Three-Dimensional Stitching of Binocular Endoscopic Images Based on Feature Points

Changjiang Zhou, Hao Yu, Bo Yuan, Liqiang Wang, and Qing Yang

Abstract: There are shortcomings of binocular endoscope three-dimensional (3D) reconstruction in the conventional algorithm, such as low accuracy, small field of view, and loss of scale information. To address these problems, aiming at the specific scenes of stomach organs, a method of 3D endoscopic image stitching based on feature points is proposed. The left and right images are acquired by moving the endoscope and converting them into point clouds by binocular matching. They are then preprocessed to compensate for the errors caused by the scene characteristics such as uneven illumination and weak texture. The camera pose changes are estimated by detecting and matching the feature points of adjacent left images. Finally, based on the calculated transformation matrix, point cloud registration is carried out by the iterative closest point (ICP) algorithm, and the 3D dense reconstruction of the whole gastric organ is realized. The results show that the root mean square error is 2.07 mm, and the endoscopic field of view is expanded by 2.20 times, increasing the observation range. Compared with the conventional methods, it does not only preserve the organ scale information but also makes the scene much denser, which is convenient for doctors to measure the target areas, such as lesions, in 3D. These improvements will help improve the accuracy and efficiency of diagnosis.

Keywords: endoscope; binocular matching; 3D stitching; registration; feature points

1. Introduction

Endoscopy has advantages such as high resolution and less trauma, in a wide range of applications in clinical diagnosis and treatment. Compared with traditional endoscopes, binocular endoscopes have three-dimensional (3D) imaging functions, which can provide surgical depth information. This helps doctors operate the endoscope accurately, efficiently and safely, and facilitates the 3D reconstruction of human organs [1,2]. The 3D reconstruction of organs, especially real-time dense 3D reconstruction, makes the picture more consistent with the real scene. It helps doctors judge the important anatomical structure and its spatial location accurately, improve the speed and safety of surgery significantly, and reduce the pain of patients, which is of great significance in clinical surgery [3,4].

Currently, the 3D reconstruction of binocular endoscopes mainly relies on a binocular matching algorithm [5]. The basic principle is to imitate the binocular stereo vision of the human eyes and then obtain the parallax value of the object point in the left and right views by the way of feature matching. Finally, build the model combined with binocular camera parameters. For example, Zhou et al. [6] extracted the blood vessels of the fundus images based on a binocular vision for stereo matching and realized the 3D reconstruction of the retinal blood vessel image. However, they are all based on single-view 3D reconstruction, and the limitations of the field of view will affect the accuracy and safety of doctors’ diagnoses and surgical operations.
The 3D reconstruction of the complete scene mainly relies on 3D stitching technology. Common methods include multi-view geometry and point cloud registration [7,8]. At present, many research teams have developed 3D reconstruction systems of endoscopic images based on multi-view geometry. For example, Mahmoud et al. [9] used the simultaneous localization and mapping (SLAM) method to map the abdominal cavity monocular image dynamically. Widya et al. [10] used structure from motion (SFM) to reconstruct gastric organs based on a gastroscopy video. This type of method is generally sparse reconstruction, with low accuracy and loss of scale information, which is not conducive to the observation and measurement of the reconstruction results in the later stage. Moreover, the method requires different viewing angles to be observed through a moving camera, which is difficult to operate in a narrow space such as human organs. In terms of point cloud registration, random sample consensus (RANSAC) is a matching algorithm based on the extraction of the 3D features of the point cloud [11]. Under normal conditions, the surface of organs such as the gastrointestinal tract is relatively smooth, and their texture features are insufficient to be used for feature matching by RANSAC. While iterative closest point algorithm (ICP), the most widely used point cloud registration method, obtains the optimal transformation matrix by constantly repeating the search. Although it is easy to understand and has a desirable result, it relies on the initial matrix heavily, which not only falls into a locally optimal solution but also requires huge computing resources [12].

In this study, we propose the 3D stitching method based on feature points. Through the detection and matching of the feature points of the endoscopic image and the registration of multiple point clouds, a complete dense 3D reconstruction of the stomach model is realized, which expands the observation range of the doctors and assists in the operation. At the same time, it can lay the foundation for later surgical navigation.

2. Methods

The specific process of the proposed 3D stitching method of endoscopic images based on feature points is demonstrated in Figure 1. First, the binocular endoscope is operated to obtain the left and right image sequences of the stomach model at different sites. Next, the semi-global matching algorithm (SGBM) is implemented to generate the disparity maps, from which the point cloud under each site can be obtained. Then they are pre-processed for outlier culling and down-sampling. At the same time, feature extraction and matching are performed on the adjacent left image sequences. The offset of the key-point in the X-axis and Y-axis can be obtained and from which the corresponding initial matrix can be calculated. Finally, the point clouds are registered and spliced through the improved ICP algorithm to achieve the 3D reconstruction of the entire stomach organ.

![Workflow of 3D stitching with binocular endoscopic images.](image)

**Figure 1.** Workflow of 3D stitching with binocular endoscopic images.

2.1. Binocular Endoscope Calibration

The internal and external parameters of the binocular camera were calculated by using Zhang Zhengyou’s checkerboard calibration method [13]. The black and white checkerboard used as the calibration board was shot in different poses within the working distance of the binocular endoscope, and the OpenCV library was used to calibrate the captured binocular images.

After the binocular calibration, 3D information can be calculated from two-dimensional images based on binocular stereo vision and the triangulation method. The principle of binocular stereo is shown in Figure 2. O_L and O_R are the optical centers of the left and right
cameras located on the same horizontal line, respectively, whose distance is the binocular baseline length. The world coordinate system $O$-XYZ is established with the optical center of the left camera as the origin, and the pixel coordinate system $o$-uv is established with the upper left corner of the left and right image planes as the origin. The coordinates of the object point $P$ in the world coordinate system are $X, Y, Z$. The vertical coordinates of the image points on the left and right image planes are equal, and the difference between the horizontal coordinates $u_0$ and $u_0$ is the disparity $d$. According to the triangle similarity relationship, the distance $Z$ from the object point $P$ to the camera can be calculated according to the triangulation formula, $Z = (f^*b)/d$, the $X$ and $Y$ coordinates of the object point $P$ can be calculated according to the camera parameters, the formula is as follows:

$$X = \frac{u - c_u}{f} \times Z$$

$$Y = \frac{v - c_v}{f} \times Z$$

![Figure 2. The geometry system of the binocular endoscope.](image)

The image plane of the binocular camera system is in the same plane and its focal length is the same under the condition of pole line correction. Thus, the reprojection matrix $Q$ of the binocular system represent the internal and external parameter information of the binocular camera and the definition is as follows:

$$Q \begin{bmatrix} x \\ y \\ d \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \\ W \end{bmatrix}$$

where $d$ is the disparity obtained by binocular matching; $x$ and $y$ are the two-dimensional coordinates in the pixel coordinate system; $X, Y, Z$ and $W$ are the corresponding unnormalized 3D coordinates and the normalized coefficient, respectively.

### 2.2. Binocular Image Acquisition and Matching

After the calibration, the binocular endoscope was operated to shooting the stomach model to obtain a total of 16 pairs of left and right images (1280 × 800) of the upper and lower halves of the stomach model. The captured image could cover the main area of the stomach model. Through the binocular matching algorithm, the disparity map was obtained to generate the corresponding point cloud. Among the many matching algorithms, SGBM has improved the cost calculation of the original semi-global matching (SGM) and added the post-processing part, which is faster than the global matching method (GC) and is more accurate than the local matching algorithm (SAD). Thus, the SGBM algorithm was selected for binocular matching [14].
Here is the procedure of SGBM. First, the left image is filtered along the horizontal direction. Then Birchfield–Tomasi algorithm is used to calculate the cost of both the left and right images. Using the multipath constraint aggregation method to calculate the effect of minimizing the global energy function, the corresponding minimum generation value in the left and right graphs can be obtained as the matching point. At last, disparity refinement is carried out, including confidence detection, subpixel interpolation and left–right consistency detection.

2.3. Point Cloud Preprocessing

Due to the texture-less, weak details and other unsatisfying areas in the endoscopic image, the disparity maps obtained by the SGBM matching algorithm still have mismatching errors, resulting in a lot of outliers in the generated point cloud [15]. These noises not only affect the visualization but also cause interference in the point cloud registration process. Therefore, it is necessary to eliminate outliers in the point cloud.

We used radius filtering to eliminate outliers. This method assumed that each valid point in the original point cloud contains at least a certain number of points in the specified radius neighborhood. The points in the original point cloud that meet the assumptions were regarded as normal points and kept, otherwise, they were regarded as noise points and removed. Since the noise deviation caused by the mismatch was very large, the method had a good removal effect on outliers.

A large amount of noise with a fixed value of the disparity was generated due to non-uniform illumination and other reasons, which was converted into a point cloud and concentrated in a specific area. Thus, the final point cloud used for stitching could be obtained by splitting the point cloud to intercept point cloud areas within the correct depth range.

The resolution of the endoscope image used in the experiment was 1280 × 800, so the number of points of point cloud generated from the disparity map reached millions. To improve the speed of registration while maintaining the shape characteristics of the point cloud, the point cloud was implemented with a voxel grid down-sampling process. By setting the voxel parameter, the current grid was represented by a bit of center of gravity in the set edge-length grid to reduce the stitching complexity.

2.4. Feature Detection and Matching

The feature point method is a common method for camera motion estimation. Feature points are composed of two parts: key points and descriptors. Key points refer to the position of the feature point in the image, and some also have information such as orientation and size. The descriptor is usually a vector that describes the information of the pixels around the key point. Commonly used feature detection algorithms include SURF, SIFT, ORB, etc. [16]. ORB is not scale-transformable and can only be applied to scenes that are directly photographed. The advantage of the SIFT algorithm is that the features are stable, and it remains invariant to rotation, scale transformation, and brightness. It also has a certain degree of stability to viewing angle transformation and noise. The disadvantage of SIFT is that the real-time performance is not satisfying and that the ability to extract feature points of smooth edges is weak. SURF is an improvement of SIFT, which uses a more efficient way to complete feature extraction and description to improves the calculation speed and robustness [17].

The point cloud of each site was generated based on the disparity map corresponding to the left image. Therefore, for adjacent two-frame peeping images, feature detection and matching were performed through the left images to estimate the movement of the camera. Considering accuracy, robustness and calculation speed, as well as the specific scene of the endoscope, we used the SURF feature method and the specific process is as follows:

- Construct the Hessian matrix to generate all the points of interest for feature extraction.
- Construct a scale space. Generate O groups of images with different scales in the L layer in the image pyramid. Two quantities (O, L) form the scale space of the Gaussian
pyramid. Given a set of coordinates \((O, L)\), an image in the Gaussian pyramid is uniquely determined to solve the problem described by the image at all scales.

- **Location of feature points.** By comparing each pixel point processed by the Hessian matrix with 26 points in the neighborhood of the two-dimensional image space and scale space, the key points can be initially located. Then filtering out the key points with weak energy and the key points of incorrect positioning to screen out the final stable feature points.

- **The main direction distribution of feature points.** In the circular neighborhood of the feature points, counting the sum of the horizontal and vertical haar wavelet features of all points in the 60-degree sector, and then rotating the sector at intervals of 0.2 radians and again counting the haar wavelet feature values in the area, finally taking the direction of the largest sector as the main direction of the feature point.

- **Generate feature point descriptors.** Taking a \(4 \times 4\) rectangular area block around the main direction of the feature point, and counting the horizontal and vertical haar wavelet features of 25 pixels in each area block. Each haar wavelet feature has 4 feature vectors and finally, a 64-dimensional feature descriptor can be obtained.

- **Feature point matching.** The matching degree is determined by calculating the Euclidean distance between two feature points. The shorter the Euclidean distance, the better the matching degree between the two feature points.

### 2.5. Multiple Point Cloud Registration

In this study, the multiple point clouds registration method was used to register the point cloud sequence for 3D stitching. The former point cloud of two adjacent sites was taken as the source point cloud, and the latter point cloud was taken as the target point cloud. ICP was used for registration. The basic idea of the ICP algorithm is to approximate the two closest points from two clouds as the same point. For two given point clouds \(P\) and \(Q\), the initial positional relationship is \([R_0 | t_0]\). Select any point of the point cloud \(P\) in the initial pose relationship \(p_i\) and the nearest point \(q_i\) in the point cloud \(Q\) as a matching point pair to establish an error function, and obtain the best value of \(R\) and \(t\) by making the error function reach the minimum. The above process is one round of iteration. Combine the obtained \(R\) and \(t\) into a new position relationship \([R_1 | t_1]\), update the position relationship of the point cloud, continue to repeat the above process, and finally achieve one of the two conditions of convergence of the error function or reaching the upper limit of the number of iterations. For the first pair of points, the error can be expressed as:

\[
ei = pi - (R pi' + t)
\]

Therefore, using the least-squares method, the error function can be expressed as:

\[
\min_{R, t} J = \frac{1}{2} \sum \left\| (pi - (R pi' + t)) \right\|^2
\]

The initial transformation matrix will affect the speed and accuracy of ICP registration greatly. Generally, the epipolar geometry method is used to recover the camera motion between two frames through the correspondence between two-dimensional image sites. In this process, \(t\) will be normalized, which directly leads to degree uncertainty, making the final reconstruction unable to carry out the 3D measurement. In this study, the initial transformation matrix was generated by the result of feature matching. According to the position changes of the matching points in the left view of the two adjacent fields of view, the average values of the offset in the X-axis and Y-axis were calculated as the initial value of the 3D transformation matrix in the translation matrix part. Since there were two adjacent sites, the transformation size of the point clouds in the X-axis and Y-axis was much larger than the change in the Z-axis and the rotation matrix.

For each pair of point clouds, two thresholds of size were set to make two registrations of coarse and fine to improve the calculation efficiency under the premise of ensuring that
the point cloud registration had a better result. The rotation part in the transformation matrix was non-linear, resulting in the ICP solution being essentially a non-linear least squares problem. Considering that the movement in a short time was very small and the rotation angle was about 0, the ICP solution was approximately converted to a linear least-squares problem.

After obtaining the pose graph, that was, the point cloud nodes and the edges of the transformation matrix containing the point cloud registration, the graph was optimized through the Bundle Adjustment algorithm [18] to reduce the accumulation of pose estimation errors during the registration process. Finally, the 3D stitching of the endoscopic images was completed, and the 3D reconstruction of the entire stomach model was realized.

3. Results

3.1. Results of System Construction and Calibration

In this paper, we established a 3D stitching experiment system for binocular endoscopes, and the structure of the system is shown in Figure 3a. The prototype system mainly includes the binocular gastroscope body, the LED light source, the signal acquisition circuit, and the image processing workstation. The scope probe (Figure 3c) contains the binocular camera system, the water and air channel, the surgical instrument channel, and the lighting system. The camera parameters are as follows: the focal length is 1059.6 pixels, the baseline distance is 5.9 mm, the CMOS size is 1.75 μm × 1.75 μm, and the camera principal point coordinates $c_x$ and $c_y$ are 633.6 pixels and 367.1 pixels, respectively. A computer with an Intel Core i5-8265U with 1.6 GHz and 1.8 GHz processors and 8GB RAM is used in this research. The algorithm is realized by PCL1.9.1 [19] in Microsoft Visual Studio 2017 in Windows 10.

![Figure 3. Three-dimensional stitching system for a binocular endoscope. (a) System diagram. (b) Binocular endoscope. (c) The tip of the binocular endoscope.](image)

When calibrating the binocular camera, a black-and-white checkerboard with a 12 mm square edge distributed by $9 \times 6$ was used as the calibration board. Within the working distance of the binocular endoscope, 15 sets of checkerboard images of different positions were taken. Finally, the resulting reprojection matrix $Q = \begin{bmatrix} 1 & 0 & 0 & -633.6 \\ 0 & 1 & 0 & -367.1 \\ 0 & 0 & 1 & 1059.6 \\ 0 & 0 & 0.17 & 0 \end{bmatrix}$. According to the reprojection matrix, we calculated that the calibration error was 0.35 pixels, which
was within the allowable error range. Based on the above results, the scope of endoscopic observation was 20–100 mm.

3.2. Results of Point Cloud Preprocessing

In the process of removing outliers with radius filtering, according to the current image size, the outliers less than 16 points in the sphere with a radius less than 15 were defined as outliers to be removed. As shown in Figure 4, the result of point cloud preprocessing after outlier removal is displayed, in which Figure 4a is the original point cloud, Figure 4b is the point cloud processed by radius outliers, and Figure 4c is the point cloud segmentation based on Figure 4b to remove the void part with depth value within a certain range. It can be seen in Figure 4 that after the outlier removal preprocessing, a large number of holes and mismatched points are removed, and the processed point cloud can reflect the 3D structure of the endoscopic image.

![Figure 4. Result of point cloud outlier removal. (a) Original point cloud. (b) Radius outlier removal. (c) Result of point cloud segmentation.](image)

In voxel down-sampling, the voxel parameter was set to 5, and the average point cloud size was reduced from 1,000,000 to 100,000. Table 1 shows the speed comparison before and after point cloud preprocessing. Time of reading refers to the reading time of the two point clouds in adjacent positions, and time of registration refers to the registration time of these two point clouds. It evidences that after pre-processing operations such as hole removal and down-sampling, especially after down-sampling, the time for point cloud registration is reduced significantly, and the real-time stitching is improved. Figure 5 demonstrates the registration results of adjacent point clouds before and after point cloud preprocessing. Figure 5a–c correspond to the original point cloud, point cloud with outlier removal, point cloud with down-sampling registration results, respectively. There are obvious mismatches in the red area, and the registration accuracy is improved after eliminating the mismatches and holes.

|                  | Time of Reading(s) | Time of Registration(s) |
|------------------|--------------------|-------------------------|
| Original point cloud | 1.02               | 72.33                   |
| Outlier removal   | 0.84               | 63.61                   |
| Outlier removal and down-sampling | 0.17 | 1.03                   |

Table 1. Comparison of registration time before and after point cloud preprocessing.
3.3. Results of Feature Detection and Matching

Figure 6 indicates the result of feature matching in the left view of two adjacent sites, where the red dots are the detected feature points, and the blue line represents the matched feature point pair. In Figure 6, the blue lines are mostly parallel lines of equal length, indicating that the movement in the X-axis and Y-axis are mainly manifested in two adjacent frames of images. The average time of feature detection and matching is 0.689 s.

3.4. Results of 3D Stitching

After the point cloud sequence is registered, the 3D reconstruction result of the binocular endoscopic image of the complete stomach model is shown in Figure 7. Figure 7a is a plan view of the stomach model, and Figure 7b is the result of the 3D stitching of the point cloud of the stomach model. In Figure 7b, the results of the 3D stitching can already display the characteristics of the main regions of the stomach model, and various textures on the model such as stomach wall folds can be displayed clearly, which is conducive to the diagnosis of the doctors. Figure 7c is the result of the field of view change. The red box is the field of view of a single image, which is 29.05°. After 16 images are spliced, the field of view is 2.20 times the original, and it achieves 63.91°, which is greatly increased. The observation range is convenient for doctors to operate.

To evaluate the accuracy of point cloud stitching, the result was registered with the scanned point cloud of the stomach model. The white part in Figure 7d is the point cloud scanned by the scanner, which can be regarded as the true value. It evidences that the two are well-matched. To know the accuracy, the root mean square error (RMS) and the relative error (RE) are computed via Equations (4) and (5). Where \(d_{o_i}\) and \(d_{s_i}\) are the depth of the one point in the reconstructed point cloud and scanned point cloud, respectively. \(d_m\) is the mean of the model surface. The accuracy provided by the binocular system for the model is \(RMS = 2.07\) mm, and the relative error is 3.18%.

\[
RMS = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (d_{o_i} - d_{s_i})^2}
\]  
\[(6)\]
\[ RE = \frac{RMS}{d_m} \]  

(7)

Figure 7e is the visual result of the two registration, and the color indicates the absolute value of the distance between the stitching point cloud and the scanning point cloud. Most of the areas are in blue, indicating that the binocular point cloud has retained the 3D features of the model after stitching.

To evaluate the research method further, the SFM method was used to perform 3D reconstruction of the same endoscopic image sequence, and the comparison of the obtained results is shown in Figure 8. The number of point clouds in the SFM reconstruction result is 27,568, while our method is 344,201, which shows its advantage in observing the details of organs. Enlarging a certain area of the organ model, this method can still reflect its 3D characteristics, and retain the scale information, allowing 3D measurement. The SFM method not only loses scale information but also makes it difficult to identify organ features due to its sparseness.
which contained 413,076 points. The point cloud of the pig stomach is shown in Figure 9.

A 3D stitching method of the binocular endoscope is proposed in this study, which shows many advantages when compared with other conventional algorithms.

The reconstruction effect in real scenes may be affected by many factors, for example, the reflection of the gastric mucosa. To verify the applicability of the method, we tried 3D stitching on the fresh stomach of a pig. A total of 10 pairs of left and right images were generated by the moving binocular endoscope. After SGBM matching and point cloud preprocessing, we stitched a complete point cloud of pig stomach by ICP registration, which contained 413,076 points. The point cloud of the pig stomach is shown in Figure 9. The results show that this method also works well in real scenes.

4. Discussion

A 3D stitching method of the binocular endoscope is proposed in this study, which shows many advantages when compared with other conventional algorithms.

The binocular endoscope mainly realizes 3D reconstruction through a binocular matching algorithm such as SGM and SGBM, which are only based on a single field of view [20]. In this research, ICP is used to splice 3D point clouds from different positions to form complete organ models. Thus, the endoscopic field of view is expanded. For example, our method can realize the 3D construction of the whole model in Figure 7c, while conventional algorithms can only realize it in a red box. At the same time, a lot of preprocessing is operated to reduce the computation, including outlier removal, point cloud down-sampling and transformation matrix calculation by feature points matching.

Figure 8. Comparison of density with other methods. (a) Spare result by SFM. (b) Dense result with scale information by our method.

Figure 9. 3D stitching of pig stomach.
Compared with the conventional 3D reconstruction algorithm SFM [21], the point clouds generated by our method are much denser and retain scale information. Point cloud with scale information shows advantages in 3D measurement [22]. Therefore, our method has great potential for clinical application, for example, measurement of polyps, which is of great significance for diagnosis.

There are also some studies on the 3D reconstruction of organs based on structured light [5]. However, structured light devices are complex and difficult to be used in a clinic setting.

Currently, this method has only been applied to models and pig stomachs. Due to the complexity of the human body structure, the clinical application of this method still needs further verification.

5. Conclusions

In this study, a novel approach for 3D stitching of binocular endoscopic images was carried out. The point cloud of different sites is obtained through binocular matching of SGBM, and the transformation matrix is calculated through feature detection and matching of adjacent sites, which is a multi-point cloud ICP registration that reduces the amount of calculation and improves accuracy. The 3D spliced endoscopic image not only expands the field of view but also retains scale information, which is conducive to reflecting the real scene of the diseased part, facilitating the operation of the doctors. The 3D splicing method was also applied to the reconstruction of a pig stomach to verify the applicability of the method. Further investigation is needed to fully understand the capabilities and limitations of the new approach.

Author Contributions: Conceptualization, L.W.; methodology, C.Z.; software, H.Y.; resources, L.W. and Q.Y.; data curation, C.Z. and H.Y.; writing—original draft preparation, C.Z.; writing—review and editing, H.Y. and B.Y.; supervision, L.W. and B.Y.; funding acquisition, L.W., B.Y. and Q.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China, grant number (No. 2019YFC0119502), the Key Research and Development Program of Zhejiang Province of China (No. 2018C03064), the Zhejiang Provincial Natural Science Foundation of China (No. LGF20F050006) and the National Natural Science Foundation of China (No. 61735017, 61822510).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The Zhejiang Lab is acknowledged.

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

References

1. Sorensen, S.M.; Savran, M.M.; Konge, L.; Bjerrum, F. Three-dimensional versus two-dimensional vision in laparoscopy: A systematic review. Surg. Endosc. 2016, 30, 11–23. [CrossRef] [PubMed]
2. Arezzo, A.; Vettoretto, N.; Francis, N.K.; Bonino, M.A.; Curtis, N.J.; Amparore, D.; Arolfo, S.; Barberio, M.; Boni, L.; Brodie, R.; et al. The use of 3D laparoscopic imaging systems in surgery: EAES consensus development conference 2018. Surg. Endosc. 2019, 33, 3251–3274. [CrossRef] [PubMed]
3. Liao, H.; Tsuzuki, M.; Mochizuki, T.; Bonino, M.A.; Curtis, N.J.; Amparore, D.; Arolfo, S.; Barberio, M.; Boni, L.; Brodie, R.; et al. Fast image mapping of endoscopic image stitchings with three-dimensional ultrasound image for intrauterine fetal surgery. Minim. Invasive. Ther. Allied. Technol. 2009, 18, 332–340. [CrossRef] [PubMed]
4. Cai, H.; Wang, R.; Li, Y.; Yang, X.; Cui, Y. Role of 3D reconstruction in the evaluation of patients with lower segment oesophageal cancer. J. Thorac. Dis. 2018, 10, 3940–3947. [CrossRef] [PubMed]
5. Sui, C.; Wu, J.; Wang, Z.; Ma, G.; Liu, Y.-H. A Real-Time 3D Laparoscopic Imaging System: Design, Method, and Validation. IEEE Trans. Biomed. Eng. 2020, 67, 2683–2695. [CrossRef] [PubMed]
6. Zhou, J.; Han, S.; Zheng, Y.; Wu, Z.; Liang, Q.; Yang, Y. Three-dimensional Reconstruction of Retinal Vessels Based on Binocular Vision. Chin. J. Med. Instrum. 2020, 44, 13–19.
7. Crandall, D.J.; Owens, A.; Snavely, N.; Huttenlocher, D.P. SfM with MRFs: Discrete-Continuous Optimization for Large-Scale Structure from Motion. *IEEE Trans. Pattern Anal. Mach. Intell.* 2012, 35, 2841–2853. [CrossRef] [PubMed]

8. Pomerleau, F.; Colas, F.; Siegwart, R. A Review of Point Cloud Registration Algorithms for Mobile Robotics. *Found. Trends Robot.* 2015, 4, 1–104. [CrossRef]

9. Mahmoud, N.; Collins, T.; Hostettle, A.; Soler, L.; Doignon, C.; Montiel, J.M.M. Live Tracking and Dense Reconstruction for Handheld Monocular Endoscopy. *IEEE Trans. Med. Imaging* 2019, 38, 79–89. [CrossRef]

10. Widya, A.R.; Monno, Y.; Okutomi, M.; Suzuki, S.; Gotoda, T.; Miki, K. Whole Stomach 3D Reconstruction and Frame Localization from Monocular Endoscope Video. *IEEE J. Transl. Eng. Health Med.* 2019, 7, 1–10. [CrossRef] [PubMed]

11. Tsai, C.-Y.; Wang, C.-W.; Wang, W.-Y. Design and implementation of a RANSAC RGB-D mapping algorithm for multi-view point cloud registration. In Proceedings of the 2013 CACS International Automatic Control Conference (CACS), Nantou, Taiwan, 2–4 December 2013; pp. 367–370. [CrossRef]

12. Sharp, G.; Lee, S.; Wehe, D. ICP registration using invariant features. *IEEE Trans. Pattern Anal. Mach. Intell.* 2002, 24, 90–102. [CrossRef]

13. Zhang, Z. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.* 2000, 22, 1330–1334. [CrossRef]

14. Zhang, H.; An, L.; Zhang, Q.; Guo, Y.; Song, X.; Gao, Q. SGBM Algorithm and BM Algorithm Analysis and Research. *Geomat. Spat. Inf. Technol.* 2016, 39, 214–216.

15. Wang, C.; Oda, M.; Hayashi, Y.; Villard, B.; Kitasaka, T.; Takabatake, H.; Mori, M.; Honma, H.; Natori, H.; Mori, K. A visual SLAM-based bronchoscope tracking scheme for bronchoscopic navigation. *Int. J. Comput. Assist. Radiol. Surg.* 2020, 15, 1619–1630. [CrossRef] [PubMed]

16. Karami, E.; Prasad, S.; Shehata, M. Image Matching Using SIFT, SURF, BRIEF and ORB: Performance Comparison for Distorted Images. *arXiv* 2017, arXiv:1710.02726.

17. Bay, H.; Tuytelaars, T.; Gool, L.V. SURF: Speeded up robust features. In *European Conference on Computer Vision*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 404–417.

18. Granshaw, S.I. Bundle Adjustment Methods in Engineering Photogrammetry. *Photogramm. Rec.* 2006, 10, 181–207. [CrossRef]

19. Aldoma, A.; Marton, Z.-C.; Tombari, F.; Wohlkinger, W.; Potthast, C.; Zeisl, B.; Rusu, R.B.; Gedikli, S.; Vincze, M. Tutorial: Point Cloud Library: Three-Dimensional Object Recognition and 6 DOF Pose Estimation. *IEEE Robot. Autom. Mag.* 2012, 19, 80–91. [CrossRef]

20. Wang, D.; Liu, H.; Cheng, X. A Miniature Binocular Endoscope with Local Feature Matching and Stereo Matching for 3D Measurement and 3D Reconstruction. *Sensors* 2018, 18, 2243. [CrossRef] [PubMed]

21. Penteck, Q.; Hein, S.; Miernik, A.; Reiterer, A. Image-based 3D surface approximation of the bladder using structure-from-motion for enhanced cystoscopy based on phantom data. *Biomed. Tech. Eng.* 2017, 63, 461–466. [CrossRef] [PubMed]

22. Sakata, S.; McIvor, F.; Klein, K.; Stevenson, A.R.L.; Hewett, D.G. Measurement of polyp size at colonoscopy: A proof-of-concept simulation study to address technology bias. *Gut* 2016, 67, 206–208. [CrossRef] [PubMed]