

## Article

# A Vascular Feature Detection and Matching Method Based on Dual-Branch Fusion and Structure Enhancement

Kaiyang Xu <sup>1,\*</sup>, Haibin Wu <sup>1,\*</sup>, Yuji Iwahori <sup>2</sup>, Xiaoyu Yu <sup>3</sup>, Zeyu Hu <sup>1</sup> and Aili Wang <sup>1</sup>

<sup>1</sup> Heilongjiang Province Key Laboratory of Laser Spectroscopy Technology and Application, Harbin University of Science and Technology, Harbin 150080, China; 2110603078@stu.hrbust.edu.cn (K.X.); 2220600027@stu.hrbust.edu.cn (Z.H.); aili925@hrbust.edu.cn (A.W.)

<sup>2</sup> Computer Science, Chubu University, Kasugai 487-8501, Japan; iwahori@isc.chubu.ac.jp

<sup>3</sup> College of Electron and Information, University of Electronic Science and Technology of China, Zhongshan Institute, Zhongshan 528402, China; yuxy@zsc.edu.cn

\* Correspondence: woo@hrbust.edu.cn

**Abstract:** How to obtain internal cavity features and perform image matching is a great challenge for laparoscopic 3D reconstruction. This paper proposes a method for detecting and associating vascular features based on dual-branch weighted fusion vascular structure enhancement. Our proposed method is divided into three stages, including analyzing various types of minimally invasive surgery (MIS) images and designing a universal preprocessing framework to make our method generalized. We propose a Gaussian weighted fusion vascular structure enhancement algorithm using the dual-branch Frangi measure and MFAT (multiscale fractional anisotropic tensor) to address the structural measurement differences and uneven responses between venous vessels and microvessels, providing effective structural information for vascular feature extraction. We extract vascular features through dual-circle detection based on branch point characteristics, and introduce NMS (non-maximum suppression) to reduce feature point redundancy. We also calculate the ZSSD (zero sum of squared differences) and perform feature matching on the neighboring blocks of feature points extracted from the front and back frames. The experimental results show that the proposed method has an average accuracy and repeatability score of 0.7149 and 0.5612 in the Vivo data set, respectively. By evaluating the quantity, repeatability, and accuracy of feature detection, our method has more advantages and robustness than the existing methods.



**Citation:** Xu, K.; Wu, H.; Iwahori, Y.; Yu, X.; Hu, Z.; Wang, A. A Vascular Feature Detection and Matching Method Based on Dual-Branch Fusion and Structure Enhancement. *Sensors*

2024, 24, 1880. <https://doi.org/10.3390/s24061880>

Academic Editor: Panicos Kyriacou

Received: 1 February 2024

Revised: 9 March 2024

Accepted: 13 March 2024

Published: 15 March 2024



**Copyright:** © 2024 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:** internal cavity features; vascular structure enhancement; vascular feature detection; MIS images; Gaussian weighted fusion; self-adaptive threshold

## 1. Introduction

The implementation of augmented reality technology in minimally invasive surgery (MIS) lies in the real-time acquisition and integration of internal cavity data by laparoscopy, feedback on its pose and depth information, and how to achieve cognitive display correlation between virtual information and real scenes, thereby reducing doctors' workload, improving spatial awareness, and expanding vision [1,2]. In response to the observation limitations of laparoscopic fixed display devices, it is necessary to use advanced 3D reconstruction technology to achieve a real-time overlay of preoperative or intraoperative 3D virtual images on the basis of 2D imaging display.

The prerequisite for reconstruction technology is accurate feature tracking. However, for narrow, lack of features, specular reflections, and poor lighting conditions in intracavity environments, feature tracking remains highly challenging [3]. Due to the fact that laparoscopy can only obtain two-dimensional information about the characteristics of the inner cavity, and cannot obtain its depth information through a single observation, when the feature is detected again during the movement of the laparoscope, the three-dimensional spatial information is calculated based on the principle of triangulation using

the disparity angle of the laparoscope at different positions. Therefore, how to obtain internal cavity features and perform data matching is a great challenge for laparoscopic 3D reconstruction [4].

Existing research indicates that there are issues in directly applying commonly used visual techniques to MIS due to the involvement of large-scale, free-form tissue deformation and constantly changing surgical scenes that are influenced by complex factors [5]. In this case, the accuracy and efficiency of tracking largely depend on the detection of visual features, which need to exhibit high repeatability under image transformation and robustness under scene changes. The inner cavity environment is very complex, with uneven lighting, high specular reflection, soft tissue deformation in the inner cavity, and often without strong edge features, which pose higher requirements for the robustness of visual processing algorithms to improve the number, repeatability, and distribution of extracted feature points in the medical image processing field.

Vascular features are the most obvious features in the lumen environment, and effective vascular structure enhancement methods can provide useful structural information for later vascular feature extraction. Although the vascular structure is complex and detailed, it shares the common characteristics of a tubular-like structure. The ideal enhancement method should be able to achieve a high and uniform response to the following factors: variable vascular morphology; intensity nonuniformity caused by blood contrast or background; other unrelated structural tissues surrounding blood vessels leading to blurred boundaries; and a large amount of complex information such as background noise [6].

Most of the curve structure enhancement methods are based on the classical Hessian matrix analysis of images, and the research of Frangi et al. has been widely recognized [7]. The method proposed by Meijering et al. further developed the use of Hessian matrices [8]. A Hessian matrix is based on second-order Gaussian derivatives and can be calculated at different scales to enhance features similar to curves, which helps distinguish specific shapes such as circular, tubular, and planar structures. In images, when dark curve structures appear relative to the background, chromaticity measurement is used to describe the image. However, the main drawback of the Hessian matrix is its sensitivity to noise [9]. Due to the large eigenvalues, the curved feature response at the connections is very small, which results in poor performance in dealing with different curved branch structures. In the field of cardiovascular structure enhancement, Huang established an improved non-local Hessian-based weighted filter (NLHWF) to suppress nonvascular structures and noise around real vascular structures in contrast images [10]. In the field of neural structure enhancement, the use of diffusion tensors has a more positive effect on vascular structure enhancement. Jerman et al. used spherical diffusion tensors to overcome the shortcomings of using Hessian matrices, but did not fully solve the problem of low intersection strength [11]. The fractional anisotropic tensor (FAT) measures the variance of eigenvalues between different structures and can detect changes in vascular anisotropy [12]. This feature enables FAT to more reliably enhance the vascular structure, ultimately resulting in a more uniform response. Alhasson proposed a new enhancement function based on FAT, which robustly maintains the relationship between isotropy and anisotropy by considering the degree of anisotropy of the target structure, in order to maintain low amplitude feature values, and successfully achieves high-quality curve structure enhancement for retinal images [13]. Dash et al. proposed a fusion method to enhance the fusion of coarse and fine blood vessels, combined with the advantage of the Jerman filter's effective response to blood vessel branches, using mean-C threshold segmentation to improve the performance of traditional curved transformation and improve the algorithm speed [14]. The above vascular structure enhancement methods are mainly used for extracting neural structures, coronary angiography, or curve structures of retinal blood vessels, and have not yet been applied in the lumen environment.

Human luminal vascular feature extraction refers to the automatic or semi-automatic extraction of key points of blood vessels from medical images, such as bifurcation points, endpoints, and intersections. The lumen environment measured by laparoscopy is often

unknown. Considering factors such as image quality, lighting conditions, background generated by image correction, uncertain environmental descriptions, and noise, robust features can provide effective and accurate two-dimensional information for three-dimensional reconstruction of the lumen environment. The traditional feature extraction methods started with Harris's definition of corners [15], and then classic methods such as the features from the accelerated segment test (FAST) [16] and scale invariant feature transform (SIFT) [17] were successively proposed. Improved algorithms for SIFT, such as the speed up robust features (SURF) algorithm [18], have better computational speed but poor stability.

Although the oriented FAST and rotated BRIEF (ORB) algorithm [19] solves the rotational invariance in binary robust independent elementary features (BRIEF), it is prone to errors caused by local homogenization of feature points and cumulative drift. Faced with the complex environment of minimally invasive surgery, traditional feature extraction methods applied in large-scale environments with sufficient light and strong edge features are no longer applicable. Therefore, Stamatia proposed an affine invariant anisotropic feature detector based on automatic feature detection (AFD), which only responds to features with low anisotropy, compensating for the shortcomings of LoG and DoG and effectively processing isotropic features [20]. Lin proposed the vessel branching circle test (VBCT) based on Frangi and verified its robustness and saliency for MIS images through monotonic mapping [21]. Subsequently, due to the presence of a certain pixel width in blood vessels, more than one branch point may be detected on a branch, leading to redundancy. Udayakumar put forward the ridgeness-based circle test (RBCT), which simplifies the structure of blood vessels into single pixels, making feature extraction of blood vessels more efficient [22]. They used the patch search process of parallel tracking and mapping (PTAM) [23] to perform feature matching by calculating the zero mean zero sum of squared differences (ZSSD). However, due to the large-scale nature of branch points, their positioning error is about  $1.6 \pm 0.7$  pixels, and the large error introduces uncertainty to subsequent 3D reconstruction. Li et al. added a decision tree based on the C4.5 algorithm to the traditional FAST algorithm, making the feature extraction performance of MIS images more stable and feature point extraction more efficient [24].

On the other hand, rough feature matching methods often have mismatches and the feature point matching methods based on certain structures are adopted according to different application scenarios. For example, matching based on lines and matching based on various stable geometric structures [25]. Davison proposed an active matching method [26], while random sample consensus (RANSAC) and its improved algorithm eliminated the occurrence of mismatches in active matching [27]. Wu propose a feature point mismatch removal method that combines optical flow, descriptor, and RANSAC, to eliminate incorrect feature point matches layer by layer through these constraints and exhibits robustness and accuracy under various interferences such as lighting changes, image blurring, and unclear textures [28]. Liu et al. proposed an improved ORB feature matching algorithm that utilizes non-maximum suppression (NMS) and retinal sampling models to solve mismatch problems in complex environments [29].

In addition to the direct impact of the accuracy of intracavity feature matching on the localization estimation of laparoscopy, the computational complexity of intracavity feature matching is also an important factor affecting the real-time performance of the entire algorithm. Puerto Souza proposed new hierarchical multiaffine (HMA) and adaptive multiaffine (AMA) algorithms to improve the feature matching performance of endoscopic images [30,31]. Zhu proposed a feature matching method for endoscopic images based on motion consensus. A spatial motion field was constructed based on candidate feature matching, and the true match was identified by checking the difference between feature motion and the estimated field, achieving a 3.7 pixel average reprojection error [32]. Li designed a global and local point-by-point matching algorithm based on feature fusion for extracting contour features from gastrointestinal endoscopic images, which used fast Fourier transform (FFT) to reduce the dimensionality of dense feature descriptors and achieve robust and accurate contour feature extraction and matching [33]. Currently, in the

medical field, most visual-based 3D reconstruction methods still rely on matching features obtained by traditional methods; even by directly estimating and fusing keyframe depth maps [34] or combining them with camera pose estimation [35] to avoid the need for feature matching. Therefore, feature extraction and matching of the inner cavity environment remains a challenging task.

The main contributions of the proposed method are summarized as follows:

1. To solve the problems of noise and false detection caused by different inner cavity environment boundary areas and highlight areas, we analyze the characteristics of the inner cavity environment, design a targeted preprocessing framework, use a self-adaptive threshold to generate binary masks, introduce a local variance threshold to automatically detect highlight areas, and use the fast marching method (FMM) method for highlight repair.
2. In order to robustly extract effective structural information for vascular features, a Gaussian weighted fusion algorithm for single-pixel enhancement of the vascular structure using dual-branch Frangi measure and multiscale fractional anisotropic tensor (MFAT) is designed. The structural similarity (SSIM) index is introduced to achieve self-adaptive Gaussian weighted fusion, solving the problems of structural measurement differences and uneven response between venous vessels and microvessels, ensuring the continuity and integrity of the vascular structure.
3. To enhance adaptability to different scale vascular structures, a dual-circumference detection of vascular features is implemented. Redundancy is reduced by introducing NMS. Feature matching of neighboring blocks of the results before and after frames is achieved by calculating the zero sum of squared differences (ZSSD).

## 2. Methods

This paper proposes a method for detecting and matching vascular features based on dual-branch weighted fusion vascular structure enhancement, as shown in Figure 1. An image preprocessing framework is designed based on the characteristics of the internal cavity environment to improve the method's performance. Binary mask images of boundary regions are generated using a self-adaptive threshold to preserve pixels with valid information in the field of view. The green channel is selected with the highest contrast and the threshold judgment and FMM are used to repair the highlighted areas. Based on the characteristics of the vascular structure, a single-pixel, dual-branch weighted fusion vascular image enhancement algorithm is designed by combining the vascular structure enhancement function based on Frangi measurement and the vascular structure enhancement function based on MFAT. Finally, the double circles detection based on the characteristics of branch points is performed and NMS to extract feature points is introduced. Then, the ZSSD is calculated for the neighboring blocks of feature points extracted from the frames before and after, followed by feature point matching.

### 2.1. Preprocessing

As medical optical equipment, endoscopes have a wide variety of types and are applied in different fields. The requirement for minimally invasive surgery limits the size of the optical field that doctors can see. Therefore, in order to maximize the doctor's field of view, a circular projection created by some laparoscopic descending optical systems is placed inside the image sensor, as shown in Figure 2. As a result, the obtained endoscopic images have noncontent areas, namely dark and uninformed boundary areas, with clear and sharp edges between them and the content areas [36]. For binocular laparoscopy, parallax correction can also result in boundary areas containing noise in the corrected image. At the same time, the large area of bleeding also lacks vascular information, so this paper defines it as a boundary area. The boundary region will generate interference information during image enhancement, so self-adaptive threshold [37] is used to generate effective mask images for different types of laparoscopic images, improving the generalization of the method.

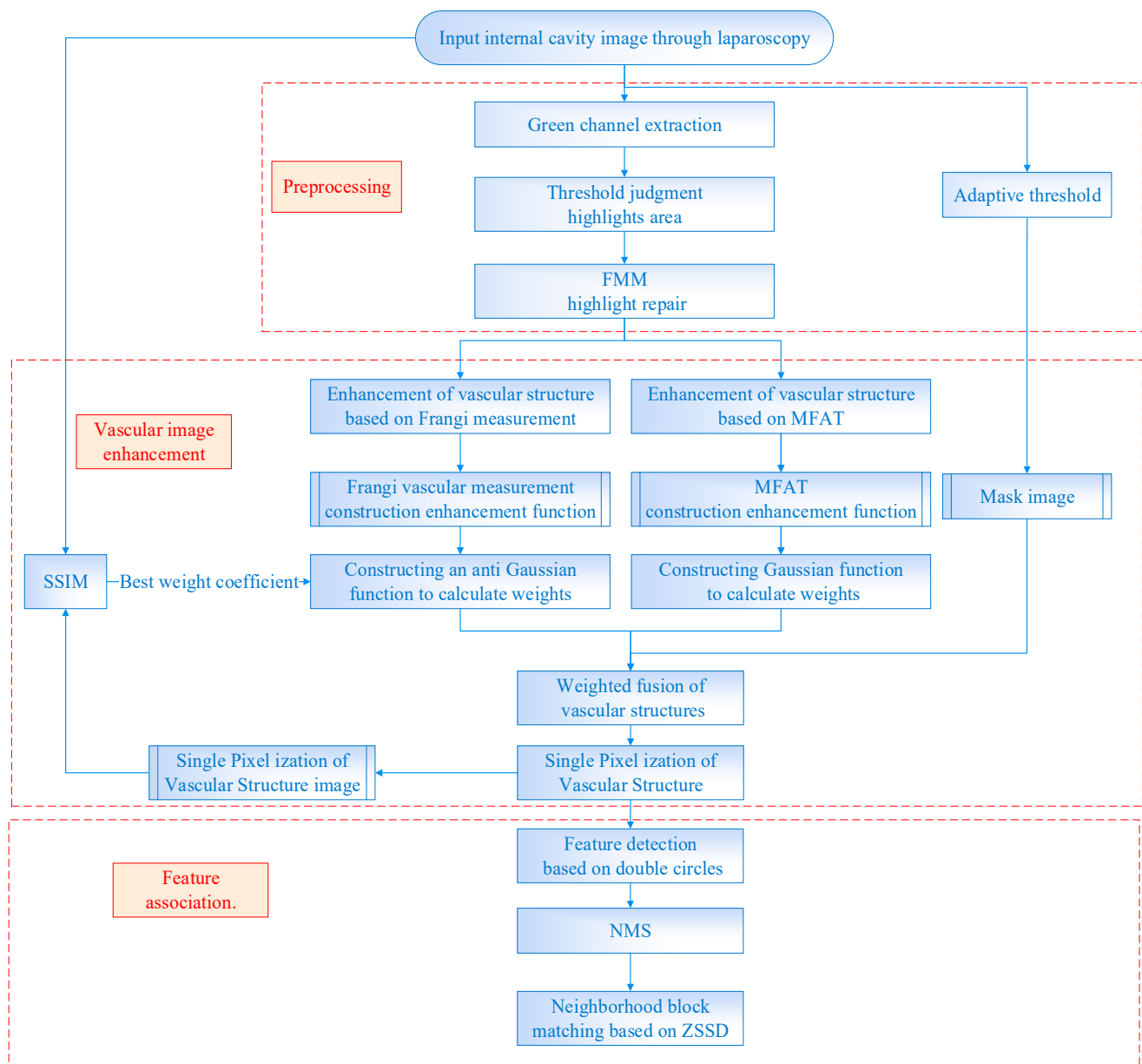


Figure 1. The flowchart of the proposed method.

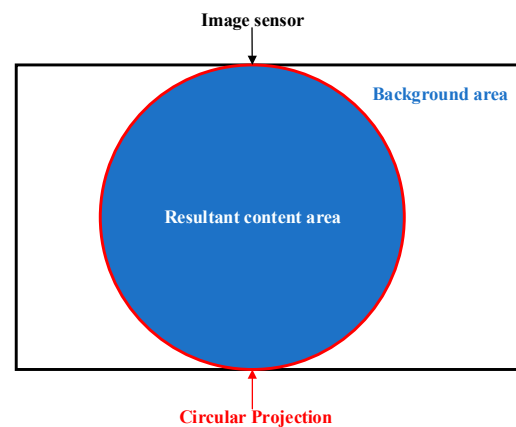


Figure 2. Laparoscopic image sensor area allocation.

For each pixel  $(x, y)$ , the window size centered on it is  $(w, h)$ , and the grayscale mean and median of all pixels in the window are calculated to obtain a binary threshold. As shown in Equation (1), pixels  $(x, y)$  that are greater than the threshold are assigned to the foreground or content area, otherwise they are the background or boundary area to obtain a masked image.

$$mask(x, y) = \begin{cases} 1 & I_m(x, y) > k \left( \frac{1}{w \cdot h} \right) \sum I(x, y) - (1 - k) \cdot median(I(x, y)), \\ 0 & otherwise \end{cases} \quad (1)$$

Among them,  $I_m(x, y)$  represents the binarization threshold of pixel points  $(x, y)$ ;  $\left( \frac{1}{w \cdot h} \right) \sum I(x, y)$  represents the average grayscale around a pixel  $(x, y)$ ;  $median(x, y)$  represents the median grayscale around a pixel  $(x, y)$ ; and  $k$  is the sensitivity coefficient used to control the threshold size. This binary mask image is applicable to various types of MIS images. We selected three typical environments from the Hamlyn data set and Vivo data set, as shown in Figure 3, where (a) represents the left image “0000000000” after inspection and correction in the Hamlyn data set Rectified01 (binocular) and its corresponding binary mask images, (b) represents the image “base” in the Vivo data set clip2 and its corresponding binary mask images, and (c) represents the image “base” in the Vivo data set clip3 and its corresponding binary mask images.

Among the three RGB channels in MIS images, the most prominent colors for blood vessels are red and green, while the background color is closer to blue. The contrast between the red and blue channels is very poor and noisy, and the corresponding channel pixels do not provide much information about blood vessels [38]. Therefore, the green channel provides the best contrast between blood vessels and the background, with blood vessels being the clearest.

Using the local variance threshold to automatically identify the highlight areas in the image, based on the special grayscale distribution and spatial distribution characteristics of the highlight areas, the threshold for each pixel point in the image is determined by calculating the variance of the surrounding pixels as follows:

$$T(x, y) = k \cdot var(R(x, y)) + C \quad (2)$$

where  $T(x, y)$  represents the threshold of the local neighborhood centered on  $(x, y)$ ,  $var(R(x, y))$  represents the pixel value variance of the local neighborhood centered on  $(x, y)$ , and  $k$  and  $C$  are scaling factors and constants, respectively, used to adjust the size and position of the threshold. Then, FMM is used to achieve highlight restoration, as shown in Figure 4.

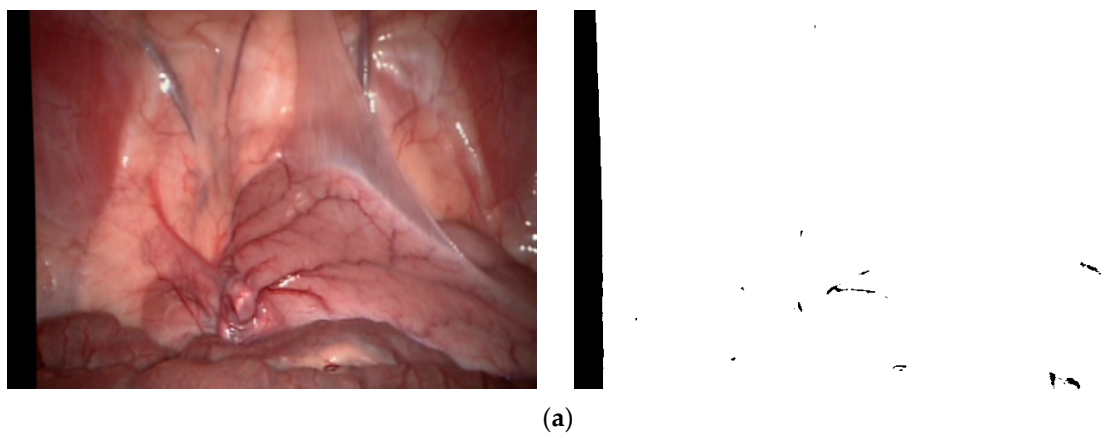
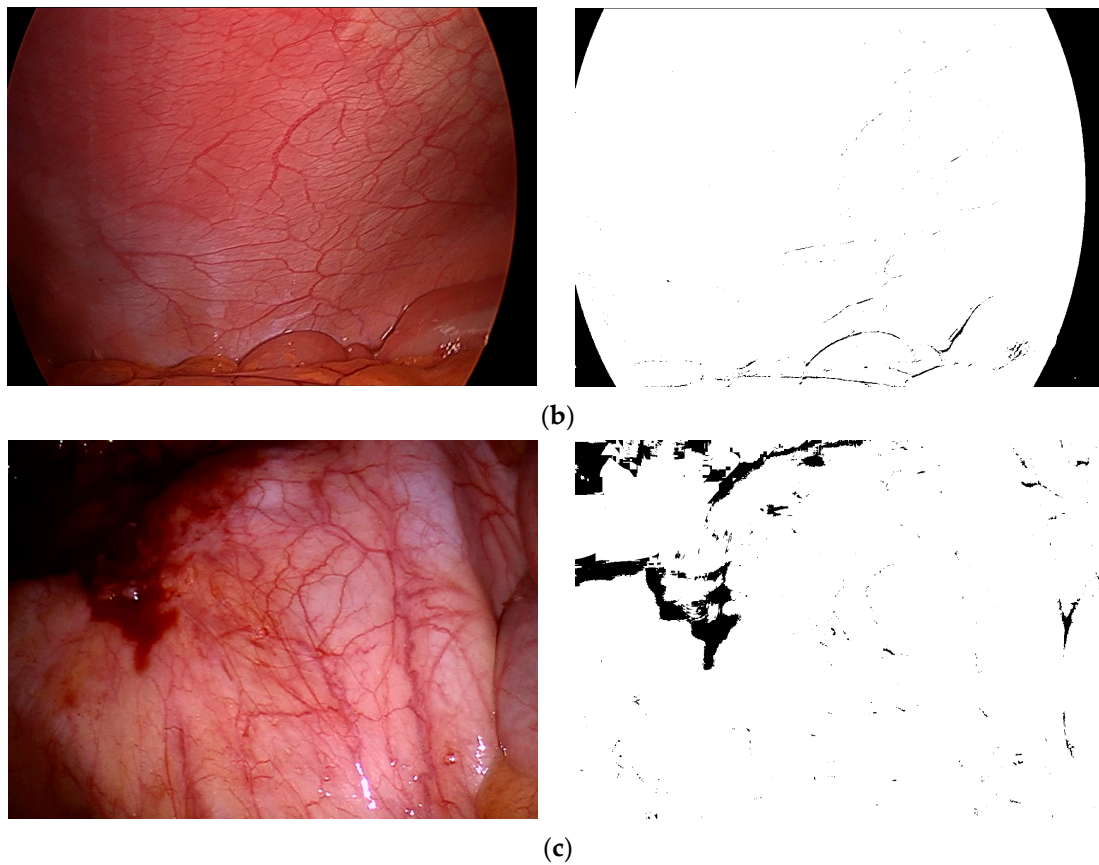
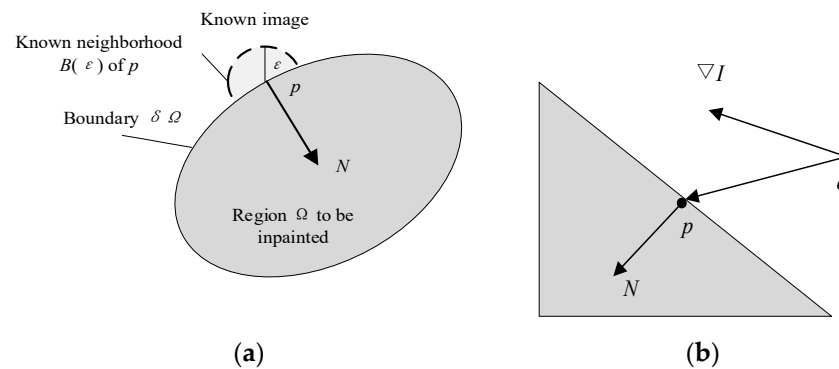


Figure 3. Cont.



**Figure 3.** Illustration of derived binary masks of images: (a) Rectified01(Hamlyn data set); (b) clip2 (Vivo data set); (c) clip3 (Vivo data set).



**Figure 4.** FMM schematic diagram: (a) repair area; (b) gradient direction.

The highlight area is used as the initial boundary condition, and new pixel values are calculated with neighboring points based on the weight of different distances to replace the values of adjacent point sets in the highlight area [39] as follows:

$$I_q(p) = I(q) + \nabla I(q)(p - q). \quad (3)$$

Among them,  $q$  is a point of  $B_\epsilon(p)$ . Due to the different influence of pixels with different distances within the neighborhood on the new pixel values of point  $p$ , it is necessary to set the respective weight for each pixel as follows:

$$I(p) = \frac{\sum_{q \in B_\epsilon(p)} w(p, q) [I(q) + \nabla I(q)(p - q)]}{\sum_{q \in B_\epsilon(p)} w(p, q)}. \quad (4)$$

$$w(p, q) = \text{dir}(p, q) \cdot \text{dst}(p, q) \cdot \text{lev}(p, q) \quad (5)$$

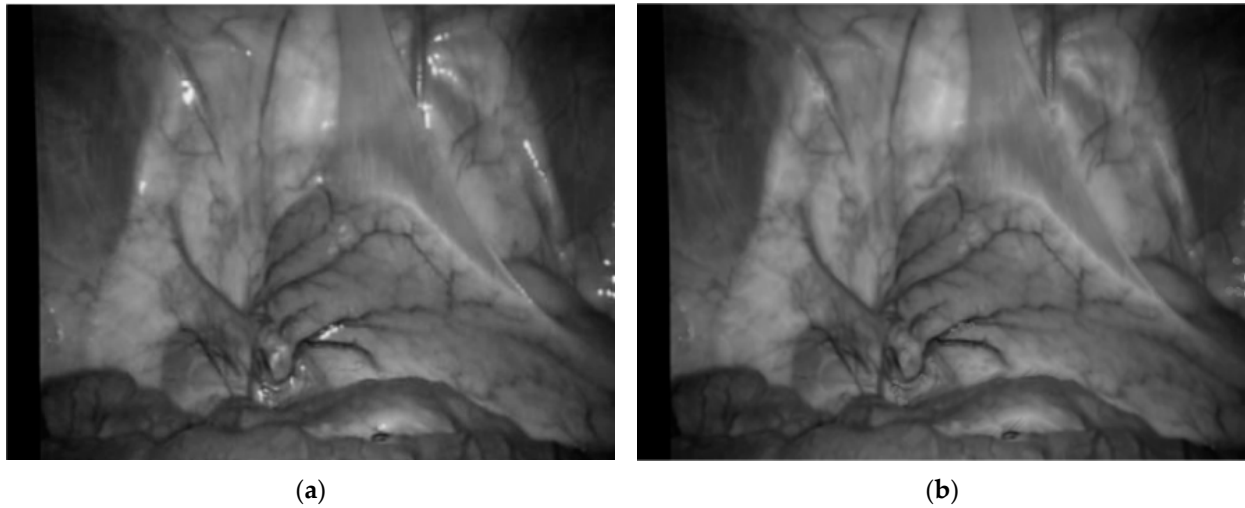
where  $w(p, q)$  is the weight function used to quantify the impact of each pixel in the neighborhood on new pixel values. The weight functions are shown in Equations (6)–(8)

$$\text{dir}(p, q) = \frac{p - q}{\|p - q\|} \cdot N(p) \quad (6)$$

$$\text{dst}(p, q) = \frac{d_0^2}{\|p - q\|^2} \quad (7)$$

$$\text{lev}(p, q) = \frac{T_0}{1 + |T(p) - Y(q)|} \quad (8)$$

Among them,  $d_0$  and  $T_0$  are distance parameters and horizontal parameters, usually taken as 1.  $\text{dir}(p, q)$  is the directional factor,  $\text{dst}(p, q)$  is the geometric distance factor, and  $\text{lev}(p, q)$  is the horizontal set distance factor. The preprocessing results are shown in Figure 5.



**Figure 5.** An example of highlight repairing: (a) original image; (b) result after repairing.

## 2.2. Vascular Structure Enhancement

### 2.2.1. Analysis of Vascular Features Based on the Hessian Matrix

The Hessian matrix is based on second-order Gaussian derivatives, where the first derivative represents the grayscale gradient and the second derivative represents the rate of change of the grayscale gradient. It can be calculated at different scales to enhance features similar to curves, which helps to distinguish specific shapes, such as circular, tubular, and planar structures. In an image, the chromaticity measurement is used to describe the image when the dark curve structure appears relative to the background. The Hessian matrix  $H$  is used to calculate the eigenvalues  $\lambda_1$  and  $\lambda_2$  pixel by pixel for  $L$ , which represent the anisotropy of image changes in the two eigenvector directions. By analyzing the symbols and sizes of Hessian eigenvalues, as shown in Table 1, the selective enhancement can be performed on local image structures independent of direction based on the shape of the structure and foreground and background brightness. The spotted structure tends to be isotropic, and the stronger the linearity, the more anisotropic the structure. This method can be used to distinguish the linear vascular structures in the inner cavity, as shown in Equation (9)

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} = \begin{bmatrix} \frac{\partial^2 L}{\partial x^2} & \frac{\partial^2 L}{\partial x \partial y} \\ \frac{\partial^2 L}{\partial y \partial x} & \frac{\partial^2 L}{\partial y^2} \end{bmatrix} \quad (9)$$

**Table 1.** The eigenvalues of the Hessian matrix.

| $\lambda_1$ | $\lambda_2$ | Directional Feature        |
|-------------|-------------|----------------------------|
| L           | L           | Background, no direction   |
| N           | N           | Noise, directionless       |
| L           | H−          | Tubular structure (bright) |
| L           | H+          | Tubular structure (dark)   |

Ideally, in a two-dimensional image, it is assumed that  $|\lambda_1| \leq |\lambda_2|$ , where  $H$  represents high eigenvalues,  $L$  represents low eigenvalues,  $N$  represents noise,  $+/-$  represents the sign of the eigenvalues, and  $\lambda_2$  represents a bright (dark) structure on a dark (bright) background. The tubular structure is a vascular target that enhances the human lumen.

### 2.2.2. Vascular Enhancement Based on the Frangi Measure Filtering

The classic Frangi filter is an enhanced filtering algorithm based on the Hessian matrix [7], where the value  $R_B$  is approximately zero when the image structure is a tubular structure; that is, a blood vessel. The Frangi vascular enhancement function uses response  $R_B$  to distinguish between speckled and tubular structures, and measures  $S$  to distinguish the background. The calculation is shown in Equations (10) and (11)

$$R_B = \left| \frac{\lambda_2}{\lambda_1} \right| \quad (10)$$

$$S = \sqrt{\lambda_1^2 + \lambda_2^2} \quad (11)$$

Then, construct an enhancement function based on  $R_B$  and  $S$  as shown in Equation (12)

$$F(x, y, \sigma) = \begin{cases} 0 & \lambda_1 \geq 0, \\ \exp(\frac{R_B^2}{2\beta^2}) \cdot (1 - \exp(\frac{-S^2}{2c^2})) & \text{otherwise} \end{cases} \quad (12)$$

where  $\beta$  is used to adjust the sensitivity of the distinguishing block and strip regions, and  $c$  is used to adjust the overall smoothness of the filtered image.

### 2.2.3. Vascular Enhancement Based on MFAT

The Frangi vascular enhancement function aims to suppress the components of circular structures, but not all vascular structures elongate. In the process of exploring three-dimensional vascular structure enhancement, image analysis is usually closely related to the formation of neural processes and dihedral differentiation. Jerman et al. analyzed multiple vascular enhancement functions and summarized the advantages and disadvantages of solving vascular structure enhancement problems based on different functions [11]. For two-dimensional images, the Frangi vascular enhancement function was simplified. Based on  $e^x \approx 1 + x$ , it was found that

$$F(x, y, \sigma) = 1 - \exp(\frac{-S^2}{2c^2}) \approx \frac{1}{2c^2} (\lambda_1^2 + \lambda_2^2) \quad (13)$$

Through analysis, it was found that the main reason for Frangi's function design, which is proportional to the size of the feature values, is to suppress noise in low-intensity and uniform-intensity image regions, where all feature values have low and similar sizes. This method usually relies on the values of eigenvalues, which leads to several problems as follows: (1) in a slender or circular structure with uniform strength, the eigenvalues are uneven, (2) the feature values vary with the intensity of the image, and (3) enhancement is uneven at different scales. Therefore, redefining characteristic values,  $-\lambda$  is for bright structures on a dark background, and  $\lambda$  is for dark structures on a bright background.

According to Meijering et al. [8], the Hessian matrix has been improved as shown in Equation (14)

$$H' = \begin{bmatrix} H_{11} + \alpha H_{22} & (1 - \alpha)H_{12} \\ (1 - \alpha)H_{12} & H_{22} + \alpha H_{11} \end{bmatrix} \quad (14)$$

Then, construct neuron measures as shown in Equations (15) and (16)

$$N = \begin{cases} \frac{\lambda_{max}}{\lambda_{min}} & \text{if } \lambda_{max} < 0, \\ 0 & \text{if } \lambda_{min} \geq 0, \end{cases} \quad \lambda_{max} = \max(\lambda_1, \lambda_2), \quad \lambda_{min} = \min(\lambda_{max}, \lambda'_1, \lambda'_2) \quad (15)$$

$$\lambda'_1 = \lambda_1 + \alpha \lambda_2, \quad \lambda'_2 = \lambda_2 + \alpha \lambda_1 \quad (16)$$

Among them,  $\lambda_{max}$  is the maximum eigenvalue and  $\lambda_{min}$  is the minimum eigenvalue, representing the normalized eigenvector and eigenvalues of  $H'$ . The above neuron measures suppress background intensity discontinuity points and darker linear structures that are immune to first-order derivatives ( $\lambda_{max}(x) \geq 0$ ). Meanwhile, the neuron measure is a scalable function that can be used as a scale parameter  $\sigma$  Gaussian kernel and is used to suppress noise. Therefore, Alhasson et al. proposed an improved diffusion tensor measure FAT based on to enhance the vascular structure in three-dimensional images [40]. We applied it to two-dimensional internal cavity images and introduced auxiliary techniques  $\lambda_3$ , which is defined as the three-dimensional shape of blood vessels in a two-dimensional image to obtain the regularized eigenvalues  $\lambda_p$  and its cutoff threshold  $\tau$ . Construction  $\lambda_{p1}$ ,  $\lambda_{p2}$  through the upper limit cutoff threshold is passed  $\tau_1$  and lower cut-off threshold  $\tau_2$ , respectively. And  $\lambda_{p1}$ ,  $\lambda_{p2}$  are used to adjust  $\lambda_3$

$$FAT_{\lambda}^{\sigma} = \sqrt{\frac{3}{2}} \sqrt{\frac{(\lambda_{p1} - \overline{D}_{\lambda})^2 + (\lambda_2 - \overline{D}_{\lambda})^2 + (\lambda_{p2} - \overline{D}_{\lambda})^2}{\lambda_{p1}^2 + \lambda_2^2 + \lambda_{p2}^2}} \quad (17)$$

$$\overline{D}_{\lambda} = \frac{\lambda_{p1} + \lambda_2 + \lambda_{p2}}{2}, \quad \lambda_p = \begin{cases} \lambda_3(x, \sigma) & \text{if } \lambda_3(x, \sigma) > \tau \max_x(\lambda_3(x, \sigma)), \\ \tau \max_x(\lambda_3(x, \sigma)) & \text{if } 0 < \lambda_3(x, \sigma) \leq \tau \max_x(\lambda_3(x, \sigma)), \\ 0 & \text{otherwise} \end{cases} \quad (18)$$

Correspondingly, to suppress background noise, construct a response  $R_{\lambda}^{\sigma}$  as shown in Equation (19)

$$R_{\lambda}^{\sigma} = \begin{cases} 0 & \text{if } \lambda_{\rho} > \lambda_{\rho} - \lambda_2 \vee \lambda_{\rho} \geq 0 \vee \lambda_2 \geq 0 \vee > \lambda_{\rho} - \lambda_2 \max_x(\lambda_{\rho} - \lambda_2), \\ 1 & \text{if } \lambda_{\rho} - \lambda_2 = \min_x(\lambda_{\rho} - \lambda_2), \\ 1 - \overline{MFAT}_{\lambda}^{\sigma} & \text{otherwise} \end{cases} \quad (19)$$

According to the concept of amplitude regularization, construct a multiscale enhancement function  $\overline{MFAT}_{\lambda}$  based on the maximum cumulative response through each scale  $\sigma$ , as shown in Equations (20) and (21)

$$\overline{MFAT}_{\lambda}^{\sigma} = \overline{MFAT}_{\lambda}^{\sigma-1} + \sigma \tanh(R_{\lambda}^{\sigma} - \sigma) \quad (20)$$

$$\overline{MFAT}_{\lambda} = \max_{\sigma}(\overline{MFAT}_{\lambda}^{\sigma}, R_{\lambda}^{\sigma}) \quad (21)$$

#### 2.2.4. Gaussian Weighted Fusion and Single-Pixel Vascular Structure

The Gaussian function has a higher weight near the center, and as the distance from the center increases, the weight gradually decreases. This paper combines two enhancement results by constructing a Gaussian weighting function to achieve compensation. The vascular enhancement result after Gaussian weighted fusion, denoted as  $I_{vessel}$ , is shown in Equation (22)

$$I_{vessel} = \frac{w_1}{w} \cdot I_{Frangi} + \frac{w_2}{w} \cdot I_{MFAT} \quad (22)$$

Among them,

$$w = w_1 + w_2 \quad (23)$$

$$w_1 = e^{-\left(\frac{(I - I_{Frangi})^2}{2 \cdot \mu^2}\right)}, w_2 = 1 - e^{-\left(\frac{(I - I_{MFAT})^2}{2 \cdot \mu^2}\right)} \quad (24)$$

Among them,  $w$  is the Gaussian weight;  $I_{Frangi}$  and  $I_{MFAT}$  are the Frangi enhancement results and MFAT enhancement results, respectively; and  $\mu$  is the parameter that controls the shape of the Gaussian function. Considering that there is no gold standard for the structure of the lumen blood vessels, the optimal weights for different types of images are obtained based on SSIM, which refers to the parameter with the best score between the vascular structure and the lumen image, thereby improving the generalization of fused image applications.

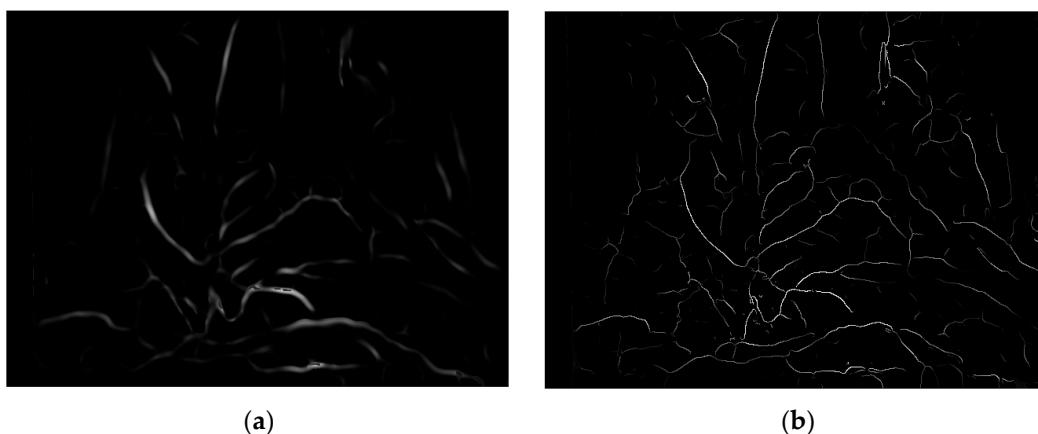
Blood vessels have a certain width, and the multipixel vascular features used for the vascular structure cannot accurately locate branch points and segments. Therefore, single-pixel vascular centerlines can be used to replace the original vascular morphology. Analyze the grayscale gradient changes of all pixels in the width direction of blood vessels, and use weighted operations to calculate the  $I_{ridge}$  of the vessel centerline, as shown in Equation (25)

$$I_{ridge}(x, y, \vartheta) = I_{vessel}(x, y, \vartheta) \cdot \frac{\text{abs}\{\text{sign}(\nabla I(x + u_2, y + v_2, \varepsilon)) - \text{sign}(\nabla I(x - u_2, y - v_2, \varepsilon))\}}{2} \quad (25)$$

where  $\nabla$  is the gradient operator,  $(u_2, v_2)^T = V_2$  and  $\varepsilon = 1.0$  is the pixel width. However, most of the blood vessel width obtained are two pixels. In order to obtain a more accurate single-pixel width ridge, it is necessary to obtain the local maximum value of each centerline pixel in the direction of the feature vector  $V_2$ , as shown in Equation (26)

$$I_{ridge}(\cdot) = \begin{cases} I_{ridge}(\cdot), & \text{if } I_{ridge}(x \pm u_2, y \pm v_2, \vartheta) < I_{ridge}(\cdot) \\ 0, & \text{otherwise} \end{cases} \quad (26)$$

where  $I_{ridge}$  represents ridge degree. As shown in Figure 6, (a) is the enhancement result of the Frangi enhancement method, and (b) is enhancement result of this paper method. The blood vessels after single-pixel transformation are thinner and have clearer branches compared to the original method blood vessel images in the human lumen.



**Figure 6.** Enhancement results of vascular structure enhancement: (a) the Frangi enhancement method; (b) the method proposed in this paper.

Algorithm 1 describes a Gaussian weighted fusion of dual-branch Frangi measures and MFAT for single-pixel enhancement of vascular structures.

**Algorithm 1**

Input: Inner cavity image  $I$ , preprocessed inner cavity image  $I_{pre}$ , masked image  $I_{mask}$

Output: Single-pixel vascular image  $I_{ridge}$

```

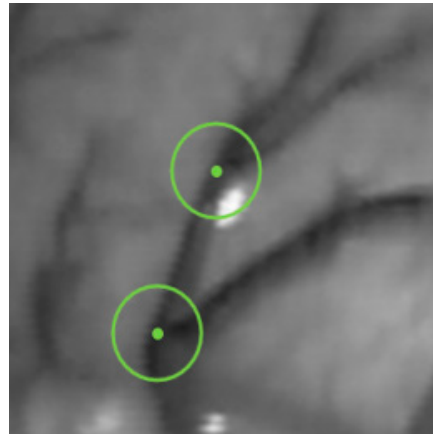
1: Branch 1 uses a blood vessel enhancement algorithm based on Frangi measure filtering on  $I_{pre}$ 
2: for each pixel  $p \in I_{pre}$  do
3:   for each scale  $\sigma$  do
4:     Calculate the Hessian matrix and its eigenvalues  $(\lambda_1, \lambda_2)$  and eigenvectors  $(I_x, I_y)$ 
5:     Calculate  $Direction$ ,  $R_B$ , and  $S$  based on eigenvalues and eigenvectors
6:     At scale  $\sigma$  Calculate the Frangi measure filtering result  $F(\sigma)$ 
7:   end for
8:   Extract the maximum response and corresponding filtering result  $\max(F(\sigma))$ 
9: end for
10: Combining  $I_{mask}$  to generate Frangi measure vascular enhancement images  $I_{Frangi}$ ,  $Direction$ 
11: Branch 2 using MFAT based vascular enhancement on  $I_{pre}$ 
12: for each pixel  $p \in I_{pre}$  do
13:   for each scale  $\sigma$  do
14:     Calculate the Hessian matrix and its eigenvalues  $(\lambda_1, \lambda_2)$  and  $Direction$ 
15:     Set  $\lambda_3 = \lambda_2$ . Construct new eigenvalues based on the eigenvalues  $\lambda_{p1}, \lambda_{p2}$ , and calculate  $FAT_{\lambda}^{\sigma}, R_{\lambda}^{\sigma}$ 
16:     At scale  $\sigma$  calculate and update MFAT results  $\overline{MFAT}_{\lambda}$ 
17:   end for
18:   Remove pixels  $\overline{MFAT}_{\lambda} < 1 \times 10^{-2}$ 
19:   Extract the maximum response and corresponding scale filtering result  $\max(\overline{MFAT}_{\lambda}(\sigma))$ 
20: end for
21: for every steps  $\mu \in (2, 4)$  do
22:   Establish a Gaussian model and calculate weights  $w_1$  and  $w_2$ 
23:   Weighted fusion of  $I_{Frangi}$  and  $I_{MFAT}$  to generate binary image  $I_{vessel}$ 
24:   Using single-pixel algorithm based on  $I$ ,  $I_{vessel}$ , and  $Direction$ 
25:   Gaussian Filter Smoothing Image  $I$ 
26:   Generate grid coordinate matrix  $[X, Y]$ 
27:   for each coordinate  $(x, y) \in [X, Y]$  do
28:     Calculate the feature direction gradient  $\nabla$  and inner product
29:     If there is a sign change in the gradient along the direction of maximum curvature, it is considered as a vascular ridge line  $mask\_ridge$ 
30:     Multiply the ridge mask with the vascular image to obtain the vascular centerline image  $I_{ridge}$ 
31:   if the single-pixel detection flag for blood vessels is true
32:     Interpolate and shift vascular images to obtain  $I_{Rshift1}$  and  $I_{Rshift2}$ 
33:     Compare to obtain a single-pixel ridge  $mask\_ridge2$  and its application
34:   end if
35: end for
36: Generate a single-pixel vascular image  $I_{ridge}$ 
37: Calculate the SSIM score between  $I$  and  $I_{ridge}$ 
38: end for
39: Select the corresponding optimal parameter  $\mu$  based on the maximum SSIM score
40: Calculate the optimal weights  $w_1$  and  $w_2$ 
41: Weighted fusion of  $I_{Frangi}$  and  $I_{MFAT}$  to generate binary image  $I_{vessel}$ 
42: Generate a single-pixel vascular image  $I_{Frangi}$  using a single-pixel algorithm based on  $I$ ,  $I_{vessel}$ , and  $Direction$ 

```

**2.3. Circle Detection of Vascular Branches**

A vascular branch point is a feature point with three or more vascular branch segments centered on it. Due to rotation and scaling, the vascular branch points can still be detected in intracavitary vessels with a single texture feature and strong robustness. Based on this definition and the principles of the FAST algorithm, a single-pixel wide blood vessel image

that has undergone image preprocessing is defined as a blood vessel branch point when there are three or more identical pixel values between the center candidate point and the point on the circumference. This method is called the circle detection of vascular branches (CDVB), and the intersection of the circumference with three or more blood vessels will cause a black and white change in the color of the pixels on the circumference, as shown in Figure 7.



**Figure 7.** Vascular branching points.

Although blood vessels have been pixelated, there is still a possibility of multiple pixels at the intersection of blood vessels and circles in the CDVB algorithm. In order to determine the accurate position of the intersection pixels, all pixel values along the circumference are counted, and the peak coordinates are taken as the intersection coordinates. Set a threshold to determine feature points, as shown in Equation (27)

$$N = \sum_{x \in \text{circle}(p)} |I_x - I_p| > \varepsilon_d \quad (27)$$

where  $P$  is a candidate detection target point with a pixel value of  $I_p$ , and  $r$  is the radius of the circle. Set a threshold  $\varepsilon_d$ , and subtract the pixel values on the circle from the pixel values at point  $P$  to obtain  $I_p - I_x$ . Compare this value with threshold  $\varepsilon_d$ . If there are  $N$  consecutive pixels with a difference greater than  $\varepsilon_d$  from  $I_p$ , then the candidate point will be judged as a feature point.

In order to enhance the recognition ability of microvascular branching points, reduce false detections, and enhance algorithm adaptability and flexibility, set different radius double circles on the candidate points for judgment. If one of the circles is successfully judged, it will be marked as a feature point and its coordinates and peak points will be recorded.

This method can accurately detect the branching points of blood vessels in the human lumen. However, there still exists uneven distribution of feature points, especially the accumulation of local feature points. There are two reasons for this type of accumulation: first, there are multiple effective corner points in a small area, and second, due to the width of blood vessels, it is highly possible to detect multiple feature points in a single effective corner area, with all but one being redundant. This situation not only increases the computational burden, but also poses a risk of reducing the accuracy of feature point matching.

For the above situations, NMS is introduced to solve the problem of uneven distribution of feature points [41]. Perform connectivity analysis based on the peak point of feature point  $P$  in the neighborhood, and obtain the score value of the score function  $V$  for

the connected component calculation. If  $P$  is the point with the highest score value in the neighborhood, it is kept:

$$V = \max \begin{cases} \sum (I_T - I_P) - \varepsilon_d & \text{if } I_T \in b, \\ \sum (I_P - I_T) - \varepsilon_d & \text{if } I_T \in d \end{cases} \quad (28)$$

where  $V$  represents the score,  $I_T$  represents the pixel value of the template point,  $b$  is the bright spot, and  $d$  is the dark spot.

#### 2.4. Neighborhood Block Branch Point Matching Based on ZSSD

Based on the feature point neighborhood block matching algorithm in PTAM, the sum of squared errors is calculated within the neighborhood range of the matching points, and the matching point pairs are determined by the value of the sum of squared errors. In the prematching graph  $S$ , take a small rectangular graph in  $S$  with an area of  $M \times N$ . The starting coordinate of the rectangular graph is the coordinate of the upper left corner  $(i, j)$ . The template that matches the small rectangular graph is found through the traversal method. Whether it matches or not needs to be determined by calculating the value of the sum of squared errors as follows:

$$D(i, j) = \sum_{s=1}^M \sum_{t=1}^N [S(i + s - 1, j + t - 1) - T(s, t)]^2 \quad (29)$$

Among them  $S(x, y)$  is the search image of size  $M \times N$ , and  $T(x, y)$  is the template image of  $M \times N$ .  $1 \leq i \leq m - M + 1$ ,  $1 \leq j \leq n - N + 1$ . The smaller the average absolute difference  $D(i, j)$ , the more similar it is. Therefore, just find the smallest  $D(i, j)$  to determine the position that can match the block.

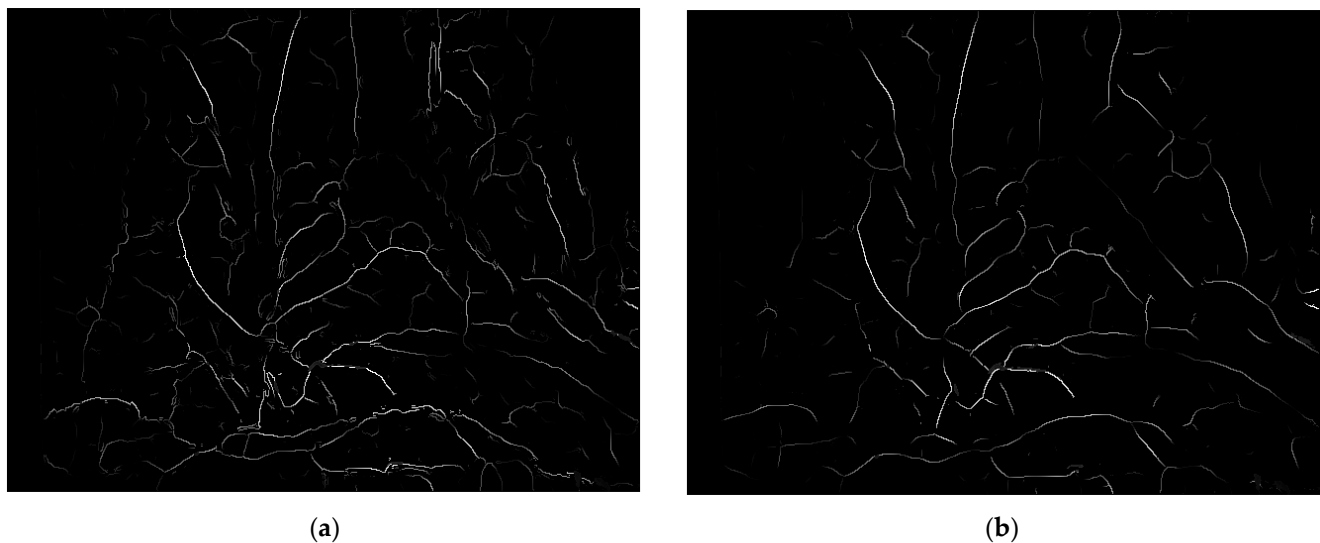
Based on the block search process, the information from previous frames is first saved by defining a search area for each map point in the current frame. For each  $P$  point saved in the frame, its corresponding point  $Q$  in the current frame can be obtained through homography mapping. The search area of  $P$  is a circular area centered on  $Q$ , with a radius of  $1/20$  of the image width. Second, select the feature points located within the neighborhood in the current frame, which are called neighboring points. Compare the  $21 \times 21$ -sized neighborhood blocks of each neighboring point with the same-sized neighborhood blocks of point  $P$ . Apply the affine warping obtained from the ground truth homography mapping to the local blocks of point  $P$  to accommodate changes in viewpoints, and calculate the ZSSD for each block. Finally, if the ZSSD value of a neighboring point is the minimum and is less than the predefined threshold of 0.02, it is considered to be matched. The condition for determining the correct pair of corresponding points is  $|Q' - Q| < 3.5$  pixel.

### 3. Results

In order to demonstrate the superiority of the algorithm proposed in this paper, the Vivo data sets were used, including abdominal wall and small flat tissue surfaces in the anterior part of the pelvis captured during laparoscopic colon surgery. The data set collection range allowed for global homology mapping, including a base image for matching with other images, manually labeled 20 well distributed feature point pairs, and ground truth homogeneous mapping. The experimental environment is Matlab R2019a, the running environment is Windows 10, and the computer configuration is Intel Core i5-5200U CPU @ 2.20 GHz. The computer brand is Asus, and the origin is Suzhou, China.

#### 3.1. Weighted Coefficient Analysis of Vascular Structure Enhancement

The analysis was conducted from both subjective and objective perspectives. When the weight was too small, there was severe distortion in the visual effect of single-pixel vascular structures. When the weight was too large, there was a significant loss of fine blood vessels. Experiments were conducted for  $\mu = 1$  and  $\mu = 5$ , as shown in Figure 8.



**Figure 8.** Weight parameter analysis: (a)  $\mu = 1$ ; (b)  $\mu = 5$ .

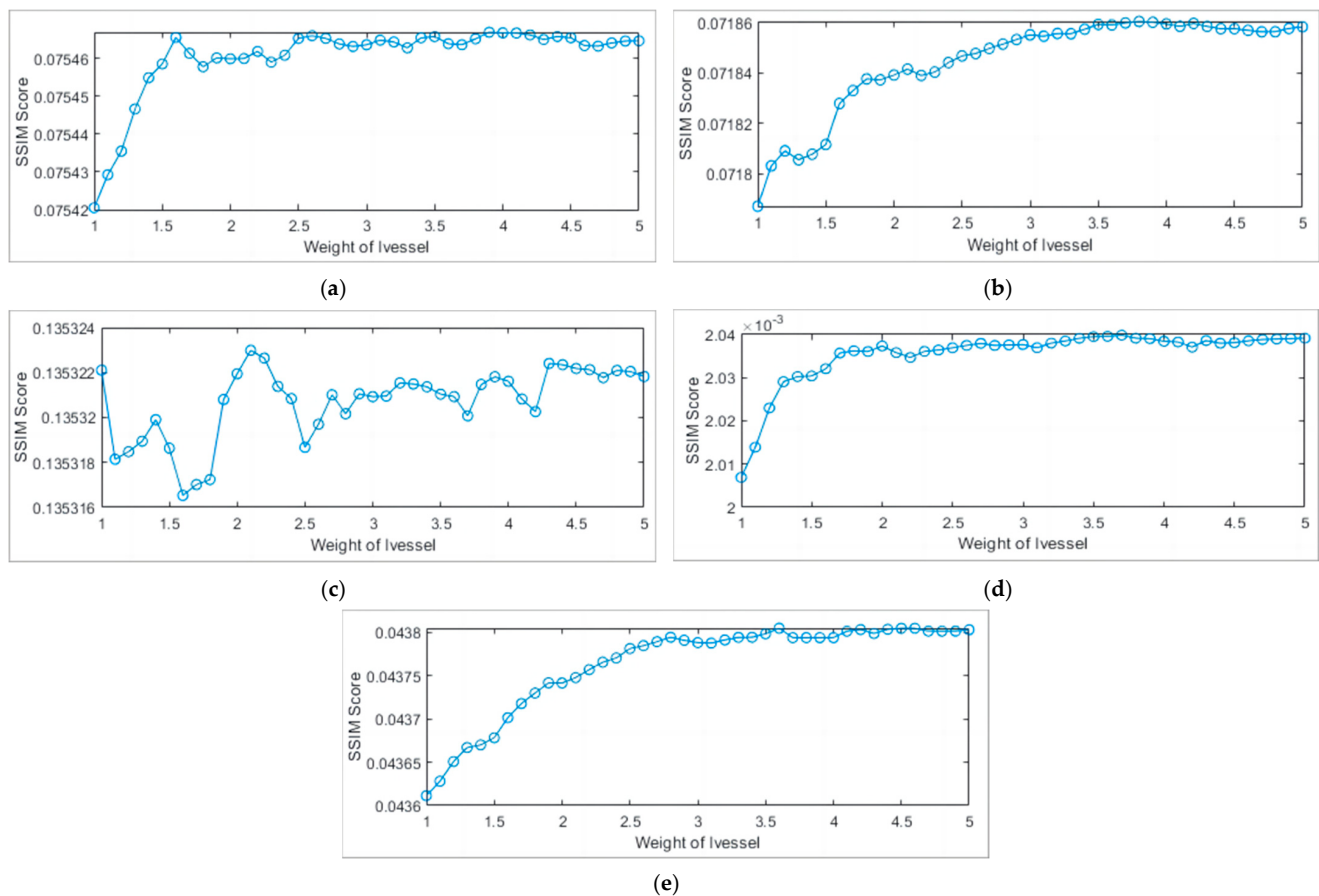
SSIM is used as an analytical indicator to reflect the similarity between the original image and the enhanced vascular structure image, as shown in Equation (30)

$$SSIM(I, I_{ridge}) = \frac{(2 \times \mu_I \times \mu_{I_{ridge}} + C_1)(2 \times \delta_{II_{ridge}} + C_2)}{(\mu_I^2 \times \mu_{I_{ridge}}^2 + C_1)(\delta_I^2 \times \delta_{I_{ridge}}^2 + C_2)} \quad (30)$$

where  $I$ ,  $I_{ridge}$  represents the inner cavity image and single-pixel vascular image,  $\mu_I$  and  $\mu_{I_{ridge}}$  is the average pixel value of  $I$  and  $I_{ridge}$ ,  $\delta_I^2$  and  $\delta_{I_{ridge}}^2$  is the variance of pixel values for  $I$  and  $I_{ridge}$ , and  $\delta_{II_{ridge}}$  is the covariance of pixel values for  $I$  and  $I_{ridge}$ .  $C_1$  and  $C_2$  are constants, ensuring that the denominator is not zero. As shown in Figure 9, although the significant difference in visual effects between the two results in low SSIM scores, the limited vascular structure information can help us evaluate the similarity between the fusion results and the original image, in order to better understand the quality of the fusion results and make necessary optimizations. This paper introduces the SSIM index to automatically obtain the current optimal parameters for Gaussian weighted fusion of images, improving the accuracy and reliability of fusion results. The highest SSIM scores for the five data sets are shown in the Table 2, and the optimal parameters are concentrated between 2 and 4. Therefore, in order to improve the algorithm speed, the parameter range is set to (2, 4).

**Table 2.** Maximum SSIM score for data set base images.

| Data Set | Best Weight Parameter for SSIM | Best SSIM Score |
|----------|--------------------------------|-----------------|
| clip1    | 3.9                            | 0.07547         |
| clip2    | 3.8                            | 0.07186         |
| clip3    | 2.1                            | 0.13532         |
| clip4    | 3.7                            | 0.00204         |
| clip5    | 3.6                            | 0.04381         |



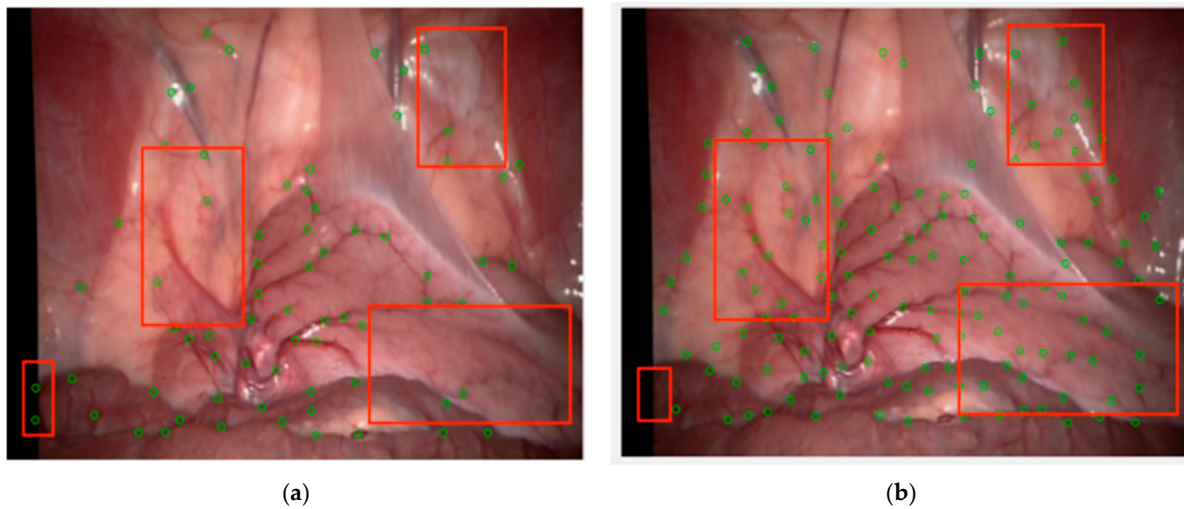
**Figure 9.** SSIM analysis: (a) clip1; (b) clip2; (c) clip3; (d) clip4; (e) clip5.

### 3.2. Analysis of Feature Points Extraction and Matching

Subjectively, taking images from the Hamlyn data set Rectified01 as an example, on the basis of enhancing the vascular structure, circles with a radius of 5 and 7 are respectively set for detection and extraction of branch points. Compared with the classical Frangi enhancement algorithm, the results are shown in Figure 10 and the green circles marked the feature points. It can be seen that the method proposed in this paper compensated for the structural measurement differences and uneven responses of the Frangi enhancement method on venous blood vessels and microvessels, suppressed the influence of boundary regions and highlights, solved the problem of generating a uniform response between different vascular structures, and effectively suppressed noise without affecting the connection as shown the regions in the red rectangles. The method proposed in this paper covers the preprocessing framework and the fusion optimization process. Although vascular enhancement is achieved through effective parallel processing, it still affects the overall running speed of the method. The average time of the RBCT is 0.95931 s, while the method in this paper increases it to 1.9515 s.

Objectively speaking, based on the feature point neighborhood block matching algorithm in PTAM, ZSSD is calculated for feature matching. Experiments are conducted on base images from five data sets and a random image. For a comprehensive evaluation, the number of branch point extracts, repeatability, average error, and variance are used as evaluation indicators. Repeatability is the percentage of points detected in images from different viewpoints. The average error is the average mean of all matching points. Variance represents the distribution of feature points, and the larger the variance, the wider the distribution range. The results are shown in Table 3 compared with the method proposed in this paper and RBCT and the adaptive FAST (AFAST) algorithm by Liu et al. [42]. The matching effect is shown in clip3 and clip4 as examples, as shown in Figure 11. The above

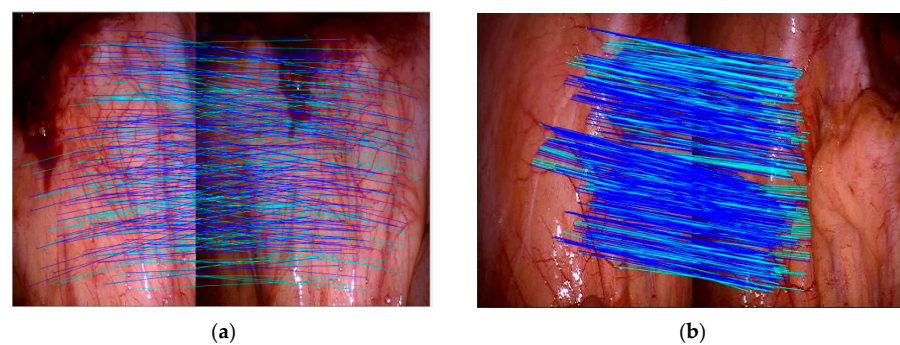
results indicate that the method proposed in this paper has significantly improved the number of extracted branch points, repeatability, and variance compared to the original method, and the average error has generally decreased.



**Figure 10.** Extraction of vascular branch points: (a) after Frangi enhancement; (b) after weighted fusion MFAT enhancement.

**Table 3.** Comparison of feature point extraction algorithms.

| Algorithms              | Data Set | Image I                 | Image J | Repeatability | Average Error | Variance |
|-------------------------|----------|-------------------------|---------|---------------|---------------|----------|
|                         |          | Number of Branch Points |         |               |               |          |
| AFAST                   | clip1    | 1368                    | 1409    | 0.162913      | 2.047352      | 0.785429 |
|                         | clip2    | 1493                    | 1628    | 0.188212      | 2.037728      | 0.840030 |
|                         | clip3    | 1023                    | 882     | 0.281159      | 1.984439      | 0.678100 |
|                         | clip4    | 801                     | 1019    | 0.134831      | 2.174954      | 0.629708 |
|                         | clip5    | 699                     | 545     | 0.294372      | 1.922701      | 0.866902 |
| RBCT                    | clip1    | 163                     | 177     | 0.361963      | 1.866042      | 0.718423 |
|                         | clip2    | 197                     | 237     | 0.441624      | 1.614276      | 0.682030 |
|                         | clip3    | 304                     | 333     | 0.579710      | 1.752783      | 0.755851 |
|                         | clip4    | 215                     | 277     | 0.302326      | 1.745650      | 0.717040 |
|                         | clip5    | 161                     | 117     | 0.440367      | 1.661636      | 0.597105 |
| OURS<br>(MFAT-<br>RBCT) | clip1    | 276                     | 278     | 0.433460      | 1.770989      | 0.670291 |
|                         | clip2    | 354                     | 378     | 0.497175      | 1.571009      | 0.691472 |
|                         | clip3    | 406                     | 376     | 0.668033      | 1.744668      | 0.783226 |
|                         | clip4    | 237                     | 315     | 0.320675      | 1.724399      | 0.739681 |
|                         | clip5    | 184                     | 158     | 0.452055      | 1.504729      | 0.690412 |



**Figure 11.** Illustration of matching vascular branching points: (a) clip3 (Vivo data set); (b) clip4 (Vivo data set).

#### 4. Discussions

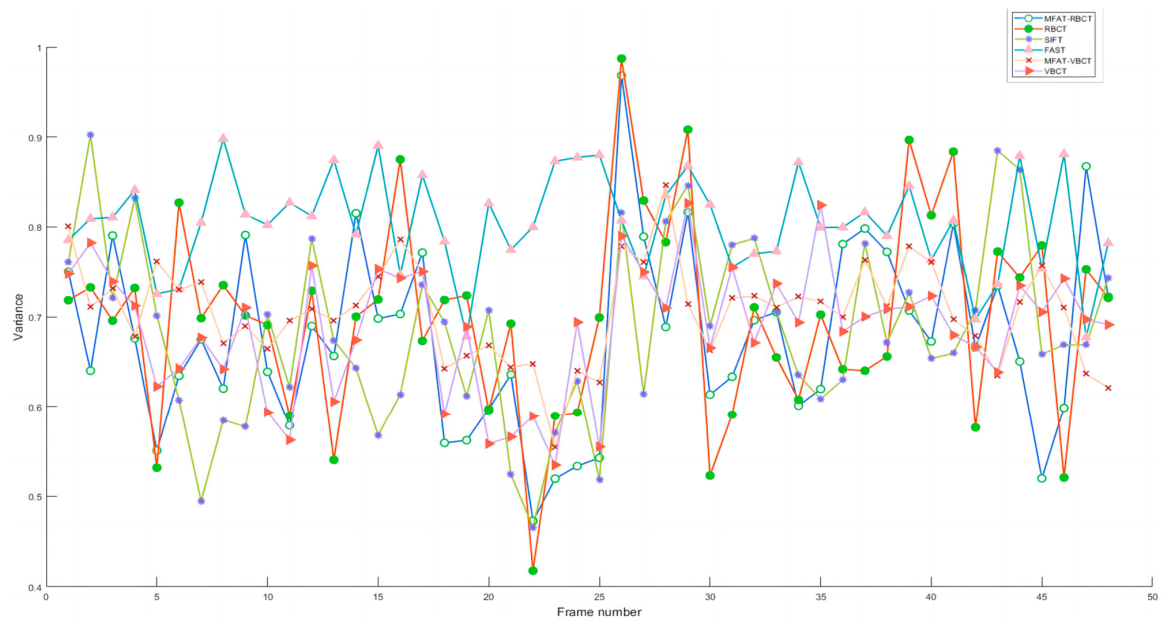
The prerequisite for reconstruction technology is accurate feature tracking, which is a well-researched topic in computer vision. Existing research indicates that there are significant issues in directly applying commonly used visual techniques to MIS due to the involvement of large-scale, free-form tissue deformation and constantly changing surgical scenes that are influenced by complex factors. In this case, the accuracy and efficiency of tracking largely depend on the detection of visual features, which need to exhibit high repeatability under image transformation and robustness under scene changes. This paper analyzes the differences between MIS images and images of the retina, neurons, etc., in response to the lack of robust and uniformly distributed features in the inner cavity space. For MIS images containing rich specular reflections, uneven lighting areas, smoke, and other factors, further research and design were conducted on a method for detecting and associating vascular features based on dual-branch weighted fusion vascular structure enhancement.

The preprocessing process analyzes the factors that lead to incorrect extraction of branch points in different internal cavity environments. The self-adaptive threshold is used to generate binary mask images and introduce them into the fusion processing stage. The threshold judgment and FMM method are used for highlight repair to suppress the influence of boundary and highlight regions. The visual effect of the experimental results shows the effectiveness of the method proposed in this paper.

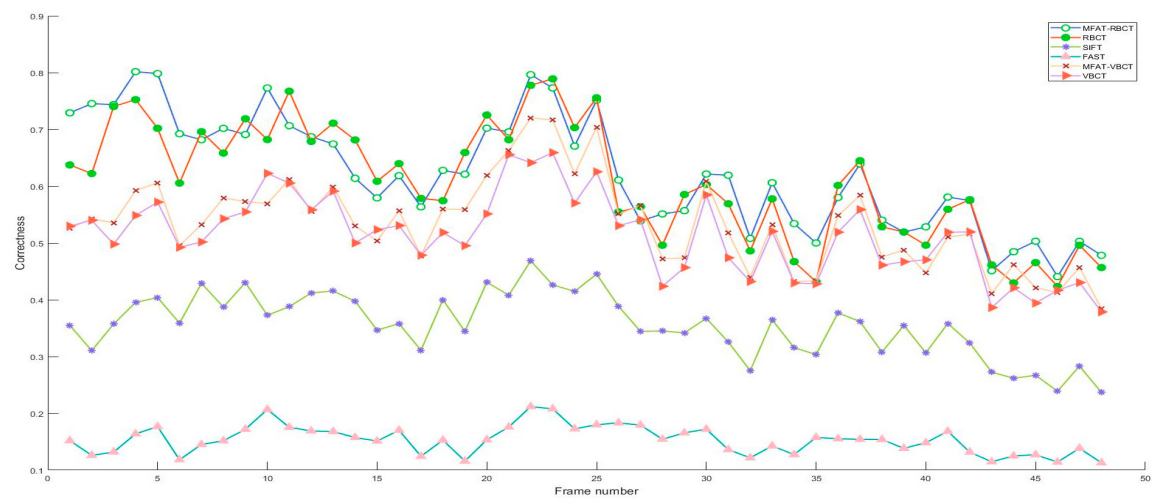
The enhancement process combines the advantages of two enhancement algorithms to design a Gaussian weighted fusion of dual-branch Frangi measure and MFAT for single-pixel enhancement of the vascular structure. The process introduced SSIM indicators as self-adaptive standards to improve algorithm accuracy and reliability, analyzed indicators to determine the weight parameter range applicable (2, 4) to the inner cavity environment, and improved the algorithm speed.

The feature extraction and associating process set up dual-circumference detection combined with the NMS algorithm to extract vascular feature points, and calculated ZSSD for branch point matching based on the feature point neighborhood block matching algorithm. Subjectively, the experimental results indicated that the algorithm proposed in this paper compensated for the structural measurement differences and uneven responses of the Frangi enhancement method to venous vessels and microvessels, providing effective structural information for later vascular feature extraction. Objectively, a detailed analysis was conducted on the methods proposed in this paper (MFAT-RBCT, MFAT-VBCT), and classical algorithms RBCT, SIFT, FAST, and VBCT. VBCT did not perform single-pixel processing on the vascular structure during the processing. The experiment is divided into frame rate analysis for all data sets and overall analysis for a single data set. As shown in Figure 12a, b, c, respectively, this study analyzed the variance, matching accuracy, and repeatability of six methods for all frame rates in data set clip1. The accuracy and repeatability of single-frame matching have been improved compared to RBCT and VBCT, and are significantly better than FAST and SIFT.

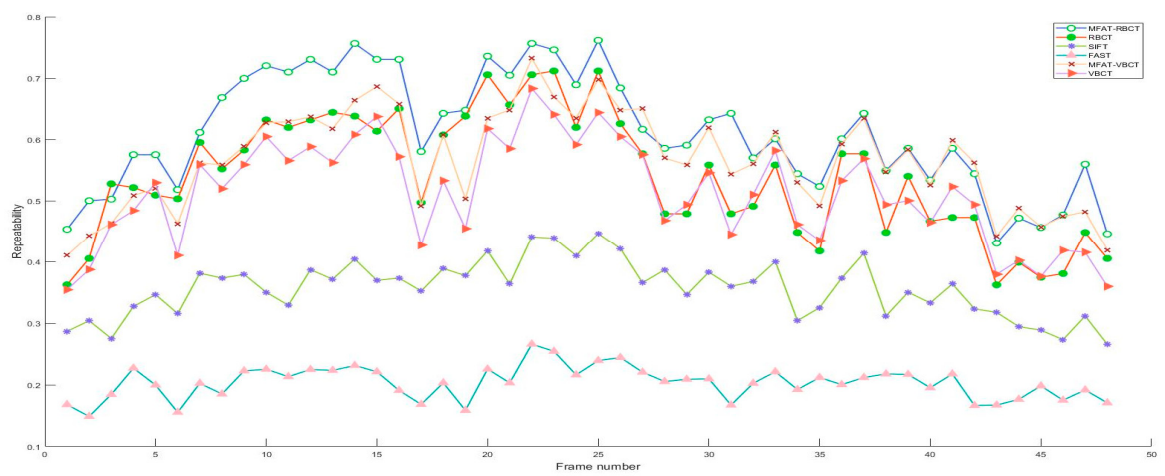
As shown in Figure 13a, b, and c, respectively, this study calculated the repeatability, matching accuracy, and number of branch points for five data sets and their synthesis. In the first four data sets, there was a significant improvement in repeatability, matching accuracy, and the number of branch points compared to RBCT and VBCT. Data set clip5 showed similar results to RBCT and VBCT due to poor image quality indicators. The results showed that compared with RBCT, the overall matching accuracy improved by 3.23%, the number of branch point detections increased by 13.08%, and the repeatability improved by 4.91%, which meant that our method has achieved good performance on five data sets, proving its generalization and robustness.



(a)

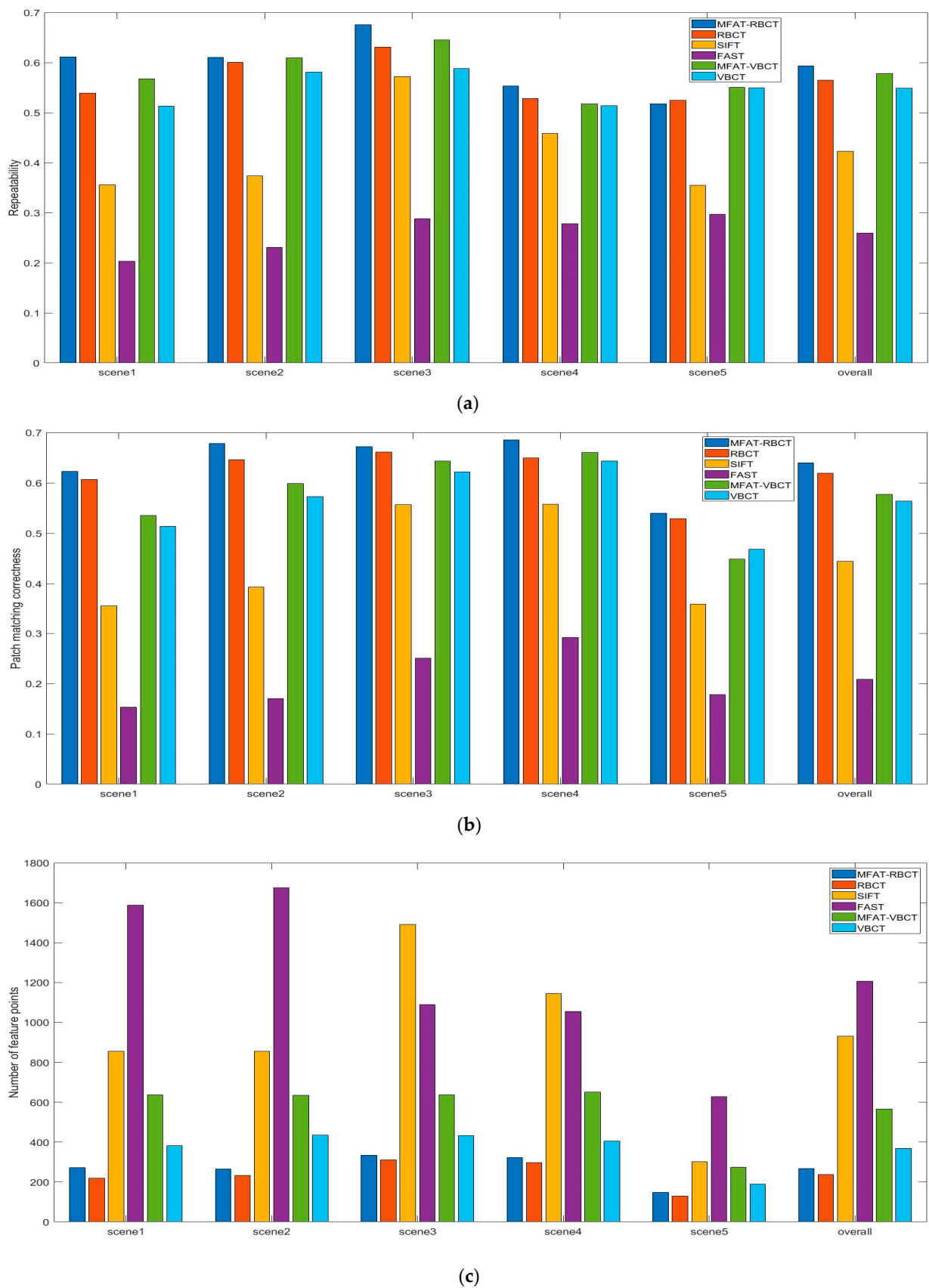


(b)



(c)

**Figure 12.** Indicator analysis: (a) variance of error of detectors; (b) correctness of matching; (c) repeatability over viewpoints.



**Figure 13.** Indicator analysis: (a) comparison of repeatability; (b) comparison of patch matching correctness; (c) comparison of number of feature points.

## 5. Conclusions

The texture features in the human lumen environment are mainly vascular morphology, and feature point extraction in this environment is based on blood vessels. Branch points of blood vessels can still be detected in lumen blood vessels with a single texture feature and strong robustness even under rotation scaling and other conditions. This study proposed a vascular feature detection and matching method based on dual-branch weighted fusion vascular structure enhancement. A universal preprocessing framework and Gaussian weighted fusion of dual-branch Frangi measure and MFAT for single-pixel enhancement of the vascular structure was designed. In addition, this study utilized dual-circle detection and NMS to extract features and calculate the ZSSD algorithm for neighborhood blocks to achieve feature correlation. The experimental results showed that compared to RBCT, the overall matching accuracy improved by 3.23%, the number of branch point detections increased by 13.08%, and the repeatability improved by 4.91%. Through quantitative analysis of multiple indicators, it has been proven that the method proposed in this paper can better handle vascular structures with robustness and generalization, and is a feature detection and matching method suitable for various lumen environments and can provide more robust feature information for subsequent 3D reconstruction.

After conducting more thorough tracking of the internal cavity features, we are exploring the possibility of incorporating this method into the visual odometry section to obtain three-dimensional information about the human internal cavity. The 3D virtual image is superimposed in real time on the imaging display to achieve an augmented reality display effect of the 3D MIS screen from the doctor's viewpoint. This provides feedback on the depth of the internal cavity and expands the doctor's surgical field of view. It also helps to reduce surgical trauma, alleviate the patient's pain, and improve the quality of the surgery. The technology seeks to promote its better application in the field of minimally invasive surgery and clinical trials.

**Author Contributions:** Conceptualization, H.W., K.X., Y.L., X.Y., Z.H. and A.W.; methodology, K.X. and Z.H.; software, K.X. and Z.H.; validation K.X. and Z.H.; writing—review and editing H.W., K.X., Y.L., X.Y. and A.W.; All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was funded by the high-end foreign experts' introduction program (G2022012010L), the Key Research and Development Plan Project of Heilongjiang (JD2023SJ19), the Natural Science Foundation of Heilongjiang Province (LH2023F034), the Reserved Leaders of Heilongjiang Provincial Leading Talent Echelon (2021), and the Major Science and Technology Projects of Zhongshan City in 2022 (2022A1020).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** [http://www.cse.usf.edu/~bingxiong/Vessel\\_Datasets.zip](http://www.cse.usf.edu/~bingxiong/Vessel_Datasets.zip) (accessed on 31 July 2014).

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Bernhardt, S.; Nicolau, S.A.; Soler, L.; Doignon, C. The status of augmented reality in laparoscopic surgery as of 2016. *Med. Image Anal.* **2017**, *37*, 66–90. [CrossRef]
2. Sampogna, G.; Pugliese, R.; Elli, M.; Vanzulli, A.; Forgione, A. Routine clinical application of virtual reality in abdominal surgery. *Minim. Invasive Ther. Allied Technol.* **2017**, *26*, 135–143. [CrossRef]
3. Allan, M.; Mcleod, J.; Wang, C.; Rosenthal, J.C.; Hu, Z.; Gard, N.; Eisert, P.; Fu, K.X.; Zeffiro, T.; Xia, W.; et al. Stereo correspondence and reconstruction of endoscopic data challenge. *arXiv* **2021**, arXiv:2101.01133.
4. Song, J.; Wang, J.; Zhao, L.; Huang, S.; Dissanayake, G. MIS-SLAM: Real-Time Large-Scale Dense Deformable SLAM System in Minimal Invasive Surgery Based on Heterogeneous Computing. *IEEE Robot. Autom. Lett.* **2018**, *3*, 4068–4075. [CrossRef]
5. Schneider, C.; Thompson, S.; Totz, J.; Song, Y.; Allam, M.; Sodergren, M.H.; Desjardins, A.E.; Barratt, D.; Ourselin, S.; Gurusamy, K.; et al. Comparison of manual and semi-automatic registration in augmented reality image-guided liver surgery: A clinical feasibility study. *Surg. Endosc.* **2020**, *34*, 4702–4711. [CrossRef] [PubMed]

6. Survarachakan, S.; Pelanis, E.; Khan, Z.A.; Kumar, R.P.; Edwin, B.; Lindseth, F. Effects of Enhancement on Deep Learning Based Hepatic Vessel Segmentation. *Electronics* **2021**, *10*, 1165. [\[CrossRef\]](#)
7. Frangi, A.F.; Niessen, W.J.; Vincken, K.L.; Viergever, M.A. Multiscale vessel enhancement filtering. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI'98: First International Conference, Cambridge, MA, USA, 11–13 October 1998; pp. 130–137.
8. Meijering, E.; Jacob, M.; Sarria, J.-C.F.; Steiner, P.; Hirling, H.; Unser, E.M. Design and validation of a tool for neurite tracing and analysis in fluorescence microscopy images. *Cytom. Part A J. Int. Soc. Anal. Cytol.* **2004**, *58*, 167–176. [\[CrossRef\]](#) [\[PubMed\]](#)
9. Jo, K.; Choi, J. Automatic Vessel Detection Technique in Laparoscopic Surgical Imaging Including Surgical Instruments. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 18–21 July 2018; pp. 774–777. [\[CrossRef\]](#)
10. Huang, Z.; Li, Q.; Hong, H.; Zhang, T.; Sang, N. Non-local Hessian-based weighted filter for 2D noisy angiogram image enhancement with cardiovascular tree structure preservation. In Proceedings of the Chinese Automation Congress (CAC), Xi'an, China, 30 November–2 December 2018; pp. 2129–2134.
11. Jerman, T.; Pernus, F.; Likar, B.; Špiclin, Ž. Enhancement of Vascular Structures in 3D and 2D Angiographic Images. *IEEE Trans. Med. Imaging* **2016**, *35*, 2107–2118. [\[CrossRef\]](#)
12. Peeters, T.H.J.M.; Rodrigues, P.R.; Vilanova, A.; ter Haar Romeny, B.M. Analysis of distance/similarity measures for diffusion tensor imaging. In *Visualization and Processing of Tensor Fields. Mathematics and Visualization*; Laidlaw, D., Weickert, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2009. [\[CrossRef\]](#)
13. Alhasson, H.F.; Alharbi, S.S.; Obara, B. 2D and 3D vascular structures enhancement via multiscale fractional anisotropy tensor. In Proceedings of the Computer Vision—ECCV 2018 Workshops. ECCV 2018, Munich, Germany, 8–14 September 2018; Lecture Notes in Computer Science. Leal-Taixé, L., Roth, S., Eds.; Springer: Cham, Switzerland, 2018; Volume 11134. [\[CrossRef\]](#)
14. Dash, S.; Verma, S.; Kavita; Khan, S.; Wozniak, M.; Shafi, J.; Ijaz, M.F. A Hybrid Method to Enhance Thick and Thin Vessels for Blood Vessel Segmentation. *Diagnostics* **2021**, *11*, 2017. [\[CrossRef\]](#)
15. Harris, C.; Stephens, M. A combined corner and edge detector. In Proceedings of the Alvey Vision Conference, Manchester, UK, September 1988; Volume 15, pp. 10–5244. [\[CrossRef\]](#)
16. Rosten, E.; Drummond, T. Machine learning for high-speed corner detection. In Proceedings of the Computer Vision—ECCV, Graz, Austria, 7–13 May 2006; pp. 430–443.
17. Lowe, D.G. Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [\[CrossRef\]](#)
18. Bay, H.; Ess, A.; Tuytelaars, T.; Van Gool, L. Speeded-up robust features (SURF). *Comput. Vis. Image Underst.* **2008**, *110*, 346–359. [\[CrossRef\]](#)
19. Calonder, M.; Lepetit, V.; Strecha, C.; Fua, P. Brief: Binary robust independent elementary features. In Proceedings of the Computer Vision—ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, 5–11 September 2010; pp. 778–792.
20. Giannarou, S.; Visentini-Scarzanella, M.; Yang, G.Z. Probabilistic Tracking of Affine-Invariant Anisotropic Regions. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 130–143. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Lin, B.; Sun, Y.; Sanchez, J.; Qian, X. Vesselness based feature extraction for endoscopic image analysis. In Proceedings of the 2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI), Beijing, China, 29 April–2 May 2014; pp. 1295–1298.
22. Udayakumar, E.; Santhi, S.; Vetrivelan, P. An Identification of Efficient Vessel Feature for Endoscopic Analysis. *Res. J. Pharm. Technol.* **2017**, *10*, 2633–2636. [\[CrossRef\]](#)
23. Klein, G.; Murray, D. Parallel tracking and mapping for small AR workspaces. In Proceedings of the IEEE and ACM International Symposium on Mixed and Augmented Reality, Washington, DC, USA, 13–16 November 2007; pp. 225–234.
24. Li, Y.; Zheng, W.; Liu, X.; Mou, Y.; Yin, L.; Yang, B. Research and improvement of feature detection algorithm based on FAST. *Rend. Lince-Sci. Fis. Nat.* **2021**, *32*, 775–789. [\[CrossRef\]](#)
25. Li, S. A review of feature detection and match algorithms for localization and mapping. In *IOP Conference Series: Materials Science and Engineering*; IOP Publishing: Bristol, UK, 2017; p. 231.
26. Davison, A.J. Active matching for visual tracking. *Robot. Auton. Syst.* **2011**, *57*, 1173–1187.
27. Barath, D.; Matas, J.; Noskova, J. MAGSAC: Marginalizing sample consensus. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 10197–10205.
28. Wu, Z.; Li, D.; Li, C.; Chen, Y.; Li, S. Feature Point Tracking Method for Visual SLAM Based on Multi-Condition Constraints in Light Changing Environment. *Appl. Sci.* **2023**, *13*, 7027. [\[CrossRef\]](#)
29. Liu, Y.; Wang, Y. A new ORB feature matching image algorithm based on Non-maximum suppression method and Retinal sampling model. In Proceedings of the 2021 International Conference on Control, Automation and Information Sciences (ICCAIS), Xi'an, China, 14–17 October 2021; pp. 729–734.
30. Puerto-Souza, G.A.; Mariottini, G.L. Wide-baseline dense feature matching for endoscopic images. In *Image and Video Technology*; Lecture Notes in Computer Science; Springer: Berlin/Heidelberg, Germany, 2014; Volume 8333, pp. 48–59.
31. Puerto-Souza, G.; Mariottini, G. Adaptive multi-affine (ama) feature matching algorithm and its application to minimally-invasive surgery images. In Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Francisco, CA, USA, 25–30 September 2011; pp. 625–633.

32. Chu, Y.; Li, H.; Li, X.; Ding, Y.; Yang, X.; Ai, D.; Chen, X.; Wang, Y.; Yang, J. Endoscopic image feature matching via motion consensus and global bilateral regression. *Comput. Methods Programs Biomed.* **2020**, *190*, 0169–2607. [\[CrossRef\]](#)
33. Li, R.; Pan, J.; Yang, Y.; Wei, N.; Yan, B.; Liu, H.; Yang, Y.; Qin, H. Accurate and robust feature description and dense point-wise matching based on feature fusion for endoscopic images. *Comput. Med. Imaging Graph.* **2021**, *94*, 102007. [\[CrossRef\]](#)
34. Ozyoruk, K.B.; Gokceler, G.I.; Bobrow, T.L.; Coskun, G.; Incetan, K.; Almalioglu, Y.; Mahmood, F.; Curto, E.; Perdigoto, L.; Oliveira, M.; et al. EndoSLAM dataset and an unsupervised monocular visual odometry and depth estimation approach for endoscopic videos. *Med. Image Anal.* **2021**, *71*, 1361–8415. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Ma, R.; Wang, R.; Pizer, S.; Rosenman, J.; McGill, S.K.; Frahm, J.M. Real-time 3D reconstruction of colonoscopic surfaces for determining missing regions. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2019, Shenzhen, China, 13–17 October 2019; Lecture Notes in Computer Science. Springer: Cham, Switzerland, 2019; Volume 11768. [\[CrossRef\]](#)
36. Budd, C.; Herrera, L.C.G.-P.; Huber, M.; Ourselin, S.; Vercauteren, T. Rapid and robust endoscopic content area estimation: A lean GPU-based pipeline and curated benchmark dataset. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **2022**, *11*, 1215–1224. [\[CrossRef\]](#)
37. Zhao, Y.; Wang, X.; Mingxing, J. Extraction method of gear interference image foreground region based on adaptive threshold. *J. Appl. Opt.* **2023**, *44*, 345–353. [\[CrossRef\]](#)
38. Upadhyay, K.; Agrawal, M.; Vashist, P. Unsupervised multiscale retinal blood vessel segmentation using fundus images. *IET Image Process.* **2020**, *14*, 2616–2625. [\[CrossRef\]](#)
39. Li, L.; Wu, H.Y.; Chen, Z. Depth image restoration method based on improved FMM algorithm. In Proceedings of the 2021 13th International Conference on Machine Learning and Computing—ICMLC 2021, Shenzhen, China, 26 February–1 March 2021; pp. 349–355.
40. Hansen, C.D.; Johnson, C.R. *Visualization Handbook*; Academic Press: Cambridge, MA, USA, 2011.
41. Teng, H.; Chatziparaschis, D.; Kan, X.; Roy-Chowdhury, A.K.; Karydis, K. Centroid Distance Keypoint Detector for Colored Point Clouds. In Proceedings of the 2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 2–7 January 2023; pp. 1196–1205.
42. Liu, Y.; Li, Y. Adaptive Fast Corner Detection Optimization Algorithm Based on Gray-Scale Mean. *Electron. Opt. Control* **2024**, *31*, 65–71+91.

**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.