eyebrows, nasolabial folds, and corners of the mouth. Drowsiness estimation is performed again using the k-NN method with the feature vector connecting the variation of the distance between two feature points and the variation of the edge intensity. The evaluation results with open tests for the same subject are summarized in Table I.

Table I. Evaluation

	Only Distance	Only Edge	Distance + Edge
Correct answer rate [%]	63.4	60.9	**71.0**
RMSE	0.68	0.74	**0.58**

Moreover, in an actual driving situation, it is necessary to generate learning data beforehand from subjects other than the driver. Therefore, we attempt to use the leave-one-out approach. Because our purpose is to provide a warning when the drowsiness level reaches 3, it is necessary to verify if level 3 or higher levels have been correctly distinguished from level 2 or lower levels. We selected one person for this test and the other nine for learning data. For evaluation, the correct answer rate of two class separations was calculated. We realized two classes of recognition with an accuracy of 82.2%.

References

Hachisuka, S., Kimura, T., Ozaki, N., Ishida, K., and Nakatani, H. 2010. Drowsiness Detection Using Facial Expression Features. *SAE International paper. 2010-01-0466.*

Irie, A., Takagiwa, M., Moriyama, K., and Yamashita, T. 2011. Improvements to Facial Contour Detection by Hierarchical Fitting and Regression. *The First Asian Conference on Pattern Recognition Oral. pp. 273–277.*