Semi-Global Stereo Matching Algorithm Based on Optimized Image Preprocessing

Huaijiao Sha
College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222, China

Miao Zhang
College of Electronic Information and Automation, Tianjin University of Science and Technology, 300222, China
E-mail: miaozhang@tust.edu.cn, 2442943837@qq.com

Abstract

This paper proposes an improved Semi-Global Matching algorithm that enhances the quality of disparity images by optimizing image preprocessing. The method applies denoising, contrast enhancement, and Sobel filtering to compute pixel gradients and highlight image edges, thereby improving image quality and reducing noise interference to enhance the clarity of feature boundary. The preprocessed images are used for the SGM algorithm, which improves matching accuracy and adaptability through adaptive sliding windows and dynamic aggregation strategies for cost calculation and aggregation. Experimental results indicate that the improved algorithm enhances the accuracy and robustness of disparity images.

Keywords: Image preprocessing, Denoising, Sobel filtering, Contrast enhancement，Dynamic aggregation

1. Introduction

Stereo matching is one of the core tasks in the field of computer vision, which is widely used in the fields of autonomous driving, robot navigation, augmented reality and 3D scene reconstruction. By estimating the disparity of a pair of corrected stereo images, the stereo matching algorithm can provide key depth information, so as to support the understanding of the scene and the interaction with the physical environment. Among many stereo matching algorithms, Semi-Global Matching (SGM) has attracted much attention because of its good balance between computational efficiency and disparity map quality, and has become a classic and widely used algorithm.

Proposed by Hirschmüller in 2005, SGM approximates global optimization through multi-directional 1D path aggregation, thereby avoiding the computational complexity associated with full global optimization [1]. However, SGM faces challenges in handling weakly textured regions (e.g., sky, walls) and occluded areas, where the lack of feature support can lead to disparity mismatches. Additionally, repetitive texture regions (e.g., lattice structures) may result in ambiguous solutions. SGM is also sensitive to dynamic lighting and noise, which can affect matching accuracy.

To address the aforementioned issues, researchers have proposed optimizations in the cost calculation and aggregation stages in recent years [2]. This paper introduces an adaptive kernel size based on Sobel filtering and a dynamic sliding window strategy. By adjusting the kernel size and weight distribution to adapt to local texture characteristics, the robustness of the matching cost is enhanced. The dynamic weight allocation, combining bilateral filtering and gradient features, further improves the adaptability to complex lighting conditions. In the cost aggregation stage, an adaptive strategy based on dynamic windows is employed to accurately smooth the cost values, enhancing the global consistency of the disparity map.

The rest of this article is organized as follows. The second section presents the image preprocessing optimization; the third section introduces the matching cost calculation and cost aggregation; the fourth section provides simulation examples to verify the effectiveness of the proposed protocol; the fifth section summarizes the main content of this paper.

2. Image Preprocessing Optimization

Image preprocessing is a critical step in the fields of computer vision and image processing [4]. It aims to perform preliminary processing on the original image to improve the effect of subsequent analysis and processing. By eliminating noise, enhancing contrast, edge detection and other operations, image preprocessing lays the foundation for tasks such as feature extraction, image segmentation, and target recognition.

2.1 Bilateral filtering for denoising processing

Bilateral Filtering is a nonlinear filter capable of simultaneously smoothing noise and preserving edges. Unlike traditional Gaussian filtering, bilateral filtering

considers not only spatial distance but also pixel value similarity, enabling edge-preserving image smoothing.
For the pixel $I(p)$ in the image, the formula for the pixel value $I'(p)$ after bilateral filtering is (1)

$$I'(p) = \frac{1}{W_p}\sum_{q\in s} G_S(\| p - q \|)\cdot G_r(|I(p) - I(q)|)\cdot I(q)\#(1)$$

2.2 Sobel filtering and adaptive kernel size

Sobel filtering is a classic edge detection method that extracts significant edge features by computing the gradients of an image in the horizontal and vertical directions, while also providing a certain level of noise resistance. Its core lies in utilizing the Sobel kernel for convolution operations, enhancing edge information and suppressing noise interference.

Although the Sobel filtering method is simple and efficient, its sensitivity to noise and limited effectiveness in processing complex images can be a drawback. The traditional Sobel filter typically uses a fixed-size kernel (e.g., 3×3), but to improve performance, the kernel size can be dynamically adjusted based on the local texture complexity of the image. In regions with complex textures, smaller kernels are selected to preserve details, while in more uniform regions, larger kernels are chosen to enhance smoothing and reduce noise interference. To achieve this, the gradient magnitude or local standard deviation of the image can be computed to assess the texture complexity of local regions, and the most appropriate kernel size is then dynamically selected based on this evaluation. This dynamic adjustment method based on local features improves the adaptability of Sobel filtering in different image regions, thus enhancing the accuracy and robustness of edge detection.

Combine the gradients in the horizontal and vertical directions to calculate the gradient magnitude (2).

$$\boldsymbol{G(x,\ y) = \sqrt{G_x^2(x,\ y) + G_y^2(x,\ y)}} \qquad (2)$$

Adaptive kernel sizes may cause uneven gradient distributions, which can be normalized through local standardization. In the filtered gradient map, for each pixel, a fixed-size window (e.g., 3×3 or 5×5) is used to compute (3) the mean μ and standard deviation σ of the gradients within the window.

$$\boldsymbol{\mu = \frac{1}{N}\sum_{i=1}^{N} G_i \qquad \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(G_{i-\mu}\right)^2}} \qquad (3)$$

Normalize the gradient values by using the local mean μ and standard deviation σ to standardize the gradient value of each pixel (4).

$$\boldsymbol{G_{norm} = \frac{G - \mu}{\sigma + \epsilon}} \qquad (4)$$

Fig. 1 is the original unprocessed input image, serving as the basis for subsequent operations. Fig, 2 illustrates the gradient magnitude of the image, with enhanced visualization achieved through normalization. Fig. 3 presents the smoothed image obtained by applying filtering to remove noise and suppress unnecessary detail variations. Fig. 4 Fig. 4 depicts the result of overlaying the gradient information onto the original image, using transparency or other visualization techniques to highlight edge features while retaining the content of the original image.

Fig. 1 Original image

Fig. 2 Gradient magnitude

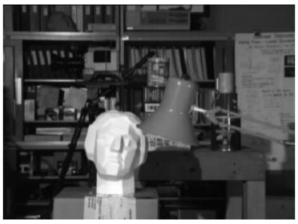

Fig. 3 Filtered image (Original)

Fig. 4 Gradient overly on original

2.3 Histogram equalization

Histogram equalization is a widely used image enhancement technique that significantly improves image contrast by adjusting the distribution of grayscale values. It is particularly effective for scenes with low original contrast. This method redistributes pixel grayscale values to achieve a more uniform distribution, thereby improving global or local contrast, enhancing visual quality, and revealing finer details.

Assuming the grayscale range of the image is [0,L−1], the grayscale equalization is performed through the following steps:

•Calculate the cumulative distribution function (CDF) of the grayscale values:

$$c(r_k) = \sum_{i=0}^{k} h(r_j) \qquad (5)$$

•Map the grayscale value r_k to the equalized grayscale value s_k:

$$s_k = \lfloor (L-1)\cdot c(r_k) \rfloor \qquad (6)$$

The image after histogram equalization in Fig. 6 has obvious contrast enhancement, especially in the area with low contrast of the original image. Fig. 7 is the histogram of the original image, which shows the distribution of the number of pixels with different gray values in the original image. After processing, the distribution of gray levels becomes more uniform (as shown in Fig. 8), the visual effect is more distinct, and the details are more prominent. Fig. 5 is the untreated control original image.

Fig. 5 Original image (Histogram)

Fig. 6 Original histogram equalized image

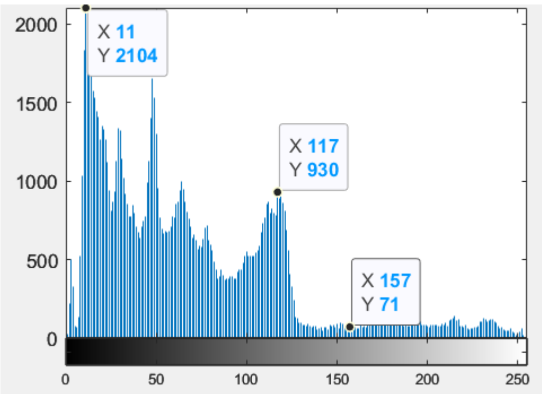

Fig. 7 Histogram of original image

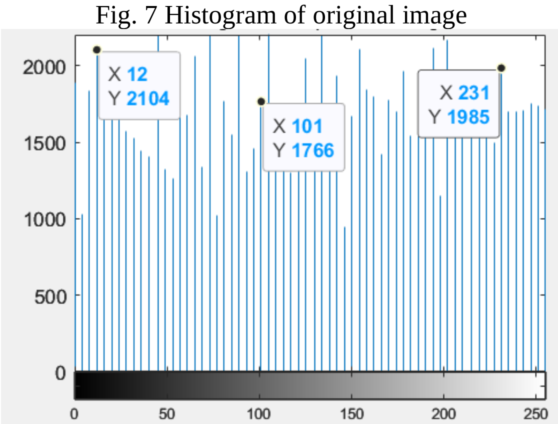

Fig. 8 Histogram of equalized image

2.4 Assigning weights

Based on the characteristics of different image regions, dynamically assign weights to balance the contributions of various preprocessing methods to image quality enhancement. The formula for weight assignment is as follows:

$$\begin{cases} w_{sobel} = f_{\text{gradient}}(\nabla \mathrm{I}) \\ w_{\text{bilateral}} = f_{\text{noise}}(\sigma_n) \\ w_{\text{histogram}} = f_{\text{contrast}}(\mu_{\text{local}}) \end{cases} \tag{7}$$

The weight for Sobel filtering (w_{sobel}) is evaluated based on local gradient variations. Calculate the gradient magnitude $G = \sqrt{(G_x^2 + G_y^2)}$ and increase the weight in regions with prominent edges (high gradient). The weight for bilateral filtering ($w_{bilateral}$) is based on image noise estimation. By analyzing the global noise level σ_n (e.g., mean absolute error or local variance), increase the filtering weight in high-noise regions. The weight for histogram equalization ($w_{histogram}$) is based on contrast distribution. Calculate the local contrast μ_{local} (e.g., the standard deviation of brightness) and assign higher weights to low-contrast regions.

The final fused result I_{final} is obtained by weighted summation of each method:

$$I_{\text{final}} = w_{\text{sobel}} \cdot I_{\text{sobel}} + w_{\text{bilateral}} \cdot I_{\text{bilateral}} + w_{\text{histogram}} \cdot I_{\text{histogram}} \tag{8}$$

3. Matching Cost Calculation and Cost Aggregation

The implementation includes the following steps:

• **Matching Cost Calculation**: Dynamically generate edge information using Sobel filtering with adaptive kernel size, and compute the initial cost matrix using a cost function based on gradient or color differences (e.g., absolute difference or mutual information).

•**Cost Aggregation**: Through a dynamic sliding window mechanism, the cost values within the window are weighted and accumulated, with the weights determined by both the spatial relationships between pixels and their color similarity. By dynamically adjusting the block Size parameter via a slider, the size of the matching window can be modified in real-time, thereby directly influencing the cost computation results. Changes in the window size present a clear trade-off in the disparity map generation process.

• **Path Optimization**: Perform multi-directional path aggregation (e.g., 8 or 16 directions in SGM) to globally smooth the matching cost, ensuring the consistency of the results.

• **Dynamic Adjustment**: During the cost aggregation process, the aggregation weights and window size are dynamically adjusted based on scene complexity, significantly improving the handling of occluded regions and boundary details. The initial cost for each pixel row is calculated using the custom $calcPixelCostBT$ function and stored in $Cbuf$. Subsequently, the cost values are weighted, accumulated, and optimized by integrating the aggregation window configuration and dynamically adjusted parameters (such as sliding window size and weight distribution), enhancing the robustness and accuracy of disparity computation.

4. Experiments and Results

4.1 Dataset and experimental setup

The experiment was conducted on a computer equipped with an Intel Core i7-9750H CPU, using the Middlebury dataset, which is widely used to validate stereo matching algorithms and contains image pairs of varying complexity. The aim of the experiment was to evaluate the effectiveness of the proposed image preprocessing methods (including Sobel filtering, adaptive kernel size, bilateral filtering, and histogram equalization) and the improved matching cost calculation and aggregation strategies in enhancing disparity estimation accuracy and robustness.

4.2 Experimental Results and Analysis

In the field of stereo vision, various algorithms have been proposed to improve the accuracy and robustness of disparity map estimation [3]. The original disparity image shown in Fig. 9 is from Hirschmüller's pioneering work, which introduced the semi-global matching (SGM) method. Fig. 10 shows the disparity map generated by the optimized SGM method.
The experimental results show that the improved image preprocessing method significantly enhances the quality and accuracy of the disparity map Fig. 10. The dynamic weight allocation strategy, combining histogram equalization and gradient features, improves the clarity of feature boundaries and enhances the algorithm's adaptability to lighting variations. By dynamically adjusting the sliding window size, weight distribution strategy, and cost calculation method, the matching accuracy is further improved. Compared to the traditional SGM algorithm, the improved method performs better in complex scenes (such as low-texture areas), producing more accurate and consistent disparity maps.

Fig. 9 Comparison of disparity maps

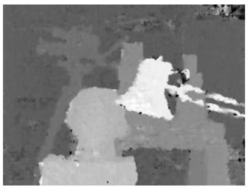

Fig. 10 Disparity map of preprocessed images

5.Conclusion

This study investigates an optimization approach in stereo matching tasks by combining Sobel filtering with adaptive kernel size, bilateral filtering, and histogram equalization. By dynamically adjusting the kernel size, it highlights image edge features and enhances the robustness of matching costs. Meanwhile, a weight allocation strategy integrates the strengths of different preprocessing methods, effectively adapting to weak texture, low contrast, and complex lighting scenarios. Experiments demonstrate that this method significantly improves the accuracy and consistency of disparity estimation. However, the experiment also has certain limitations. Although the improved algorithm shows significant enhancements in accuracy and robustness, some shortcomings remain. Despite the introduction of a dynamic weight allocation method based on local image characteristics, the precision of dynamic adjustment still leaves room for improvement in certain complex scenarios, particularly under extreme lighting conditions and large-scale texture variations.

References

1. Hirschmüller Heiko, Stereo processing by semiglobal matching and mutual information, *IEEE transactions on pattern analysis and machine intelligence*, 2008, 30(2):328-41.
2. Deng Changgen, Liu Deyuan, Zhang Haodong, et al., Semi-Global Stereo Matching Algorithm Based on Multi-Scale Information Fusion, *Applied Sciences*, 2023,13(2):1027-1027.
3. Tippetts, B., Lee, D.J., Lillywhite, K. et al., Efficient stereo vision algorithms for resource-limited systems, *Journal of Real-Time Image Processing*, 2015,10(1):163-174.
4. W. Mao, M. Wang, J. Zhou et al., Semi-Dense Stereo Matching Using Dual CNNs, *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2019, pp. 1588-1597.

Authors Introduction

Mr. Huaijiao Sha

He received his Bachelor's degree in New Energy Science and Engineering from Hebei North University in 2023. Currently, he is a master's student at Tianjin University of Science and Technology, and his research area is machine vision.

Ms. Miao Zhang

She is a postgraduate tutor of Tianjin University of Science and Technology. In 2019, she received a doctorate from the University of Windsor, Ontario, Canada. The research direction is intelligent algorithms design filters, the control system design of industrial robots and control theory.