Research on Image Enhancement Algorithm Based on Artificial Intelligence

Jie Liu*, Yuanyuan Peng

1School of Information Engineering, China University of Geoscience Beijing, Beijing, China, 100083

*Corresponding author e-mail: liujie@cugb.edu.cn

Abstract. With the continuous development of social science and technology, people have higher and higher requirements for image quality. This paper integrates artificial intelligence technology and proposes a low-illuminance panoramic image enhancement algorithm based on simulated multi-exposure fusion. First, the image information content is used as a metric to estimate the optimal exposure rate, and the brightness mapping function is used to enhance the V component, and the low-illuminance. The image and the overexposed image are input, the medium exposure image is synthesized by the exposure interpolation method, and the low illumination image, the medium exposure image and the overexposure image are merged using a multi-scale fusion strategy to obtain the fused image, which is corrected by a multi-scale detail enhancement algorithm. After the fusion, the details are enhanced to obtain the final enhanced image. Practice has proved that the algorithm can effectively improve the image quality.

Keywords: Image Enhancement, Multi-exposure Fusion, Exposure Interpolation, Image Information Content

1. Introduction
With the continuous development of social science and technology, people have higher and higher requirements for image quality [1-3]. Due to the low-light environment and limited camera equipment, the image has problems such as low brightness, low contrast, high noise, color distortion, etc., which will not only affect the aesthetics of the image and the human visual experience, but also reduce the use of normal lighting images. The performance of visual tasks. In order to effectively improve the quality of low-light images, scholars have proposed many low-light image enhancement algorithms, which undergo three stages of grayscale transformation, retinal cortex theory, and deep neural network [4-7]. And at night, shadows, etc. The quality of panoramic images collected under poor lighting conditions will deteriorate, which is mainly manifested in the overall brightness of the image, low contrast, and color deviation. It has a serious impact on the visual effect of the panoramic image and is a subsequent computer vision processing task (Such as image segmentation, target tracking, target recognition, etc.) bring certain difficulties. Therefore, the development of low-light panoramic image enhancement algorithm research is of great significance to the field of machine vision [8-10].
In this paper, artificial intelligence technology is integrated, and algorithm research is carried out for the needs of image enhancement. A low-illumination panoramic image enhancement algorithm based on simulated multi-exposure fusion is proposed to improve the quality of the image.

2. **Low-light image enhancement algorithm based on fusion**

In order to improve the visualization quality of low-illuminance panoramic images, to solve the problems of weak brightness, low contrast and unclear details of panoramic images collected under low-illuminance conditions, this paper proposes a low-illuminance panoramic image based on simulated multi-exposure fusion Enhanced algorithm, the algorithm flow chart is shown in Figure 1.

![Algorithm flow](image)

**Figure 1.** Algorithm flow

It can be seen from Figure 1 that the algorithm in this paper mainly includes four modules: overexposure image generation, medium-exposure image generation, multi-scale fusion and multi-scale detail enhancement.

2.1. **Shallow up and down sampling structure of the decomposition network**

Different from the commonly used deep U-Net (U-Net) structure and Retinex-Net's simple stacked convolutional layer, the decomposition network of the algorithm in this paper is a shallow up-sampling structure composed of convolutional layers and channel cascade operations. The sampling layer has only 4 layers, and the network training is simpler. Experiments show that when using this up-sampling structure to transform the image scale, the down-sampling operation discards noise-containing pixels to a certain extent, achieving the purpose of noise reduction, but at the same time it will cause the image Blur.

In the shallow up-sampling structure, first, a 9×9 convolutional layer is used to extract the features of the input image Slow. Then, a five-layer convolutional layer with ReLU as the activation function is used to transform the image scale, and learn the reflection component and illumination The features of the components. Finally, the two layers of convolutional layers and the Sigmoid function are used to map the learned features into the reflection map Rlow and the illumination map Ilow and then output.

For the constraint loss of the decomposition network, the algorithm in this paper uses Retinex-Net's reconstruction loss lcon, constant reflectivity loss IR, and illumination smoothing loss IL. In addition, in order to further reduce noise in the decomposition network, add denoising loss ld. Therefore, The total loss is as follows:

\[
l = l_{con} + \lambda IL + \lambda d_1 + \lambda d_2
\]

Among them, λ1, λ2, λ3 are weight coefficients, which are used to balance the loss components. For the selection of L1, L2 norm and Structural Similarity (Structural Similarity, SSIM) loss, when it comes to image quality tasks, L2 norm and human Vision does not have a good correlation with the perception of image quality, and it is easy to fall into a local minimum during training. Although SSIM can learn image structural features better, it is less sensitive to errors in smooth areas, causing color deviation. Therefore, The algorithm in this paper uses the L1 norm to constrain all losses.

The results Rlow and Rnormal output in the decomposition network can be reconstructed with the light map to form a new image, and the reconstruction loss is as follows:
\begin{equation}
I_{\text{recon}} = \sum_{i=\text{low, normal}} W_i R_{\text{low}} \circ I_i - S_i \bigg\| + \sum_{j=\text{low, normal}} W_j R_{\text{normal}} \circ I_j - S_j \bigg\|
\end{equation}

Where represents the pixel-by-pixel multiplication operation. When i is low or j is normal, the weight coefficient $W_1=W_2=1$, otherwise $W_1=W_2=0.001$. For paired images, using larger weights can make the decomposition network better Learning the features of the paired images. For paired image pairs, using larger weights can make the decomposition network better learn the features of the paired images.

The constant reflectance loss IR is based on the color constancy of Retinex theory. It is mainly used in the decomposition network to constrain the consistent reflectance of different illumination images:

\begin{equation}
I_R = \| R_{\text{low}} - R_{\text{normal}} \|
\end{equation}

For the illumination smoothing loss II, this paper adopts the structure-perceived smoothing loss. This loss uses the reflection component gradient term as the weight. In the area where the image gradient changes greatly, the illumination becomes discontinuous, so that the light map with smooth brightness can retain the image Structural information, then

\begin{equation}
I_I = \| \Delta_{\text{low}} \exp \left( -\lambda_g \Delta R_{\text{low}} \right) \| + \| \Delta_{\text{normal}} \exp \left( -\lambda_g \Delta R_{\text{normal}} \right) \|
\end{equation}

Among them, $\Delta$ represents the sum of the horizontal and vertical gradients of the image, and $\lambda_g$ represents the balance coefficient.

The total variation (TV) of a noisy image is greater than that of a noiseless image, and image noise can be reduced by limiting TV. However, in image enhancement, limiting TV is equivalent to minimizing the gradient term. Inspired by the theory of TV minimization, this article introduces reflection The gradient term of the component is used as the loss to control the reflected image noise, so it is called the denoising loss:

\begin{equation}
I_d = \lambda \| \Delta R_{\text{low}} \|
\end{equation}

When the $\lambda$ value increases, the noise decreases and the image will be blurred. Therefore, the choice of weight parameters is very important. After experimental research, it is found that when the weight $\lambda=0.001$, the image obtains a better visual effect.

### 2.2. Enhance the attention mechanism of the network

In response to the color distortion problem, the attention mechanism module is embedded in the enhancement network. It is worth noting that, unlike other complex attention modules, the attention mechanism module is composed of simple convolutional layers and activation operations, and does not require powerful hardware Equipment, there is no need to train multiple models and a large number of additional parameters. In the process of light adjustment, it can reduce the feature response to irrelevant background, activate only the features of interest, and improve the algorithm’s ability to process image details and sensitivity to pixels, Instruct the network to adjust the brightness of the image and preserve the structure of the image.

The input of the attention module is image features $\alpha_i$ and $\beta_i$, and the output is image feature i, i=1,2,3, which represents the serial number of the attention mechanism module. $\alpha_i$ is the image feature output by the down-sampling layer, and $\beta_i$ is the up-sampling layer The output features of these two image features respectively carry different brightness information. After the two pass through the attention module, the response of brightness-independent features (such as noise) is reduced, so that the output feature carries more brightness information and is input to the next The up-sampling layer improves the network’s ability to learn brightness features.

$\alpha_i$ and the reconstructed $\beta_i$ respectively pass through an independent $1\times1$ convolutional layer, and perform an additive operation before ReLU activation. They pass through the $1\times1$ convolutional layer
and the Sigmoid function in turn, and finally multiply with $\beta_i$ element by element. The result is channel-cascaded with $\alpha_i$. In this propagation process, the attention mechanism can fuse image information of different scales, while reducing the response of irrelevant features, and enhancing the network's ability to adjust brightness.

Independent of the constraint loss of the decomposition network, the enhancement network adjusts the degree of illumination based on the assumptions of local consistency and structure perception. In addition to the loss of the constraint enhancement network in Retinex-Net, the algorithm in this paper is aimed at Retinex-Net. The color deviation that occurs increases the color loss and therefore enhances the network loss:

$$L = L_{\text{recon}} + L_i + \mu L_c$$

(6)

Among them, $L_{\text{recon}}$ is the reconstruction loss of the enhanced image,

$$L_{\text{recon}} = \|S_{\text{normal}} - R_{\text{low}} \circ I_n\|$$

(7)

$L_I$ represents the structure-perceived smoothing loss, $L_c$ represents the color loss of this article, and $\mu$ represents the balance coefficient. The definition of $L_{\text{recon}}$ represents the distance term between the enhanced image and its corresponding normal-illuminated image. The structure-perceived smoothing loss $L_I$ is similar to the smoothing loss of the decomposition network, but different. It is that in the enhanced network, $I_{en}$ uses the gradient of $R_{\text{low}}$ as the weight coefficient:

$$L_i = \|\lambda \exp(-\lambda \Delta R_{\text{low}})\|$$

(8)

In addition, this article adds color loss $L_c$ to measure the color difference between the enhanced image and the normal illuminated image. First, Gaussian blur is used on the two images to filter out the high-frequency information such as texture and structure of the image, leaving low-frequency parts such as color and brightness. Calculate the mean square error of the blurred image. The blur operation allows the network to more accurately measure the color difference of the image while limiting the interference of texture details, and further learn color compensation. The color loss is

$$L_c = \|F(S_{\text{en}}) - F(S_{\text{normal}})\|^2$$

(9)

Among them: $F(x)$ represents the Gaussian blur operation, $x$ represents the image to be blurred. This operation can be understood as each pixel of the image takes the average of the neighboring pixels with the normal distribution weight, so as to achieve the blur effect, and $S_{\text{en}}$ is the enhanced image, $S_{\text{normal}}$ is the corresponding normal lighting image.

$$F(x(i,j)) = \sum_{k,l} \alpha(i+k,j+l)G(k,l)$$

(10)

$G(k,l)$ represents the weight coefficient that obeys the normal distribution. In the convolutional network, $G(k,l)$ is equivalent to a fixed-size convolution kernel.

$$G(k,l) = 0.053 \exp\left(\frac{k^2+l^2}{6}\right)$$

(11)

3. Multi-scale fusion

In order to obtain a better image enhancement effect, this paper uses a multi-scale fusion strategy to fuse low-light images, medium-exposure images and over-exposure images. The fusion frame can be expressed as:
\[ I_i(x,y) = \sum_{k=1}^{K} Y_i \{ W_i(x,y) \} L_i \{ E_i(x,y) \} \]  

(12)

In the formula: YI and LI respectively represent the Gaussian pyramid of the first/layer and the Laplacian/pyramid of the lth layer, and \( \frac{W_i}{\sum_{k=1}^{K} W_k} \) is the normalized weight, and E1, E2 and E3 are the low-illumination image, the medium-exposure image and the overexposure image, respectively. Exposing the image, a large number of experiments in different scenes show that the 5-layer pyramid decomposition usually achieves the best results, so the value of l is set to 5 in this paper. This article uses the above method to generate medium and overexposure images for a low-illumination image.

For the low-illumination image E1, it is hoped that it can effectively enhance the poorly exposed areas in the image while retaining the well-exposed areas in the image; compared with E1 and E2, the over-exposed image E3 loses the image details while at the same time. It can show more effective image content information. For this reason, this paper adopts the Sigmoid function based on the illumination component to set the weights of E1 and E3. A large amount of statistical data shows that the pixel value distribution of a well-exposed image approximately satisfies a Gaussian distribution with a mean value of 0.5 and a variance of 0.25, so we use the Gaussian distribution function to set the weight of the medium-exposure image E2. In order to balance the Gaussian distribution function and the Sigmoid function, this paper proposes an improved brightness weight function, which is defined as follows:

\[ W_1 = \frac{1}{1 + e^{-6L_1 + 3}} \]  

(13)

\[ W_2 = e^{\frac{1(L_2 - 0.5) + 3}{0.25}} \]  

(14)

\[ W_3 = \frac{1}{1 + e^{6L_3 - 3}} \]  

(15)

Where L1, L2 and L3 represent the illumination components of E1, E2 and E3, respectively. In order to obtain the illumination components, this article transfers the images E1, E2, E3 from the RGB color space to the HSV color space, and obtains the brightness component of the image, and then uses a weighted least square filter (Weighted Least Square, WLS) that can maintain the edges of the image. Smoothing and filtering is performed on the V component to obtain the illumination component.

![Figure 2. Brightness weight function](image-url)
In Figure 2, the abscissa is the illumination component of the image, and the ordinate is the weight. The red, green, and blue curves represent the brightness weight functions of low-illuminance images, medium-exposure images, and over-exposure images, respectively. By appropriately assigning weights to the pixel values of the three images with different exposure levels, the fused image achieves a good balance between enhancing the brightness and avoiding overexposure.

In the process of Gauss-Laplace pyramid decomposition and reconstruction of the image, as the number of pyramid layers increases, part of the image details will be lost, and reducing the number of pyramid layers will cause halo artifacts in the fusion result. In order to enrich the image details, this paper adopts a multi-scale Gaussian filtering algorithm to enhance the image details while avoiding halo artifacts.

First, a multi-scale Gaussian filter is used to smooth and filter the fused image to obtain 3 different Gaussian blurred images, as shown in equation (16):

\[ B_1 = G_1 \ast I', B_2 = G_2 \ast I', B_3 = G_3 \ast I' \]  

(16)

Secondly, extract fine details D1, intermediate details D2 and coarse details D3 for the image; as shown in equation (18):

\[ D_1 = I' - B_1, D_2 = B_1 - B_2, D_3 = B_2 - B_3 \]  

(17)

Then D1, D2 and D3 are weighted and fused to obtain the detail image D*, as shown in equation (18):

\[ D^* = (1 - w_1 \times \text{sgn}(D_1)) \times D_1 + w_2 \times D_2 + w_3 \times D_3 \]  

(18)

4. Experimental results and analysis

In order to verify the effect of this algorithm on low-light panoramic image enhancement, this paper selects 6 different low-light panoramic images in different scenes for experiments, using NPE algorithm, LIME algorithm, SRIE algorithm, L algorithm, BIMEF algorithm, RetinexNet algorithm and this paper. The algorithms are processed separately, and the experimental results are compared and analyzed.

In order to objectively evaluate the processing results of different algorithms, this paper uses Lightness Order Error (LOE) and Structure Similarity Index (SSIM) as objective evaluation indicators to evaluate the processing results of the methods proposed in this paper. The image brightness distortion is defined as:

\[ \text{LOE} = \frac{1}{m} \sum_{i=1}^{m} RD(x) \]  

(19)

Where: RD(x) represents the relative order difference between the original image and the enhancement result, and x represents the image pixel. RD(x) is defined as:

\[ RD(x) = \sum_{y=1}^{m} U(L(x), L(y)) \oplus U(L'(x), L'(y)) \]  

(20)

Among them: m is the number of pixels, \( \oplus \) is the exclusive OR operation, L(x) and L'(x) respectively represent the maximum value of pixel x in the original image and the enhanced result image. For U(x,y), the default return value is 1. If x>y, the return value is 0. For the enhancement result, the smaller the LOE value of the image, the better the brightness naturalness is maintained, and the lower the brightness distortion rate.

Compared with most of the comparison algorithms, for low-illumination panoramic images in different scenes, the LOE index of the algorithm in this paper is smaller, indicating that the brightness
of the image enhanced by the algorithm in this paper is natural and the brightness distortion rate is lower.

![Image](image1.png)

**Figure 3.** Comparison of LOE objective evaluation results of different algorithms

It can be seen from Figure 3 that for grayscale panoramic images in different scenes, the LOE index of the enhanced image of this algorithm is lower than that of other comparison algorithms, and the result of brightness distortion rate is better, indicating that the algorithm of this paper is enhancing the naturalness and robustness of grayscale images. The stickiness is better.

Structural similarity (SSIM) is an important indicator to measure whether the image structure is distorted. For low-illumination panoramic images in different scenes, the SSIM index of the enhanced image is higher than that of most other comparison algorithms, indicating that the proposed algorithm can improve the brightness of the image while maintaining the original structure of the image. For grayscale panoramic images in different scenes, the SSIM index of the proposed algorithm is better, which shows that while maintaining the original structure of the image.

In order to objectively evaluate the processing results of different algorithms, this paper combines the Natural Image Quality Evaluator (NIQE) and the Blind/Rsfesncslss Images Spatial Quality Evaluator, BRISQUE ) As an objective evaluation indicator.

The method in this paper performs well in various performance indicators. Although not all objective evaluation indicators are the highest value, the values of the indicators are within the normal range, indicating that the algorithm while maintaining the image structure. As well as detailed information, and the color of the enhanced image is more natural.

5. Conclusion

In order to solve the problem of low-light panoramic image enhancement, this paper proposes a low-light panoramic image enhancement algorithm based on simulated multi-exposure fusion, and verifies the effectiveness of this algorithm through subjective visual perception and objective index evaluation. The experimental results show that the algorithm is effective Effectiveness.

References

[1] Park S, Yu S, Moon B, et al. Low-light image enhancement using variational optimization-based retinex model [J]. IEEE Transactions on Consumer Electronics, 2017, 63(2):178-184.
[2] Goldstein T, Xu L, Kelly K, et al. The STONE Transform: Multi-Resolution Image Enhancement and Real-Time Compressive Video [J]. IEEE Trans Image Process, 2015, 24(12):5581-5593.
[3] Lore K G, Akintayo A, Sarkar S . LLNet: A Deep Autoencoder Approach to Natural Low-light Image Enhancement [J]. Pattern Recognition, 2017, 61:650-662.
[4] Fu X, Liao Y, Zeng D, et al. A Probabilistic Method for Image Enhancement With Simultaneous Illumination and Reflectance Estimation [J]. IEEE Transactions on Image
Shribak M, Larkin K G, Biggs D. Mapping optical path length and image enhancement using quantitative orientation-independent differential interference contrast microscopy [J]. Journal of Biomedical Optics, 2017, 22(1): 160-166.

Luo X, Zeng T, Zeng W, et al. Comparative analysis on landsat image enhancement using fractional and integral differential operators [J]. Computing, 2020, 102(1):247-261.

He D, Sun X, Liu M, et al. Image enhancement based on intuitionistic fuzzy sets theory [J]. Iet Image Processing, 2016, 10(10):701-709.

Matsuda T, Ono A, Sekiguchi M, et al. Advances in image enhancement in colonoscopy for detection of adenomas [J]. Nature Reviews Gastroenterology & Hepatology, 2017, 14(5):1-8.

Kwon S, Lee H, Lee S. Image enhancement with Gaussian filtering in time-domain microwave imaging system for breast cancer detection [J]. Electronics Letters, 2016, 52(5):342-344.

Montanini R, Quattrocchi A, Piccolo S A. Active thermography and post-processing image enhancement for recovering of abraded and paint-covered alphanumeric identification marks [J]. Infrared Physics & Technology, 2016, 78(2):24-30.