Yawning Recognition based on Dynamic Analysis and Simple Measure

Michał Ochocki and Dariusz Sawicki
Warsaw University of Technology, Institute of Theory of Electrical Engineering, Measurements and Information Systems, Warsaw, Poland

Keywords: Worker Fatigue, Yawning Detection, Facial Landmarks.

Abstract: Nowadays drivers fatigue is amongst significant causes of traffic accidents. There exist many academic and industrial publications, where fatigue detection is presented. Yawning is one of the most detectable and indicative symptoms in such situation. However, yawning identification approaches which have been developed to date are limited by the fact that they detect a wide open mouth. And the detection of open mouth can also mean talking, singing and smiling, what is not always a sign of fatigue. The research aims was to investigate the different situations when the mouth is open and distinguish situation when really yawning occurred. In this paper we use an algorithm for localization of the facial landmarks and we propose a simple and effective system for yawning detection which is based on changes of mouth geometric features. The accuracy of presented method was verified using 80 videos collected from three databases: we have used 20 films of yawning expression, 30 films of smiling and 30 films with singing examples. The experimental results show high accuracy of proposed method on the level of 93%. The obtained results have been compared with the methods described in the literature – the achieved accuracy puts proposed method among the best solutions of recent years.

1 INTRODUCTION

1.1 Motivation

Fatigue and sleep symptoms are today very important factors for psychology of work and safety determination. The National Sleep Foundation (National Sleep Foundation, 2017) has been conducting research in this field for many years. Fatigue and falling asleep are not only a simple work-related problem. In many situations; for many professions they lead to accidents causing deaths (Osh in figures, 2011). Driver is one of such profession, where fatigue consequences could be dangerous.

Today many types of cars are equipped with fatigue sensors (Bosch at CES, 2017). Fatigue tests are carried out using many different fields of driver activity (Friedrichs and Bin Yang, 2010). The most popular are eye movements, eye blinking (especially frequency of closing state) and yawning. In some systems physical activity is also analyzed. In many systems the yawning analysis is associated effectively with the eye tracking (Weiwei, et al., 2010) or with eye blinking analysis (Rodzik and Sawicki, 2015). It is physiologically natural. There exist many publications, where the surveys of techniques for fatigue recognition we can find (Friedrichs and Bin Yang, 2010, Coetzerm and Hanckem, 2009, Sigari, et al., 2014). The industrial publications (Bosch at CES, 2017, Fatigue Risk Assessment, 2017) can be also treated as interesting reviews of discussed problem. These documents show how quickly the technology of employee security is being developed and implemented.

Many systems for fatigue recognition are based on yawning detection. It is commonly believed that identification of yawning is very simple. After all, it is just detection of a wide open mouth. But people open their mouths also while talking, laughing, and especially while singing. We are not sure whether considering a driver singing at work is an exaggeration. But the question is important: does every mouth opening for a long time means yawning? How to distinguish yawning from singing or laughing in simple way?

Ochocki M. and Sawicki D.
Yawning Recognition based on Dynamic Analysis and Simple Measure.
DOI: 10.5220/0006497701110117
In *Proceedings of the International Conference on Computer-Human Interaction Research and Applications (CHIRA 2017)*, pages 111-117
ISBN: 978-989-758-267-7

1.2 The Aim of the Article

The aim of this study was to develop a simple and effective system for yawning detection. We have tried to analyze different situation when the mouth is open (for different purposes) and distinguish situation when yawning is really occurred. Many advanced methods used to recognize yawning are known from the literature of the subject. Our goal was to try to find a simple geometric measure that would give comparable results.

2 YAWNING DETECTION IN THE PREVIOUS WORKS

In almost all systems, where the yawning is detected as the fatigue symptom, the recognition methods are built in similar way. The camera (one or sometimes two) registers head/face image. Using image analysis the mouth is recognized and its state (close or open) is detected. In case of open state of mouth we have to decide if it is yawning.

Alioua, Amine and Rziza in paper (Alioua, et al., 2014) have applied support vector machine (SVM) for face extraction and circular Hough transform (CHT) for mouth detection. Then, using CHT they identified yawning. The advantage of this solution is that such system does not require any training data. Wang and Shi in (Wang and Shi, 2005) have used Kalman filter at the state of face region tracking. In this algorithm mouth is extracted from proper region. It is simple and effective solution. Saradadevi and Bajaj in paper (Saradadevi and Bajaj, 2008) used Support Vector Machine in analysis of face/mouth image. After training SVM allowed distinguishing between "Normal Mouth" and "Yawning Mouth". Ibrahim at al. in paper (Ibrahim, et al., 2015) implemented local binary pattern based on color differences. Algorithm measures the mouth opening and identifies yawning. This paper is the first one where Authors also analyzed situation, when the yawning mouth is covered by hand for a period of time. In the solution presented in paper (Li, et al., 2009) two different cameras have been used: Camera A records the driver's head according to face position. Image from Camera B is used for detection of driver's mouth using Haar-like features.

The Viola-Jones algorithm (Viola and Jones, 2001) is very popular in extraction of the face region. It detects face correctly, when the face is tilted up/down to about ±15 degrees and left/right to about ±45 degrees (Azim et al., 2009). Authors of paper (Abtahi, at al., 2013) have implemented the Viola-Jones algorithm for face and mouth detection, they have used the back projection theory for measuring the rate of mouth changes in yawning detection. This idea have been developed and presented, by practically the same team of Authors, in paper (Omidyeganeh et al., 2016). The region of mouth is found by Viola-Jones algorithm. Then the mouth region is segmented by k-means clustering method. Manu in paper (Manu, 2016) presented system where the Feature Vector containing states of left eye, right eye and mouth was used in training and classification. Papers (Omidyeganeh et al., 2016) and (Manu, 2016) are the latest published conference documents concerning the yawning recognition, which we found during our work.

Yawning detection is not so simple task. Especially if we try to compare yawning to singing, smiling and other states when the mouth is open. The simplest solution used in algorithms for fatigue recognition is to analyze the height of opened mouth. In the work (Kumar and Barwar, 2014) the contour of open mouth is analyzed, Y coordinate values greater than a certain threshold, informs about yawning. The Authors of the paper (Alioua, et al., 2014) propose a distinction between slightly open mouth and widely open mouth using circular Hough transform (CHT). Although very elegant, it is a time consuming solution. Area analysis of the open mouth can give good result. It is so convenient that after detection of opening, we can mark the visible interior of the mouth on the basis of the difference in color. Such solution is proposed by the Authors of the work (Omidyeganeh et al., 2016). After highlighting the area of mouth, the Authors specify the number of pixels in the image of mouth opening. If it is greater than proper threshold, the yawning state is detected. Similar method is used in (Li, et al., 2009) but additionally Authors of this paper analyzed the ratio between height and width of the area recognized as the mouth opening.

3 PROPOSED ALGORITHM FOR YAWNING RECOGNITION

In frequently used methods to identify yawning (and blinking as well), this phenomenon is analyzed "holistically" – taking into account the features of the entire face. It is important in this situation to accurately distinguish the relevant elements (mouth or eyes and mouth) in the properly identified face. Hence, in such algorithms there is an independent condition that the face is recognized in a "reasonable" way (Li, et al., 2009).

A typical problem in this task is to analyze the image of face that is tilted and / or rotated. The Author of paper (Manu, 2016) pointed out that his system does not always work properly, when head was tilted left or right. Authors often pay attention to this problem by proposing additional algorithms for the normalization of the position associated with the "compensation" of deviation from the horizontal line (Wang and Shi, 2005, Hasan et al., 2014). The Authors of the paper (Wang and Pendeli, 2005) examine the degree of mouth openness by considering the trigonometric function (cos) dependent on the rotation of the mouth. Similarly, the height of the mouth opening and the angle of deviation from the horizontal line are taken into account in (Azim et al., 2009).

In order to separate the problems of analysis as well as to improve efficiency and accuracy of yawning identification, we propose a two-stage recognition algorithm.

- **Facial Identification and Separation of the Mouth Shape.** The aim of this stage is to obtain the shape of the mouth from an image of head. In addition, the resulting shape of the mouth is normalized regardless of lighting conditions, and regardless of the head position (tilting up - down and left - right). But of course, in a certain range of correct work, i.e. when the entire mouth is recorded despite the tilt.

- **Yawning Detection.** The aim of this stage is to identify yawning with particular regard to the differences for other states of mouth opening.

3.1 Facial Identification and Separation of the Mouth Shape

In the face recognition and analysis an algorithm for localization of the facial landmarks was applied (Ochocki and Sawicki, 2016). This algorithm was originally prepared for emotion recognition, however can be used for any analysis of the face geometry. On the other hand, the original aim (landmarks detection) guarantees a high recognition accuracy of face features. In the algorithm preliminary analysis has been used. In this part the set of proper regions of interest (ROI) is identified. It is based, mainly, on analysis of features in color space which corresponds to each region.

In order to separate the mouth area the properties of *CIE-Lab* space is used. *CIE-Lab* is a standardized color space *Lab* by the International Commission on Illumination. *Lab* color model is designed to illustrate the difference between the colors seen through the human eye. The color is represented by three coordinates *L*, *a* and *b*. Component *L* represents the luminance, *a* determines the color from green to magenta, *b* color from blue to yellow. In the algorithm the *a* and *b* components of *Lab* color model is used for mouth segmentation.

The whole algorithm can locate basic 21 landmarks of the face (Figure 1) with a very good average accuracy of 95.62% (Ochocki and Sawicki, 2016). This is comparable to the results of the best algorithms of this type, which were published between 2005 and 2015.

Figure 1: The facial landmarks analyzed in the algorithm for localization. The free face chart from (Want a free mua face chart? 2010).

In the used algorithm, landmarks identification is divided into four independent groups - the facial region (oval), the eye region, the mouth region and the region of the nose. For the purposes of yawning analysis, we modified the algorithm so that it only identifies the landmarks of the mouth. The modified algorithm achieved an average accuracy of 95.2%. The result of the algorithm's operation in the form of a normalized image of the mouth is shown in Figure 2.

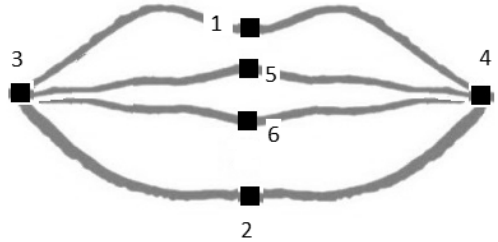

Figure 2: Landmarks of the mouth identified by the modified algorithm for face analysis.

3.2 Yawning Detection

To develop an effective detector, the set of analyzed features should be extracted. For this purpose, the selection of features was performed, i.e. the process of identifying the appropriate set of geometric changes in the shape of mouth.

In the yawning detection algorithm, we use the coordinates of the landmarks obtained from facial analysis (Figure 2.). We have introduced D_{yx} (1) as a simple measure of the mouth opening. The value of the D_{yx} measure is the ratio of the height of mouth to their width in a given frame of video, determined using coordinates of landmarks. The width of mouth is calculated as a simple distance between proper X coordinates (points 3 and 4). Height is calculated as a simple distance between proper Y coordinates (points 1 and 2).

$$D_{yx} = \frac{Y_D}{X_D} \qquad (1)$$

where:
D_{yx} – the measure of opening, Y_D – the height of mouth (distance between points 1 and 2 – in Figure 2.), X_D – the width of mouth (distance between points 3 and 4 – in Figure 2.).

The graph of the dynamic changes of the D_{yx} value in normalized time values by the min-max method is presented in Figure 3. This is a graph for the three cases of opening the mouth. The X axis indicates the normalized elapsed time from the beginning (0) to the end (1) of action (facial event). This approach allows comparing actions of different (real) duration of times.

From the analysis of the graph shown in Figure 3, it can be concluded that the cases of opening mouths can be easily distinguished on the basis of the maximum D_{yx} value. Therefore, at the stage of classification, feature d expressed by formula (2) was extracted from the sets of measure's changes D_{yx}.

$$d = \frac{\max(D_{yx})}{\min(D_{yx})} \qquad (2)$$

where:
d – the dynamic measure of opening, D_{yx} – the measure of opening.

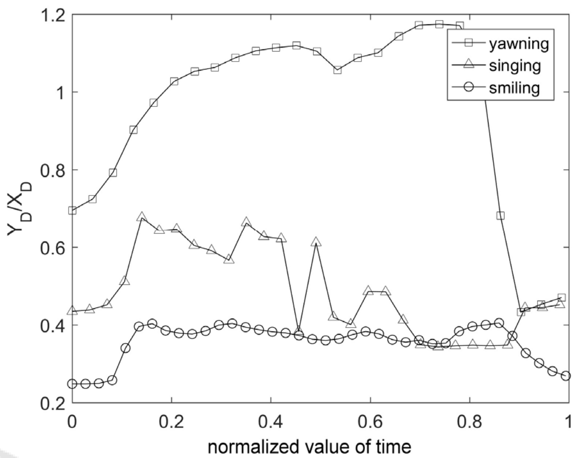

Figure 3: An example of a graph which represents change of D_{yx} measure for the three different discussed types of opening the mouth.

4 TESTS OF THE PROPOSED ALGORITHM

We have conducted experiments using three independent databases:

- **Smiling.** In the study of smile expressions we have used UvA-NEMO Smile Database (Dibeklioglu et al., 2012). This database was created to analyze the dynamics of genuine (sincere) and false (posed) smile as part of the Live Science program. UvA-NEMO Smile Database consists of 1240 color videos at resolution of 1920x1080px in controlled, unified lighting environment. It contains 597 videos of genuine (sincere) smile and 643 videos of false (posed) smile, recorded from 400 people (185 women and 215 men). Genuine and posed smile of each subject was selected from approximately five minutes of recordings There were people of different races, genders and ages. Only the videos of genuine smile were used in our research.
- **Yawning.** A colorful video from YouTube (The Yawn-O-Meter, 2015) at resolution of 1280x720px was used to analyze the yawning phenomenon. This videos contains

expressions of yawning recorded in unified lighting conditions. The videos was divided into fragments - each fragment contains yawning expression of another person. The database contains 20 yawning videos from people of different races, genders and ages.

- **Singing.** We have used the Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) (Livingstone, 2012). RAVDESS was created to analyze singing and speaking in 8 different emotions – neutral, calm, happy, sad, angry, and anxious. The database contains more than 7,000 color videos from 24 actors (12 women and 12 men).

Changes of the value of *d* parameter depending on discussed cases are presented in Figure 4 in the form of boxplot graph. This approach allows determining the characteristic of the *d* distribution for each case. The analysis was carried out using:

- 20 films of yawning expression;
- 30 films of smiling expression;
- 30 films with singing examples.

The graph of *d* distribution contains three areas of high separability. Therefore, this attribute has been subjected to a final assessment of suitability by maximization of the classification efficiency. For this purpose, a simple classification method was chosen. The distinguishing feature *d* was compared with threshold value based on the data collected during the training phase.

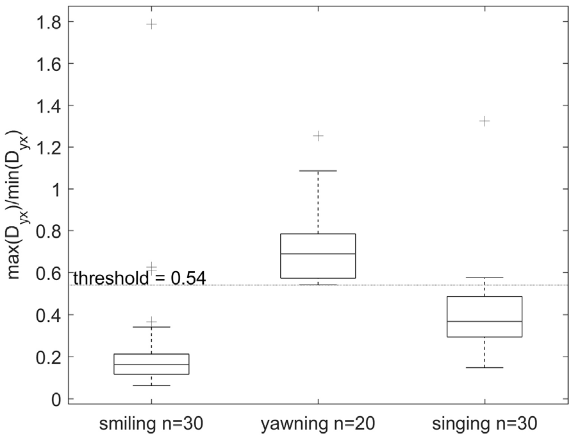

Figure 4: The graph of *d* distribution for three discussed actions.

For the optimal classification threshold of 0.54 (Figure 4), all samples with expressing yawning (20 out of 20) were correctly classified. Three of 30 smile samples and two of 30 singing samples were also classified as yawning. In this way, the overall result of the classification was 93.75%.

We have compared the result achieved in our solution with the results known from the publication. The set of results is presented in Table 1. It is worth emphasizing the fact that we used the typical, standard video databases. Databases that are

Table 1: Results of tests – comparison to the previous works.

	Author(s)	Number of videos/ participants	Accuracy
1.	Saradadevi, M. and Bajaj, P. (Saradadevi and Bajaj, 2008)	337	81% - 86%
2.	Li, L., Chen, Y., Li, Z. (Li, et al., 2009)	4	95%
3.	Abtahi, S., Shirmohammadi, S., Hariri, B., Laroche, D., Martel ,L. (Abtahi, et al., 2013)	20	60%
4.	Alioua, N., Amine, A., Rziza, M. (Alioua, et al., 2014)	6	98%
5.	Hasan, M., Hossain, F., Thakur, J.M., Podder, P. (Hasan, et al., 2014)	150	95%
6.	Omidyeganeh, M., Shirmohammadi, S., Abtahi, S., Khurshid, A., Farhan, M., Scharcanski, J., Hariri, B., Laroche, D., Martel, L. (Omidyeganeh, at al., 2016)	107	65% - 75%
7.	Manu, B.N. (Manu, 2016)	4	94%
8.	**Proposed method in this paper**	80	93.75%

commonly used to test the facial image analysis. Because we considered three independent cases (yawning, smiling, singing), so we used three similar (numerically comparable) data groups - 20 videos of yawning, 30 videos of smiling and 30 videos of singing. Perhaps 80 videos are not the basis for generalizing the results. But we can show the published methods that are tested on 4, 6 or 10 cases (videos or participants) only (see Table 1). In this context, our experiments confirm the correctness of the proposed method.

By analyzing of previous publications, it can be concluded that high accuracy (94% or higher) was achieved primarily for experiments in very small groups: 4 - 6 videos / participants. If the groups were larger (20 videos / participants or more) then the accuracy was on the level of 60% - 86%. The paper (Hasan, et al., 2014) is an exception: 150 tests were conducted to give 95% as the value of accuracy. However, in this solution two independent cameras and advanced, complicated algorithm have been used. In this context it can be concluded that our method achieves very good results, which exceeded the known solutions.

5 SUMMARY

The yawning detection algorithm has been presented in the paper. The algorithm recognizes the yawning and distinguishes it from singing and smiling. The introduced method is based on a two-step recognition procedure. In the first step there is segmentation of the mouth area in order to obtain 6 facial landmarks which describing the shape of the mouth. The purpose of the second stage is to detect yawning with particular regard to differences for other states of mouth opening.

The proposed solution simplifies the detection procedure of yawning considerably. The effect of the first stage is the normalized image of the mouth. So there is neither problem with geometric distortions nor a deviation from the horizontal line, which have always been a problem in solutions known from publications. On the other hand, the use of ROI and facial landmarks gives the possibility of proposing a simple measure of the mouth opening. The classification of mouth stage (and thus yawning detection) by analyzing the value of proposed measure is more effective than calculating the sum of pixels of the mouth opening - often used in known solutions. On the other hand, the size of the mouth is a strictly individual feature and additionally depends on videos resolution. Thus, there may be problems in yawning identification based on the size of the opened mouth. This problem does not occur at all if we use facial landmarks and our measure.

The aim of the work has been achieved. We have managed to propose a geometric measure that allows distinguishing yawning from other mouth openings (smiles and singing). This measure takes into account the dynamics of changes in the shape of the mouth and in addition is simple and effective.

The introduced method has given very good results. This has been confirmed experimentally. Effectiveness of the presented algorithm was verified using 80 videos collected from three independent databases. The developed method achieved value of accuracy on the level of 93.75%. Comparison of the obtained results with the methods described in the literature has been also carried out. The achieved accuracy of 93.75% for 80 video samples puts proposed method among the best solutions of recent years.

ACKNOWLEDGEMENTS

This paper has been based on the results of a research task carried out within the scope of the fourth stage of the National Programme "Improvement of safety and working conditions" partly supported in 2017–2019 --- within the scope of research and development --- by the Ministry of Labour and Social Policy. The Central Institute for Labour Protection -- National Research Institute is the Programme's main coordinator.

REFERENCES

Abtahi, S., Shirmohammadi, S., Hariri, B., Laroche, D., Martel, L. 2013. A yawning measurement method using embedded smart cameras. In: *Proc. of 2013 IEEE International Instrumentation and Measurement Technology Conference*. 1605-1608. doi: 10.1109/ I2MTC.2013.6555685

Alioua, N., Amine, A., Rziza, M. 2014. Driver's Fatigue Detection Based on Yawning Extraction. International *Journal of Vehicular Technology*. Vol. 2014, Article ID 678786, 7 pages. doi: 10.1155/2014/678786

Azim, T., Jaffar, M.A., Ramzan, M., Mirza, A.M. 2009. Automatic Fatigue Detection of Drivers through Yawning Analysis. In: Ślęzak D., Pal S.K., Kang BH., Gu J., Kuroda H., Kim T. (eds) *Signal Processing, Image Processing and Pattern Recognition. Communications in Computer and Information Science*. Vol 61. Springer. doi: 10.1007/978-3-642-10546-3_16

Bosch at CES 2017: New concept car with Driver Monitor. http://www.bosch-presse.de/pressportal/de/en/bosch-at-ces-2017-new-concept-car-with-driver-monitor-84050.html (retrieved March 31, 2017)

Coetzerm R.C., Hanckem G.P. 2009. Driver Fatigue Detection: A Survey. In: *Proc. of AFRICON '09* Nairobi, Kenya, 1-6, doi: 10.1109/AFRCON.2009.5308101

Dibeklioglu H., Salah A.A., Gevers T. 2012. Are You Really Smiling at Me? Spontaneous versus Posed Enjoyment Smiles, In: *Proc. European Conference on Computer Vision (ECCV 2012)*. 525-538..

Fatigue Risk Assessment. Gain visibility to hidden risk factors to see the full impact of fatigue and distraction on your operations. CATERPILAR (2017) http://www.cat.com/en_US/support/operations/frms/fra.html (retrieved April 10, 2017)

Friedrichs, F., Bin Yang 2010. Camera-based Drowsiness Reference for Driver State Classification under Real Driving Conditions. In: *Proc. of 2010 IEEE Intelligent Vehicles Symposium*, San Diego USA, 101-106. doi: 10.1109/IVS.2010.5548039

Hasan, M., Hossain, F., Thakur, J.M., Podder, P. 2014. Driver Fatigue Recognition using Skin Color Modeling. *International Journal of Computer Applications*, 97 (16), 34-40.

Ibrahim, M.M., Soraghan, J.J., Petropoulakis, L., Di Caterina, G. 2015. Yawn analysis with mouth occlusion detection. *Biomedical Signal Processing and Control*. 18, 360-369. doi: 10.1016/j.bspc.2015.02.006

Kumar, N., Barwar, N.C 2014. Detection of Eye Blinking and Yawning for Monitoring Driver's Drowsiness in Real Time. *International Journal of Application or Innovation in Engineering & Management (IJAIEM)*. 3 (11), 291-298.

Li, L., Chen, Y., Li, Z. 2009. Yawning detection for monitoring driver fatigue based on two cameras. In: *Proc. of 12th International IEEE Conference on Intelligent Transportation Systems*, 1-6. doi: 10.1109/ITSC.2009.5309841

Livingstone, S.R., Peck, K., Russo, F.A. 2012. RAVDESS: The Ryerson Audio-Visual Database of Emotional Speech and Song. Paper presented at: *22nd Annual Meeting of the Canadian Society for Brain, Behaviour and Cognitive Science (CSBBCS)*, Kingston, ON

Manu B.N. 2016. Facial features monitoring for real time drowsiness detection. In: *Proc. of 12th International Conference on Innovations in Information Technology (IIT)*, doi: 10.1109/INNOVATIONS.2016.7880030

National Sleep Foundation. http://sleepfoundation.org/ (retrieved April 11, 2017)

Ochocki, M., Sawicki, D. 2016. Facial landmarks localization using binary pattern analysis, *Przeglad Elektrotechniczny*, 92 (11), 43-46. doi: 10.15199/48.2016.11.11

Omidyeganeh, M., Shirmohammadi, S., Abtahi, S., Khurshid, A., Farhan, M., Scharcanski, J., Hariri, B., Laroche, D., Martel, L. 2016. Yawning Detection Using Embedded Smart Cameras. *IEEE Transactions on Instrumentation and Measurement*. 65 (3), 570-582. doi: 10.1109/TIM.2015.2507378

Osh in figures 2011: Annex to Report: Occupational Safety and Health in the Road Transport sector: An Overview. National Report: Finland. In: *Report of the European Agency for Safety and Health at Work*. Luxembourg: Publications Office of the European Communities.

Rodzik, K., Sawicki, D. 2015. Recognition of the human fatigue based on the ICAAM algorithm. In: *Proc. of the 18th International Conference on Image Analysis and Processing (ICIAP 2015). LNCS 9280* Springer. 373-382. doi: 10.1007/978-3-319-23234-8_35

Saradadevi, M., Bajaj, P. 2008. Driver Fatigue Detection Using Mouth and Yawning Analysis. International *Journal of Computer Science and Network Security*, 8 (6), 183-188.

Sigari, M-H., Pourshahabi, M-R., Soryani, M., Fathy, M. 2014. A Review on Driver Face Monitoring Systems for Fatigue and Distraction Detection. *International Journal of Advanced Science and Technology*, 64, 73-100. doi: 10.14257/ijast.2014.64.07

The Yawn-O-Meter (How Long Can You Last?) https://www.youtube.com/watch?v=AJXX4vF6Zh0 (retrieved February 30, 2015)

Viola, P., Jones, M. 2001. Rapid Object Detection using a Boosted Cascade of Simple Features. *Computer Vision and Pattern Recognition (CVPR)*. 1, 511-518. doi: 10.1109/CVPR.2001.990517

Wang, T., Pendeli, S. 2005. Yawning detection for determining driver drowsiness. In: *Proc of IEEE Int. Workshop VLSI Design & Video Tech*. 373-376. doi: 10.1109/IWVDVT.2005.1504628

Wang, T., Shi, P. 2005. Yawning detection for determining driver drowsiness. In: *Proc. of 2005 IEEE International Workshop on VLSI Design and Video Technology*, 373-376. doi: 10.1109/IWVDVT.2005.1504628

Want a free mua face chart? 2010. http://smashinbeauty.com/want-a-free-mua-face-chart/ Retrieved 25 January 2013

Weiwei Liu. Haixin Sun, Weijie Shen 2010. Driver Fatigue Detection through Pupil Detection and Yawning Analysis. In: *Proc. of 2010 International Conference on Bioinformatics and Biomedical Technology (ICBBT)*, Chengdu China, 404-407. doi: 10.1109/ICBBT.2010.5478931