

Article

Low-Illumination Road Image Enhancement by Fusing Retinex Theory and Histogram Equalization

Yi Han ¹, Xiangyong Chen ¹, Yi Zhong ^{1,*}, Yanqing Huang ², Zhuo Li ², Ping Han ¹, Qing Li ³ and Zhenhui Yuan ⁴

¹ School of Information Engineering, Wuhan University of Technology, Wuhan 430070, China

² SAIC GM Wuling Automobile Co., Ltd., Liuzhou 545007, China

³ Peng Cheng Laboratory, Shenzhen 518066, China

⁴ Department of Computer and Information Science, Northumbria University, Newcastle Upon Tyne NE1 8ST, UK

* Correspondence: zhongyi@whut.edu.cn; Tel.: +86-27-8785-8005

Abstract: Low-illumination image enhancement can provide more information than the original image in low-light scenarios, e.g., nighttime driving. Traditional deep-learning-based image enhancement algorithms struggle to balance the performance between the overall illumination enhancement and local edge details, due to limitations of time and computational cost. This paper proposes a histogram equalization–multiscale Retinex combination approach (HE-MSR-COM) that aims at solving the blur edge problem of HE and the uncertainty in selecting parameters for image illumination enhancement in MSR. The enhanced illumination information is extracted from the low-frequency component in the HE-enhanced image, and the enhanced edge information is obtained from the high-frequency component in the MSR-enhanced image. By designing adaptive fusion weights of HE and MSR, the proposed method effectively combines enhanced illumination and edge information. The experimental results show that HE-MSR-COM improves the image quality by 23.95% and 10.6% in two datasets, respectively, compared with HE, contrast-limited adaptive histogram equalization (CLAHE), MSR, and gamma correction (GC).



Citation: Han, Y.; Chen, X.; Zhong, Y.; Huang, Y.; Li, Z.; Han, P.; Li, Q.; Yuan, Z. Low-Illumination Road Image Enhancement by Fusing Retinex Theory and Histogram Equalization. *Electronics* **2023**, *12*, 990. <https://doi.org/10.3390/electronics12040990>

Academic Editor: Chiman Kwan

Received: 28 January 2023

Revised: 12 February 2023

Accepted: 15 February 2023

Published: 16 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: low illumination; image enhancement; Retinex theory; histogram equalization; image fusion

1. Introduction

With the development of automatic driving technology, computer vision methods are based on simulated human vision, and they are also used to carry out important sensing tasks in multiple automatic driving scenarios, such as object detection, semantic road segmentation, etc. Due to changes in ambient light, such as day and night, the visibility of the images varies significantly. If the computer vision algorithm needs to ensure stable performance under different lighting conditions, it should cover all lighting scenes as much as possible during training. This will undoubtedly require more time and human labor resources in collecting the dataset, as well as training based on this dataset. Image enhancement is an effective solution to solve the above problems. The night image is enhanced by the day image, which can greatly enhance the information perception of computer vision and human vision. Through image enhancement, image characteristics such as brightness, contrast, signal-to-noise ratio, edge sharpness, and color accuracy are improved [1,2], and the feature differences between night and day images are further reduced. This increases the degree of image aggregation in the feature space, which is beneficial to the training and reasoning processes of visual deep learning networks. Traditional image enhancement methods are based on mathematical computations that do not need training in advance. This can save computing power for computationally constrained automated driving applications. Such methods can be used as the data preprocessing

module of the deep-learning-based automatic driving computer vision tasks for images with low illumination at night.

Land and McCann proposed and developed the Retinex theory [3,4]. Retinex theory regards the image as the superposition of two components: illumination and reflectance. Illumination is the influence of ambient light in the imaging process. The reflectance component represents the natural properties of the objects in the image and is not affected by other factors. The purpose of the Retinex algorithm is to separate the reflectance component of the object from the image, removing the effects of ambient lighting. In night image enhancement, the Retinex algorithm can obtain the reflectance component of night image objects and remove unfavorable illumination conditions.

Many image enhancement algorithms are derived from Retinex theory. These algorithms separate only the reflectance components of the object and ignore the illumination. This normally leads to poor imaging results. Reflectance components pay more attention to the high-frequency information, such as the edge texture of the image, but lack information on color and brightness. This is not conducive to enhancing contrast and producing proper brightness. In addition, better image enhancement requires more careful manual parameter adjustment to guarantee high performance [5]. This limitation makes it difficult to generalize the algorithm based on the Retinex theory in practice.

HE (histogram equalization) has been widely used for image brightness enhancement [6]. It expands the existing gray levels of the original image to the whole gray level (0–255). For example, for night images, the overall image style is dark, and the gray level is concentrated in a small gray level range. HE can significantly improve the image brightness by expanding the gray level distribution to the entire gray level. The classic HE increases brightness by evenly distributing the entire gray level. The average brightness of the enhanced image changes dramatically. However, if there are both over-light and over-dark areas in one image, HE will map the pixel brightness in the two areas to medium-level brightness. A bright pixel may be mapped to the same medium brightness as a dark pixel, resulting in the loss of image edge details [7]. Additionally, HE expands the gray level of the image from 0–50 to 0–255, meaning that the brightness of pixels with the same gray level will be different after expansion. This brings high-frequency noise to the enhanced image.

DCT (discrete cosine transform) is similar to DFT (discrete Fourier transform) but only operates with real numbers. Compared with DFT, DCT has better aggregation for certain information. In the image field, images are often processed by DCT and IDCT (inverse discrete cosine transform) in the frequency domain. The low-frequency signal of the image mainly corresponds to the slowly changing information, such as color and brightness. High-frequency signals correspond to rapidly changing information in images, such as the edges. Ordinary high-pass filters or low-pass filters can only achieve image smoothing or sharpening [8]. The image enhancement algorithm based on Retinex theory retains more edge information, but the visual effect of the image depends on fine parameter adjustment. HE enhances the lighting information better, but the edge information is lost. The image frequency domain transformation is performed via DCT, combining with the advantages of the Retinex and HE. HE enhances the image brightness, and more edge information is retained by the Retinex algorithm.

This paper proposes HE-MSR-COM, which combines the low-frequency information of HE-enhanced images with the high-frequency information of MSR-enhanced images. The low-frequency information of HE-enhanced images can provide enhanced illumination, to ensure a better visual experience. While the high-frequency information of MSR can retain more edge details, it will improve the quality of the image, e.g., contrast, mean gradient, etc. This method can filter the high-frequency noise brought by HE and achieve performance balance and optimization with overall illumination and edge details. This paper mainly focuses on enhancing low-illumination images. Images under rainy and foggy weather can be categorized as low-illumination images and can use the same processing method proposed in this paper. The salt and pepper noise introduced by these typical

weather conditions needs further image processing steps, such as image noise reduction, which is not within the scope of this paper.

The structure of this paper is as follows: Section 2 introduces the development of different research directions and related work on night image enhancement. In Section 3, the relevant theoretical basis is introduced, and the research method of this paper is proposed. Section 4 describes the selection of the dataset and experimental evaluations, as well as the analysis of the experimental results. Section 5 summarizes the performance of the proposed algorithm and indicates future research directions.

2. Related Works

2.1. Retinex Theory

Many low-light image enhancement algorithms have been developed based on Retinex theory. Jobson et al. improved the Retinex theory and proposed SSR (single-scale Retinex) [9] and MSR (multiscale Retinex) [10]. These methods simply assume that the illumination is smooth and the reflectance components are unsmooth. The Gaussian low-pass filter (LPF) and logarithm operation are used to estimate the illumination of the image. The gradient and region size in different images are different. SSR needs to strike a balance between overall illumination estimation and local image detail performance. MSR uses different weights for several linear LPFs to estimate illuminance. This can ensure the balance of performance in the overall illumination and local image details. Wang et al. [11] proposed a low-illumination color image enhancement algorithm based on the Gabor filter and Retinex theory. The algorithm extracts the illumination component from the HSI (hue, saturation, intensity) color space of the original image. The authors enhanced the illumination component using MSRCR (multiscale Retinex with color restore) to obtain the enhanced illumination component and illuminated images. Additionally, the original image of the RGB space is enhanced using the SSR algorithm. Then, the illuminated image and the enhanced image are weighted and fused for better performance. Traditional Retinex-based algorithms use Gaussian filters (GSFs) to estimate illumination. However, GSFs cannot adapt themselves to different backgrounds in images, which is the main reason why they cannot accurately estimate illumination [12]. Tao et al. [13] replaced the GSF with a region covariance filter (RCF), which depends on the covariance matrix of local image features for each pixel. As a result, the RCF is adaptive to different pixels in an image and can estimate illumination more accurately than the GSF. The RCF Retinex algorithm increases contrast, cancels noise, and enhances detail compared with GSF Retinex algorithms. However, the calculation of RCF Retinex is time-consuming and impractical.

The performance of these methods often depends on the careful selection of the parameters of the filters and their corresponding weights, and most of these parameters require human-involved decisions, which are time- and human-labor-intensive and are impractical for real-time applications such as night image enhancement in autonomous driving.

2.2. Histogram Equalization

Histogram equalization (HE) is used to enhance contrast and improve image quality. Yeong Kim [14] considered that the original HE algorithm would cause the loss of edge information and, therefore, reduce the image contrast. They proposed bi-histogram equalization (BBHE) to enhance the image contrast. The average value of illumination is used as a threshold to distinguish dark and bright areas. The HE algorithm is used in both a bright area and a dark area to reduce the loss of edge information. However, this results in unbalanced overall distribution of illumination in the enhanced image. Chen et al. [15] believe that the median illumination is more appropriate as the threshold instead of the average illumination. Therefore, they proposed dualistic sub-image histogram equalization (DSIHE) to prevent over-light or over-dark areas from affecting the threshold, and their experiments proved that the median is more statistically significant. Ooi et al. [16] proposed bi-histogram equalization with a plateau level (BHEPL), which reduces the processing time compared to BBHE. Ooi et al. [17] proposed quadrant dynamic

histogram equalization (QDHE), which divides the histogram into four (quadrant) sub-histograms based on the input image's median value. It reduces noise amplification and over-enhancement. Salah et al. [18] proposed a combination of gamma correction and the retinal filter (gamma-HM-COMP), which preserves the contrast between the gray levels of the original pixels, thereby preserving more edge information. Tan et al. [19] proposed a background-brightness-preserving HE (BBPHE) based on nonlinear histogram equalization. This method divides the image into background regions and non-background regions. It can enhance the brightness of the whole image and preserve the edge information of the object as much as possible. Adaptive histogram equalization (AHE) is a commonly used method that calculates the local gray histogram of images to obtain more local details and improve contrast. Shome et al. [20] proposed a contrast-limited AHE (CLAHE) to overcome the problem that AHE will overamplify the noise in the same area of the image. On the other hand, Lin et al. [21] proposed averaging histogram equalization (AVHEQ) for color images. This algorithm divides the original image into sub-images and equalizes them independently. It proposes a new mathematical algorithm to determine the optimal threshold and achieves better performance compared with conventional methods such as BBHE, DSIHE, and BHEPL. Chen et al. [22] used a fast guide filter to decompose the image into a base layer and a detail layer. The plateau equalization (PE) enhances the detail and the background separately, increasing the contrast of the detail. Kwan et al. [23] used a second-order histogram matching algorithm that enhances 16-bit infrared video contrast. This optimizes the possible information loss caused by using processed 8-bit infrared video. The performance of this method has been improved in the target detection using You Only Look Once (YOLO) and classification using a residual network (ResNet). Liao et al. [24] proposed an innovative box filtering method by combining the mean and median filtering techniques to achieve the balance between noise removal and edge preservation.

HE-based algorithms are popular because they are easy to implement and fast to process. However, these algorithms also have various limitations, such as adding noise to the output image and increasing the contrast of the background rather than the object in the image. The direct stretching on the gray level also causes the loss of edge information, resulting in a fuzzy edge. Much research has been carried out to prevent the loss of edge information. However, this issue is more complicated to solve in complex illumination scenes.

2.3. Data-Driven Methods

Recently, many image enhancement methods have been combined with deep learning. These methods use a data-driven approach to enhance the night image adaptively based on a priori trained model. CNN (convolutional neural network) is a typical approach that employs supervision training of a large number of labeled datasets and has shown good adaptability to different scenes. It is a resource-intensive task to collect the required datasets that contain a large number of paired low-light and normal-light images as sample data and label data, respectively. LLNet [25] is trained by pseudo-labels generated by random gamma correction. These unreal labels are given by the traditional image enhancement algorithm, which limits its enhancement effect. Due to the cost of the dataset and the poor generalization ability of CNNs, this method often results in artifacts and unnatural images.

Methods based on unsupervised GANs (generative adversarial networks) do not require a large number of paired images as a training set. These methods can mitigate the cost of collecting labeled datasets. EnlightenGAN [26], a low-light image enhancement algorithm based on an unsupervised GAN, uses unpaired low-light and normal-light data as the dataset. However, the performance of GAN methods is highly affected by the selection of the dataset. GAN methods can produce unpredictable outputs. Some produces features that fool the discriminator and are regarded as the correct result, which is actually an unsatisfactory result.

Qu et al. [27] adopted deep learning to compensate for the defects of traditional image enhancement methods. However, these methods rely heavily on datasets with perfect

scenes for training. It is challenging to allocate adequate computing resources to image enhancement in real-time automatic driving applications.

3. Method

The structure of the proposed HE-MSR-COM Algorithm 1 contains three main parts, including the MSR enhancement module, HE enhancement module, and frequency-domain fusion module. The MSR and HE enhancement modules are responsible for obtaining the edge and illumination enhancement information of the image, respectively, which can be seen in Figure 1. The frequency-domain fusion module is used to adaptively unify the edge and illumination information by deriving weights for different scenarios.

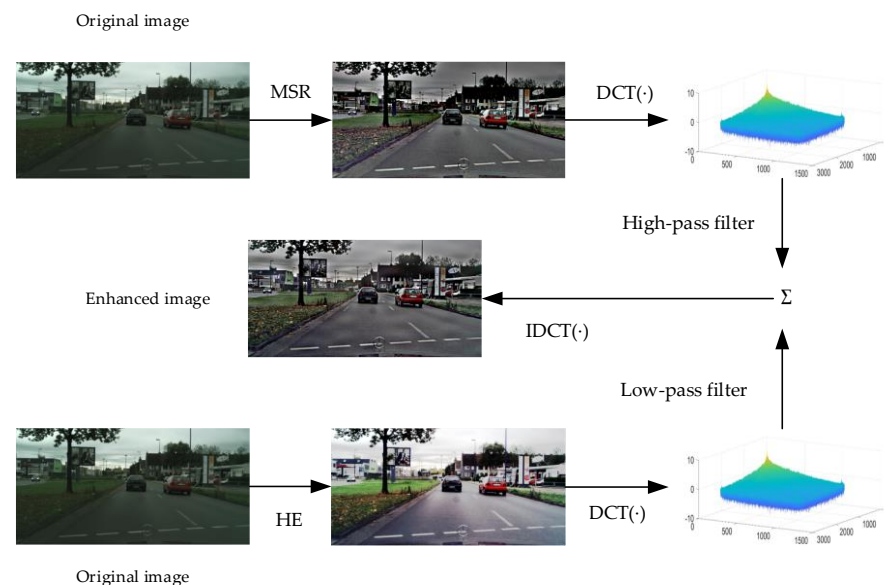


Figure 1. Overview of frequency–domain fusion based on MSR and HE.

3.1. MSR Image Enhancement

Retinex theory is based on the idea that images are a combination of illumination and reflectance. The theory of Retinex is shown in Figure 2.

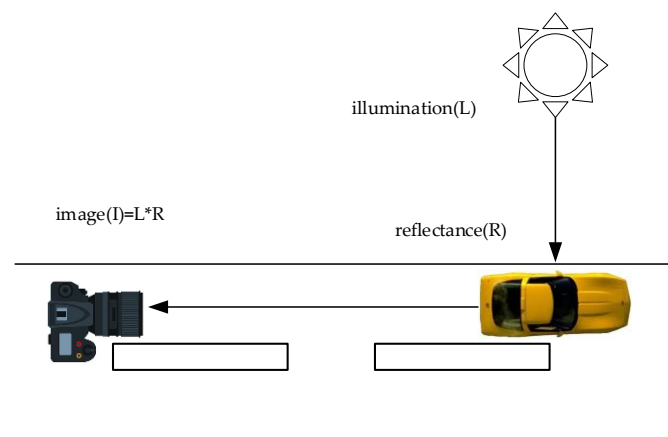


Figure 2. Sketch of Retinex theory.

Retinex theory can be defined as follows:

$$I = L * R \quad (1)$$

where I is the original image, L is a matrix of illumination, and the matrix R represents the reflectance components of the object in I . The operation $*$ is the matrix multiplication of the corresponding elements. Illumination L is a dynamic result of a series of different light sources, such as clear daytime lighting, nighttime street lighting, and other common lighting environments. The reflectance component R represents the key information for humans or computers to understand the semantics of the images. MSR separates the reflectance components to reduce the interference of the dynamic lighting environment with the image semantics. This allows the observers to better understand the image. It is difficult to calculate the reflectance component R directly. By first estimating the illumination L , R can be computed indirectly by $R = I/L$. The MSR can be defined as follows:

$$I(x, y, c) = L(x, y, c) \times R(x, y, c) \quad (2)$$

$$R(x, y, c) = I(x, y, c) / L(x, y, c) \quad (3)$$

$$\log(R(x, y, c)) = \log(I(x, y, c)) - \log(L(x, y, c)) \quad (4)$$

The image is composed of multiple pixels; x and y are the two-dimensional coordinates of the image pixels, and c is the channel of the image. If the image is gray, then c is 1, representing the gray channel. If it is a color image, c is 1, 2, or 3, representing the R, G, and B color channels, respectively. Equation (4) is the logarithmic form of Equation (3).

It is assumed that the illumination component L changes slowly on different objects, while the object reflectance R changes significantly at the edges of the objects. Therefore, the common method is to estimate the slowly changing illumination L by the Gaussian filtering method in the spatial domain. The Gaussian filter is used to estimate the illumination by calculating a weighted average of a pixel and its surrounding pixels. L can be estimated as follows:

$$L(x, y, c) = I(x, y, c) \times G(x, y, c) \quad (5)$$

$$G(x, y, c) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (6)$$

The parameter σ is a key parameter of the Gaussian filter, which determines the filtering scale when estimating the illumination. Selecting a large value is not conducive to local illumination estimation. A small value of σ would defeat the original purpose of the hypothesis and would not be conducive to estimating the overall illumination. Therefore, MSR estimates the illumination by using three different scales: large, medium, and small. The accurate illumination is determined by the weighted average value.

3.2. HE Image Enhancement

The grayscale distribution histograms of over-light or over-dark images are concentrated in the area of high or low brightness, respectively. The grayscale distribution histograms of images with normal lighting are evenly distributed within the overall gray value range. HE mainly uses the CDF (cumulative distribution function) to shift the gray/brightness of the image to ensure that it is distributed uniformly within the overall gray value range, which is similar to that of a normal lighting image.

For the original gray image $I(x, y)$, there are N pixels whose value range is $[P_{min}, P_{max}]$. The brightness is divided into L discrete levels with a range of $[0, L - 1]$. The original histogram of the image is obtained by (7). CDF is defined by (8).

$$H(k) = \frac{n_k}{N}, \quad \text{for } 0 \leq k \leq L - 1 \quad (7)$$

$$\text{CDF}(k) = \sum_{i=0}^k H(k) \quad (8)$$

where $H(k)$ is the PDF (probability density function) of the pixel with a brightness of k , and also the histogram height of the pixel with a brightness of k , while n_k is the number of pixels with a brightness of k .

$$P_{out} = CDF(P_{in}) \times (L - 1) \quad (9)$$

HE pixel brightness mapping is defined in (9). P_{in} is the input pixel brightness, while P_{out} is the output brightness of the corresponding pixel. For the color images, the three channels (RGB) can be enhanced by the above HE. The grayscale distribution before and after HE enhancement is shown in Figure 3.

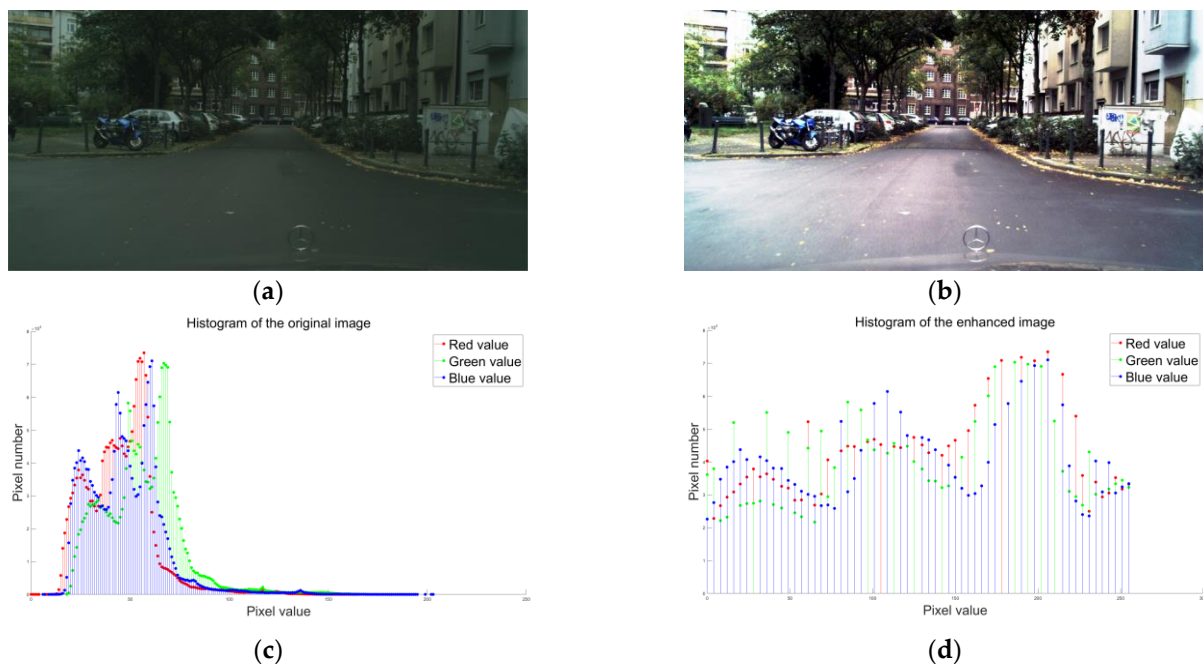


Figure 3. HE enhancement demo: (a) Original image. (b) HE-enhanced image. (c) RGB histogram of the original image. (d) RGB histogram of the HE-enhanced image.

The original night image has low brightness, and its pixel brightness is concentrated in a small range, resulting in poor visibility. After the enhancement, the image brightness is evenly distributed in the value area, and the overall image brightness increases noticeably.

3.3. Image Fusion

MSR can separate the object reflectance components of the image and, thus, retain edge information. The visibility of the MSR-enhanced image is limited, as MSR eliminates the illumination components and only keeps the reflectance components. HE directly modifies the gray value of pixels to achieve better enhancement in illumination, but it also introduces high-frequency noise to the enhanced image. Direct conversion on the gray level will also cause the loss of edge information, which is the key information for semantic segmentation in autonomous driving. The image illumination and color information are mainly in the low-frequency range, while the edge information is mainly in the high-frequency range. The enhancement effect of MSR is more remarkable in the high-frequency range, but it is not stable in the low-frequency range. Conversely, HE can effectively enhance the low-frequency information, but it also causes high-frequency noise and loss of edge information—mainly located in the high-frequency range of the image. The proposed HE-MSR-COM combines the above two methods by using DCT to generate high-quality images that include the high-frequency information from MSR and the low-frequency information from HE.

The proposed HE-MSR-COM uses the high-frequency information of the MSR-enhanced image to obtain the clear edge information and uses the low-frequency information of the HE-enhanced image to obtain the enhanced illumination. HE-MSR-COM overcomes the disadvantages of MSR-enhanced images, such as halo and poor visibility. It also overcomes the shortcomings of HE-enhanced images, such as blurred edges and high-frequency noise. The fusion of MSR and HE processes in the proposed HE-MSR-COM is defined in (10).

$$I_{out} = IDCT(\alpha(I) \times DCT(I_{MSR}) * mask_{MSR} + \beta(I_{HE}) \times DCT(I_{HE}) * mask_{HE}) \quad (10)$$

where I_{out} is the output enhanced image; I_{MSR} and I_{HE} denote the MSR- and HE-enhanced images, respectively; $DCT(\cdot)$ is the discrete cosine transform, and $IDCT(\cdot)$ is the inverse discrete cosine transform; $mask_{MSR}$ is the high-pass filter, and $mask_{HE}$ is the low-pass filter; $*$ represents the multiplication of the corresponding positions of two matrices of the same size; $\alpha(I)$ is an edge-adaptive coefficient, which is a function of the original input image I ; $\beta(I_{HE})$ is an adaptive coefficient as a result of a function of the image illumination.

The frequency-domain diagram after DCT transformation is shown in Figure 4. The low-frequency information is concentrated near the origin of the coordinates, and the high-frequency information is distributed in other areas. The frequency-domain filter design is shown in Figure 5.

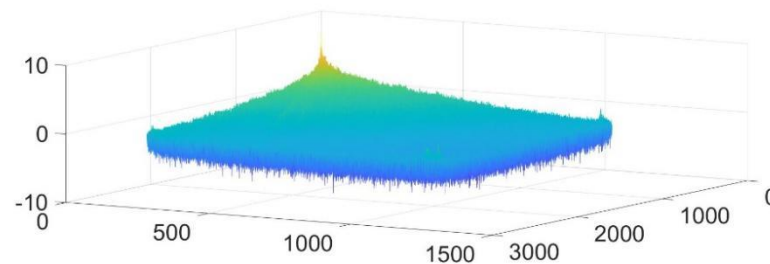


Figure 4. DCT transform spectrum diagram.

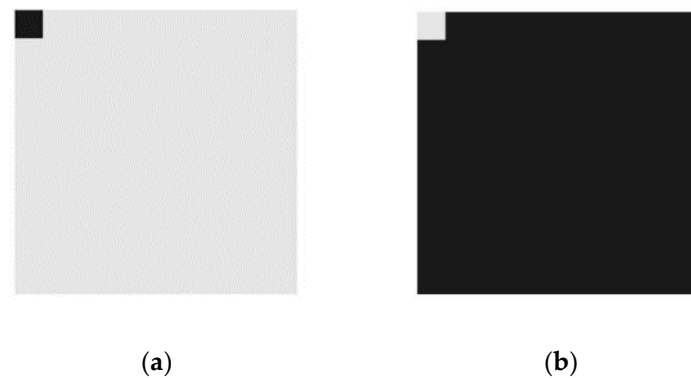


Figure 5. The filter is a logical matrix of the same size as the original image. The matrix element value of the gray part is 1, and the matrix element value of the black part is 0. (a) High-pass filter. (b) Low-pass filter.

The mean gradient is an evaluation of edge information, defined as follows:

$$g(I) = \frac{1}{(M-1) \times (N-1)} \times \sum_{i=1}^{M-1} \sum_{j=1}^{N-1} \sqrt{\frac{(I(i,j) - I(i+1,j))^2 + (I(i,j) - I(i,j+1))^2}{2}} \quad (11)$$

where M, N define the size of the image, I is the image, and i, j are the coordinates of the pixels.

$\alpha(I)$ is determined by the edge information of the original image, and it is defined as follows:

$$\alpha(I) = \alpha \times g(I) / \text{mean}(g) \quad (12)$$

where $\text{mean}(g)$ represents the mean gradient values of the selected images in the dataset, and $g(I)$ is the mean gradient of the current image; α is an adjustable parameter in a range of 0.8–1.2. If α is too small, the edge information will be lost. If α is too large, the edge of the object will be too bright, and the enhanced image will not be natural.

$\beta(I_{HE})$ is determined by the HE-enhanced image. It is used to adjust for excessive enhancement effects that HE may bring. It is defined as follows:

$$\beta(I_{HE}) = \beta \times \text{mean}(\text{mid}(I_{day})) / \text{mid}(I_{HE}) \quad (13)$$

where $\text{mid}(\cdot)$ is the median brightness of an image, which is a statistical function to reasonably judge the brightness distribution of an image. I_{day} is a subset of the normal illuminated images in the dataset. The subset can be selected manually from daylight images or automatically selected according to the calculated brightness values of the images. β is an adjustable parameter in a range of 0.7–1.0. The image is over-dark if β is less than 0.7, and over-bright if β is larger than 1, which both degrade the image's visibility. HE tends to have excessive enhancement, so a value less than 1 is generally selected. γ is a mean memory parameter that is used to update $\text{mean}(g)$ and $\text{mean}(\text{mid}(I_{day}))$ with an additive contribution rate of the current image. If γ is too small, the mean values change slowly and reduce the adaptability, and if γ is too large, the enhanced performance becomes unstable.

The filter parameters of mask_{MSR} and mask_{HE} are mainly determined by prior knowledge of the dataset that contains both day and night images.

Algorithm 1 HE-MSR-COM

Input: Low-light input image I ;

Output: Enhanced image I_{out} ;

Initialization:

$\text{mean}(g)$ is the mean gradient obtained from the sampling data of the dataset;

I_{day} samples from selected normal lighting images;

Mean memory parameter γ ;

Calculate MSR weight parameter α , HE weight parameter β ;

Dataset sampling to obtain prior filter parameters mask_{MSR} , mask_{HE} .

1: **while** (**Input** $\neq \emptyset$) **do**

2: Update $\text{mean}(g)$ by $\text{mean}(g) = \gamma \times g(I) + (1 - \gamma) \times \text{mean}(g)$

3: **if** (I is normal illumination image) **then**

4: Update $\text{mean}(\text{mid}(I_{day}))$ by

$\text{mean}(\text{mid}(I_{day})) = \gamma \times \text{mid}(I) + (1 - \gamma) \times \text{mean}(\text{mid}(I_{day}))$;

5: **else**

6: Estimate initial illumination L via (5), (6);

7: Estimate reflectance $R(I_{MSR})$ via (4)

8: Obtain HE-enhanced image I_{HE} via (7), (8), (9);

9: Calculate weight parameters $\alpha(I)$ via (12);

10: Calculate weight parameters $\beta(I_{HE})$ via (13);

11: Fuse enhanced image via (10) to obtain I_{out} ;

12: **end if**

13: **end while**

4. Experiments

4.1. Datasets

From GTA5 [28] and Cityscapes [29], driving images with low light and normal lighting were selected as data sources for the experiment. The GTA5 dataset contains

24,966 high-resolution composite images and is a commonly used dataset for semantic segmentation training in the field of autonomous driving. The Cityscapes dataset consists of 25,000 street images from 50 different cities, collected using different devices under varying lighting conditions.

4.2. Evaluation Metrics

There are two main ways to evaluate the performance of enhanced images: subjective evaluation and objective evaluation. Subjective evaluation is based on human vision and involves human interaction. Objective evaluations are performed by different defined mathematical metrics based on image information. In this paper, entropy, mean gradient, PSNR (peak signal-to-noise ratio), and contrast ratio are used to evaluate the enhanced image.

Entropy is a common objective metric of image quality evaluation. It reflects the richness of an image. In general, the greater the entropy of the image, the richer the information, and the better the quality. It is defined as follows:

$$E(I) = - \sum_{i=0}^{L-1} P(i) \times \log_2(P(i)) \quad (14)$$

where $P(i)$ is the probability of the pixels with gray level of i in the image, and L is the pixel's gray dispersion level—generally 256.

PSNR is used to measure the distortion degree of the enhanced image. The larger the PSNR, the more semantic information the enhanced image retains and the less noise it introduces. It is defined as follows:

$$MSE = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [I(i,j) - K(i,j)]^2 \quad (15)$$

$$PSNR = 20 \times \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right) \quad (16)$$

where M, N represent the size of the image. I is the original image, and i, j are the coordinates of the pixels. K is the enhanced image. MAX_I is the maximum pixel value—for general RGB images, it is 255.

Contrast ratio usually shows the sharpness of an image. The higher the contrast, the higher the resolution of the image. It is defined as follows:

$$C(I) = \sum_{\delta} \delta(i,j)^2 P_{\delta}(i,j) \quad (17)$$

where $\delta(i,j)$ is the gray difference between adjacent pixels, and $P_{\delta}(i,j)$ is the probability of pixels with a gray difference of δ .

CE (comprehensive evaluation): for the above four evaluation metrics, the maximum value is taken as 100%, and the CE of each algorithm is calculated. It is defined as follows:

$$CE(I) = \left(\frac{E(I)}{MAX_E} + \frac{g(I)}{MAX_g} + \frac{PSNR(I)}{MAX_{PSNR}} + \frac{C(I)}{MAX_C} \right) / 4 \quad (18)$$

4.3. Experimental Results and Analysis

Frequency components in different spectral ranges are separated from the dataset, and their corresponding mean gradients are calculated. The experimental results are shown in Table 1. Thus, the values of the filter parameters of $mask_{MSR}$ and $mask_{HE}$ are adjusted.

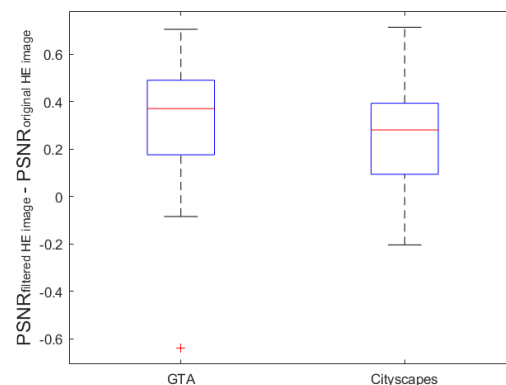
Table 1. Mean gradients of different frequency components.

Spectrum	Mean Gradient
[0, 4)	0.022
[2, 8)	0.062
[4, 16)	0.145
[8, 32)	0.292
[16, 64)	0.550
[32, 128)	0.982
[64, 256)	1.574
[128, 512)	2.151
[256, 1024)	2.322
[1024, 2048)	0.078

Filter parameters' selection ranges are determined based on the algorithm characteristics of HE and MSR. HE has advantages in low-frequency information enhancement, while MSR is better at high-frequency information enhancement. Therefore, the basic parameter selection ranges can be determined, and the violation of this range will lead to poor enhancement results.

It can be noted from Table 1 that the edge information is mainly concentrated in the frequency range (16, 1024), so $mask_{MSR}$ selects these image frequency components in this range to obtain the edge information of the image, while $mask_{HE}$ selects the low-frequency component of the image (0, 16) for generating the enhanced illumination information of the image.

The PSNR of the original HE-enhanced image was compared with that of the filtered HE-enhanced image. As shown in Figure 6, the ordinate is the filtered PSNR minus the original PSNR. The PSNR of the filtered image is larger than that of the original image on both datasets. This proves that a larger PSNR of the image can be obtained by using frequency filtering of the high-frequency noise introduced by the HE enhancement.

**Figure 6.** PSNR comparison between the filtered HE-enhanced image and the original HE-enhanced image.

The next step of the proposed method is combining the filtered HE-enhanced result with the MSR-enhanced result using a set of designed weights. Therefore, using a HE-enhanced image with a higher PSNR contributes to a better final result after fusing with the MSR result. The original HE-enhanced image is not used in the subsequent fusion process.

More experiments are underway to tackle more accurate selection of the filter parameters and the weight parameters (α and β) for a better performance by using optimization algorithms, such as genetic algorithms. These results will be presented in a subsequent paper.

α is an adjustable weight of edge information. The larger α is, the larger the weight of the edge information will be. β is used to adjust the enhanced brightness. β values are generally 0.7–1.0, because HE tends to produce over-bright enhanced results. Five criteria

have been adopted to study the performance of different selections of α and β , as illustrated in Table 2.

Table 2. Enhancement performance of different parameter combinations.

	Entropy	Mean Gradient	PSNR	Contrast Ratio	CE
$\alpha = 0.8, \beta = 0.7$	7.30	3.59	63.71	65.25	77.53%
$\alpha = 1.0, \beta = 0.7$	7.35	4.44	63.48	100.06	87.89%
$\alpha = 1.2, \beta = 0.7$	7.38	5.24	63.17	139.12	98.69%
$\alpha = 0.8, \beta = 0.85$	7.45	3.60	61.82	64.95	77.27%
$\alpha = 1.0, \beta = 0.85$	7.50	4.43	61.68	98.92	87.42%
$\alpha = 1.2, \beta = 0.85$	7.52	5.21	61.50	136.43	97.86%
$\alpha = 0.8, \beta = 1.0$	7.70	3.59	58.83	63.17	76.51%
$\alpha = 1.0, \beta = 1.0$	7.72	4.36	58.78	94.31	85.82%
$\alpha = 1.2, \beta = 1.0$	7.73	5.07	58.73	127.94	95.24%

The experimental results show that when $\alpha = 1.2, \beta = 0.7$, it can obtain the optimal CE on the Cityscapes dataset. Thus, the two weights $\alpha = 1.2, \beta = 0.7$ were selected for the follow-up experiments. The mean memory parameter γ was selected as 0.02 for all of the above experiments. The definitions of α and β take into account the content differences between different images in the dataset. Using adaptive weights makes the enhancement results more stable across different images. Ordinary HE and MSR algorithms can be regarded as methods with weights of 0 or 1, so α and β values close to 1 were selected to prevent over-enhancement. Since HE tends to over-enhance, three discrete values of less than or equal to 1 were selected for β . The high-frequency information processed by MSR can allow for a larger range in selecting α values. Therefore, three groups of representative values of alpha and beta were selected in the experiment. Other parameter values may result in better performance, which will be further studied in our future work. This paper mainly shows that HE and MSR can obtain better enhancement results in frequency-domain combination. CE is a comprehensive consideration of a variety of indicators, so it was selected as the primary evaluation factor. Entropy sometimes cannot accurately represent the image quality. For example, HE usually equally distributes the pixel values, which can theoretically produce the maximum entropy, but there is still room for improvement in the HE enhancement results. The CE result of the selected combination of α and β was extremely close to the highest PSNR result, with a difference of only 0.85%.

The enhancement results of different algorithms—HE, CLAHE, MSR, GC, and HE-MSR-COM (our algorithm)—were compared. The enhanced performance is shown in Figure 7.

All of the methods are implemented via MATLAB programming. MSR and GC were implemented by code in MATLAB. HE and CLAHE were implemented by calling `histeq()` and `adapthisteq()`, respectively, using library functions provided by MATLAB.

The HE-enhanced image has blurred edges, and the image brightness is over-enhanced. CLAHE makes up for HE's over-enhancement issue, but there is still a loss of edge information. MSR enhancement produces sharp edges, but their visibility depends on time-consuming manual adjustment of parameters. The image enhancement results are not stable when using the same parameters for different images. GC directly maps the pixel values nonlinearly. It maps over-light or over-dark pixels to medium brightness. GC achieves good visibility and robust enhancement in brightness, but the image edge information has a great loss and suffers from a foggy effect. The proposed HE-MSR-COM retains the advantages of HE in better adaptive illumination enhancement and those of MSR in adaptive edge information enhancement. The highest visual enhancement performance was obtained in the above evaluations.



(1) Enhanced GTA image, subjective evaluation.



(2) Enhanced Cityscapes image, subjective evaluation

Figure 7. Results comparison: (a) original image; (b) HE; (c) CLAHE; (d) MSR; (e) GC; (f) HE-MSR-COM.

The performance comparisons of different algorithms on the GTA5 dataset are shown in Table 3 and Figure 8.

Table 3. Mean performance comparison in GTA5.

Method	Original	HE	CLAHE	MSR	GC	HE-MSR-COM
Entropy	6.52	7.80	7.24	3.47	6.81	7.48
Mean Gradient	2.25	5.85	5.34	4.12	2.40	9.88
PSNR	Inf ¹	57.60	61.83	61.70	60.14	61.45
Contrast Ratio	47.14	200.15	140.64	242.51	38.78	424.19
CE	54.40%	74.89%	70.02%	60.79%	54.51%	98.84%

¹ The PSNR score of the original image was deemed to be 100%.

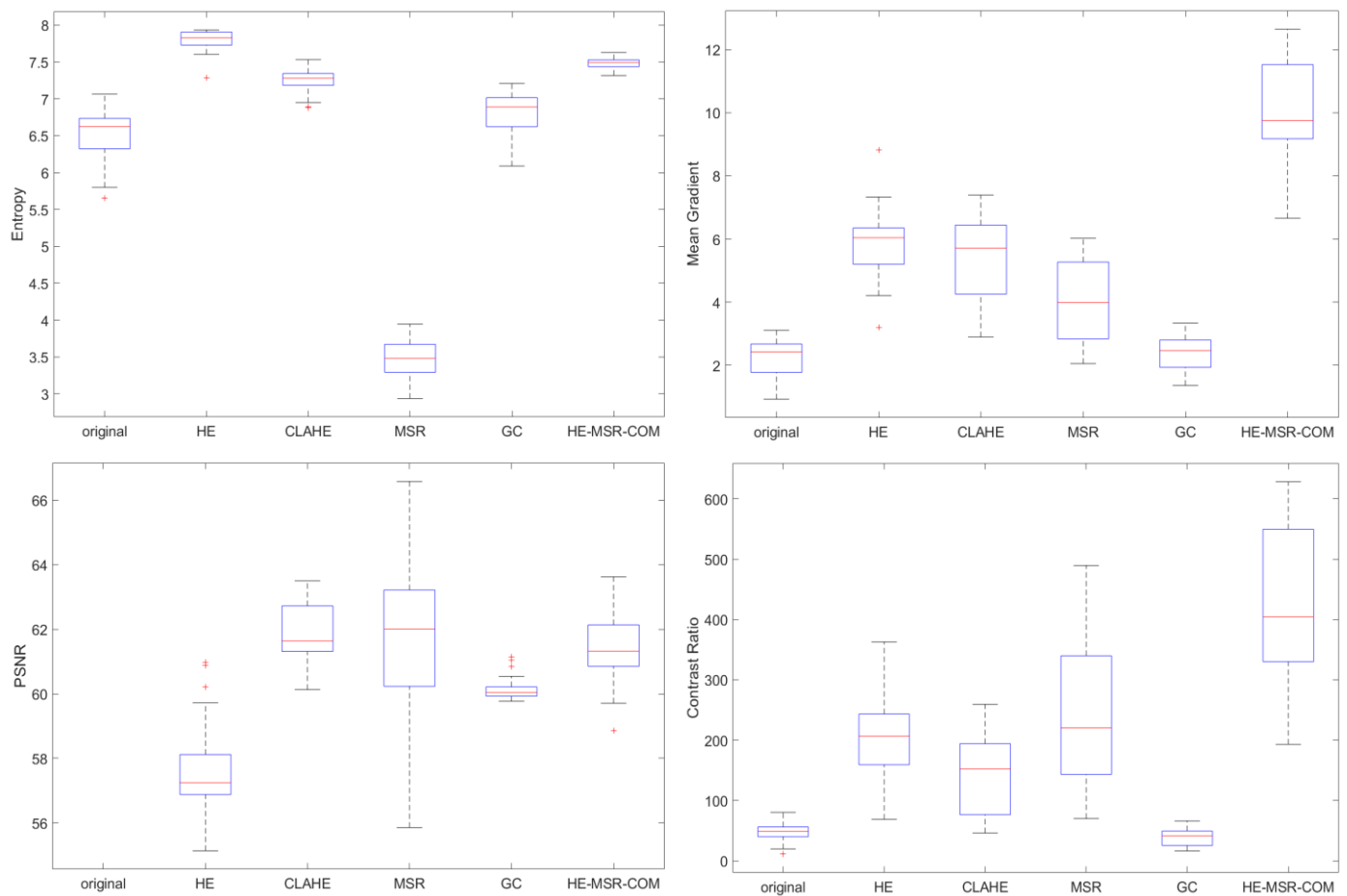


Figure 8. Evaluation distribution in GTA5.

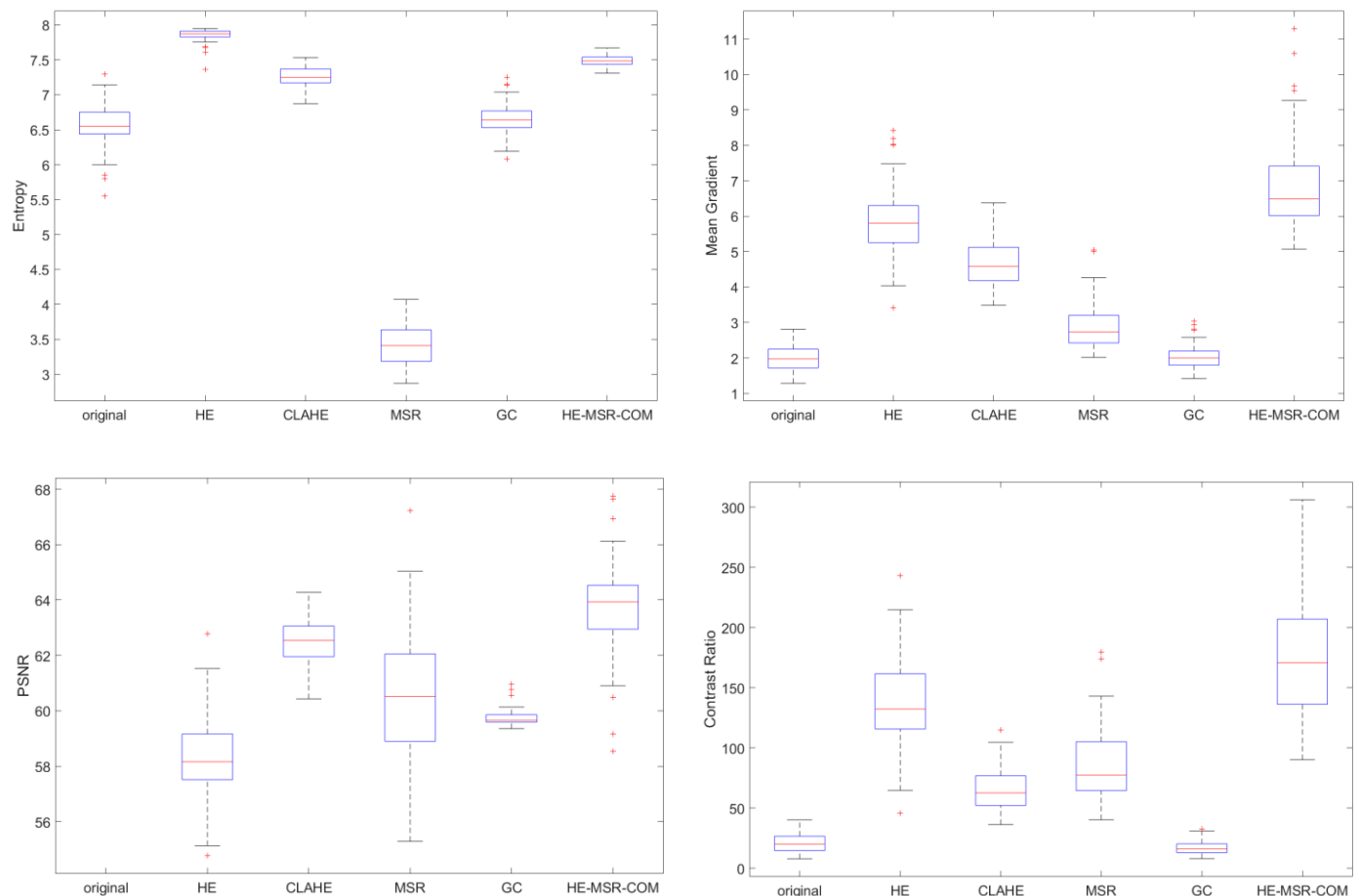
Figure 8 presents the results of five different image enhancement methods under four evaluation metrics, including entropy, mean gradient, PSNR, and contrast ratio. The “+” symbols represent the outliers. The upper and lower boundaries of the dashed line are the maximum and minimum values, respectively. The upper and lower boundaries of the box are the upper and lower quartiles, respectively, and the inner red line is the median value. The boxplots can adequately represent the distribution of enhanced image performance. As shown in Figure 8, HE achieves the maximum entropy by evenly distributing the grayscale of the pixels. HE-MSR-COM retains the advantages of HE in low-frequency information with respect to brightness and color, so HE-MSR-COM has the second-highest score in entropy. HE-MSR-COM magnifies the advantages of MSR by adaptive weight and achieves the highest mean gradient. HE-MSR-COM overcomes the problem of high-frequency noise caused by HE and achieves the second-highest PSNR, which is almost equal to the highest PSNR. HE-MSR-COM has the highest average value in contrast ratio, but it also brings large variance. However, the subjective evaluation shows that the contrast ratio of HE-MSR-COM is significantly enhanced, and it receives the highest score in the CE, as indicated in Table 3. HE-MSR-COM shows optimized performance in many evaluation metrics on the GTA5 dataset.

The performance comparison of different algorithms on the Cityscapes dataset is shown in Table 4 and Figure 9.

Table 4. Mean performance comparison in Cityscapes.

Method	Original	HE	CLAHE	MSR	GC	HE-MSR-COM
Entropy	6.55	7.85	7.25	3.42	6.66	7.48
Mean Gradient	1.98	5.81	4.71	2.93	2.05	6.88
PSNR	Inf ¹	58.34	62.50	60.52	59.76	63.74
Contrast Ratio	20.68	135.90	66.42	85.75	16.88	176.62
CE	55.98%	88.23%	74.14%	57.42%	54.47%	98.83%

¹ The PSNR score of the original image was deemed to be 100%.

**Figure 9.** Evaluation distribution in Cityscapes.

As shown in Figure 9, HE and HE-MSR-COM have the highest and second-highest score in entropy, respectively. HE-MSR-COM has the highest mean gradient in all methods. HE-MSR-COM removes noise as much as possible through frequency-domain filtering and obtains the highest PSNR. HE-MSR-COM has the highest average value in contrast ratio, but it also brings large variance. Because the images of Cityscapes come from different devices, the contrast between images is different, and the high variance is consistent with the actual images. Our method also receives the highest score in the CE, as indicated in Table 4. HE-MSR-COM has advanced performance in entropy, mean gradient, PSNR, and contrast ratio in many real scenes of the Cityscapes dataset.

In the above GTA5 and Cityscapes datasets, HE achieved stable enhancement performance in different datasets, but its PSNR was low due to the introduction of high-frequency noise and over-enhancement. MSR requires manual adjustment of parameters to achieve optimal performance. Using the same parameters in different datasets brings performance instability. HE obtained the highest CE score in both datasets—23.95% and 10.6% higher than those of the second-highest method, respectively.

4.4. Discussion

The experiment compared subjective evaluation with objective evaluation. The results reveal that HE is simple and reliable, but this comes at the expense of losing edge information and excessively enhancing brightness. On the other hand, the stability of MSR is low. The optimal enhancements usually require manual adjustment of the parameters independently. It is difficult to adapt fixed parameters for complex realistic conditions. HE-MSR-COM uses HE-enhanced images to obtain good brightness enhancement with an adjustable weight and excellent subjective visual evaluation performance, and MSR is used in the proposed method to obtain the edge information, achieving good performance under the different metrics such as contrast ratio and mean gradient. Frequency-domain processing can effectively combine the advantages of HE and MSR, as well as avoiding the problems of high-frequency noise caused by HE and unstable brightness in the low-frequency domain of MSR. Based on the experimental performance, HE-MSR-COM showed stable and superior enhancement performance on different datasets. HE-MSR-COM is simple, reliable, and efficient, and it can be used as a preprocessing module for low-illumination images for most visual algorithms.

5. Conclusions

This paper proposes a method combining HE with MSR, called HE-MSR-COM. The image enhancement method proposed in this paper focuses on enhancing the low-illumination image, which is used as an image preprocessing module. When the image is collected by the autopilot system, it is first processed by the image enhancement before it is taken as an input for the subsequent autonomous driving visual tasks, such as semantic segmentation, target detection, etc. We aim to improve performance in the visual tasks of autonomous driving by providing a higher quality of the visual image. Our experiments showed that HE-MSR-COM has the advantages of both HE and MSR, enabling it to achieve higher performance and balance in the overall illumination and edge details. The HE-MSR-COM night image enhancement algorithm has two advantages: (1) The enhanced illumination component is obtained from HE-enhanced image. The low-pass filter in the frequency domain retains the advantage of enhanced illumination and removes the high-frequency noise. This successfully ensures good adaptive illumination. (2) The enhanced reflectance component is obtained from the MSR-enhanced image. The high-pass filter in the frequency domain retains the advantage of enhanced edge information. This successfully reserves more semantic information. HE-MSR-COM achieves excellent night image enhancement performance. It can be embedded into common visual algorithms for autonomous driving to improve their visual detection performance in night scenes.

In the future, HE-MSR-COM will be deployed in autonomous driving semantic segmentation networks. The night image enhancement can be further evaluated and optimized by combining it with practical autonomous driving visual algorithms in real night scenes.

Author Contributions: Conceptualization, X.C. and Y.H. (Yi Han); methodology, Y.H. (Yi Han), X.C., and Y.Z.; software, X.C., Y.H. (Yi Han), and Y.Z.; validation, X.C., P.H., Y.H. (Yanqing Huang), and Z.L.; formal analysis, X.C., Y.Z., Y.H. (Yanqing Huang), and Z.L.; investigation, X.C., Z.L., Z.Y., and Q.L.; resources, X.C., P.H., and Y.H. (Yi Han); data curation, X.C., Y.H. (Yi Han), and Z.L.; writing—original draft preparation, X.C. and Y.H. (Yi Han); writing—review and editing, Y.H. (Yi Han), Z.Y., and Q.L.; visualization, X.C. and Y.H. (Yi Han); supervision, Y.H. (Yi Han), P.H., and Z.Y.; project administration, Y.Z.; funding acquisition, Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by a grant from the National Natural Science Foundation of China (Grant No. 61801341). This work was also supported by the Research Project of Wuhan University of Technology Chongqing Research Institute (No. YF2021-06).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The simulation data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Wang, W.; Wu, X.; Yuan, X.; Gao, Z. An experiment-based review of low-light image enhancement methods. *IEEE Access* **2020**, *8*, 87884–87917. [\[CrossRef\]](#)
- Sobhahi, R.A.; Tekli, J. Comparing deep learning models for low-light natural scene image enhancement and their impact on object detection and classification: Overview, empirical evaluation, and challenges. *Signal Process. Image Commun.* **2022**, *109*, 116848. [\[CrossRef\]](#)
- Land, E.H. The retinex theory of color vision. *Sci. Am.* **1977**, *237*, 108–129. [\[CrossRef\]](#) [\[PubMed\]](#)
- Land, E.H.; McCann, J.J. Lightness and retinex theory. *JOSA* **1971**, *61*, 1–11. [\[CrossRef\]](#) [\[PubMed\]](#)
- Wei, C.; Wang, W.; Yang, W.; Liu, J. Deep retinex decomposition for low-light enhancement. *arXiv* **2018**, preprint. arXiv:1808.04560.
- Gonzalez, R.C.; Woods, R.E. *Digital Image Processing*, 2nd ed.; Prentice Hall: Upper Saddle River, NJ, USA, 2002; p. 793.
- Dhal, K.G.; Das, A.; Ray, S.; Gálvez, J.; Das, S. Histogram equalization variants as optimization problems: A review. *Arch. Comput. Methods Eng.* **2021**, *28*, 1471–1496. [\[CrossRef\]](#)
- Sande, K.V.D.; Gevers, T.; Snoek, C. Evaluating color descriptors for object and scene recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *32*, 1582–1596. [\[CrossRef\]](#)
- Jobson, D.J.; Rahman, Z.; Woodell, G.A. Properties and performance of a center/surround retinex. *IEEE Trans. Image Process.* **1997**, *6*, 451–462. [\[CrossRef\]](#)
- Jobson, D.J.; Rahman, Z.; Woodell, G.A. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Trans. Image Process.* **1997**, *6*, 965–976. [\[CrossRef\]](#)
- Wang, P.; Wang, Z.; Lv, D.; Zhang, C.; Wang, Y. Low illumination color image enhancement based on Gabor filtering and Retinex theory. *Multimed. Tools Appl.* **2021**, *80*, 17705–17719. [\[CrossRef\]](#)
- Yang, X.; Jian, L.; Wu, W.; Liu, K.; Yan, B.; Zhou, Z.; Peng, J. Implementing real-time RCF-Retinex image enhancement method using CUDA. *J. Real-Time Image Process.* **2019**, *16*, 115–125. [\[CrossRef\]](#)
- Tao, F.; Yang, X.; Wu, W.; Liu, K.; Zhou, Z.; Liu, Y. Retinex-based image enhancement framework by using region covariance filter. *Soft Comput.* **2018**, *22*, 1399–1420. [\[CrossRef\]](#)
- Kim, Y.T. Contrast enhancement using brightness preserving bi-histogram equalization. *IEEE Trans. Consum. Electron.* **1997**, *43*, 1–8.
- Chen, S.D.; Ramli, A.R. Minimum mean brightness error bi-histogram equalization in contrast enhancement. *IEEE Trans. Consum. Electron.* **2003**, *49*, 1310–1319. [\[CrossRef\]](#)
- Ooi, C.H.; Kong, N.S.P.; Ibrahim, H. Bi-histogram equalization with a plateau limit for digital image enhancement. *IEEE Trans. Consum. Electron.* **2009**, *55*, 2072–2080. [\[CrossRef\]](#)
- Ooi, C.H.; Isa, N.A.M. Quadrants dynamic histogram equalization for contrast enhancement. *IEEE Trans. Consum. Electron.* **2010**, *56*, 2552–2559. [\[CrossRef\]](#)
- Salah-ELDin, A.; Nagaty, K.; ELArif, T. An enhanced histogram matching approach using the retinal filter's compression function for illumination normalization in face recognition. In Proceedings of the International Conference Image Analysis and Recognition, Póvoa de Varzim, Portugal, 25–27 June 2008; pp. 873–883.
- Tan, T.L.; Sim, K.S.; Tso, C.P. Image enhancement using background brightness preserving histogram equalisation. *Electron. Lett.* **2012**, *48*, 155–157. [\[CrossRef\]](#)
- Shome, S.K.; Vadali, S.R.K. Enhancement of diabetic retinopathy imagery using contrast limited adaptive histogram equalization. *Int. J. Comput. Sci. Inf. Technol.* **2011**, *2*, 2694–2699.
- Lin, S.C.F.; Wong, C.Y.; Rahman, M.A.; Jiang, G.; Liu, S.; Kwoka, N.; Shi, H.; Yu, Y.H.; Wu, T. Image enhancement using the averaging histogram equalization (AVHEQ) approach for contrast improvement and brightness preservation. *Comput. Electr. Eng.* **2015**, *46*, 356–370. [\[CrossRef\]](#)
- Chen, Y.; Kang, J.U.; Zhang, G.; Zhang, G.; Cao, J.; Xie, Q.; Kwan, C. Real-time infrared image detail enhancement based on fast guided image filter and plateau equalization. *Appl. Opt.* **2020**, *59*, 6407–6416. [\[CrossRef\]](#)
- Kwan, C.; Gribben, D. Target Detection and Classification Improvements using Contrast Enhanced 16-bit Infrared Videos. *Signal Image Process. Int. J. (SIPIJ)* **2021**, *12*. [\[CrossRef\]](#)
- Liao, K.C.; Wu, H.Y.; Wen, H.T. Using Drones for Thermal Imaging Photography and Building 3D Images to Analyze the Defects of Solar Modules. *Inventions* **2022**, *7*, 67. [\[CrossRef\]](#)
- Lore, K.G.; Akintayo, A.; Sarkar, S. LLNet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognit.* **2017**, *61*, 650–662. [\[CrossRef\]](#)
- Jiang, Y.; Gong, X.; Liu, D.; Cheng, Y.; Fang, C.; Shen, X.; Yang, J.; Zhou, P.; Wang, Z. Enlightengan: Deep light enhancement without paired supervision. *IEEE Trans. Image Process.* **2021**, *30*, 2340–2349. [\[CrossRef\]](#) [\[PubMed\]](#)

27. Qu, H.; Yuan, T.; Sheng, Z.; Zhang, Y. A pedestrian detection method based on yolov3 model and image enhanced by retinex. In Proceedings of the International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Beijing, China, 13–15 October 2018; pp. 1–5.
28. Richter, S.R.; Vineet, V.; Roth, S.; Koltun, V. Playing for data: Ground truth from computer games. In Proceedings of the European Conference on Computer Vision (ECCV), Amsterdam, The Netherlands, 11–14 October 2016; pp. 102–118.
29. Cordts, M.; Omran, M.; Ramos, S.; Rehfeld, T.; Enzweiler, M.; Benenson, R.; Franke, U.; Roth, S.; Schiele, B. The cityscapes dataset for semantic urban scene understanding. In Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 26–30 June 2016; pp. 3213–3223.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.