# Perception of Drowsiness based on Correlation with Facial Image Features

Yugo Sato[1] Takuya Kato[1] Naoki Nozawa[1] Shigeo Morishima[2]

[1]Waseda University [2]Waseda Research Institute for Science and Engineering

**Keywords:** Drowsiness detection, face evaluation, feature learning, correlation coefficient, k-nearest neighbor algorithm

**Concepts:** •**Computing methodologies** → *Perception;* Perceptron algorithm;

## 1 Introduction

This paper presents a video-based method for detecting drowsiness. Generally, human beings can perceive their fatigue and drowsiness through looking at faces. The ability to perceive the fatigue and the drowsiness has been studied in many ways. The drowsiness detection method based on facial videos has been proposed [Nakamura et al. 2014]. In their method, a set of the facial features calculated with the Computer Vision techniques and the k-nearest neighbor algorithm are applied to classify drowsiness degree. However, the facial features that are ineffective against reproducing the perception of human beings with the machine learning method are not removed. This factor can decrease the detection accuracy.

Thus, we calculate a correlation coefficient between each facial feature and drowsiness degree's transition to estimate which facial features are close to the assessors' perception. Using our method, we achieve to reproduce their perception more accurately than the previous work. Moreover, the result shows that eyes and a mouth are one of the visual factors which correlate with their perception.

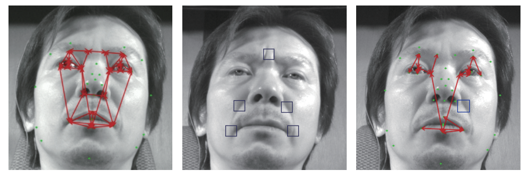
## 2 Our Approach

### 2.1 Drowsiness Classifier's Learning

To begin with, we film videos of subjects' faces to create datasets for the drowsiness classifier's learning. Simultaneously, the subjects play a driving game in a monotonous racetrack as an example of a simple task which induces their drowsiness. To these videos' every 5 seconds, two assessors evaluate each subject's drowsiness degree divided into five levels as follows, "level 1: not sleepy", "level 2: slight sleepy", "level 3: sleepy", "level 4: rather sleepy", "level 5: very sleepy". We assign these levels as correct labels for the classifier's learning. In this paper, these procedures were followed to create 10 subjects' datasets.

Subsequently, we calculate two kinds of facial features based on the previous method [Nakamura et al. 2014] to track the subjects' facial transitions caused by drowsiness. The first feature is the variation of distance between facial feature points in 38 segments from their awake to drowsy state (see Figure 1(a)). This feature is calculated in that their faces change such as their cheek's loosening and their eyes' gradually closing. The second feature is the variation of edge intensity in 5 areas from their awake to drowsy state (see Figure 1(b)). This feature is introduced because some subjects have deep wrinkles in their drowsy state.


SAP '16, July 22-23, 2016, Anaheim, CA, USA
ISBN: 978-1-4503-4383-1/16/07
DOI: http://dx.doi.org/10.1145/2931002.2947705



(a) Distance parameters (b) Filtering areas (c) Selected features

**Figure 1:** *Calculate and select facial features.*

Finally, we calculate a correlation coefficient between each facial feature and the subjects' drowsiness degree's transitions to estimate which features correlate with the assessors' perception. We consider that the features whose correlation coefficient is higher than an average of all features' calculated value, are effective against the drowsiness detection (see Figure 1(c)). The drowsiness classifier learns only these effective features. We apply the k-nearest neighbor algorithm in the same learning method as the previous method to compare their performance.

### 2.2 Evaluation

We attempt to use the leave-one-out cross-validation of 10 subjects in learning and testing. The input subject's drowsiness degree is classified into 5 levels every 5 seconds. Furthermore, we calculate the time difference between a detected and a grand truth time at level 3, since traffic accidents increase when their drowsiness degree is level 3 or higher in the driving game.

## 3 Result and Future Work

Table 1 shows the time difference of 10 subjects. In this paper, we select the facial features that correlate with the subjects' drowsiness transitions. As a result, our method can reproduce the assessors' perception more accurately than the previous work. However, the accuracy depends on the subject because of a facial features' individuality. For instance, some faces do not have deep wrinkles in both their awake and drowsy state. In our method, we select the facial features based on the average of the correlation coefficient for all subjects. Therefore, the accuracy of the subjects whose facial features are different from the average is low. As our future work, we need to consider the selection method with each subject.

Furthermore, the select facial features that correlate with the assessors' perception are distributed in the subjects' eyes and mouths, and connected with each other part (see Figure 1(c)). These results imply that their eyes and mouth are one of the visual factors when they perceive drowsiness through looking at faces. In addition, they can also perceive their drowsiness from the positional relation between their eyes and mouth. We aim to experiment with more assessors to reproduce human begins' perception more accurately.

**Table 1:** *RMSE of the time difference (10 subjects).*

| [Nakamura et al. 2014] | Our method |
|---|---|
| 376.3 s | 349.5 s |

## References

Nakamura, T., Maejima, A., and Morishima, S. 2014. Driver Drowsiness Estimation from Facial Expression Features Computer Vision Feature Investigation Using a CG Model. In *Proc. VISAPP*, vol. 2, 207–214.