Gabor/PCA/SVM-Based Face Detection for Driver's Monitoring

Djamel Eddine Benrachou, Brahim Boulebtateche, and Salah Bensaoula
University Badji Mokhtar, Department of electronic, Annaba, Algeria
djamelben.univ@gmail.com, {bbouleb, bensaoula_salah}@yahoo.fr

***Abstract*—Driver fatigue cause each year a large number of road traffic accidents, this problem sparks the interest of research to move towards development of systems for prevention of this phenomenon. This article implements a face detection process as a preliminary step to monitor the state of drowsiness on vehicle's drivers. We propose an algorithm for pre-detection based on image processing and machine learning methods. A Gabor filter bank is used for facial features extraction. The dimensionality of the resulting feature space is further reduced by PCA technique and then follows a classification of Face/No Face classes using Support Vector Machine (SVM), for face detection. The simulation results on both databases namely PIE and ORL datasets show the efficiency of this approach.**

***Index Terms*— drowsiness, car driver, face detection, gabor filter, PCA, SVM classifier**

I. Introduction

Driver's drowsiness causes each year a large number of road traffic accidents. Statistics show that 10% to 20 % of accidents overall road traffic are due to the decrease level of driver's alertness [1].

The hypovigilance reduce the capacity to react, judge and analyze information and it is often caused by fatigue and/or drowsiness. However fatigue and drowsiness are different. The first one refers to a cumulative process producing difficulty to pay attention while the second one concerns the inability to stay awake. Therefore, it is important to monitor driver's vigilance level and issue an alarm when he is not paying attention.

Monitoring driver's responses are approached by a lot of methods, sensing physiological characteristics, driver operations, or vehicle response. These methods work well and give good indicators of vigilance state.

Recently, we can find detection systems using vehicle embedded cameras [2]. These systems analyze visual cues generated by the drowsiness such as eye blinking, the driver's gaze or positioning of the driver's head [3] because the decline of the head could be a good indicator of drowsiness.

This paper investigates the ability of Gabor representation and Support Vector Machine for visual features extraction and captures the important information by discriminatory method for face detection task. The idea is to decompose a face image into different spatial frequencies (scales) and orientations where salient discriminant features may appear.

Dimensionality reduction is adopted by PCA technique to create low dimensional features vectors for more convenient processing. SVM is used to extract relevant information from this low dimensional training data in order to construct a robust specific classifier. This method has been tested on two available AT&T (ORL) and (PIE) Databases of human faces. The statistical evaluation is presented for two different databases using both SVM's kernels namely linear and Gaussian kernels, implemented separately in order to detect the presence of a face or not.

A. Proposed Algorithm

The use of non-intrusive drowsiness detection methods requires several processing modules. In the proposed approach a first step of extracting essential features of the face detection is performed by applying the Gabor's representation on the image database. The advantage of this representation is that it allows us better spatial-frequency features localization. For the separation of features obtained by the Gabor filter bank, we use an SVM classifier.

Dimensionality reduction is applied using PCA to create low dimensional features vectors for more convenient processing. For the separation of the reduced features obtained by Gabor filter bank, we use an SVM classifier. The classification will be followed by a dynamic neural network module (TDNN). This phase of decision takes into account the dynamics of yawning and blinking. The structure of this algorithm is illustrated in "Fig. 1".

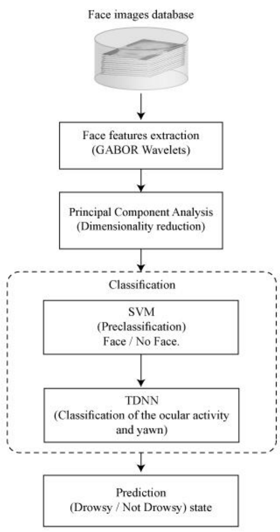


Figure 1. Flowchart of the detection algorithm (detecting the state of driver's drowsiness by analysis of visual cues).

Manuscript received October 15, 2012; revised December 22, 2012.

doi: 10.12720/joace.1.2.115-118

B. Feature Extraction

The features extraction step consists in transforming the input raw data into meaningful information. Thus, we obtain a reduction of the decision space which may accelerate the processing time. The Gabor filters provide a simultaneous representation in spatial and frequency domain. This representation is an optimal tool used for the purpose of local features extraction. It is efficient because it produces the same operating principle of simple cells in visual cortex of the mammal's brain and also properties in multi-directions, optimal for measuring local spatial frequencies. Based on these advantages, the Gabor representation is widely used in applications of image analysis and applications of face recognition [4], [5] and extraction of features such as facial expressions.

The family of complex Gabor wavelets could be represented as follows:

$$\psi_{\mu,\nu}(x) = \frac{\left\|k_{\mu,\nu}\right\|^2}{\sigma^2}\exp\left(\frac{-\left\|k_{\mu,\nu}\right\|^2\left\|x\right\|^2}{2\sigma^2}\right)\left[\exp\left(ik_{\mu,\nu}.x\right)-\exp\left(-\frac{\sigma^2}{2}\right)\right] \tag{1}$$

where σ is the standard deviation of Gaussian kernel, μ and ν define the orientation and scale of Gabor filter kernels and wave vector $k_{\mu,\nu}$ can be represented as:

$$k_{\mu,\nu} = k_\nu \exp(i\phi_\mu) \tag{2}$$

where $k_\nu = \frac{k_{\max}}{f^\nu}$ and $\phi_\mu = \frac{\mu\pi}{8}, k_{\max}$ is the maximum frequency, spacing factor between the kernels and the frequency domain.

The representation of the Gabor wavelet of the face image is the result of the convolution product of the input image with the family of Gabor kernels defined in (1).

$$G_{\mu,\nu}(x,y) = I(x,y) * \psi_{\mu,\nu}(x,y) \tag{3}$$

where * is the convolution operator. Consequently, the image $I(x,y)$ could be represented by the Gabor wavelets;

$$\left\{G_{\mu,\nu}(x,y), \nu = 0,...,4; \mu = 0,...,7\right\} \tag{4}$$

We applied the Gabor filters bank with eight orientations and five central frequencies

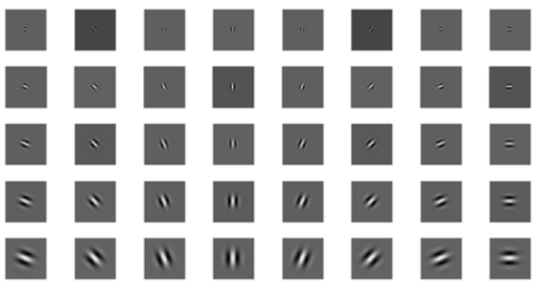


Figure 2. Representation of the Gabor wavelets with eight orientations and five center frequencies.

C. Dimentionality Reduction

Principal component analysis is applied for the dimensionality reduction task, knowing that the features vector obtained from the application of the Gabor representation resides in high dimensional space and learning in high dimensional space in not efficient since number of training examples cannot match the dimensionality to attain a good level of performance from a viewpoint of computation time. Therefore, our aim is to identify the lower dimensional subspace that spans the high-dimensional feature space and containing mainly the most useful information. Dimensionality reduction can be achieved by projection the high dimensional image into a lower dimensional image using projection basis which is optimal in mean-squared error sense. Principal component analysis (PCA) is a decorrelation technique to derive the desired orthonormal projection basis from high-dimensional data. Orthonormal projection basis is derived from finding the eigenvectors of the covariance matrix that captures the important features with expressive information.

D. Support vector Machine

Support vector machines (SVM's) are a very popular method for binary classification. The support vector classifier chooses one particular solution, the classifier which separates the classes with maximal margin.

SVM's has many advantages. A unique global optimum for its parameters can be found using standard optimization software. Nonlinear boundaries can be found without much extra computational effort.

Mathematically, this is an optimization problem that seeks to find a linear classifier $g(x) = \theta^T x_i + b$ minimizing a cost function given by, [6]

$$L = \frac{1}{2}\left\|\Theta\right\|^2 + \sum_{i=1}^{n}\alpha_i\left(y_i\left[\Theta^T x_i + b\right] - 1\right), \alpha_i \geq 0 \tag{5}$$

where we have a set of training examples X_i ; $i = \{1, 2, ..., n\}, n$ is the number of training examples. Each example is labeled $y_i \in \{1, -1\}$ indicating membership of each example for a specific class. In our case, the two specific classes are represented as FACE/NO FACE class respectively.

$$L = \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} y_i y_j \alpha_i \alpha_j x_i x_j \tag{6}$$

L should be minimized with respect to Θ and b , and maximized with respect to the Lagrange multipliers α_i. The so-called dual form of this optimization problem This formulation of the support vector classifier covers only a linear classifier for separable data. In our case, input space is mapped into high dimensional feature space, we use nonlinear kernels in the hope to get better linear separation of the training data.

To construct nonlinear decision boundaries, we used the Gaussian kernel with $\sigma^2 I$ as weighting matrix (the radial basis function kernel, RBF kernel is used in practice):

$$L=\sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}y_iy_j\alpha_i\alpha_jK\left(x_ix_j\right) \quad (7)$$

where K is a specific kernel.

An RBF kernel is given by

$$K\left(x_ix_j\right)=\exp\left(-\gamma\left\|x_i-x_j\right\|^2\right) \quad (8)$$

And $\gamma=\frac{1}{\sigma^2}$

Experimental results reported in this document use publicly available SVM library known as LIBSVM [7]. In this work, we use two kernel function, linear and RBF kernels, for evaluation in order to get the best classification of training data for face detection task.

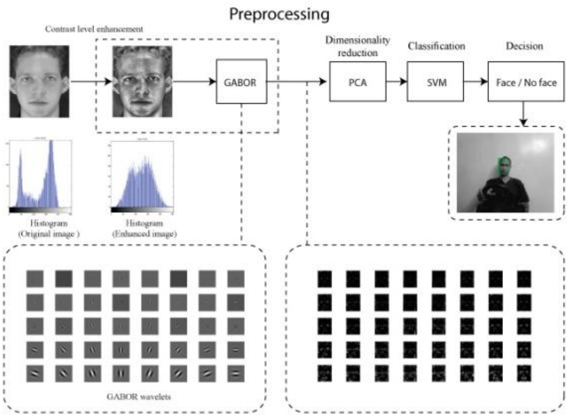


Figure 3. Flowchart presenting features extraction by gabor wavelets and SVM classification.

E. PIE Database

The CMU pose, illumination and expressions dataset, namely PIE database of human faces was collected between October and December 2000, it contains 41 368 images of 68 peoples. Each person's image is captured under 13 different poses, 43 different illumination conditions and with 4 several different expressions such neutral expression, to smile, to blink and talk. These particular expressions are supposed to be the four most common «expressions» in normal life.

Figure 4. "CMU Pose, Illumination, and Expression (PIE) database", "THE ROBOTICS INSTITUTE", carnegie mellon university.

F. ORL Database

The ORL database contains a set of face images, commonly used in the context of face recognition. This database contains various images of 40 distinct subjects, taken under different conditions, changes in lighting and facial expressions (eyes open / closed, smiling / not smiling), with presence or absence of particular features (glasses / no glasses). These images are captured with a dark homogeneous background and ten different poses for each individual.

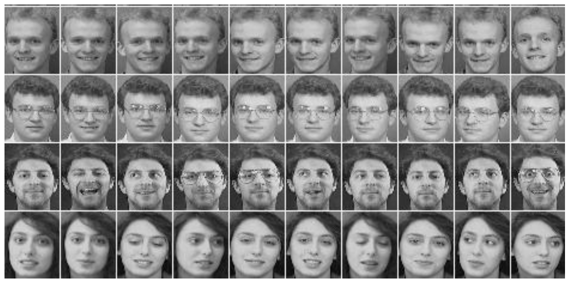
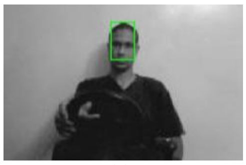


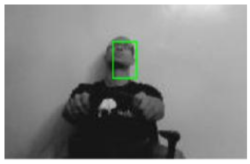
Figure 5. "ORL Database of Faces", "AT&T laboratories cambridge," cambridge university computer laboratory, "the digital technology group", 1992-1998.

II. Experimental Results

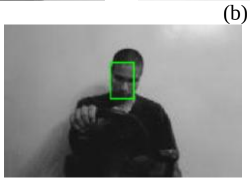
In this part a comparison performance between two SVM's kernels, namely linear and RBF kernel evaluated on the two different datasets, AT&T database and PIE dataset. The results of the face detection in different postures are illustrated in "Fig 6".



(a)



(b)



(c)

Figure 6. Face detection of the car driver in three different postures (a), (b), (c).

The results are presented as statistical measure such True Positive (TP), False Positive (FP), True Negative (TN), False Negative (FN) and F-score for a final evaluation, see Table I.

Our system for driver's face detection part is evaluated with both SVM's Kernels for two datasets, we use 100 face images for each datasets for training and testing part, for ORL-database this system provides success accuracy rate of 91,34% with an SVM's Gaussian kernel and 89,94% with a linear kernel. On PIE-dataset it yields 93, 07% with a Gaussian kernel and 91, 96% with linear kernel.

TABLE I. Statistical Results for Linear and RBF Kernel in Face Detection

Database	Algorithm	SVM's kernel	TP	TN	FP	FN	F.Score
ORL-Database	GABOR+ PCA+SVM	Linear	258	64	9	27	0,9348
		RBF	255	72	12	19	0,9427
PIE-Database	GABOR+ PCA+SVM	Linear	258	64	9	20	0,9473
		RBF	260	76	10	15	0,9542

The PIE-database is more challenging than AT&T dataset and gives better classification results for face detection task, because of larger lighting variations due to non-uniform illumination source. These conditions are nearer to the real world's conditions.

III. Conclusion

Face detection is a preliminary step in car driver drowsiness monitoring. In this work, we presented the potential of Gabor representation with Support Vector Machines for face detection task. The application of the

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, 2004, pp.78-83.

[2] M. C. Weng, C. T. Chenan, and H. C. Kao, "Remote Surveillance System for Driver Drowsiness in Real-time Using Low-cost Embedded Platform," i*n Proc. IEEE International Conference on Vehicular Electronics and Safety Columbus*, OH, USA. September 22-24, 2008.

[3] E. M. Chutorian, A. Doshi, and M. M. Trivedi, "Head Pose Estimation for Driver Assistance Systems: A Robust Algorithm and Experimental Evaluation," in *Proceedings 2007 IEEE Intelligent Transportation Systems Conference Seattle*, WA, USA, Sept. 30-Oct. 3, 2007.

[4] L. B. Ayinde and Y. H. Yang, "Face recognition approach based on rank correlation of Gabor filtered images," *Pattern Recognition*, vol. 35, 2002.

[5] C. J. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fischer discriminant model for face recognition," *IEEE Trans. Image Processing*, vol. 11, no. 4, pp. 467–476, 2002.

[6] V. Vapnik, "The Nature of Statistical Learning Theory," *Springer-Verlag*, ch. 5, pp. 133-137, 1995.

[7] C. C. Chang and C. J. Lin. (2001). "LIBSVM: a library for Support Vector Machine." [Online]. Available : http://www.csie.ntu.edu.tw/~cjlin/libsvm/

Djamel Eddine Benrachou was born in Annaba City in 1986, he received the bachelor's degree in Automatic from the faculty of Engineering Science Badji Mokhtar Annaba in 2008, and the academic master's degree in Automatic and Signals in 2010. He currently is a Ph.D student at Laboratory of Automatic and Signals- Annaba (LASA), faculty of engineering science Badji Mokhtar Annaba. His research interests are car driver's drowsiness detection.

Salah Bensaoula was born in Annaba City in 1959. He received the Engineer Degree of State from National Polytechnique of Algies in 1983, the DEA degree from Clermont-Ferrand (France) and the Doctorate Degree from Saint-Etienne University (France) in 1984 and 1987, respectively. His main interests of research include fault detection and isolation in industrial systems and man-machine communication. Gabor representation to extract features is followed by dimensionality reduction using PCA technique. An SVM classifier can then be trained using the feature vectors in order to build a robust specific classifier for driver's face detection. The compound formed by Gabor filters, PCA analysis and SVM classifier gives acceptable results and makes it very feasible to pursue the next task, namely the dynamic classification of the visual cues for drowsiness prevention.