Fatigue Expression Recognition Algorithm Based on Reconstructed LBP-HOG (LBP-RHOG) Feature

Yaru Qi¹, Binjie Hu¹*, Jinguang Qiu¹ and Huijiao Lin²

¹School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China.
²China Research Institute of Radiowave Propagation.

Email: eeyaruqi@mail.scut.edu.cn; eebjiehu@scut.edu.cn; qjg.scut@gmail.com; 827145746@qq.com

Abstract. This paper aims at two problems in fatigue expression recognition. First, texture features extracted by LBP (Local Binary Patterns) descriptor are limited and can not effectively describe the edge and direction information of image. Second, structural features extracted by HOG (Histogram of Oriented Gradient) descriptor are redundant and its computational complexity is high. To fill the gaps of these two problems, we proposed a reconstructed LBP-HOG (LBP-RHOG) algorithm which extracted texture spectrum features and edge features from LBP operator and reconstructed HOG operator respectively and obtain fusion information by fusing these two features. To better evaluate the recognition performance, we complete simulation under a self-built fatigue expression database. The results show that our method has low computational complexity and high recognition rate, and can identify fatigue state well.

1. Introduction

Recently, with the rapid development of highway transportation and the increase of the number of vehicles, the incidence of road traffic accidents has been increasing year by year, causing huge property losses and casualties to countries around the world. Among them, driver fatigue and distraction are important factors leading to a large number of accidents. Therefore, how to effectively detect driver fatigue has become the focus and hot spot of researchers.

Research on driver fatigue identification can be divided into three categories: 1) vehicle-based methods; 2) behavior-based methods; 3) physiological signal-based methods[1][2]. In physiological methods, driver fatigue is detected by using physiological signals of the body, such as electroencephalogram (EEG) for brain activity, electrooculogram (EOG) for eye movement and electrocardiogram (ECG) for heart rate[3]. Recent studies have shown that compared with other methods, fatigue detection using physiological signals (especially EEG signals) is more reliable and accurate [4]. However, invasive measurements of physiological signals can impede driving, especially during long-term driving. The vehicle-based approaches collect signal data (e.g. steering wheel angle, Lane position, speed, acceleration and braking) from vehicular sensors to assess driving behavior. Although it is convenient to collect vehicle signals, these methods are not real-time for detecting driver fatigue. The behavior-based approaches use the on-board camera to capture the driver's face and rely on visual analysis to monitor driver behavior, including eye closure, blinking, yawning, head posture, gestures, facial expressions, and so on[7,8]. The visual-based behavioral analysis systems are attractive to the automotive industry because of these measures are effective and reliable for predicting driver fatigue. What's more, they are not intrusive to drivers.
The facial expression recognition is one of the most common visual-based behavioral analysis systems. It is closely related to people's mental state, such as anger, sadness, fatigue. The facial expression recognition can be used to detect driver fatigue effectively and accurately. The Existing methods applied to facial expression recognition mainly include: geometric-based correlation algorithms and appearance-based algorithms.

Geometric methods consider some predefined geometric points, which represent the shape and location of permanent facial features, such as eyebrows, eyes, nose and mouth. Lanitis A et al. proposed the Active Shape Model (ASM) which considers the shape and appearance. It also considers the effect of change with lighting, personal image, 3D poses, and facial expressions. Tsalakanidou et al. extended the ASM algorithm to develop a 3D face tracker that can handle measurement uncertainty and data loss due to occlusion and sensor errors[10]. Although the algorithm can handle measurement uncertainty and data loss, it can not distinguish similar expressions.

Appearance-based methods consider facial appearance change such as wrinkles. The most popular methods at present include: LBP (Local Binary Patterns) [11], HOG (Histogram of Oriented Gradients) [12], HOE (Histogram of Spatiotemporal Orientation Energy) [13] and Gabor Wavelets [14] and other feature extraction methods. Caifeng Shan et al. empirically evaluated face independent expression recognition based on local binary patterns of statistical local features in document [15]. Based on several databases, it checked several machine learning methods systematically. A large number of experiments have shown that LBP features are efficient for facial expression recognition. In addition, they also studied the LBP features of low-resolution facial expression recognition. In the experiment, it was observed that the LBP features are stable and robust in the useful range of low-resolution facial images. To capture dynamic texture information, Zhao and Pietikäinen extended the LBP neighborhood from a two-dimensional plane to a three-dimensional space and name it VLBP. Compared with LBP, VLBP takes into account changes with entire time domain [16]. It is also proved that the VLBP feature is invariant to translation and rotation and is robust to monotonic grayscale variations. Zhao and Pietikäinen also simplified the VLBP to reduce the complexity of the VLBP by considering only the co-occurrence of LBP-TOP. The texture is modeled using cascaded LBP histograms extracted from TOP-XY, XT and YT [16]. Chihang Zhao et al. proposed a novel and effective combined feature extraction method based on pyramid directional gradient histogram (PHOG) and contour wave transform (CT) for describing fatigue expressions of vehicle drivers. Random subspace ensemble (RSE) of linear perception (LP) classifier was used to classify three predefined categories of fatigue expressions, namely wakefulness table. Expressions, Moderate Fatigue Expressions and Severe Fatigue Expressions [17]. Junkai Chen et al. extended LBP neighborhood from two-dimensional plane to three-dimensional space, and proposed a new feature descriptor named HOG from three orthogonal planes (HOG-TOP) to represent the dynamic characteristics of video sequences [18]. Although the accuracy of this method is high, the computational complexity is very high.

This paper proposes a reconstructed LBP-HOG (LBP-RHOG) algorithm. The algorithm not only combines the texture features of LBP with the HOG edge structure information, but also reduces the computational complexity and feature redundancy. In addition, the performance of the algorithm is superior to LBP, HOG, and LBP-HOG algorithms in complex environments.

The rest of the paper is organized as follows. In Section 2 LBP and HOG are explained. In Section 3 we present the overview of proposed method. Section 4 shows the data set, experiments and results analysis. Section 5 gives a conclusion.

2. System Model

2.1. LBP(Local Binary Patterns)
LBP (Local Binary Pattern) was first proposed by T. Ojala, M. Pietikäinen, and D. Harwood in 1994 for texture feature extraction. It can be used to describe the local texture features of an image, has the advantages of rotation invariance and gray invariance.

The original LBP operator is defined as a window with 3*3, and the gray value of eight adjacent pixels is compared with the threshold value of the central pixel of the window. If the surrounding pixel value is greater than the center pixel value, the position of the pixel is marked as 1, otherwise 0. In this
way, 8 points in the 3*3 neighborhood can be compared to produce an 8-bit binary number (usually converted to a decimal number, i.e., LBP code, a total of 256 types), that is, the LBP value of the center pixel of the window is obtained, and this value is used to reflect texture information for this area. As shown in Figure 1(a), its value is \((11001111)_2\).

In order to adapt to the texture features of different scales and achieve the requirements of gray scale and rotation invariance, Ojala et al. improved the LBP operator. Extend the 3×3 neighborhood to any neighborhood and replace the square neighborhood with a circular neighborhood. The improved LBP operator allows for any number of pixels within a circular neighborhood of radius \(R\). As shown in Figure 1(b), its value is \((11001111)_2\).

For LBP operators with \(P\) samples in a circular area of radius \(R\), \(2^P\) modes will be generated. Obviously, as the number of sample points in the neighborhood set increases, the type of binary pattern increases dramatically. Maenpaa et al. proposed an LBP operator with rotational invariance. Constantly rotating the circular neighborhood gives a series of initially defined LBP values. Take the minimum value as the LBP value of the neighborhood. Ojala proposed using a "uniform pattern" to reduce the dimension of the LBP operator's pattern, which has a dimension of \(P \times (P-1)+3\). Subsequently, the rotation invariant equivalent mode is proposed.

![Figure 1. (a) Rectangle LBP operator; (b) Circular LBP operator](image)

2.2. HOG (Histogram of Oriented Gradient)

Histogram of Oriented Gradient (HOG) is a feature descriptor used for object detection in computer vision and image processing. It constructs features by calculating and counting the gradient direction histogram of the local area of the image. The implementation process of the HOG feature extraction algorithm is as follows:

**Step 1:** Image grayscale: For color images, convert RGB components into grayscale images

**Step 2:** Gamma correction. In the case where the image illuminance is not uniform, the image brightness can be adjusted by Gamma correction. In practice, Gamma can be standardized by square root or logarithm method. Square root is used in this paper.

\[
i(x, y) = I(x, y)^{\gamma_{\text{gamma}}} 
\]

Where \(I(x, y)\) represents the pixel at the \((x, y)\) point of the image.

**Step 3:** Calculate the gradient (including size and direction) of each pixel of the image.

\[
G_x(x, y) = I(x+1, y) - I(x-1, y) 
\]

\[
G_y(x, y) = I(x, y+1) - I(x, y-1) 
\]

\[
G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} 
\]

\[
\alpha = \tan^{-1}\left(\frac{G_y(x, y)}{G_x(x, y)}\right) 
\]
Where \( G_x(x, y) \) , \( G_y(x, y) \) represent the horizontal and vertical gradient values, respectively; \( G(x, y) \) is the gradient in point \( (x, y) \), and \( \alpha \) is the gradient direction.

**Step 4:** Divide the image into small cells; count the gradient histogram of each cell. Each pixel in a cell votes for an orientation-based histogram channel. Each ticket is weighted, and this weight is calculated based on the gradient of the pixel. Combine several cells into a block. Intra-block normalized gradient histogram.

**Step 5:** The HOG feature descriptor of the image is obtained by concatenating the HOG feature descriptors of all the blocks in the image image.

### 3. Overview of Proposed Method

This paper proposes a reconstructed LBP-HOG (LBP-RHOG) algorithm. The LBP-RHOG algorithm does not make any changes to the LBP operator, while a method of reconstructed HOG based on statistical features is proposed. This method not only retains the structural features of the image, but also reduces the dimension of HOG. The emphasis of reconstructed HOG is to calculate the mean, variance, skewness, kurtosis and entropy of each Block eigenvector. The process is as follows:

**Step 1:** Standardized gamma space and color space.

**Step 2:** Calculate the image gradient.

**Step 3:** Divide the image into several cells. The gradient direction histogram is counted in each cell. Divide all gradient directions into 9 bins (i.e. 9-dimensional feature vectors) as the horizontal axis of the histogram, and the gradient value corresponding to the angular range as the vertical axis of the histogram.

**Step 4:** The Block consists of 2 x 2 cells. Each block is a 36-dimensional vector. As shown in formula (6), it builds a matrix \( F \) with the block vector as the basic unit, where \( A_{11}, A_{12}, \ldots, \) etc, represents the 36-dimensional vector of the corresponding Block on the image.

\[
F = \begin{bmatrix}
A_{11} & A_{12} & \cdots & A_{1m} \\
A_{21} & A_{22} & \cdots & A_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
A_{n1} & A_{n2} & \cdots & A_{nm}
\end{bmatrix}
\]  

(6)

**Step 5:** Reconstruct matrix. The matrix \( F \) can be reconstructed in the direction of the row or in the direction of the column as shown in (7). The following steps will take \( F_1 \) as an example.
Step 6: Calculate five statistical characteristics of the reconstructed matrix: mean $\mu$, variance $\sigma$, skewness $\gamma$, kurtosis $\kappa$, entropy $H$. The calculation methods are as shown in formula (8) - (11).

$$
\mu_j = \frac{\sum_{i=1}^{n} a_{ij}}{n}
$$

(8)

$$
\sigma_j = \sqrt{\frac{\sum_{i=1}^{n} (a_{ij} - \mu_j)^2}{n-1}}
$$

(9)

$$
\gamma_j = \frac{\sum_{i=1}^{n} (a_{ij} - \mu_j)^3}{n(n-1)} \left[ \frac{1}{n} \sum_{i=1}^{n} (a_{ij} - \mu_j)^2 \right]^\frac{3}{2}
$$

(10)

$$
\kappa_j = \frac{\sum_{i=1}^{n} (a_{ij} - \mu_j)^4}{\left( \sum_{i=1}^{n} (a_{ij} - \mu_j)^2 \right)^2}
$$

(11)

Formula (8) to (11) can be used to calculate the mean, variance, skewness, and kurtosis of the jth column. Where $a_{ij}$ denotes the i-th row and the j-th column in the matrix $F_1$. The image entropy of the jth column calculates in (12), where $p_{ij}$ represents the probability that the value $a_{ij}$ appears in the jth column.

$$
H_j = -\sum_{i=1}^{n} p_{ij} \log_2(p_{ij})
$$

(12)

Step 7: Combine LBP with reconstructed HOG to form feature vectors with texture information, structure information and edge information. The reconstructed features are shown in (13):

$$
Feature = \{Feature_{LBP}, Feature_{HOG}\} = \{Feature_{LBP}, [mean(F_1)], [var(F_1)], [skewness(F_1)], [kurtosis(F_1)], [entropy(F_1)]\}
$$

(13)

4. Performance Evaluation and Simulation

4.1. Data Set

In this experiment, a self-built fatigue expression library was used. Its characteristics are as follows:

- Skin color: yellow, white, black.
- Age group: old, young and middle-aged, children.
4.2. Experimental Result

The experiment of fatigue recognition mainly compares the four modes of LBP (original mode, rotation invariant mode, uniform mode, rotation invariant equivalent mode), four modes of pyramid LBP, HOG, direct fusion LBP-HOG, performance of the LBP-RHOG method. Among them, PLBP is the method mentioned in [19]. The simulation experiment uses svm to train and classify the extracted feature vectors.

In this experiment, the cell size of LBP operator is 3*3, the sampling point P is 8, and the radius R is 1. The cell size of HOG is 8*8. In all HOG-related methods, the size of the cell is 8*8, and the block is composed of 2*2 cells.

At this point, for an image of 130*200, the feature dimension obtained by HOG is about 12960 dimension; while The dimension of RHOG operator is: the number of cells in a Block × the number of gradient intervals in the cell ×5 (statistical number) =180 dimensions, which is much smaller than the 12,960 dimensions obtained by HOG. The experimental results are shown in Figure 3:

![Figure 3. Experimental results](image)

In table I and Fig.3, HOG does not have four modes, which used only to compare results; "ri" represents the rotation invariant mode; "u2" refers to the uniform mode; "riu2" refers to the rotation invariant uniform mode.

It can be seen from the results that the recognition rate of the proposed algorithm is significantly higher than that of other algorithms in the four LBP modes. The reason is that this method not only combines texture information, structure information and edge information, but also solves the problem of uneven illumination through Gamma correction. Moreover, we use several statistical features in reconstructed HOG, where the mean is an indicator that reflects the general trend of the population or the concentration trend of the distribution, the variance is used to characterize the variation or dispersion of the overall distribution, the degree of indicators, skewness and kurtosis are indicators that reflect the overall distribution pattern, and the entropy describes the average amount of information of the image source. These statistics highly refine the structural information of the image, eliminating redundancy.

Then the time of feature extraction by each algorithm is compared, as shown in Table 1. The computer configuration that takes time to compute feature extraction is: Intel(R)Core(TM)i3-4170 CPU@3.70GHz; RAM: 6 GB.
Table 1. Average time of feature extraction of a frame by different algorithms

|       | ori (s) | ri (s) | u2 (s) | riu2 (s) |
|-------|---------|--------|--------|----------|
| LBP   | 0.004   | 0.087  | 0.088  | 0.077    |
| PLBP  | 0.148   | 0.807  | 0.303  | 0.258    |
| HOG   | 0.018   | 0.017  | 0.018  | 0.018    |
| LBP-HOG | 0.020   | 0.123  | 0.108  | 0.108    |
| LBP-RHOG | 0.019   | 0.099  | 0.089  | 0.089    |

Since the two operators are merged, it can be seen from Table 1 that the average feature extraction time of LBP-HOG and LBP-RHOG algorithms is higher than that of HOG and LBP. However, LBP-RHOG adopts the way of reconstruction HOG, which greatly reduces the feature dimension, so the average time consumption is lower than the LBP-HOG. In summary, LBP-RHOG can not only effectively identify facial fatigue facial expression, but also combine the advantages of the two algorithms, and use the reconstruction method to reduce the feature dimension and reduce the feature extraction time.

5. Conclusion

In this paper, a reconstructed LBP-HOG (LBP-RHOG) algorithm based on statistical features is proposed. This proposed algorithm use LBP operator and reconstructed HOG algorithm to obtain texture information, structure information and the edge information. It can be seen from the above section that the proposed algorithm not only reduces the feature dimension and computation cost to a certain extent, but also improves the recognition rate. In the future work, we will improve the algorithm for wearing glasses and having occlusion.

6. Acknowledgment

This work was supported by the national natural science foundation of China (NSFC) under Grant (61871193) and by the key project of guangdong natural science foundation under Grant (2018B030311049).

7. References

[1] Sahayadhas, Arun, K Sundaraj, M Murugappan, “Detecting Driver Drowsiness Based on Sensors: A Review”, Sensors 12.12(2012):16937-16953.
[2] Sigari M H, Pourshahabi M R, Soryani M, et al, “A Review on Driver Face Monitoring Systems for Fatigue and Distraction Detection”, International Journal of Advanced Science & Technology, 2014, 64(5): 73-100.
[3] Kar S, Bhagat M, Routray A, “EEG signal analysis for the assessment and quantification of driver’s fatigue”, Transportation Research Part F Traffic Psychology & Behaviour, 2010, 13(5): 297-306.
[4] Akin M, Kurt M B, Sezgin N, et al, “Estimating vigilance level by using EEG and EMG signals”, Neural Computing & Applications, 2008, 17(3): 227-236.
[5] Liang Y, Reyes M L, Lee J D, “Real-Time Detection of Driver Cognitive Distraction Using Support Vector Machines”, IEEE Transactions on Intelligent Transportation Systems, 2007, 8(2): 340-350.
[6] Wu B F, Chen Y H, Li Y F, et a, “Reasoning-Based Framework for Driving Safety Monitoring Using Driving Event Recognition”, IEEE Transactions on Intelligent Transportation Systems, 2013, 14(3): 1231-1241.
[7] Ji Q, Yang X, “Real-Time Eye, Gaze, and Face Pose Tracking for Monitoring Driver Vigilance”, Real-Time Imaging, 2002, 8(5): 357-377.

[8] Ohn-Bar E, Trivedi M M, “Hand Gesture Recognition in Real Time for Automotive Interfaces: A Multimodal Vision-Based Approach and Evaluations”, IEEE Transactions on Intelligent Transportation Systems, 2014, 15(6): 2368-2377.

[9] Lanitis A, Taylor C J, Cootes T F, “Automatic interpretation and coding of face images using flexible models”, IEEE Transactions on Pattern Analysis & Machine Intelligence, 2002, 19(7): 743-756.

[10] Filareti Tsalakanidou, Sotiris Malassiotis, “Real-time 2D+3D facial action and expression recognition”, Pattern Recognition, 2010, 43(5): 1763-1775.

[11] Huang D, Shan C, Ardabilian M, et al, “Local Binary Patterns and Its Application to Facial Image Analysis: A Survey”, IEEE Transactions on Systems Man & Cybernetics Part C Applications & Reviews, 2011, 41(6): 765-781.

[12] Dalal, N, Triggs B, “Histograms of oriented gradients for human detection”, IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, San Diego, CA, 2005, pp. 886–893.

[13] Wang L M, Qiao Y, Tang X, “Motionlets: Mid-level 3D Parts for Human Motion Recognition”, Computer Vision and Pattern Recognition. IEEE, 2013: 2674-2681.

[14] Lyons M, Amakatsu S, Kamachi M, et a, “Coding Facial Expressions with Gabor Wavelets”, IEEE International Conference on Automatic Face and Gesture Recognition, 1998. Proceedings. IEEE, 2002: 200-205.

[15] Shan C, Gong S, Mcowan P W, “Facial expression recognition based on Local Binary Patterns: A comprehensive study”, Image and Vision Computing, 2009, 27(6): 803-816.

[16] Zhao G, Pietikäinen M, “Dynamic texture recognition using local binary patterns with an application to facial expressions”, IEEE Transactions on Pattern Analysis & Machine Intelligence, 2007, 29(6): 915-928.

[17] Zhao C, Zhang X, Zhang B, et al, “Driver's fatigue expressions recognition by combined features from pyramid histogram of oriented gradient and contourlet transform with random subspace ensembles”, Iet Intelligent Transport Systems, 2013, 7(1): 36-45.

[18] Chen J, Chen Z, Chi Z, et al, “Dynamic texture and geometry features for facial expression recognition in video”, IEEE International Conference on Image Processing. IEEE, 2015: 4967-4971.

[19] Bin Jie Hu,Ya Ru Qi, “Fatigue expression recognition method based on image pyramid local binary mode”, CN Patent 1083334876A, 2018.5.