Pedestrian Detection Using Quaternion Gradient Based Weber Local Descriptor

GUOYUN LIAN [ID]
School of Artificial Intelligence, Shenzhen Polytechnic, Shenzhen 518000, China

e-mail: lianguoyun@szpt.edu.cn

ABSTRACT In the past decades, pedestrian detection has attracted more attention in many practical applications. In this paper, a novel pedestrian detection method using Quaternion Gradient based Weber Local Descriptor (QGWLD) is proposed. Unlike many other local pedestrian detectors that only extracted features from the gray-scale image, which ignored the color information, a new powerful and robust local pedestrian detector is developed, which integrates the advantages of both the color and texture information and can be used to address the pedestrian detection problem well. As we know, the quaternion can entirely characterize a color object well and the WLD feature consisting of the gradient orientation and differential excitation has acquired a good performance for pedestrian detection. Therefore, combining the WLD feature with the quaternion representation, the QGWLD pedestrian detector is presented. Firstly, the quaternion gradient representation is performed on the color image in the sliding window, and then the Weber Local Descriptor (WLD) is extracted over the quaternion gradient feature map. After that, the QGWLD histogram is constructed to characterize the sliding window. Experimental results on INRIA and PennFudanPed pedestrian databases validate the effectiveness of the proposed QGWLD pedestrian detector. Comparing with the similar pedestrian detectors, the proposed QGWLD pedestrian detector performs better.

INDEX TERMS Pedestrian detection, quaternion gradient, Weber local descriptor, video surveillance.

I. INTRODUCTION

Pedestrian detection has attracted considerable attention in many practical applications [1]–[4], such as advanced robotics, automotive systems, visual surveillance, and human activity understanding etc. However, detecting human in images or videos has been proven to be a challenging task because of the wide variability in appearance due to clothing, articulation, body shape, partial occlusion and illumination conditions, particularly in moving cameras or cluttered backgrounds. The reliable pedestrian detection methods are still far from being used in practical applications. Therefore, a more discriminative and robust pedestrian detector is expected to improve its performance.

With the last years, texture information plays an important role in computer vision [46], [47]. Many pedestrian detection algorithms based on texture feature extraction have been

The associate editor coordinating the review of this manuscript and approving it for publication was Mohammad Shorif Uddin [ID].

proposed and great progress has been achieved [5]–[8]. Existing feature extraction approaches can be roughly categorized into data-driven features and hand-crafted features.

Recently, Deep Learning models (CNNs) have achieved great success in object detection [9]–[14]. Naturally, they have also been applied to pedestrian detection by using the learned deep feature to improve the detection performance [15]–[22], [37], [39]. While these pedestrian detection methods using Deep Learning models have improved the detecting accuracy, they demand more complex architectures and higher computational costs, and also need special hardware to run in reasonable time.

Based on hand-crafted features, many pedestrian detection methods have been proposed, such as Haar features proposed by Papageorgiou and Poggio [23], scale invariant feature transform (SIFT) descriptor [24], edge templates [25], covariance descriptor [26], Weber local descriptor (WLD) [27], adaptive contour features [28] and histograms of oriented gradients (HOG) [29]. In these feature extraction methods,

the performance of the HOG feature for pedestrian detection has shown great advantage. Based on the HOG feature framework, many HOG feature variants have been proposed for pedestrian detection. The feature combination of the HOG feature and optical flow feature by Dalal *et al.* was proposed to characterize the pedestrian. Another HOG variant was the combination of the HOG, Haar wavelets and optical flow feature by Wojek *et al.* was proposed to represent the pedestrian. Also Walk *et al.* proposed a feature representation combining the HOG, HOF and CSS as the pedestrian features [30]. Another approach with the combination (HOG-LBP) of the HOG and local binary pattern (LBP) feature was proposed by Wang *et al.* for pedestrian detection [31], [32]. These pedestrian detection methods based on hand-crafted features have achieved competitive results with low computational complexity.

In this paper, a hand-crafted feature based pedestrian detection algorithm with low computational complexity is explored. A novel and effective differential statistic feature with color information is proposed to discover the discriminative local feature patterns for pedestrian detection. The proposed feature exhibits the advantages in computation efficiency and discriminating power. The main contributions of this work can be summarized into three aspects:

- A novel and robust local pedestrian detector is developed, which integrates the advantages of both the color and texture information and can be used to address the pedestrian detection problem well.
- A quaternion gradient representation is explored on the color image, which can effectively characterize a color pedestrian well.
- Comparing with the similar pedestrian detectors, the proposed method achieves a good performance, which is validated by extensive experiments on two popular color pedestrian databases, i.e., INRIA and PennFudanPed pedestrian database.

II. RELATED WORKS

There is a long history of research on pedestrian detection with a vast literature. Before the emergence of Deep Learning models, most state-of-the-art detection approaches based on the hand-crafted features have been widely used to achieve a good performance under a sliding window strategy [23]–[32]. In these approaches, most works have focus on designing powerful feature representations. Among these pedestrian detection methods, the HOG and its many variants have gained a lot of success and popularity [29], [33]–[36]. However, these methods based on the HOG and its variants only consider the gradient orientations. Different from the HOG feature, the Weber local descriptor (WLD) method consisting of the gradient orientation and differential excitation has acquired better performance for pedestrian detection [27].

It is known that the original HOG method has shown success for pedestrian detection in gray image. And in most pedestrian detection approaches based on hand-crafted features, the features have been extracted from the gray-scale image, which ignore the color information. However, the color information is very important for pedestrian detection in color image. To consider the color information, some pedestrian detection methods based on color features have been proposed, such as color histograms [38], color self-similarity (CSS) [30], and combining the HOG feature with color information feature (CHOG) [32]. But in these methods, the extracting features are only from each separate color channel. The separate gradients for each channel are calculated and the one with the largest norm is taken as the pixel's gradient vector. However, if the features are extracted from the each color channel respectively or just to gray image, the relations between each color channel component are ignored [29]. As we know, it is very important for this relation information to represent a color image.

A preliminary version of this article has appeared in [48], in which the quaternion gradient orientation of a color pixel was defined and used to construct the QHOG descriptor. Similar to the HOG method [29], the gradient orientation was only considered in the QHOG approach.

Inspired by the quaternionic representation of a color pixel in a holistic way and the success of the WLD feature for pedestrian detection, in this paper, a very powerful and robust local pedestrian detector is presented, which integrates the advantages of both the WLD feature and quaternion gradient representation of color information and can be used to address the pedestrian detection problem well. More specifically, in the proposed method, the quaternion gradient is firstly utilized to represent the color image. Then the WLD feature is computed based on the quaternion gradient map. Here, this pedestrian detector is called as ''Quaternion Gradient based Weber Local Descriptor'' (QGWLD). Experimental results show that the histogram of QGWLD can be used as a proper color pedestrian detector.

III. PROPOSED APPROACH

In pedestrian detection, the robust feature representation strongly influences the pedestrian detecting performance. In this section, the proposed method: Quaternion Gradients based Weber Local Descriptor is presented. The objective of this work is to develop an effective and efficient detector for color pedestrian detection.

A. QUATERNION GRADIENTS OF COLOR IMAGE

For color pedestrian detection, how to characterize a pedestrian object well in a color image is a key problem. As we know, the holistic analysis of all color components can better represent the color information, as compared to separate analysis of color components. Fortunately, the quaternion representation of a color pixel has been obtained good performance in color image processing [42].

The quaternion is a four-dimensional complex number with one real part and three imaginary parts, which was introduced by Hamilton [40]. It can be represented as follows:

$$q = a + ib + jc + kd \qquad (1)$$

here a, b, c and d are real numbers, i, j and k are complex operators, and they satisfy:

$$i^2 = j^2 = k^2 = ijk = -1$$
$$ij = -ji = k$$
$$jk = -kj = i$$
$$ki = -ik = j$$

In this complex form, a denotes the real part, and $ib + jc + kd$ denotes the imaginary part. Here, $S(q)$ and $V(q)$ are used to represent the real part and imaginary part, respectively, then the quaternion can be rewritten as:

$$q = S(q) + V(q) \qquad (2)$$

Like the two-dimensional complex number, any quaternion can also be represented in polar form:

$$q = |q| e^{\mu\theta} = |q| (\cos\theta + \mu\sin\theta) \qquad (3)$$

here μ is a unit pure quaternion, and $0 \le \theta \le \pi$, which can be obtained by:

$$\mu = \frac{V(q)}{|V(q)|} \qquad (4)$$

$$\theta = \arctan\left(\frac{V(q)}{S(q)}\right) \qquad (5)$$

The quaternion can be effectively used in processing color image in spatial and complex domains [42]. The three-component imaginary part of the quaternion allows for processing the color as one unit. In the RGB model, a color pixel can be represented in the quaternion space as:

$$q = a + ir + jg + kb \qquad (6)$$

here r, g, and b are denoted as red, green, and blue components of the color pixel respectively, i, j and k are pure quaternion units, and the real part of the quaternion can be set as the grayscale pixel: $a = (r + g + b)/3$ or considered as $a = 0$.

Many gradient operators based on the grayscale image have been proposed, which can process the gray value. When processing the color image, these gradient operators can be used by operating separately each color component. And the component-wise edge detection in color image can be obtained. It is clear that not all edges can be found by separately processing the color components by a single gradient.

For effectively processing the color image, the quaternion gradient operator was introduced by Grigoryan etc. [42]. The windowed convolution of the color image with a quaternion mask was defined as:

$$H_{n,m} = (H_e)_{n,m} + (i(H_i)_{n,m} + j(H_j)_{n,m} + k(H_k)_{n,m}) \qquad (7)$$

with the right-side multiplication as

$$QG_{n,m} = q_{n,m} * H_{n,m} = \sum_{n_1=-L_1}^{L_1} \sum_{m_1=-L_2}^{L_2} q_{n-n_1, m-m_1} H_{n_1, m_1} \qquad (8)$$

where $(2L_1 + 1)(2L_2 + 1)$ be the size of the mask.

When the quaternion components are equal, that is, the convolution mask can be represented as:

$$H_{n,m} = (1 + (i + j + k))(H_e)_{n,m} \qquad (9)$$

and the real mask H_e is defined as the gradient operator G_x and G_y in the X and Y directions, respectively. Thus, two quaternion gradient operators are considered:

$$H_x = (1 + (i + j + k))G_x \text{ and } H_y = (1 + (i + j + k))G_y \qquad (10)$$

Based on the two operators, the quaternion gradient of the color image can be represented by the gradients of the X and Y directions, that is:

$$QG_x = q * H_x \text{ and } QG_y = q * H_y \qquad (11)$$

B. QUATERNION GRADIENTS BASED WEBER LOCAL DESCRIPTOR

Weber local descriptor (WLD) proposed by Chen *et al.* has obtained promising performance in texture recognition, human face detection and pedestrian detection [41], [27]. The original WLD feature is extracted from just gray image. If the pedestrian detection is in a color image, it obviously ignored the color information. In order to effectively characterize a color image, the quaternion gradient is presented in subsection III.A. In this subsection, the quaternion gradient based Weber local descriptor is proposed, which is used to detect pedestrian in a color image.

1) OVERVIEW OF WLD

WLD is based on the Weber's Law which is a psychological law [41]. The law reveals the universal influence of the background stimulus I on human's sensitivity to the intensity increment ΔI and claims that the so-called Weber's fraction is a constant which is known as Weber's Law. Therefore, WLD consists of two components, one is differential excitation (ξ) and the other is orientation (θ).

Differential excitation $\xi(x_c)$ of a current pixel x_c is computed by the pixel and its neighbor pixels. According to [41], it is a function of the ratio between the relative intensity differences of the current pixel against its neighbors and the intensity of the current pixel. The relative intensity difference of the current pixel against its neighbors is calculated using a filter:

$$v_s^{00} = \sum_{i=0}^{p-1} (\Delta x_i) = \sum_{i=0}^{p-1} (x_i - x_c) \qquad (12)$$

where x_i ($i = 0, 1, 2, \cdots, p-1$) represents the i-th neighbor of x_c and p is the number of the neighbor pixels (here, 3×3 square neighborhood is considered). Therefore, the differential excitation of the current pixel $\xi(x_c)$ is computed as:

$$\xi(x_c) = \arctan\left[\frac{v_s^{00}}{x_c}\right] = \arctan\left[\sum_{i=0}^{p-1}\left(\frac{x_i - x_c}{x_c}\right)\right] \qquad (13)$$

The orientation is the gradient orientation of the pixel, which is computed as:

$$\theta(x_c) = \arctan\left(\frac{v_s^{11}}{v_s^{10}}\right) \tag{14}$$

where v_s^{11} and v_s^{10} are respectively calculated as:

$$v_s^{11} = x_5 - x_1, \; and \; v_s^{10} = x_7 - x_3 \tag{15}$$

After calculating the gradient orientation $\theta(x_c)$ and differential excitation $\xi(x_c)$ of each pixel x_c, the 2D histogram is obtained by:

$$WLD_{2D}(r, t) = \sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\delta(\xi(x_{i,j}), r)\delta(\theta(x_{i,j}), t) \tag{16}$$

with

$$\delta(x, y) = \begin{cases} 1, & x = y \\ 0, & otherwise \end{cases}$$

where $M \times N$ denotes the image size, $x_{i,j}$ is the pixel on location (i, j) in the image, $r = 0, 1, \cdots, R - 1$, $t = 0, 1, \cdots, T - 1$, T is the dominant gradient orientation and R denotes the number of the differential excitation bins. The 2D WLD histogram is constructed as illustrated in Fig. 1.
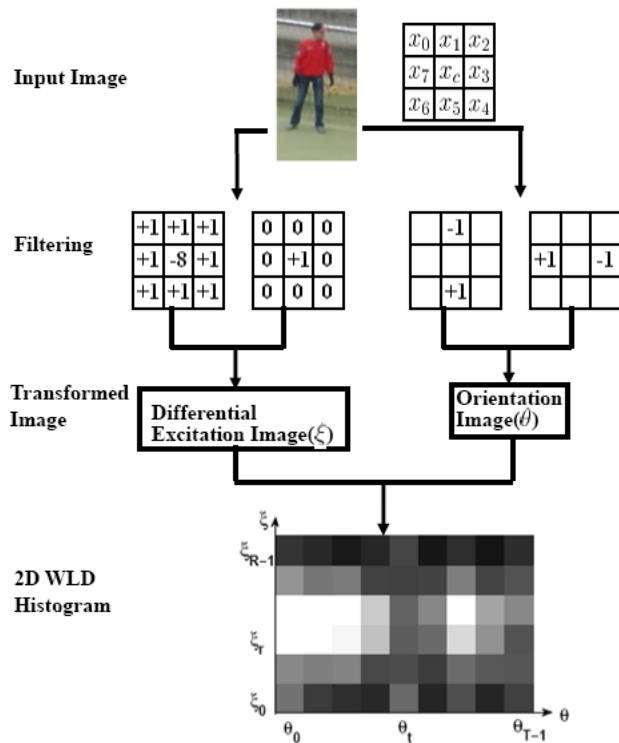

FIGURE 1. Illustration of the construction of the 2D WLD histogram.

2) PROPOSED QGWLD DETECTOR

It is known that the original HOG method has shown great success for pedestrian detection in gray image. If the pedestrian detection is in a color image, the separate gradients for each channel are calculated and the one with the

largest norm is taken as the pixel's gradient vector. However, if the HOG feature is extracted from the each color channel respectively or just gray image, the relations between each color channel component are ignored [29]. In this section, the quaternion number is utilized to represent each color pixel, then a quaternionic representation map can be obtained, and then based on the map, the quaternion gradients (QG_x and QG_y) of the color image described in section III.A can be computed. After this, the magnitude map of their real part is calculated as:

$$QG = |re(QG_x)| + |re(QG_y)| \tag{17}$$

Then, the two components of WLD, that is, differential excitation (ξ) and orientation (θ) for each pixel are extracted from the magnitude map QG. Here, the WLD feature computed from the quaternion gradient map QG is called as QGWLD feature. As the above mentioned, the 2D QGWLD histogram can be calculated as follows:

$$QGWLD_{2D}(r, t) = \sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\delta(\xi(QG_{i,j}), r)\delta(\theta(QG_{i,j}), t) \tag{18}$$

In order to reduce the complexity of the pedestrian detector, the 2D $QGWLD_{2D}$ histogram was simply encoded into a 1D $QGWLD_{1D}$ histogram for pedestrian detection, which is similar as mentioned in the WLD method proposed by Chen *et al.* [41]. The encoding process is illustrated in Fig. 2, in which given the 2D $QGWLD_{2D}$ histogram, each row is extracted as a 1D sub-histogram $SH_r(r = 0, 1, \ldots, R - 1)$. The sub-histogram SH_r represents the differential excitation segment with the similar differential excitation values. Finally, the all R sub-histograms are concatenated and then constructed the 1D $QGWLD_{1D}$ pedestrian detector. The extracted feature integrates the advantages of both the WLD feature and quaternionic representation, which can be used to address the color pedestrian detection problem.
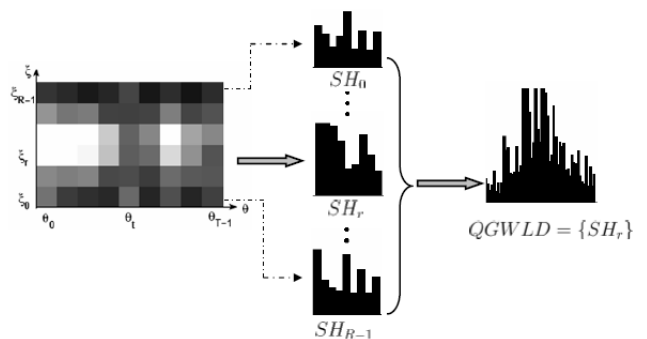

FIGURE 2. Illustration of the QGWLD pedestrian detector for a given color image.

In the WLD method [41], it was suggested that the differential excitation bins R was set as $R = 6$, which approximately simulate the variances of low, middle or high frequency. As we all known, the variant regions with high frequency

TABLE 1. Weights for a QGWLD histogram.

	SH_0	SH_1	SH_2	SH_3	SH_4	SH_5
Weights	0.2688	0.0852	0.0955	0.1000	0.1018	0.3487

are usually paid more attention than the flat regions with low frequency in a given image. For a pixel $x_{i,j}$, the differential excitation value $\xi(x_{i,j})$ can be obtained by using Equation (13). And the authors in [16] thought that if $\xi(x_{i,j}) \in \xi_0$ or ξ_5, the variance near $x_{i,j}$ is of high frequency; if $\xi(x_{i,j}) \in \xi_1$ or ξ_4, it is of middle frequency; if $\xi(x_{i,j}) \in \xi_2$ or ξ_3, it is of low frequency. According to the different weights for the different differential excitation computed from the texture classification, they have gained a better performance for texture classification. And also, they have validated the effectiveness of the weights for face detection. The weights provided by Chen *et al.* [41] are shown in Table 1. In our experiments, the same weights in the Table 1 were used for pedestrian detection. Fig. 3 shows the comparison results of pedestrian detection between the $QGWLD$ with weights ($QGWLD - W$) and without weights ($QGWLD$) on two datasets. From the experimental results, it can be concluded that the $QGWLD - W$ performs better. In the following experiments, the $QGWLD - W$ is used as the pedestrian detector and for simple representation, the $QGWLD$ term is adopted to substitute for the $QGWLD - W$. The main parts of processing the color image are shown in the block-diagram of Fig. 4. The color image is considered in the RGB model.

In the popular HOG method for pedestrian detection [29], the pedestrian detecting window was divided into small cells and the local HOG histogram was extracted from each cell. And then the all HOG histograms from the cells in the detecting window are concatenated. Finally the concatenating HOG histogram was used to characterize the current window. Inspired by the HOG method [29], the same way is also adopted to construct the $QGWLD$ histogram to characterize the detecting window for pedestrian detection. Similar to [29], the contrast-normalization is utilized to alleviate the effectiveness of the illumination, shadowing etc. In the contrast-normalization process, a block (2×2 cells) is used to compute the $QGWLD$ histogram and the histogram results are used to normalize all of the cells in the block. In the detecting window, the blocks are extracted in a scanning fashion, which typically overlap by one or more cells. Finally, the proposed pedestrian detector is formed by concatenating all blocks to characterize the detecting window.

IV. SIMILARITY MEASUREMENT

Many similarity measurement methods based on histogram matching have been presented. In the previous pedestrian detection methods, the Non-linear kernels classifiers, such as AdaBoost [43] or MPLBoost classifier [36] were proposed which have shown some good performance for pedestrian
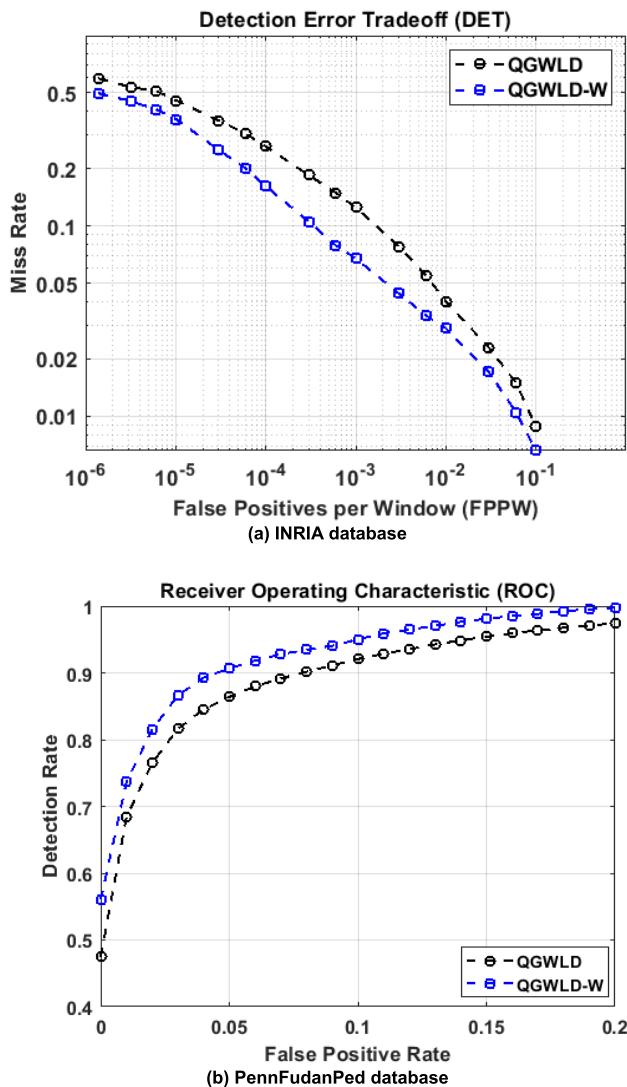

(a) INRIA database

(b) PennFudanPed database

FIGURE 3. The comparison results between the QGWLD-W detector (with weight) and QGWLD detector (without weight) on two database.
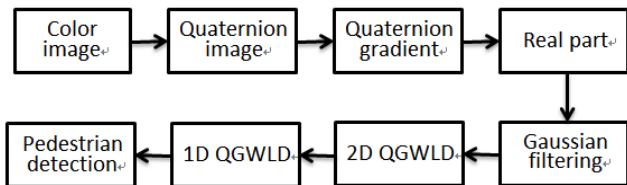

FIGURE 4. The block-diagram of color image pedestrian detection.

detection. However, these classifiers usually required more time for classifying a pedestrian example, which was intractable in practice. Analyzing the similarity measurement methods for pedestrian detection, the Support Vector Machine (SVM) approach has been proven to be a popular choice in real-time condition because of the logarithmic computing time or the approximate constant time [44].

In this paper, the purpose of this work is to validate the effectiveness of the proposed detector comparing with the similar detectors. Therefore, the similarity measurement is chosen as the same one of the similar detectors [27], [29], [31]. Similar to [27], [29], [31], the linear SVM classifier is adopted in all of the comparing experiments.

V. EXPERIMENTAL RESULTS

In this section, in order to validate the effectiveness of the proposed *QGWLD* detector for pedestrian detection, two groups of experiments are conducted. The experiments are performed on two popular color pedestrian Benchmark database, that is, INRIA database [29] and PennFudanPed database [45]. The comparative experiments are conducted with the similar pedestrian detectors based on hand-crafted extracting feature, that is, the MWLD [27], HOG [29] and HOG-LBP pedestrian detector [31].

A. DATABASE PREPARATION

The INRIA pedestrian database introduced in [29] is a color pedestrian database and has a wide variety of variable appearance, illumination changes, articulated poses and complex backgrounds, which show that it is very difficult for pedestrian detection. In INRIA pedestrian database, the standard sample size for training and testing is 128×64 pixels. Similar to [45], the training set consists of 1237 positive samples and 3891 negative samples, which are obtained from the training set of INRIA database. The first test set used for experiments contains 589 positive samples and 453 negative images which are all from the test set of INRIA database. The second test set has 423 positive images, which are constructed from the PennFudanPed dataset, and 453 negative samples are the same as the first test set, which from the test set of INRIA dataset. To maintain consistency with the standard sample in INRIA database, the 423 positive images from the PennFudanPed dataset are also resized as 128×64 pixels. Table 2 summarizes the specifications of both training and test sets. Some sample images are shown in Fig. 5. In the pedestrian detection process, the test images are scanned using a sliding detection window of size 128×64 pixels. The detection window is shifted 8 pixels across the image.

**TABLE 2.** Specifications of training and test sets.

	Training set	Test set #1	Test set #2
Positive samples	1237	589	423
Negative samples	3891	453	453

B. EXPERIMENTAL RESULTS ON INRIA DATABASE

In this experiment, the Detection Error Tradeoff (DET) curve was used to report the experimental result, that is, miss rate ($\frac{FalseNeg}{TruePos+FalseNeg}$) versus false positives per window (FPPW)

(a) Training set　　　　　(b) Test set

FIGURE 5. Sample images for experiments.
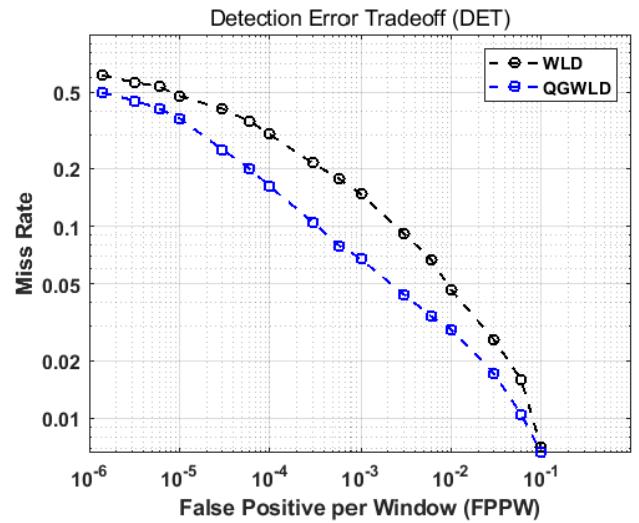
FIGURE 6. The experimental comparison between the proposed QGWLD pedestrian detector with the WLD detector on INRIA database.

is utilized to characterize the results on per-window classification accuracy. As mentioned in the previous pedestrian detection method [27], in our experiment, the image scanning window is also divided into small cells, for each cell, the *QGWLD* histogram is calculated. Then for all cells in the window, the *QGWLD* histograms are obtained and then concatenated into a histogram to describe the current window. Similar to [27], the cell size with 16×16 pixels is also adopted. The proposed detector is integrated the advantages of both the quaternion representation of a color image and the WLD feature. Comparing to the WLD feature which is extracted from the gray-scale image, the *QGWLD* feature which is extracted from the color image performs better. Fig. 6 shows their comparison result. Also, an extensive experiment comparing the proposed *QGWLD* pedestrian detector with the MWLD pedestrian detector [27], the HOG pedestrian detector [29] and the HOG-LBP pedestrian detector [31] is carried. Fig. 7 shows the experimental results with the other pedestrian detection methods which are similar to the proposed *QGWLD* method. From this figure, it can be seen that the proposed *QGWLD* method performs better than
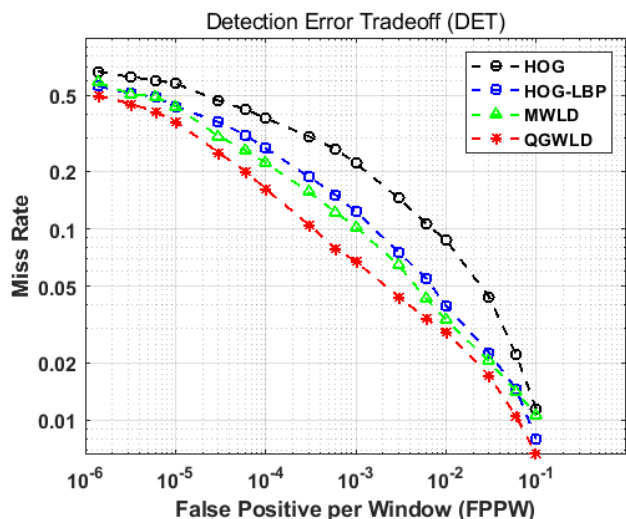
FIGURE 7. The experimental comparison between the proposed QGWLD pedestrian detector with the other similar pedestrian detectors on INRIA database.

FIGURE 9. The experimental comparison between the QGWLD detector with the other detectors on PennFudanPed database.

the other similar pedestrian detection methods. In quantitative analysis, comparing to the HOG pedestrian detector, the detection rate is improved by 21.83% at FPPW $=10^{-4}$. Also comparing to the similar detector, MWLD pedestrian detector [27], an improvement of 6% at FPPW $=10^{-4}$ is obtained using the proposed *QGWLD* pedestrian detector.
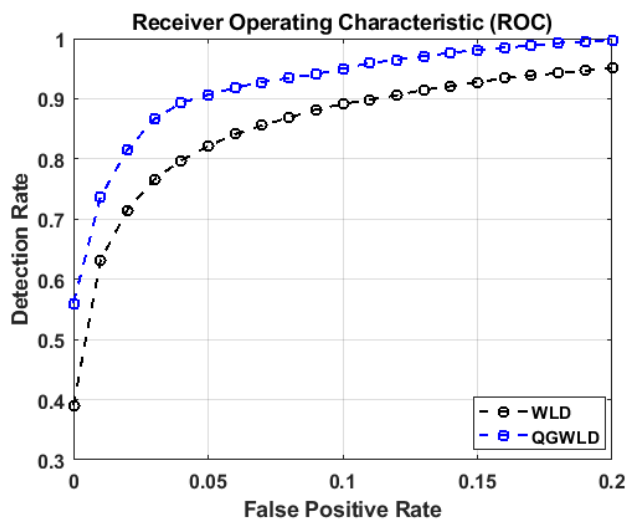

FIGURE 8. The experimental comparison between the QGWLD detector with the WLD detector on PennFudanPed database.

C. EXPERIMENTAL RESULTS ON PENNFUDANPED DATABASE

In this work, the training and testing datasets are the same as the experimental set in [45]. Fig. 8 shows that the QGWLD pedestrian detector also performs better than the WLD detector. Also, Fig. 9 shows the performance comparison between the proposed QGWLD pedestrian detector and
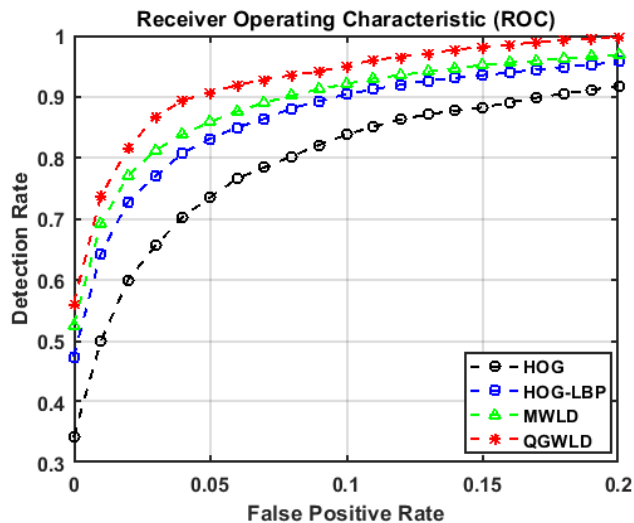
the MWLD detector [27], HOG detector [29] and HOG-LBP detector [31] on PennFudanPed database. From this figure, one can also conclude that the proposed QGWLD pedestrian detector still outperforms the MWLD, HOG and HOG-LBP methods. In quantitative analysis, the QGWLD pedestrian detector achieves a detection rate of 90.67% at FPR $=0.05$ on the testing set, while the HOG method, the HOG-LBP method and the MWLD method achieved 73.42%, 83.05% and 85.77% respectively. That is, the proposed QGWLD detector for pedestrian detection obtains 17.25%, 7.62% and 4.9% improvements in the detection rate over the HOG, HOG-LBP and MWLD method respectively.

D. RUNTIME COMPARISON

An extra experiment is conducted to compare the execution time of the proposed QGWLD pedestrian detection approach with other methods on the INRIA dataset. All the experiments are implemented with Intel Core i7 CPU at 3.5GHz and NVIDIA Titan GPU. Since some methods depend heavily on hardware configuration and program working environment, it is very hard to compare their computational complexity objectively. Here, the speed performance is compared among these pedestrian detection methods in terms of average computation time per one image in the test process. Table 3 shows the comparative results of their performance including the execution time, log-average miss rates and hardware specification. From Table 3, it can be seen that the detecting performance of the proposed approach is slightly inferior to other methods, but their gap is acceptable. However, the executive time of the proposed QGWLD pedestrian detection approach is rather competitive and it also does not need the support of the extra GPU hardware. Therefore, the proposed approach can be used in fast real-time pedestrian detecting applications.

TABLE 3. Comparison of log-average miss rate [%] and average computation time [ms].

Method	Log-average miss rate (%)	Computation time (ms)	GPU
MS-CNN [39]	6.21	497	O
SA-FastRCNN [15]	5.85	995	O
DSBF [7]	10.39	420	X
DLSF [8]	10.59	126	X
Proposed QGWLD	16.24	26	X

VI. CONCLUSION

In this paper, an effective and efficient feature descriptor for pedestrian detection in a color image is proposed, which integrates the advantages of both the color and texture information. Inspired by the literature that the quaternion representation of all color components to characterize a color image in a holistic way and the great success of the WLD feature for pedestrian detection, a quaternion gradient based Weber local descriptor (QGWLD) is constructed for pedestrian detection in a color image. Two groups of experiments are conducted on two popular color pedestrian database, that is, INRIA database and PennFudanPed database respectively. Experimental results validate the effectiveness of the proposed QGWLD pedestrian detector. Comparing with the similar pedestrian detectors, the proposed QGWLD detector performs better.

How to extract effective and efficient features for real-time pedestrian detection under real-world conditions is still a challenging topic. In the proposed QGWLD approach, the executive time is rather competitive and it also does not need the support of the extra GPU hardware, which is suitable for the real-time pedestrian detecting applications. However, comparing to other pedestrian detection methods based on Deep Learning models (CNNs), the detecting performance of the proposed QGWLD approach is slightly inferior. In future, more study should be explored for the effective feature extracting method for pedestrian detection under real-time and real-world conditions. The tradeoff between the real-time detection and accuracy should be explored, e.g. the combining methods between the effective handcrafted features and the CNN features. Also, in this paper, the RGB color space is only considered for comparing experiments. More study will be explored in other color spaces, such as HSV, HIS, Lab, etc.

REFERENCES

[1] Y. Yuan, X. Lu, and X. Chen, "Multi-spectral pedestrian detection," *Signal Process.*, vol. 110, pp. 94–100, May 2015.

[2] N. K. Ragesh and R. Rajesh, "Pedestrian detection in automotive safety: Understanding state-of-the-art," *IEEE Access*, vol. 7, pp. 47864–47890, 2019.

[3] S. Alfasly, B. Liu, Y. Hu, Y. Wang, and C.-T. Li, "Auto-zooming CNN-based framework for real-time pedestrian detection in outdoor surveillance videos," *IEEE Access*, vol. 7, pp. 105816–105826, 2019.

[4] P. Dollar, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: An evaluation of the state of the art," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 4, pp. 743–761, Apr. 2012.

[5] R. Benenson, M. Omran, J. Hosang, and B. Schiele, "Ten years of pedestrian detection, what have we learned?" in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 613–627.

[6] S. Zhang, R. Benenson, M. Omran, J. Hosang, and B. Schiele, "Towards reaching human performance in pedestrian detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 973–986, Apr. 2018.

[7] H.-K. Kim and D. Kim, "Robust pedestrian detection under deformation using simple boosted features," *Image Vis. Comput.*, vol. 61, pp. 1–11, May 2017.

[8] C. Zhu and Y. Peng, "Discriminative latent semantic feature learning for pedestrian detection," *Neurocomputing*, vol. 238, pp. 126–138, May 2017.

[9] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 580–587.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1904–1916, Sep. 2015.

[11] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1440–1448.

[12] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 91–99.

[13] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 779–788.

[14] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 21–37.

[15] J. Li, X. Liang, S. Shen, T. Xu, J. Feng, and S. Yan, "Scale-aware fast R-CNN for pedestrian detection," *IEEE Trans. Multimedia*, vol. 20, no. 4, pp. 985–996, Apr. 2018.

[16] Q. Hu, P. Wang, C. Shen, A. van den Hengel, and F. Porikli, "Pushing the limits of deep CNNs for pedestrian detection," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 6, pp. 1358–1368, Jun. 2018.

[17] C. Lin, J. Lu, G. Wang, and J. Zhou, "Graininess-aware deep feature learning for pedestrian detection," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 745–761.

[18] J. Mao, T. Xiao, Y. Jiang, and Z. Cao, "What can help pedestrian detection?" in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 3127–3136.

[19] G. Brazil, X. Yin, and X. Liu, "Illuminating pedestrians via simultaneous detection and segmentation," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 4950–4959.

[20] Y. Tian, P. Luo, X. Wang, and X. Tang, "Pedestrian detection aided by deep learning semantic tasks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1904–1912.

[21] X. Zhang, S. Cao, and C. Chen, "Scale-aware hierarchical detection network for pedestrian detection," *IEEE Access*, vol. 8, pp. 94429–94439, 2020.

[22] C. Fei, B. Liu, Z. Chen, and N. Yu, "Learning pixel-level and instance-level context-aware features for pedestrian detection in crowds," *IEEE Access*, vol. 7, pp. 94852–94944, 2019.

[23] C. Papageorgiou and T. Poggio, "A trainable system for object detection," *Int. J. Comput. Vis.*, vol. 38, no. 1, pp. 15–33, 2000.

[24] J. Vourvoulakis, J. Kalomiros, and J. Lygouras, "Fully pipelined FPGA-based architecture for real-time SIFT extraction," *Microprocessors Microsyst.*, vol. 40, pp. 53–73, Feb. 2016.

[25] D. M. Gavrila, "A Bayesian, exemplar-based approach to hierarchical shape matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 8, pp. 1408–1421, Aug. 2007.

[26] O. Tuzel, F. Porikli, and P. Meer, "Pedestrian detection via classification on Riemannian manifolds," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 10, pp. 1713–1727, Oct. 2008.

[27] G. Lian, J. Lai, and Y. Yuan, "Fast pedestrian detection using a modified WLD detector in salient region," in *Proc. Int. Conf. Syst. Sci. Eng.*, Jun. 2011, pp. 564–569.

[28] W. Gao, H. Ai, and S. Lao, "Adaptive contour features in oriented granular space for human detection and segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1786–1793.

[29] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2005, pp. 886–893.

[30] S. Walk, N. Majer, K. Schindler, and B. Schiele, "New features and insights for pedestrian detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 1030–1037.

[31] X. Wang, T. X. Han, and S. Yan, "An HOG-LBP human detector with partial occlusion handling," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep. 2009, pp. 1–8.

[32] W. R. Schwartz, A. Kembhavi, D. Harwood, and L. S. Davis, "Human detection using partial least squares analysis," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Oct. 2009, pp. 24–31.

[33] X. Wang, T. X. Han, and S. Yan, "An HOG-LBP human detector with partial occlusion handling," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep. 2009, pp. 32–39.

[34] N. Dalal, B. Triggs, and C. Schmid, "Human detection using oriented histograms of flow and appearance," in *Proc. ECCV*, 2006, pp. 428–441.

[35] P. Felzenszwalb, D. McAllester, and D. Ramanan, "A discriminatively trained, multiscale, deformable part model," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

[36] C. Wojek, S. Walk, and B. Schiele, "Multi-cue onboard pedestrian detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2009, pp. 794–801.

[37] F. B. Tesema, H. Wu, M. Chen, J. Lin, W. Zhu, and K. Huang, "Hybrid channel based pedestrian detection," *Neurocomputing*, vol. 389, pp. 1–8, May 2020.

[38] P. Dollar, Z. Tu, P. Perona, and S. Belongie, "Integral channel features," in *Proc. Brit. Mach. Vis. Conf.*, 2009, pp. 1–11.

[39] Z. Cai, Q. Fan, R. Feris, and N. Vasconcelos, "A unified multi-scale deep convolutional neural network for fast object detection," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 354–370.

[40] W. Hamilton, *Elements of Quaternions*. London, U.K.: Longmans Green, 1866.

[41] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikäinen, X. Chen, and W. Gao, "WLD: A robust local image descriptor," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1705–1720, Sep. 2010.

[42] A. M. Grigoryan and S. S. Agaian, "Retooling of color imaging in the quaternion algebra," *Appl. Math. Sci., Int. J.*, vol. 1, no. 3, pp. 23–39, 2014.

[43] P. Sabzmeydani and G. Mori, "Detecting pedestrians by learning shapelet features," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.

[44] S. Maji, A. C. Berg, and J. Malik, "Classification using intersection kernel support vector machines is efficient," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

[45] L. Wang, J. Shi, G. Song, and I. Shen, "Object detection combining recognition and segmentation," in *Proc. Asian Conf. Comput. Vis.*, 2007, pp. 189–199.

[46] L. Armi1 and S. Fekri-Ershad, "Texture image analysis and texture classification methods—A review," *Int. Online J. Image Process. Pattern Recognit.*, vol. 2, no. 1, pp. 1–29, 2019.

[47] S. Fekriershad and F. Tajeripour, "Color texture classification based on proposed impulse-noise resistant color local binary patterns and significant points selection algorithm," *Sensor Rev.*, vol. 37, no. 1, pp. 33–42, Jan. 2017.

[48] G. Lian, "Pedestrian detection using quaternion histograms of oriented gradients," in *Proc. IEEE Int. Conf. Power, Intell. Comput. Syst. (ICPICS)*, Shenyang, China, Jul. 2020, pp. 415–419, doi: 10.1109/ICPICS50287.2020.9202071.

GUOYUN LIAN received the M.Sc. degree in computer science from the Kunming University of Science and Technology, Kunming, China, in 2008, and the Ph.D. degree in information science and technology from Sun Yat-sen University, Guangzhou, China, in 2011.

He was a Postdoctoral Researcher with Shenzhen University, Shenzhen, China. He is currently an Associate Professor with the School of Artificial Intelligence, Shenzhen Polytechnic, Shenzhen. His current research interests include computer vision, pattern recognition, digital image processing and analysis, and visual surveillance.

● ● ●