Gabor-based dynamic representation for human fatigue monitoring in facial image sequences

Xiao Fan [a,b,*], Yanfeng Sun [a], Baocai Yin [a], Xiuming Guo [c]

[a] Beijing Key Laboratory of Multimedia and Intelligent Software, College of Computer Science and Technology, Beijing University of Technology, Beijing 100124, China
[b] Cogent Beijing R&D Center, Beijing 100089, China
[c] National Engineering Research Center for Information Technology in Agriculture, Beijing 100097, China

ARTICLE INFO

ABSTRACT

Human fatigue is an important reason for many traffic accidents. To improve traffic safety, this paper proposes a novel Gabor-based dynamic representation for dynamics in facial image sequences to monitor human fatigue. Considering the multi-scale character of different facial behaviors, Gabor wavelets are employed to extract multi-scale and multi-orientation features for each image. Then features of the same scale are fused into a single feature according to two fusion rules to extract the local orientation information. To account for the temporal aspect of human fatigue, the fused image sequence is divided into dynamic units, and a histogram of each dynamic unit is computed and combined as dynamic features. Finally, AdaBoost algorithm is exploited to select the most discriminative features and construct a strong classifier to monitor fatigue. The proposed method was tested on a wide range of human subjects of different genders, poses and illuminations under real-life fatigue conditions. Experimental results show the validity of the proposed method, and an encouraging average correct rate is achieved.

1. Introduction

According to the estimated data from the National Highway Traffic Safety Administration (NHTSA, 2005), 100,000 police-reported crashes are directly caused by driver fatigue each year, which result in an estimated 1550 deaths, 71,000 injuries, and $12.5 billion losses. In China, driver fatigue resulted in 3056 deaths in vehicular accidents in 2004, and caused 925 deaths in highway accidents that amounted to about 14.8%. Human fatigue has become an important factor for many traffic accidents. Therefore, it is essential to develop novel methods for monitoring human fatigue in order to improve transportation safety.

1.1. Previous work

In fact, there have been many attempts to achieve reliable fatigue monitoring for reducing the number of automobile accidents due to human fatigue in the last decade. These methods can be divided into three major categories as follows.

1.1.1. Driver physiological parameters

This method focuses on measuring physiological changes of drivers. It can accurately, validly, and objectively to determine fatigue and sleep of the drivers. A significant effort has been made to measure them in laboratory. The popular physiological parameters include electroencephalogram (EEG) (Abdul-Latif et al., 2004; Parikh and Micheli-Tzanakou, 2004; Wu et al., 2004a,b; Lin et al., 2005a,b), electrocardiogram (ECG) (Hayashi et al., 2005), EOG (Galley and Schleicher, 2004), and electromyography (EMG) Bonato et al., 2001. EEG is found to be useful in determining the presence of ongoing brain activity, and its measures have been used as the reference point for calibrating other measures of sleep and fatigue. Abdul-latif et al. (2004) found the mean RMS of EEG bands were increased during fatigue compared to the RMS value in the case of relaxation before fatigue, and the RMS value was seen to be greatest in the beta band and lowest in the gamma band. In literature (Parikh and Micheli-Tzanakou, 2004), Alpha waves (8–13 Hz) are observed with increasing amplitude when fatigue. Wu et al. (2004a,b) describes a system that combines EEG power-spectrum estimation, principal component analysis (PCA), and fuzzy neural network model to estimate/predict drivers' drowsiness level in a driving simulator. Lin et al. (2005a) proposed a system that combines EEG power spectra estimation, independent component analysis (ICA) and fuzzy neural network models to estimate

* Corresponding author. Address: Beijing Key Laboratory of Multimedia and Intelligent Software, College of Computer Science and Technology, Beijing University of Technology, Beijing 100124, China. Fax: +86 10 67396568.
  E-mail address: wolonghongni@yahoo.cn (X. Fan).

drivers' cognitive state in a dynamic virtual-reality-based driving environment. Lin et al. (2005b) developed a drowsiness-estimation system based on EEG by combining ICA, power-spectrum analysis, correlation evaluations, and a linear regression model to estimate a driver's cognitive state when he/she drives a car in a virtual-reality-based dynamic simulator. Unfortunately, most of these physiological parameters are obtained intrusively, making them unacceptable in practical applications.

1.1.2. Vehicle based performance

Fatigue can also be characterized by the behaviors of the vehicle that a driver operates. The vehicle based performance methods detect the behaviors of the drivers by monitoring the transportation hardware systems under the control of the drivers, such as steering wheel movements (Takei and Furukawa, 2005), driver's grip force on the steering wheel (Thum et al., 2003), speed, acceleration, lateral position, turning angle, changing course, braking and gear changing, etc. Thum et al. (2003) described an automobile driver fatigue detection method by monitoring the driver's grip force on the steering wheel, based on the variation in steering grip force due to fatigue or loosing alertness. In Takei and Furukawa (2005), the chaos theory was applied to explain the changes of steering wheel motion. If there is chaos in the motion, a strange trajectory called attractor can be found by applying the Takens' theory of embedding. The chaos characteristics are used to estimate a driver's fatigue. While these methods may be implemented non-intrusively, they are subject to several limitations, including the vehicle type, driver experiences, and driving conditions.

1.1.3. Driver physical conditions

These methods focus on detecting driver's physical changes during drowsiness by image-processing techniques. People in fatigue exhibit certain visual behaviors that are easily observable from changes in facial features. Visual behaviors that typically reflect a person's fatigue level include slow eyelid movement, smaller degree of eye openness (or even closed), frequent nodding, yawning, gaze (narrowness in the line of sight), sluggish in facial expression, and sagging posture. These image-processing based methods use optical sensors or video cameras to get visual fatigue cues.

Many efforts have been reported in the literatures on developing image-processing fatigue monitoring systems. When fatigue, the frequency and time of eye closed would increase. Much attention is paid to eye's features for fatigue detection. In 1998, based on the data of the Federal Highway Administration (Dinges et al., 1998), percentage of eyelid closure (PERCLOS) (Dinges and Grace, 1998) was taken as the most reliable and valid measure of a person's alertness level among several drowsiness detection measures. Liu et al. (2002) incorporated Kalman filtering and mean shift to track eyes, extracted eye's motion information as driver features. Hamada et al. (2003) extracted the driver's stage of drowsiness by means of the blink measurement with motion picture processing. Wang et al. (2003) used Gabor wavelets to extract texture features of drivers' eyes, and used neural network classifier to identify drivers' fatigue behavior. The doze stage was judged when the area of the iris becomes below a threshold (Miyakawa et al., 2004). Dong and Wu (2005) decided whether the driver was fatigue by detecting the distance of eyelids. Wang and Qin (2005) combined gray scale projection, edge detection with Prewitt operator and complexity function to judge whether the driver had his eyes closed. Fan et al. (2008) extracted LBP features of eye areas and used AdaBoost algorithm to determine whether a driver was fatigue.

When fatigue, people often yawn. Mouth features are extracted to detect fatigue (Wang et al., 2004; Wang and Shi, 2005). Wang et al. (2004) took the mouth region's geometric features to make up an eigenvector as the input of a BP ANN, and they acquired the BP ANN output of three different mouth states that represent normal, yawning or talking state, respectively. Wang and Shi (2005) represented the openness of the mouth by the ratio of mouth height to width, and detected yawning if the ratio was above 0.5 in more than 20 frames. Lu and Wang (2007) used directional integral projection to locate the midpoint of nostrils, recognized yawn by calculating the vertical distance between the midpoint of nostrils and the chin. To acquire accurately or reliably fatigue monitoring with the change in time, environment, or different persons, systems that can extract multiple visual cues which typically characterize the alertness level of a person and systematically combine them have been introduced (Bergasa et al., 2006; Zhu and Ji, 2004). Study shows that the performance of methods based on driver physical conditions is comparable with those methods using physiological signals. The major benefits of the visual measures using computer vision technologies are that they can be acquired non-intrusively.

Among those different methods, the best detection accuracy is achieved with methods that measure physiological parameters. Requiring physical contact with drivers (e.g., attaching electrodes), the methods based on driver physiological parameters are intrusive, causing annoyance to drivers. Good results have also been reported with methods that monitor driver physical conditions. These methods are non-intrusive and become more and more practical and popular with the rapid development of camera and computer vision technology. Most of these methods are spatial approaches. The visual features obtained from a single face image are used for classification. Although spatial approaches can achieve good recognition in some cases, they do not model the dynamics of fatigue and therefore do not utilize all information available in facial image sequences.

In facial expression recognition, according to psychologists (Bassili, 1979), an analysis for an image sequence produces more accurate and robust facial expression recognition. The facial motion is fundamental to the facial expression recognition. Therefore, more attention (Zhao and Pietikainen, 2007; Yang et al., 2007; Tong et al., 2007) has been shifted particularly towards modeling of dynamic facial expressions.

Human fatigue is a cognitive status that is developed over time. Dynamic features which capture the temporal pattern should be the optimal features to describe fatigue. To account for the temporal aspect of human fatigue, Ji et al. (2006) introduced a probabilistic framework based on dynamic bayesian networks (DBN) for modeling and inferring human fatigue by integrating information from various sensory data and certain relevant contextual information. States of nodes in a DBN satisfy the Markovian condition that is, the state at time t depends only on its immediate past. The dynamic fatigue model integrates the fatigue evidences spatially and temporally, therefore, leading to a more robust and accurate fatigue modeling and inference. But, in nature, no dynamic features are extracted in the system. In summary, there is limited research in extracting dynamic features from image sequences for fatigue monitoring. High accuracy in fatigue monitoring is still a challenge due to the complexity and variety of facial dynamics.

1.2. Overview of the proposed method

In this paper, the attention is focused to the dynamics of fatigue and using the spatial and temporal information from continuous image sequences to monitor the fatigue. To account for the temporal character of human fatigue and the multi-scale character of different facial behaviors, a novel Gabor-based dynamic representation with feature level fusion is proposed to monitor human fatigue from image sequences.

Fig. 1 gives an overview of the architecture of the proposed method. The system can be divided into training process and test