

Facial-Video-Based Drowsiness Estimation Technology for Operation on Low-End IoT Devices

TSUJIKAWA Masanori

Abstract

This paper introduces NEC's drowsiness estimation technology based on facial video data. This system can operate on low-end IoT devices that use fewer computation resources. Drowsiness in humans is clearly expressed via changes in facial expression — especially changes in eyelid movement — and can be determined with high accuracy by capturing the closing duration of the eyelids and the increase in blinking rate. In order to capture fast blinking, however, it is necessary to precisely extract the movement of the eyelids. This bears a heavy computation cost. NEC has developed a method to capture slowly fluctuating eyelids (eyelid variability) as a sign of drowsiness, which makes it possible to accurately determine drowsiness even with eyelid movements extracted with only one third the detail of conventional methods. Since this system can easily be installed and operated on low-end devices, it should be applicable across a broad range of scenarios.

Keywords



drowsiness estimation, facial video data, eyelid variability, low frame rate, IoT device, work style reform, productivity improvement

1. Introduction

Efforts to improve work productivity through work style reform promoted by the Japanese government have included visualizing the way people work. The visualization makes it possible that areas needing improvement can be identified and solutions developed. One warning sign closely connected to a drop-off in productivity is a decreased arousal level — what is commonly called drowsiness. Visualizing drowsiness is critical to ensuring that workers are in the optimal state — both physically and mentally — to maintain and improve productivity.

Developing technology to estimate drowsiness¹⁾⁻⁵⁾ has been of particular interest to the automotive industry as drowsiness is one of the leading causes of deadly accidents. It is possible to accurately estimate drowsiness by extracting detailed eyelid movements from facial video data and analyzing the movement of the eyelids¹⁾⁻⁴⁾. However, a lot of processing power is required to do this since the facial video data has to be processed at a high frame rate — over 15 fps³⁾ — and this is not something that can easily be done with low-end devices.

At NEC, we have developed drowsiness estimation tech-

nology using facial video data that can even operate on low-end IoT devices⁵⁾. Based on a method of capturing the eyelid variability — a newly discovered sign of drowsiness is discussed (Section 2). This technology has proven effective in evaluations, accurately estimating drowsiness from facial video data with a frame rate that is a third (5 fps) of that used in conventional systems (Section 3). This system promises to provide an effective and affordable tool for estimating drowsiness in the workplace and anywhere else where arousal level is important (Section 4).

2. Drowsiness Estimation Technology Using Facial Video Data

Most of the research so far conducted into drowsiness estimation has centered around the automotive industry. Examples of drowsiness estimation using various methods are shown in **Table**.

Asking subjects to assess their own drowsiness is problematic as it requires them to assess their drowsiness frequently, which is distracting and inconvenient. Estimating drowsiness based on how the subject is operating machinery such as a steering wheel or brakes, while quite effective, is not broadly applicable except in

Table Estimation of drowsiness using various methods.

Category	Estimation method	Accuracy	Convenience	Device cost
Subjectivity	Self-assessment	⊙	× Demanding due to need for frequent assessment	⊙
Behavioral information	Machinery operation, etc.	△ Low compatibility	⊙	○
Physiological information	Pulsation	△ Significant differences between individuals	△ User must wear sensor	○
Physiological information	Eyelid movement	⊙	⊙	△ Requires high-speed image data processing

⊙ Excellent ○ Good △ Fair × Poor

cars, trucks, and so on. Using the pulse rate as a surrogate for estimating drowsiness falls short due to wide differences between individuals and because it requires the subject to wear a sensor. The only system that offers both accuracy and convenience in a broadly applicable solution is the method in which eyelid movement is estimated and analyzed. Individual differences are relatively inconsequential and there is no need to wear a sensor so it is convenient, reliable, and can be used in a wide range of scenarios, not just automobiles. Despite all the positives, the system suffers from one major drawback; equipment cost is relatively high because high-speed image data processing is required.

In Section 2.1, we describe the flow of facial-video-based drowsiness-level estimation, illustrating why high-speed image data processing is required. Then in Sec. 2.2, we'll explain how our newly discovered method of tracking eyelid variability — which are a sign of drowsiness — has made high-speed image data processing unnecessary.

2.1 Drowsiness-level estimation flow

Fig. 1 shows the flow of estimating drowsiness from facial video data. The flow is as follows: (1) the positions of the eyes are estimated from facial image data and time series information on how wide the eyelids are open is obtained; (2) the feature for drowsiness is calculated according to the movement of the eyelids (time series data); and (3) drowsiness is estimated with an estimation model created using machine learning.

The feature for drowsiness — that is, the time in which

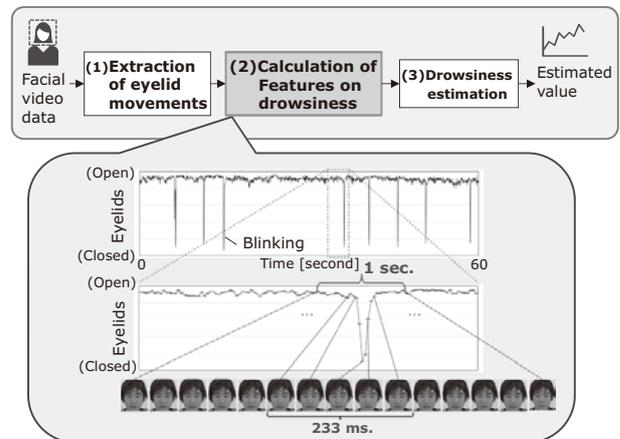


Fig. 1 Flow of drowsiness-level estimation using facial video data.

the eyelids are closed (eye closure duration)¹⁾ — and the feature for blinking²⁾³⁾ are often used. Eye closure duration is the most commonly used feature. However, on its own, this cannot provide optimal estimation accuracy. To achieve high-accuracy estimation, features such as the blink occurrence frequency, blink movement velocity, and distributed time intervals are often used. Nonetheless, as is obvious from the movement of the eyelids shown in Fig. 1, the blink of an eye can occur as fast as several hundred milliseconds. To capture this kind of rapid movement, the facial image data must be processed at a high frame rate of over 15 fps. This pushes up the cost of the equipment required and reduces the applicability of the technology.

2.2 Eyelid variability

To achieve high-precision drowsiness-level estimation even with a low-end IoT device, we directed out attention to eyelid variability as a sign of drowsiness. Since the eyelids move slowly, the movement can be reliably captured from video even at a low frame rate, making high-speed image data processing unnecessary. Now, let's look at the two types of features for eyelid variability that NEC has discovered.

The first one is time variability as shown in **Fig. 2**. This is a feature that expresses the difficulty of keeping the opening and closing of the eyelids constant when the subject becomes drowsy. The fluctuation in the conditions in which the eyelids are open (light gray line in Fig. 2) is set to be the feature for time variability. Fig. 2 clearly demonstrates that this kind of variability increases when the subject is feeling drowsy. Since this time variability is a slow movement, the features can sufficiently be captured from facial video data at a slow

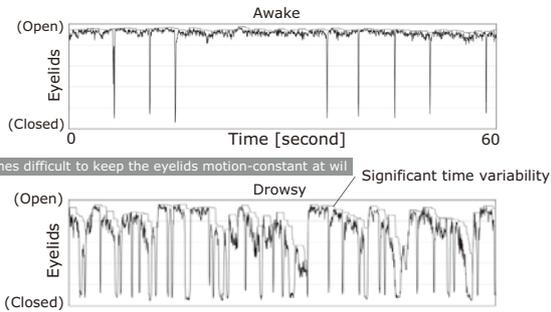


Fig. 2 Example of eyelid movement (time variability).

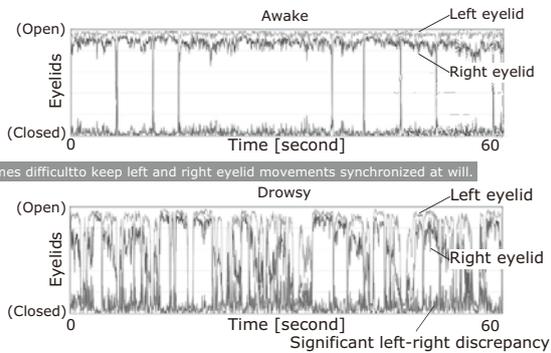


Fig. 3 Example of eyelid movement (difference between left and right).

frame rate of about 5 fps.

The second feature is based on the difference between the left and right eyelids as shown in **Fig. 3**. This is a feature that captures the difficulty of coordinating the movements of the left and right eyelids when the subject becomes drowsy. The difference between the left and right eyelids (black line in Fig. 3) is set to be the feature here. Fig. 3 shows that the difference between the left and right eyelids increases as the subject becomes drowsier. Because eyelid variability also involves slow motions in the case of left-right discrepancy, this value can be derived from facial video data at a low frame rate just like for time variability.

3. Evaluating Our Drowsiness-level Estimation Technology

We conducted an evaluation of our drowsiness-level estimation technology to assess the usefulness of the features related to eyelid variability. Evaluation data was drawn from video recordings of 29 people totaling 41 hours. Participants were fully informed of the nature and purpose of the tests and consent was obtained. The resolution of the collected facial video data was set to 640 x 480 pixels and the frame rate to 30 fps. The degrees

of drowsiness were labeled based on Kitajima et al.'s paper⁴⁾ ranging from level 1 (no signs of drowsiness at all) to level 5 (appears very drowsy). The levels were defined using the mean values of the 5-level labels judged by our well-trained labelers. For the evaluation indices of estimation precision of drowsiness, we used the correlation coefficients (0.0–1.0) between estimated values and the above drowsiness labels. In order to compare the conventional technology with our newly developed technology, we used eyelid closure time, blink occurrence frequencies, blink movement velocities, and distributed time intervals as features for the conventional system, while additionally using the two types of eyelid variability described in Sec. 2.2 above for the new system. To compare estimation accuracy between the two systems, we changed video frame rate from 30 fps to 3 fps.

Fig. 4 shows the evaluation result. We confirmed that the new technology operating at a frame rate of 5 fps could achieve estimation accuracy comparable to that achieved when facial video data was processed at 15 fps with the conventional technology. The reason that high accuracy can be maintained at such a low frame rate is that the eyelid movement is so slow that it is still possible to effectively capture the features. When the new technology is used, the volume of data and the amount of processing required for drowsiness-level estimation can be reduced by two thirds. This means that even low-end IoT devices can be used to achieve high-precision drowsiness-level estimation (**Fig. 5**).

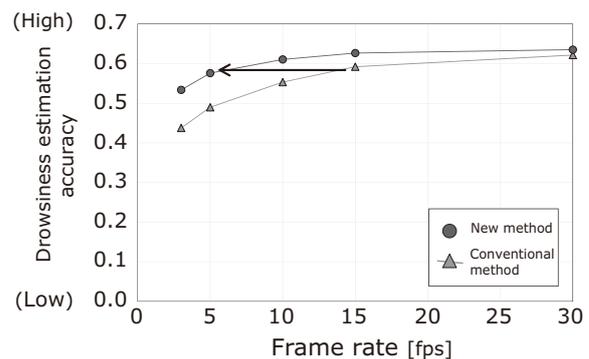


Fig. 4 Results of drowsiness-level estimation evaluation.

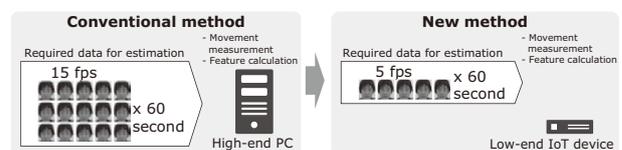


Fig. 5 Comparison of data and devices necessary for estimation.

4. Practical Examples of Our Drowsiness-level estimation System

We developed a trial system that would estimate drowsiness in real time using the newly developed eyelid variability feature. Below we describe a couple of examples of this system in action.

4.1 Compact PC

Fig. 6 shows an example of how the system performs when run on a compact PC. We used a very compact PC with dimensions of about 115 x 110 x 35 mm. Four USB cameras were connected to the PC's four USB ports. Facial video data (640 x 480 pixels, 5 fps) for four people was input from the cameras and processed simultaneously. We were able to successfully estimate each person's level of drowsiness. The ability to process the data for four people and accurately assess their levels of drowsiness with a single compact PC means that the cost of drowsiness-level estimation systems can be significantly reduced.

4.2 Smartphone

Fig. 7 shows an example using a smartphone. We

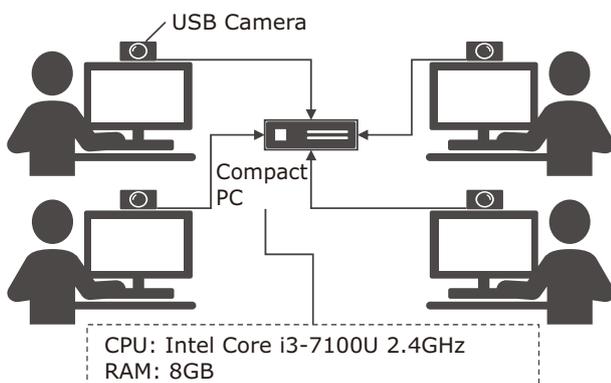


Fig. 6 Compact PC.

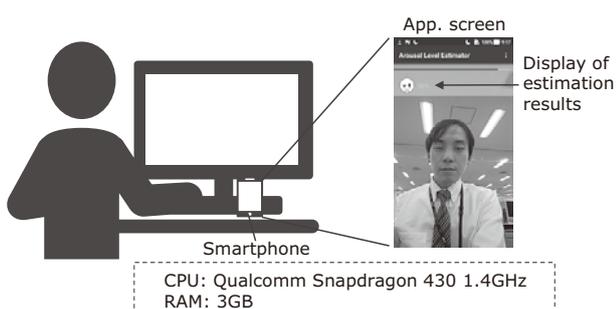


Fig. 7 Low-cost smartphone.

were able to estimate drowsiness by processing facial image data (640 x 480 pixels, 5 fps) input from the smartphone's built-in camera. We used a relatively inexpensive smartphone priced at about 20,000 Japanese yen (as of November 2018). Processing drowsiness-level estimation with such a smartphone alone can contribute to the reduction of equipment cost.

5. Conclusion

In this paper, we have demonstrated how our new drowsiness-level estimation technology can be used on low-end IoT devices. Offering low overhead and high accuracy, this system works by tracking the slow eyelid variability that occurs when a person is drowsy and struggling to keep their eyes open. Unlike other symptoms of drowsiness such as blinking, eyelid variability can be monitored accurately at a much slower frame rate of 5 fps (one third the rate of conventional systems). With less video data to process, estimation can be performed accurately and efficiently using low-cost, easily available devices such as smartphones and compact PCs. The system also avoids the drawbacks of other systems such as intrusiveness and inconvenience. We believe our system offers businesses an ideal solution to support workplace productivity by ensuring that workers are functioning at peak levels, both mentally and physically. Under the partnership in various fields, we plan to continue our efforts to streamline this technology and make even more affordable and reliable.

* Intel Core is a trademark of Intel Corporation and/or its affiliates in the U.S. and other countries.

* Snapdragon is a registered trademark of Qualcomm, Inc.

* All other company names and product names that appear in this paper are trademarks or registered trademarks of their respective companies.

Reference

- 1) David F. Dinges et al.: PERCLOS: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance, U.S. Department of Transportation, Federal Highway Administration, Publication No. FHWA-MCRT-98-006, 1998.10
https://rosap.nhtl.bts.gov/view/dot/113/dot_113_DS1.pdf
- 2) Takuhiro Omi: Detecting Drowsiness with the Driver Status Monitor's Visual Sensing, DENSO TECHNICAL REVIEW, Vol.21, pp.93-102, 2016
<https://www.denso.com/jp/ja/innovation/technology/dtr/v21/paper-10.pdf>
- 3) Antoine Picot et al.: Comparison between EOG and high frame rate camera for drowsiness detection, 2009 Workshop on Applications of Computer Vision (WACV), 2009.12
- 4) Hiroki Kitajima et al.: Prediction of Automobile Driver Sleepiness (1st Report, Rating of Sleepiness Based on Facial Expression and Examination of Effective Predictor Indexes of Sleepiness), Transaction of the JSME (in Japanese) (C), vol. 63 no. 613, pp.3059-3066, 1997.9
https://www.jstage.jst.go.jp/article/kikaic1979/63/613/63_613_3059/_article/-char/en
- 5) Masanori Tsujikawa et al.: Drowsiness Estimation from Low-Frame-Rate Facial Videos using Eyelid Variability Features, 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018.7

Authors' Profiles

TSUJIKAWA Masanori

Principal Researcher
Biometrics Research Laboratories
IEEE member
IEICE member

The details about this paper can be seen at the following.

Related URL:

Daikin and NEC successfully demonstrate effectiveness of temperature stimulation for improving intellectual productivity in the workplace

https://www.nec.com/en/press/201807/global_20180725_03.html

Physiological information recognition: Drowsiness-level (Arousal-level) estimation from facial videos

<https://www.nec.com/en/global/rd/technologies/201801/index.html>

Information about the NEC Technical Journal

Thank you for reading the paper.

If you are interested in the NEC Technical Journal, you can also read other papers on our website.

Link to NEC Technical Journal website

Japanese

English

Vol.13 No.2 Social Value Creation Using Biometrics

Remarks for Special Issue on Social Value Creation Using Biometrics
Committed to Supporting Social Values via Biometrics

Papers for Special Issue

Commitment to Biometrics NEC Is Promoting

Bio-IDiom — NEC's Biometric Authentication Brand
The Future Evolution and Development of Biometrics Studies
Privacy Measures of Biometrics Businesses

Services and Solutions That Leverage Biometrics

The Western Identification Network: Identification as a Service in a Federated Architecture
Use of Face Authentication Systems Associated with the "My Number Card"
Face Recognition Cloud Service "NeoFace Cloud"
NEC Enhanced Video Analytics Provides Advanced Solutions for Video Analytics
New In-Store Biometric Solutions Are Shaping the Future of Retail Services
ID Service Providing Instantaneous Availability of User's Desired Financial Services
Biometrics-Based Approach to Improve Experience from Non-routine Lifestyle Fields
Construction Site Personnel Entrance/Exit Management Service Based on Face Recognition and Location Info
The Importance of Personal Identification in the Fields of Next-Generation Fabrication (Monozukuri)

Core Technologies and Advanced Technologies to Support Biometrics

How Face Recognition Technology and Person Re-identification Technology Can Help Make Our World Safer and More Secure
Advanced Iris Recognition Using Fusion Techniques
Advanced New Technology Uses New Feature Amount to Improve Accuracy of Latent Fingerprint Matching
Safety, Security, and Convenience: The Benefits of Voice Recognition Technology
Ear Acoustic Authentication Technology: Using Sound to Identify the Distinctive Shape of the Ear Canal
Automatic Classification of Behavior Patterns for High-Precision Detection of Suspicious Individuals in Video Images
Facial-Video-Based Drowsiness Estimation Technology for Operation on Low-End IoT Devices

NEC Information

NEWS

2018 C&C Prize Ceremony



Vol.13 No.2
April 2019

Special Issue TOP