Driver Fatigue Detection Using Mouth and Yawning Analysis

Mandalapu Saradadevi[†] **Dr. Preeti Bajaj**[††]

B.D.College of Engineering, Sevagram, India G.H.Raisoni College of Engineering, India

Summary
Driver fatigue is an important factor in large number of accidents. There has been much work done in driver fatigue detection. This paper presents driver fatigue detection based on tracking the mouth and to study on monitoring and recognizing yawning. The authors proposed a method to locate and track driver's mouth using cascade of classifiers proposed by Viola-Jones for faces. SVM is used to train the mouth and yawning images. During the fatigue detection mouth is detected from face images using cascade of classifiers. Then, SVM is used to classify the mouth and to detect yawning then alert Fatigue.

Key words:
Driver fatigue, Fatigue detection, Driver monitoring system, Yawning, Support Vector Machine (SVM).

1. Introduction

The increasing number of traffic accidents due to a diminished driver's vigilance level has become a serious problem for society. Statistics show that 20% of all the traffic accidents are due to drivers with a diminished vigilance level [1]. Furthermore, accidents related to driver hypo-vigilance are more serious than other types of accidents, since sleepy drivers often do not take correct action prior to a collision. For this reason, developing systems for monitoring driver's level of vigilance and alerting the driver, when he is drowsy and not paying adequate attention to the road, is essential to prevent accidents. The prevention of such accidents is a major focus of effort in the field of active safety research.

In the last decade many researchers have been working on the development of the driver monitoring systems using different techniques. Driver's state of vigilance can be characterized by driver performance with a focus on the vehicle behavior or based on driver's physiological / physical state. The best accurate detection techniques are based on physiological phenomena of drivers like brain waves, heart rate, pulse rate and respiration [2]. These techniques are intrusive, since they need to attach some electrodes on the drivers body and causing annoyance to them. So this paper is focused on non-intrusive methods of fatigue detection mainly on measures of the driver's state.

People in fatigue show some visual behaviors easily observable from changes in their facial features like eyes, head, mouth and face [5]. Computer vision can be a natural and non-intrusive technique to monitor driver's vigilance. Faces as the primary part of human communication have been a research target in computer vision for a long time. The driver fatigue detection is considered as one of the most prospective commercial applications of automatic facial expression recognition [3]. Automatic recognition (or analysis) of facial expression consists of three levels of tasks: face detection, facial expression information extraction, and expression classification. In these tasks, the information extraction is the main issue for the feature based facial expression recognition from an image sequence [4]. It involves detection, identification and tracking facial feature points under different illuminations, face orientations and facial expressions.

Some common assumptions in previous face related works were: frontal facial views, constant illumination, and the fixed lighting source. Unfortunately these assumptions are not realistic. In the application of real world facial expression understanding, one has to consider at least three issues: capturing the full features in a variety of lighting conditions and head motion, multiple and non rigid object tracking, and the self-occlusion of features.

The process of falling asleep at the wheel can be characterized by a gradual decline in alertness from a normal state due to monotonous driving conditions or other environmental factors; this diminished alertness leads to a state of fuzzy consciousness followed by the onset of fatigue [6]. The critical issue that a fatigue detection system must address is the question of how to accurately and early detect fatigue at the initial stage. Possible non intrusive techniques for detecting fatigue in drivers using computer vision are

- Methods based on eye and eyelid movements
- Methods based on head movement
- Methods based on mouth opening

The authors have chosen methods based on mouth opening and yawning. Many of the previous researches deal with yawning detection focuses their methods on geometric features of the mouth. [7]. There are some disadvantages in yawning detection using geometric features of the

Manuscript received June 5, 2008.
Manuscript revised June 20, 2008.

mouth. First, left and right mouth corners are obvious feature points, but the lip positions are difficult to detect precisely. At the same time, lips move more acutely, which makes the lip detection more difficult. Third, geometric features are liable to pose and have more difference for individual. Here a method is proposed to locate and track a driver's mouth using cascade of classifiers proposed by Viola-Jones for faces. SVM is used to train the mouth and yawning images. During the fatigue detection mouth is detected from face images using cascade of classifiers. Then, SVM is used to classify the mouth regions to detect yawning then alert Fatigue.

2. Related Work

Haisong Gu et al. [3] proposes a graph-based reliability propagation to tackle the occlusion problem and verify the tracking results and their experimental results show validity of their active approach to real-life facial tracking under variable lighting conditions, head orientations, and facial expressions.

Xiao Fan et al. [7] gives to locate and track a driver's mouth movement using a CCD camera to study on monitoring and recognizing a driver's yawning and their experiment results show that Gabor coefficients are more powerful than geometric features to detect yawning and the average recognition rate is 95% which has more than 20% improvement.

Paul Viola et al. [8] have described a face detection framework that is capable of processing images extremely rapidly while achieving high detection rates as a process for training an extremely simple and efficient classifier which can be used as a "supervised" focus of attention operator and they present a set of experiments in the domain of face detection.

J. Nesvadba et al. [9] provides the potential of the Cassandra Framework's modular approach – using SUs for individual services - in combination with face-related content analysis algorithms and their framework provides an easy-to-use prototyping environment enabling the real-time execution of efficient and heterogeneous face-related algorithms, such as omni-directional face detection, pose estimation and face tracking in a distributed environment.

Yoav Freund et al. [10] introduces the boosting algorithm AdaBoost, and explains the underlying theory of boosting, including an explanation of why boosting often does not suffer from overfitting as well as boosting's relationship to support-vector machines and also they describe some of the basic underlying theory of boosting, including an explanation of why it often tends not to overfit.

Christopher J.C. Burges [11] give numerous examples and proofs of most of the key theorems and they how Support Vector machines can have very large (even infinite) VC dimension by computing the VC dimension for homogeneous polynomial and Gaussian radial basis function kernels and describes linear Support Vector Machines (SVMs) for separable and non-separable data, working through a non-trivial example in detail.

Chih-Wei Hsu et al. [12] proposes a simple procedure, which usually gives reasonable results and also they do not intend to solve challenging or difficult problems and they briefly introduce SVM basics which are necessary for explaining our procedure.

Theodoros Evgeniou et al. [13] discussed well known as well as emerging learning techniques such as Regularization Networks and Support Vector Machines which can be justified in term of the same induction principle and they overview the main concepts of Statistical Learning Theory, a framework in which learning from examples can be studied in a principled way.

Boser, B. E. et al. [15] developed a training algorithm that automatically tunes the capacity of the classification function by maximizing the margin by training examples and class boundary, optionally after removing some atypical or meaningless examples from the training data.

Cortes, C. et al. [16] construct a new type of learning machine, the so-called support-vector Network and also they compare the performance of the support-vector network to various classical learning algorithms that all took part in a benchmark study of Optical Character Recognition.

Yoshihiro Takei et al. [19] discussed the method to estimate a driver's fatigue through steering motion. They applied the Chaos theory to explain the change of steering wheel motion.

3. Mouth Extraction Algorithm

For Mouth detection the authors have used the detection algorithm proposed by Paul Viola and Micheal.J.Jones [8] used for Face detection by using cascade of classifiers. The overall proposed system is depicted in Figure 1

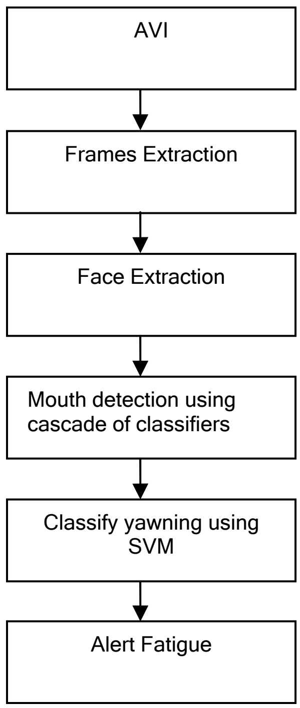


Fig. 1 Proposed System

3.1 Mouth Detection Algorithm

This mouth detection procedure classifies images based on the value of simple features. There are many motivations for using features rather than the pixels directly [8]. This image-based detection algorithm works on uncompressed images and has proven to be robust under various lighting conditions. The method is based on a cascade of boosted classifiers of simple Haar-wavelet like features on different scales and positions. The features are brightness and contrast-invariant and consist of two or more rectangular region pixel-sums that can be efficiently calculated by the canny integral image [9]. The feature set is over complete and an adaptation of the AdaBoost [10] learning algorithm is proposed to select and combine features into a linear classifier. To speed up detection a cascade of classifiers is used such that every classifier can reject an image. All classifiers are trained to reject part of the candidates such that on average only a low amount of features are used per position and scale. After all possible mouth candidates are obtained, a grouping algorithm reduces groups of mouth candidates into single positive detections.

A weak classifier thus consists of a feature (f), a threshold (θ) and a polarity (p) indicating the direction of the inequality:

$$h(x, f, p, \theta) = \begin{cases} 1 & if \;\; pf(x) < p\theta \\ 0 & otherwise \end{cases}$$

Here x is a 24 × 24 pixel sub-window of an image. Figure 2 displays the training of classifiers for mouth detection.

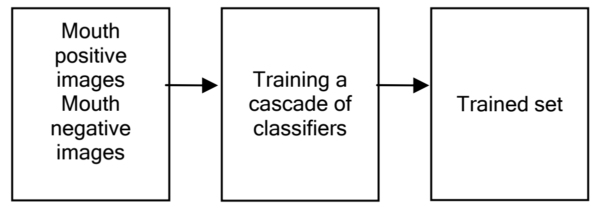
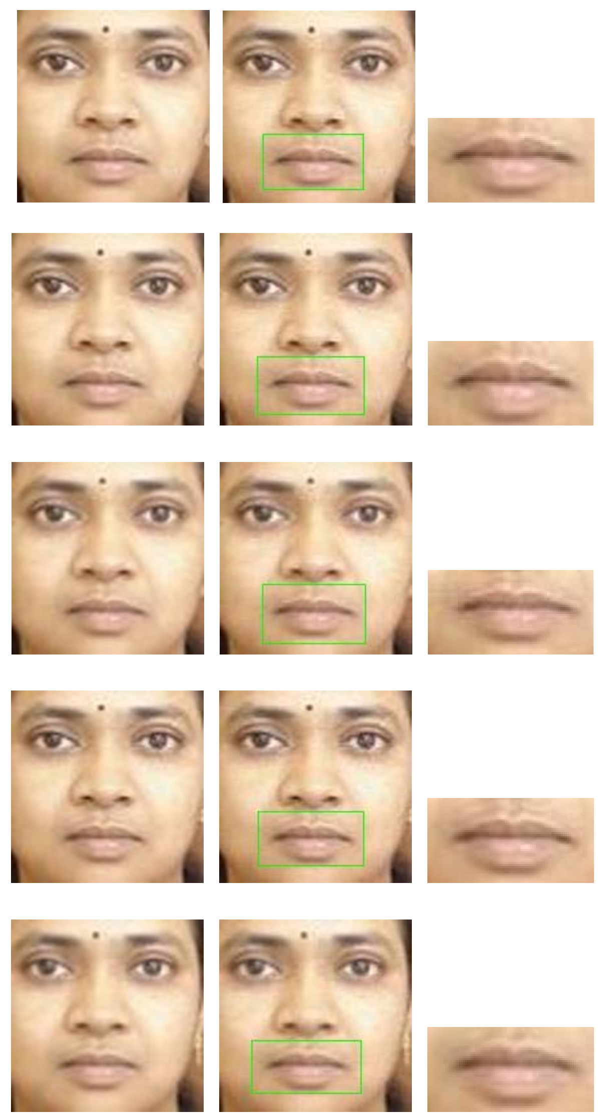


Fig. 2 Training of Classifiers

Figure 3 displays the list of Images used, Mouth Detection and extracted Mouths.



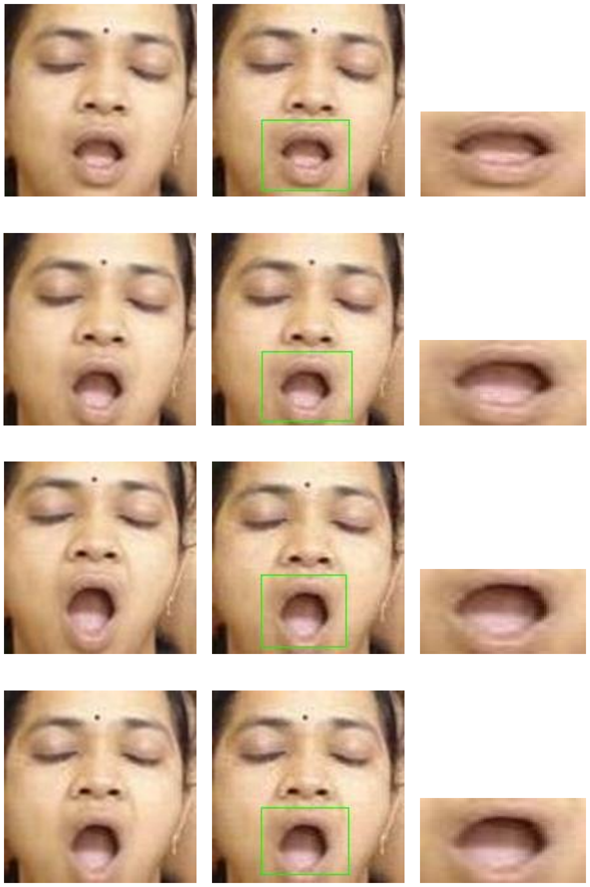


Fig. 3 Mouth Extraction from the Face Images

4. Fatigue Detection

Fatigue Detection has got two phases, one is training phase and the other one is detection phase which discussed below.

4.1 Support Vector Machine

SVM (Support Vector Machine) is a useful technique for data classification [11] - [14]. A classification task usually involves with training and testing data which consist of some data instances. Each instance in the training set contains one “target value" (class labels) and “several attributes" (features). The goal of SVM is to produce a model which predicts target value of data instances in the testing set which are given only the attributes. Figure 4 depicts the training of SVM and Figure 5 and Figure 6 lists the Training set of mouths used for SVM.

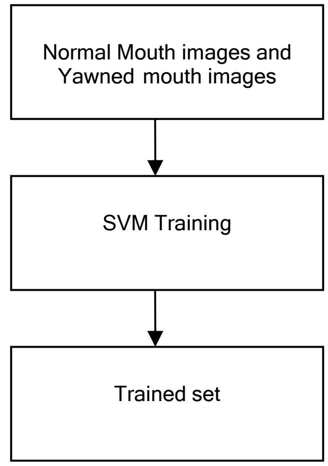


Fig. 4 SVM Training

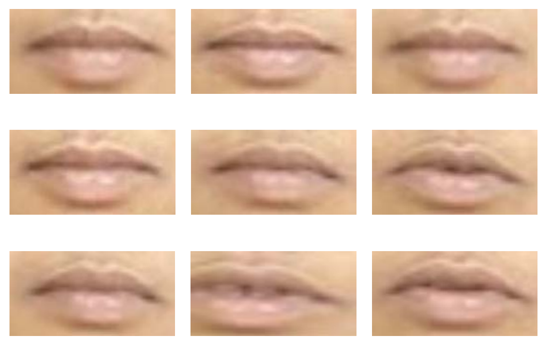


Fig. 5 Normal Mouths for SVM Training

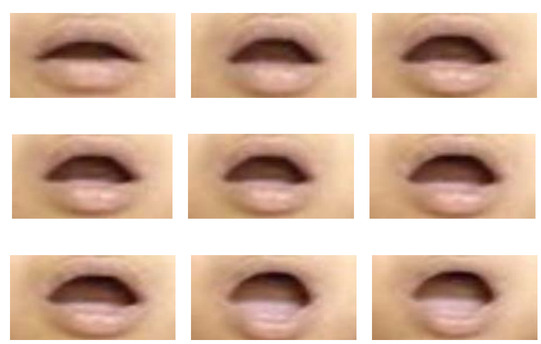


Fig. 6 Yawning Mouths for SVM Training

Given a training set of instance-label pairs (xi, yi), $i = 1,\ldots,l$ where $xi \varepsilon R^n$ and $y\varepsilon\{1,-1\}^l$, the support vector machines (SVM) [15]-[16] require the solution of the following optimization problem:

$$\min_{w,b,\xi} \quad \frac{1}{2} w^T w + C \sum_{i-1}^{l} \xi_i$$

$$subject \ \ to \quad y_i(w^T\phi(x_i)+b) \geq 1-\xi_i,$$
$$\xi_i \geq 0.$$

Here training vectors x_i are mapped into a higher (maybe infinite) dimensional space by the function. Then SVM finds a linear separating hyperplane with the maximal margin in this higher dimensional space. C > 0 is the penalty parameter of the error term. Furthermore, $K(x_i, x_j) \equiv \phi(x_i)^T \phi(x_j)$ is called the kernel function. Here we have used Radial basis function (RBF) kernel and given by

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0$$

Here γ is kernel parameter.

The following are the steps involved in the training of SVM.

Steps:

- Transform data to the format of an SVM software
- Conduct simple scaling on the data
- Consider the RBF kernel $K(x, y) = e^{-\gamma\|x-y\|^2}$
- Use cross-validation to find the best parameter C and γ
- Use the best parameter C and γ to train the whole training set.

SVM requires that each data instance is represented as a vector of real numbers. Hence, if there are categorical attributes, we first have to convert them into numeric data. Scaling them before applying SVM is very important. [17] explains why the data is scaled while using Neural Networks and most of considerations also apply to SVM. The main advantage is to avoid attributes in greater numeric ranges dominate those in smaller numeric ranges. The RBF kernel nonlinearly maps samples into a higher dimensional space, so it, unlike the linear kernel, can handle the case when the relation between class labels and attributes is nonlinear. The second reason is the number of hyperparameters which influences the complexity of model selection. The polynomial kernel has more hyperparameters than the RBF kernel. Finally, the RBF kernel has less numerical difficulties. One key point is $0 < K_{ij} \leq 1$ in contrast to polynomial kernels of which kernel values may go to infinity $(\gamma x_i^T x_j + r > 1)$ or zero $(\gamma x_i^T x_j + r < 1)$ while the degree is large. Moreover, we must note that the sigmoid kernel is not valid (i.e. not the inner product of two vectors) under some parameters [18]. There are several advantages of SVM's. The most important advantage is that during the training process, only a few vectors out of the training set are selected to become support vectors. This reduces the computational cost and provides a better generalization. Another advantage is that there are no local minima in the quadratic program, so the found solution is always the optimum of the given training set. Finally the main advantage is that the solution is not dependent on start conditions unlike neural networks.

5. Results

The authors collected few videos and selected about 20 yawning images and more than 1000 normal images as the data set. From each video, 10 yawning images and 10 normal images are given to cascade of classifiers for training. Same samples of Video are trained by SVM to classify normal and yawing mouths. Then the proposed approach is tested and the results of classification are presented in table 1.

Table1: Results of classification of Normal and yawning mouths

State	Normal	Yawning	Correct Rate
Normal	260	40	86%
Yawning	7	30	81%

6. Conclusion

In this paper the authors have proposed a method to locate and track a driver's mouth using cascade of classifiers training and mouth detection. They also trained the mouth and yawning images using SVM. Finally, SVM is used to classify the mouth regions to detect yawning then alert Fatigue. The experimental results show that proposed method gives better results than methods using geometric features. The proposed method detects yawning, alert fatigue earlier, and will facilitate to make drive safer. In future, the authors will capture more video clips to train and test the proposed method. The main goal is to develop a system to combine more features including mouth features, eye features and head tracking to monitor driver fatigue.

References

[1] L.M. Bergasa, J. Nuevo, M.A. Sotalo, and M. Vazquez, "Real-time system for monitoring driver Vigilance," in Proc. Intelligent Vehicle Symp., Parma, Italy, pp.78-83, 2004.

[2] J. Healey, and R. Picard, "Smart Car: detecting driver stress," in Proc.15th Int.Conf. Pattern Recognition, Barcelona, Spain, vol. 4, pp. 218-221, 2000.

[3] H. Gu, Q. Ji, and Z.W. Zhu, "Active facial tracking for fatigue detection," in Proc.Workshop on Application of Computer Vision, Orlando, pp. 137-142, 2002.

[4] Quong Wang, Jingyu Yang, Mingwu Ren, Yujie Zheng, "Driver Fatigue Detection: A Survey", in Proc Intelligent

control and Automation, Dalion, China, pp 8587- 8591, 2006.
[5] "AWAKE - System for Effective Assessment of Driver Vigilance and Warning According to Traffic Risk Estimation" from http://www.awake-eu.org/back.html
[6] Ueno, H.. Kaneda, M.. Tsukino, M. "Development of drowsiness detection system" IEEE, Vehicle Navigation and Information Systems Conference, 1994. Proceedings. 1994 Volume, Issue, 31 Aug-2 Sep 1994.
[7] Xiao Fan Bao-Cai Yin Yan-Feng Sun "Yawning Detection for Monitoring Driver Fatigue" IEEE Machine Learning and Cybernetics, 2007 International Conference on 19-22 Aug. 2007.
[8] Viola, P. and Jones, M.J., Robust real-time face detection. International Journal of Computer Vision. v57 i2. 137-154.
[9] J. Nesvadba, A. Hanjalic, P. Fonseca, B. Kroon, H. Celik, E. Hendriks, 'Towards a real-time and distributed system for face detection, pose estimation and face-related features', Invited Paper, Proc. Int. Con. on Methods and Techniques in Behavioral Research, Wageningen, The Netherlands, 2005.
[10] Y. Freud and R. Schapire, A short introduction to Boosting, J. of Japanese Society for AI, 14(5): 71~780, 1999
[11]Christopher J. C. Burges. A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery, 2(2):121{167, 1998.
[12] Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin. A practical guide to support vector classification. Technical report, Department of Computer Science and Information Engineering National Taiwan University Taipei 106, Taiwan, 2003.
[13] Theodoros Evgeniou, Massimiliano Pontil, and Tomaso Poggio. Statistical learning theory: A primer. International Journal of Computer Vision, 38(1):9-13, 2000.
[14]T. Hastie, R. Tibshirani, and J. Friedman. The Elements of Statistical Learning. Springer, 2001. HAS t 01:1 1.Ex.
[15] Boser, B. E., I. Guyon, and V. Vapnik (1992). A training algorithm for optimal margin classifiers. In Proceedings of the Fifth Annual Workshop on Computational Learning Theory, pp. 144{152. ACM Press.
[16] Cortes, C. and V. Vapnik (1995). Support-vector network. Machine Learning 20,273-297.
[17] Sarle, W. S. (1997). Neural Network FAQ. Periodic posting to the Usenet newsgroup comp.ai.neural-nets.
[18] Vapnik, V. (1995). The Nature of Statistical Learning Theory. New York, NY: Springer-Verlag.
[19] Takei, Y. Furukawa, Y.: Estimate of driver's fatigue through steering motion. In: Man and Cybernetics, 2005 IEEE International Conference. (Volume: 2, On page(s): 1765-1770 Vol. 2)

Mandalapu Saradadevi received B.Tech degree in Electronics & Communication Engineering from V.R.Siddhardha Engineering college, Vijayawada and M.E in Electronics & Communication Engineering from Andhra university ,Visakhapatnam. She is currently pursuing PhD from RTM Nagpur University, Nagpur, India. She has over 3 publications in refereed International Conferences and Journals. She has been with B.D.College of Engineering, Sevagram since 1994. She also served in the faculty of K.L.College of Engineering, Vijayawada. She is member of Board of studies in RTM Nagpur university.

Dr. Preeti Bajaj holds a PhD degree in Electronics Engineering. Currently she is working as Principal, G.H.Raisoni College of Engineering, Nagpur, India. Her research interests include various domains like Soft Computing and optimizations, Hybrid Systems & Applications of Fuzzy logic in areas of Intelligent Transportation Systems such as Driver Monitoring/ Fatigue Systems. Her professional societies affiliation include Member- Institute of Engineers, Senior Member- IEEE, LM-ISTE, ILM-UK, LM-CSI. She has chaired many technical sessions at International Conferences in India and abroad. She is also a recipient of Best Teacher Award in the year 2001-02. She worked as reviewer for many International Conferences and Journals in India and abroad. Currently she is also working as General Chair for First International Conference on Emerging Trends in Engineering & Technology(ICETET-08) at GHRCE Nagpur, India. She has over 30 publications in refereed International Conferences and Journals. She has offered consultancy to National Highways Authority of India for designing and proposing various Intelligent Transportation projects on National Highways.